A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers

Lauridsen, Carsten Ammitzbøl

Publication date:
2012

Document Version
Pre-print: Det originale manuskript indsendt til udgiveren. Artiklen har endnu ikke gennemgået peer-review (fagfællesbedømmelse) og redigering.

Link to publication

Citation for published version (APA):
# Index

Preface and Acknowledgements..............................................................................2  
List of papers in the PhD ......................................................................................4  
List of Abbreviations...............................................................................................5  
Introduction.............................................................................................................7  
Aim..........................................................................................................................13  
Material and Methods.............................................................................................14  
  Study I....................................................................................................................14  
  Study II..................................................................................................................15  
  Study III................................................................................................................15  
  Study Population (Study II and III)......................................................................23  
Results....................................................................................................................24  
Discussion.................................................................................................................27  
Limitations of the studies.......................................................................................32  
Perspectives and Conclusion..................................................................................34  
Figures.....................................................................................................................36  
References...............................................................................................................41  
Summary in English.................................................................................................45  
Summary in Danish..................................................................................................50  
First Paper...............................................................................................................55  
Second Paper..........................................................................................................79  
Third Paper.............................................................................................................104
Preface and Acknowledgements

This PhD thesis is based on three studies performed at the Departments of Diagnostic Radiology at Herlev and Odense University Hospitals from 2008 to 2011.

The PhD project was granted by Metropolitan University College (DK), University College Nordjylland (DK), Odense University Hospital (DK), Copenhagen University Hospital Herlev (DK) and the Danish Association of Radiographers.

First of all I would like to thank my professional advisor Philippe Lefere MD for all the support that he has given me throughout the course of this project. Without his radiological insight, experience and support, the study of this thesis would have been difficult or impossible to carry out. Thanks for your patience and dedication to the research.

Second of all I would like to thank my main advisor Jens Karstoft MD, PhD, MPO for his valuable comments and assistance with manuscript editing.

I also want to thank the radiographers - Bo Holm Andersen, Jeanette Overgaard Ege, Mette Andersen Stove, Inge Hansen, Anne Seierøe and Henriette Raaschou, who went through the training course with great enthusiasm and achieved great skills in interpreting CTCs.
Also thanks to Randi Brinckmann Wiencke (Dean at the Metropolitan University College) for encouraging me to start as a PhD student and for the support through the whole study period.

Furthermore I would like to thank Professor Dr. Med. Henrik S. Thomsen, Henrik Nørgaard MD, Sven Adamsen MD, Stefaan Grysperdt MD, Karsten Nielsen Dr. Med., Cecile Skaarup MS, and Steven Hageman MD for their worthy advice on study design and manuscript editing.

I also wish to thank statistician Dr. Oke Gerke for his patience and for assistance with statistical analysis.

Also thanks to my colleagues at the radiography education at Metropolitan University College for their support and inspiration.

I also wish to thank all the patients participating in the study despite the difficulties and stressful situation that some of the patients were under due to their disease.

Also thanks to the staff at the Departments of Surgical Gastroenterology and the staff at the Departments of Diagnostic Radiology at Herlev and Odense University Hospitals.

Finally I would like to thank Helene Iversen MD, for her eternal support, love and cooking during the project period. Thanks to my two sons Oscar and Oliver for supporting me on their own way and giving me perspective in life.
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen

List of papers in the PhD

I. Lauridsen C, Lefere P, Gerke O, Gryspeerdt S.

II. Lauridsen C, Lefere P, Gerke O, Hageman S, Gryspeerdt S.
   Comparison of the Diagnostic performance of CT colonography interpreted by Radiologists and Radiographers.

III. Lauridsen C, Lefere P, Gerke O, Gryspeerdt S.
   Analysis of the false negative and false positive findings in CTC interpreted by radiographers and radiologists.

Supervisors:

Jens Karstoft, associate professor MD, Phd, MPO.
Head of radiology department at Odense University Hospital
Dk 5000 Odense, Denmark.

Philippe Lefere MD.
VCTC- Virtual Colonoscopy Teaching Centre
Akkerstraat 32c
B-8830 Hoogelee
Belgium

List of Abbreviations

OC: Optical Colonoscopy
CTC: Computed Tomography Colonoscopy
MDCT: Multi Detector Computed Tomography
MRC: Magnetic Resonance Colonography
DCBE: Double Contrast Barium Enema
FS: Flexible Sigmoidoscopy
FOBT: Fecal Occult Blood Test
RS: Rectoscopy
2D: Two dimensional
3D: Three dimensional
CRC: Colorectal Cancer
CAD: Computed Aided Detection
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen

SOS: Satisfaction of search (Satisfaction of a target in a radiological search can reduce the detectability of a second target).

PPV: Positive predictive value

NPV: Negative predictive value

CI: Confidence interval
Introduction

Colorectal Cancer and current diagnostic Modalities

Cancer in the colon and rectum (CRC) is the third most frequently occurring type of cancer in Denmark, and CRC is the second leading cause of cancer related death [1].

The estimated number of new cases and deaths due to CRC were 4026 and 1835 respectively in 2007 [2,3]. In comparison to other European countries, Denmark has the third highest and the second highest incidence of CRC for women and men respectively [4].

Compared to the other Nordic countries, Denmark has the second-highest incidence of CRC after Norway, and the mortality from CRC in Denmark is the highest in the Nordic countries [5].

Most CRC`s are adenocarcinomas and arise from benign adenomatous polyps that develop [3] slowly over 10-20 years [6-9], and the malignant transformation is related to polyp size. The risk of malignancy is approximately 1 % for adenomas smaller than 10 mm, however the risk for adenomas larger than 10 mm increases to 15 % and for adenomas larger than 20 mm it increases up to 40 % for transforming into a malignancy within 10 years [9,10].

At present time several diagnostic modalities are used for colonic evaluation including RS, FS, OC, FOBT, DCBE, MRC and CTC. However OC is considered to
be the gold standard due to a high diagnostic accuracy and the option of therapeutic intervention. Although OC is an excellent examination, it still has several disadvantages in form of serious complications, incomplete procedures and the need for sedation and post procedural monitoring.

Introduction to CT-Colonography

CTC has attracted multidisciplinary attention as a minimally invasive structural evaluation of the entire colon and rectum for the detection of polyps and cancers.

CTC represents a modified CT-examination in a patient who has undergone bowel preparation and colonic distension, in which the images are interpreted using advanced 2D and 3D display techniques. The examination was first introduced in 1994 by Vinning [11], and since the introduction there have been rapid advancements in CTC-technology. MDCT now permits image acquisition of thin 1- to 2-mm slices of the entire large intestine well within breath-hold imaging times. Computer imaging graphics allow for visualization of 3D endoscopic flight paths through the inside of the colon, which are simultaneously viewed with interactive multiplanar 2D images. The integrated use of the 3D and 2D techniques allows for ease of polyp detection as well as characterization of lesion density and location.

There have been several meta-analyses of CTC accuracy [12-15] (table 1) which have analyzed studies which included low-prevalence subjects and
increased prevalence subjects or symptomatic patients. These meta-analyses showed 85–93 % sensitivity on a per-patient basis for larger polyps and specificity of 95 % or greater. There has certainly been variability of accuracy among several individual studies (table 1). Two studies (Cotton et al. and Rockey et al.) that have been widely quoted [16,17], cast doubts on the accuracy of CTC, reporting poor per-patient sensitivity, even for lesions greater than 10 mm in size. The methodology of both of these studies has been criticized, particularly, with regard to the Cotton trial, the lack of experience of the reporting radiologists.

However other studies showed better results. A landmark study, employing meticulous methodology, and one of the few large studies of CTC in asymptomatic average risk individuals, published by Pickhardt et al. [7], reported excellent accuracy for CTC, equal or better than OC. This group’s results have been attributed to state of the art scanning and software, using 3D as the primary read, and fecal tagging. Similarly, the published American College of Radiology Imaging Network (ACRIN) multicentre trial [18] reported excellent sensitivity for large adenomas and cancers. A further study by Kim et al. compared the diagnostic yield from parallel OC and CTC screening programs and found similar detection rates of advanced neoplasia in the two groups [19].

Interpretation
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen

Some studies have documented wide inter-reader variability between radiologists interpreting CTC-examinations [20]. One solution to prevent wide inter-reader variability could be using two radiologists to double-read every CTC-examination, as double-reading of CTC-examinations has been shown to improve the overall performance of CTC significantly [21]. However, double-reading by radiologists, increases interpretation time and thereby increases professional workload and total cost and creates significant logistical challenges to busy diagnostic imaging departments. The shortage of radiologists makes this an impractical long-term solution, and therefore the diagnostic imaging departments should ensure efficient work routines and assessment of the possibilities for assigning tasks to personnel groups with a shorter education [3].

For that reason it could be interesting to examine if radiographers could learn to interpret intraluminal findings at CTC with a high degree of diagnostic accuracy determining if they could play a role in interpreting CTC-examinations.

Until now, most research has focused on the technical capabilities of CTC; however, it is increasingly being realized that reader experience and training are equally important [22].

In terms of how teaching cases are selected and presented, there is a wide disparity [23], but there is an emerging consensus that there is a variable learning curve associated with interpreting CTC-findings. This implies that
interpretation performance improves as the number of interpreted cases increases [24].

At the time of writing, there is no evidence-based guidelines for training in interpreting CTC-examinations, although the literature suggest that the interpretation of at least 50 validated CTC-cases for trained readers, and 75 cases for novice readers may be required to achieve high levels of performance [22,25-27]. Particularly three studies [23,25,28] investigate the performance of radiographers as CTC-readers besides one [29].

Table 2 shows three cases where non-experienced radiographers in interpreting CTC-examinations started training in reporting CTC-examinations. It shows the number of training cases, number of radiographers involved, and sensitivity per-polyp for the clinically relevant polyp sizes. The studies had different training schemes, the patients underwent different preparation techniques and none of the studies included patients who had an OC and a CTC on the same day.

Indications for CTC

The most widely accepted clinical indication is incomplete or failed OC. An incomplete colonoscopic examination is defined as a failure to intubate the cecum. Incomplete OC may be the result of poor bowel preparation, and patient intolerance to the procedure, spasm, or colonic obstruction caused by a
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen

neoplastic or non-neoplastic stenosis. The CTC-examination can be performed on the same day directly after OC, and thus, no additional bowel preparation is needed [30]. CTC can complete the colonic evaluation in the majority of patients, which can also analyze the cause of incomplete endoscopy [31,32]. Patients with a history of incomplete OCs are at higher risk for failure of a second attempt. Therefore, CTC, rather than a second attempt at OC, may be recommended.

When unsuspected extracolonic findings such as aneurysms and other cancers are considered, CTC screening appears to dominate OC, being both more clinically efficacious and cost effective [33]. Beyond complications and cost issues there are a number of other limitations to OC screening. Postprocedural recovery time and a chaperone for transportation are required even for negative cases. A number of surveys indicate that many patients would prefer less invasive screening options [34-37].

Optical Colonoscopy

Optical Colonoscopy

OC is an invasive procedure using a colonoscope for direct visualization of the mucosal surface throughout the length of the colon to the cecum. The colonoscope is capable of air and water insufflations, irrigation, suction and passage of biopsy forceps and polypectomy snares and permits biopsy and removal of suspicious lesions during the examination. Patients must undergo
full oral bowel preparation including a liquid diet one day or more before the examination. Although not always required, intravenous sedation and pain control are standard to avoid discomfort of this invasive examination [10].

OC clearly represents the therapeutic gold standard for colorectal evaluation [16,38], but it has a number of important limitations as a primary screening tool. Chief among these are that OC represents the most invasive and expensive initial screening test [33]. OC with biopsy or polypectomy is associated with increased risk for complications but perforation may also occur during colonoscopies without biopsies [39].

Perforation at screening OC occurs in approximately 1 in 500 to 1000 cases (0.1-0.2 %) [19,39].

Compared to OC the perforation risk of the colon is approximately 0,005 % for CTCs [40-46].

Aim

The purpose of this PhD thesis was to investigate if radiographers without any CTC-experience could achieve a sufficient level of diagnostic performance in interpreting CTCs.

Furthermore the aim of this study was to examine the reviewer performance of trained radiographers in comparison with that of experienced radiologists with OC as the reference standard and to assess the pitfalls in CTCs interpreted by trained radiographers and experienced radiologists.
Material and Methods

Study I

The study went on from May 2008 to October 2009. A total of five radiographers participated in the study and they were all inexperienced with CTC-examinations. They participated in a training program interpreting CTC arranged by two expert radiologists with great experience in teaching CTC-interpretation. The training encompassed a 3-day workshop including interpretation of 50 cases with normal and pathologic shape of the colon. Furthermore, the training contained 75 CTC-cases from their local department. The cases were read by all the readers independently using CT-workstations at their local department. The examinations were sent to the expert readers whose interpretation was considered the golden standard. The radiographers received feedback from the expert readers via email and frequent group conferences. Moreover the radiographers went through a test of 20 CTC-examinations created by the training center. Only findings inside the colon were included in the interpretation of the examinations.
Study II

Study II was a prospective study running from September 2008 to November 2010. A total of 126 symptomatic patients referred to OC at two hospitals went through CTC and OC on the same day.

70 patients from hospital A and 56 patients from hospital B underwent a colonic preparation. The preparation was different at the two hospitals and was probably not adequate in all cases particularly at Hospital B.

Four radiographers trained in CTC and two experienced radiologists individually interpreted the cases. They were blinded to each others` findings and to the results from OC.

The OCs were performed by experienced gastroenterologists or by a gastroenterologist fellow under supervision.

The readers annotated their results in a document including a screendump of the lesion. Sensitivity, specificity, and positive and negative predictive values were calculated per-polyp and per-patient including 95 % confidence interval.

Study III

Study III was a prospective study running between September 2008 and November 2010. The patients (126) (same as in study II) were examined at
two centers and went through CTC and OC the same day. Four trained radiographers and two experienced radiologists in interpreting CTCs participated in the study.

At the two centers the patients underwent different preparation before the OC- and CTC-examinations. One of the radiologists used a different workstation than the five other readers for reading the cases. The number, locations, and the reasons of the false positive and false negative findings were calculated and analyzed.

**CT-Colonography protocol**

All the examinations were performed using a 64-channel multislice CT scanner (Hospital A, Brilliance Philips Medical Systems, The Netherlands; Hospital B Lightspeed, General Electric Medical Systems, France). Scans were obtained at 50 mAs (Hospital A) and 40 mAs (Hospital B) with 120 kV. Patients were examined in supine and prone positions with identical scanning parameters for both positions: collimation 64x0.625-slice thickness: 1 mm – increment: 1 mm-rotation time 0.5 seconds.

Image processing and interpretation at the Diagnostic Radiology department at hospital A were performed with the use of a CT-workstation (Extended Brilliance workspace 3.5, Philips, The Netherlands) provided with dedicated
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen

CTC-software and allowing 2D and 3D reading of the colon. This system was used by the radiographers and by one radiologist.
The other radiologist interpreted the examinations on a Vitrea workstation (Vital Images, Minnetonka, USA).
Due to local technical limitations of the workstation, simultaneous projection of the supine and prone acquisition allowing fast comparison between both acquisitions was impossible.

**Patient preparation for CTC**

In study II and III, the patients went through a bowel purgation including a low fiber diet, 2 L of polyethylene glycol solution and fecal tagging.
In these studies the patients received fecal tagging of different doses. At hospital A, the tagging encompassed 100 ml of ionic iodinated contrast (Gastrografin® 370 mgI/ml, Bracco Diagnostics, Princeton, USA) soluted in 400 ml of water self-administered by the patients the day before the CTC-examination.
At hospital B, the fecal tagging comprised 20ml of non-ionic iodinated contrast (Iomeron® 300 mgI/ml, Bracco Diagnostics) soluted in 200 ml of water managed by the patients themselves the day before the CTC.
Before the CTC-examination began, patients at hospital A were administered hyoscine butylbromide (Buscopan®, Boehringer Ingelheim, Germany) which
minimised the bowel peristalsis and colonic spasm. At hospital B there was no use of medicine for bowel relaxation.

At both hospitals, all patients underwent automated colonic distension using carbon dioxide using a CO₂ injector (Bracco PROTOCO 2 L, E-Z-EM TM).

**Optical Colonoscopy protocol**

In study II and III, OCs were performed with a standard endoscope (Olympus® CF – Q1; 160DL; Olympus Europe Ltd., Hamburg, Germany) by a senior gastroenterologist or by a gastroenterology fellow under guidance of experienced staff. Before the OCs, all patients were administered 2.5-7.5 mg of midazolam (Dormicum®, Roche, Basel, Schweiz) and 0.05-0.1 mg fentanyl (Fentanyl®-Janssen, Janssen Pharmaceuticals, Titusville New Jersey, USA). The size, morphologic shape, segmental location, and distance of the lesions from the anal margin were annotated in a protocol by the gastroenterologist who performed the examination.

When executing the OCs, the endoscopists were blinded to the CTC-findings. Discrepancies in the results of the lesion-matching were adjudicated by a third expert reader.

**Inclusion and Exclusion Criteria for patients**
Inclusion criteria for the training and test cases in study I was a completely performed CTC.

In study II and III, the inclusion criteria were referral for OC, age ≥18 years, and the capability of giving written or verbal agreement.

Exclusion criteria were inflammatory bowel disease, colostomy after colorectal surgery, colorectal biopsy performed within 72 hours, and/or polypectomy within two weeks prior to the CTC, and/or known pregnancy.

**Ethical Considerations**

In study II and III, all patients participated after oral and written informed consent and in accordance with the Helsinki-II declaration. The Institutional Review Board approved the study protocol under the following ID-number H-A-2007-0066.

**Data recording and databases**

Data collected from patient files and during CTCs and OCs were annotated on paper study forms and filled in a database (Microsoft Excel version 2007,
Microsoft Corporation, Redmond, Wash., USA), and analysed by using Stata/MP 11.1 (StataCorp, Texas 77845 USA). Paper files were assigned a consecutive ID number different from the patient`s CPR (Central Personal Register) number. The link was available to the author when the patient inclusion and the data registration were recorded. Study forms, CTC paper prints, and the computer with the database were kept in lockable rooms at the hospital during the whole study period.

**Statistical Considerations**

In study I five radiographers participated in the training and interpretation of CTCs. In study II and III, four radiographers and two radiologists interpreted the CTCs. Colorectal polyps were registered on both patient basis and polyp basis, stratified according to the respective size categories, i.e. polyps ≥6 mm and ≥10 mm, respectively. Carcinomas were included in the calculation and analysed as polyps, but were described and discussed separately as well.

**Definitions and classifications**
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen

To be considered a match between findings by the radiographers and the radiologists and OC, a given polyp had to be assessed as appearing within the same segment or in adjacent segments. Also the two recorded diameters had to be the same within a 50 % margin of error.

*True Positive (TP)*
A patient was considered to be a true positive if the same patient with at least one true positive polyp also identified by OC was detected correctly by the reader.
A polyp was considered to be a true positive if the same polyp also identified by OC in the respective size category were detected by the reader.

*True Negative(TN)*
A patient in who no polyps had been identified by the reader was considered to be true negative if in the same patient no polyps were detected by the OC.

*False Positive (FP)*
A patient was considered to be a false positive when at least one wrongly diagnosed polyp was detected by the reader and no true positive polyp was identified by OC.
A polyp was considered to be a false positive if the polyp in the respective size category was wrongly identified by the reader.
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen

False Negative (FN)
A patient was considered to be a false negative when at least one polyp was not detected by the reader and no true positive polyp was identified by the OC. A polyp was considered to be a false negative if the polyp in the respective size category was not detected by the reader.

When calculating the diagnostic performance of the readers, OC and an expert reader with an experience of ≥6000 examinations served as the standard of reference. Sensitivity, specificity, and PPV were evaluated by means of point estimates and respective 95 % confidence intervals (95 % CI) in study II and III.

Definitions for the Diagnostic Tests

Sensitivity:
The number of positives (polyps or patients with at least one polyp) identified by the reader relative to the number (polyps or patients with at least one polyp) detected by OC.

Specificity:
The number of patients with no polyps identified by the readers relative to the number of patients without polyps detected by OC.

**Predictive values:**

**PPV**

The number of patients with at least one true positive polyp among all patients with positive polyps correctly detected by the readers.

**NPV:**

The number of patients without polyps who were correctly detected by the readers.

**Study Population (Study II and III)**

A total number of two CTCs per hospital per week were pre-booked in the study period. Totally, 350 patients were scheduled for an OC and asked to participate in the study via mail. Of these, 192 (101 from hospital A and 91 from hospital B) did not wish to participate after receiving the first letter and information by telephone. Of the 158 patients who responded positively after pre-inclusion by telephone, 32 recalled their consent after having received a second letter including detailed information about CTC. Thus 126 patients went through CTC and signed a written consent document before the examination.
started (figure 1). Of the 126 patients, 39 (3 from hospital A and 36 from hospital B) were excluded by the expert reader due to deficient preparation or unsatisfactory distension of the colon. Finally, 87 CTCs were acceptable for interpretation.

Results

Study I

Five radiographers interpreted the training cases which consisted of 75 patients. Nine were excluded due to inadequate preparation or deficient distension of the colon.

In the training cases the radiographers achieved an overall per-polyp sensitivity of 57 % (95 % CI 46.1-67.9) and 69.1 % (95 % CI 50.6-87.5) for lesions ≥6 mm and ≥10 mm respectively.

Overall per-patient sensitivity, specificity and PPV were 86.4 % (95 % CI 76.7-96.1), 85.4 % (95 % CI 77-93.9) and 78.3 % (95 % CI 64.9-91.7), respectively.

Two colorectal tumors were seen by all the readers and were categorized as polyps ≥10 mm. Histology showed adenocarcinoma.

In the test cases overall per-polyp sensitivity was 80.7 % (95 % CI 69.5-92) and 94.7 % (95 % CI 85.6-100*) for lesions ≥6 mm and ≥10 mm, respectively. Overall per-patient sensitivity, specificity and PPV were 92.9 %
(95 % CI 83.1-100*), 64 % (95 % CI 13.1-100*) and 87.8 % (95 % CI 71.7-100*), respectively.

There was a statistically significant improvement in per-polyp sensitivity for lesions ≥6 mm in the test cases.

**Study II**

A total of 40 lesions were discovered in 22 patients and the lesions included four masses and 36 polyps (figure 1). Among the 40 lesions, 28 had a sessile morphology, 6 pedunculated, 2 flat and 4 were categorized as mass.

There were no clinically relevant complications due to OC or CTC.

For the radiographers, overall per-polyp sensitivity using bootstrapping was 60.3 % (95 % CI 50.3-70.3) and 60.7 % (95 % CI 42.2-79.2) for polyps ≥6 mm and ≥10 mm, respectively.

For the radiologists, overall per-polyp sensitivity was 59.2 % (95 % CI 46.4 – 72.0) and 69.0 % (95 % CI 48.1 – 89.6) for polyps ≥6 mm and ≥10 mm, respectively.

No statistically significant differences in the overall per-polyp sensitivity were found in detection rates between radiologists and radiographers. For polyps ≥10 mm there was a higher difference compared to polyps ≥6 mm.
Overall sensitivity per-patient with polyps ≥6 mm using bootstrapping was 76.2 % (95 % CI 61.4-91.0) and 76.2 % (95 % CI 61.7-90.6) for the radiographers and radiologists, respectively.

Overall specificity per-patient with polyps ≥6 mm using bootstrapping were 81.4 % (95 % CI 73.7-89.2) and 81.1 % (95 % CI 73.8-88.3) for the radiographers and the radiologists, respectively.

There was no statistically significant difference in the overall per-patient sensitivity between the radiographers and the radiologists.

**Study III**

There were six incomplete OCs. In these cases the CTCs were compared with the deficient OC examination, but only with the colon segments that had been examined with both technologies. The six incomplete OCs included two cases with stenosing masses, one case with a polyp in the ascending colon and three cases which showed no polyps.

All of the above-mentioned lesions were detected with CTC. One rectal mass 7.5 cm from the anal margin with the size of 17 mm was initially missed by the OC in the rectum. This mass was detected by five out of six readers. All six readers missed a 9 mm neuroendocrine tumor located at the ileo-cecal valve. For polyps ≥6 mm the false negative rate was 39.7 % and 40.8 % (table 1, study III) for the four radiographers and the two radiologists,
respectively. The radiographers had a 26.8 % higher false negative rate for polyps ≥10 mm than the radiologists (table 1, study III).

For both groups of readers, the most frequent location of the false negative lesions was in the left hand side of the colon (table 2, study III).

In total, 30.1 % of the missed lesions were categorized as multiple lesions in one patient for both the radiographers and the radiologists.

The number of false positive findings in total were 131 lesions for both groups of readers. The false positive rate was per-patient was similar for both the radiographers and the radiologists.

The most common reason for the false positive findings was mentioned as stool including 79.4 % of these lesions.

**Discussion**

In study I, we used a tele-training program to train five radiographers and they all achieved a good diagnostic performance in a test with statistically significant improvement for sensitivity compared to the training cases.

Overall sensitivity for polyps ≥10 mm was 94.7 % (95 % CI 85.6-100*) which is comparable to the criteria for the participating radiologists in the large Acrin [18] multicenter study including 2531 patients.
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen

Compared to other studies the training method used in study I showed good results. In two other studies by Jensch and Bodily [25,28] they achieved a per-polyp sensitivity for polyps ≥10 mm at 66 % and 74 % respectively.

With an overall PPV for polyps ≥6 mm at 87 % (95 % CI 71.7-100*) which indicates the rate of the false positive findings, the results from study I were good compared to the Acrin study [18] and the study by Jensch [28] with a PPV per-polyp ≥6 mm at 40 % and 37.5 %, respectively.

Two large multicenter studies (Pickhardt et al. (1233 patients) and Johnson et al. (2531 patients)) [7,18] and one smaller study by Graser et al. (307 patients) [47] including asymptomatic patients who had CTC and OC on the same day showed an excellent performance for CTC (table 1). Two large multicenter studies (Cotton et al. (614 patients) and Rockey et al (615 patients)) [16,17] also including symptomatic patients who had CTC and OC on the same day showed less impressive results (table 1).

This wide scatter in sensitivity was attributed to technical differences with regard to the preparation, different hardware used for data acquisition, different methods used for primary interpretation, different reader training, and different reader experience.

As for breast cancer screening, double-reading of CTC by two radiologists has shown to be an effective solution to improve inter-reader variability [21]. This solution, however, is costly and logistically impractical in daily clinical practice and legitimates exploration of alternative solutions.
As an alternative, radiographers trained in CTC could be considered. If sufficient experience of these radiographers is obtained and validated, their interpretation of CTC under supervision of a radiologist could be considered. In study II, we examined the diagnostic performance characteristics of CTC of trained radiographers and experienced radiologists in 87 consecutively enrolled symptomatic outpatients. The overall sensitivity per-patient with polyps ≥6 mm was exactly the same for both radiographers and radiologists (76.2 %) and this probably demonstrates the impact of the training of the radiographers in study I. The overall specificity per-patient with polyps ≥6 mm for the radiographers and the radiologists was 81.1 % and 81.4 % respectively (figure 4 study II).

In three multicenter trials (table 1), the overall per-patient sensitivity and specificity with polyps ≥6 mm were in a range from 78-89 % and 47-88 % respectively. Particularly the sensitivity was higher compared to the results in study II (figure 3 and 4 study II) and a probable reason could be failed preparation in some cases (figure 2 and 3).

In comparison with two other studies including trained radiographers and radiologists interpreting CTCs, study II showed comparable results. In the study by Bodily [25] the radiographers and the radiologists obtained an overall per-patient sensitivity and specificity at 70 % versus 84 % and 80 % versus 74 % respectively. In the study by Jensch et al. [28] two trained radiographers and two radiologists obtained a sensitivity and specificity per-patient with polyps ≥6 mm at 87 % versus 81 % and 67 % versus 71 %, respectively.
In study II the radiographers and the radiologists achieved exactly the same per-patient sensitivity (76.2 %). This probably confirms the efficacy of the training of the radiographers in our study.

For calculation of the inter-reader variability, the kappa value is the accepted statistic method, and in study II we calculated lower values of sensitivity per-patient inter-reader agreement between the two experienced radiologists at $K=0.42$ compared to the four radiographers at $K=0.69$. As described by Altman et al. [48] these kappa values are equivalent to moderate and good for the radiologists and the radiographers respectively.

The reason could be that the radiologists had different education in CTC compared to the radiographers who all went through the training mentioned in study I.

The results of the inter-reader agreement are similar to a study by Burling et al. [49] which showed an inter-reader agreement between the reference standard and CAD-assisted radiographers at $K=0.72$.

In study II and III 126 patients were examined in two hospitals and went through OC and CTC the same day. The two hospitals had dissimilar procedures for preparation. At hospital B, they did not use suppressant medicine for bowel relaxation and a lesser amount of fecal tagging material. This method can probably explain the exclusion of 36 out of 56 patients from hospital B.

In study III, the false negative rate per-polyp was calculated (table 1, study III) and the results were favourable compared to a study by Doshi et al [50]
and the study by Rockey et al. [17], which demonstrated a false negative rate per-polyp at 48.7 % and 51 % for lesions ≥6 mm. In study III, the false negative rate for polyps ≥6 mm was 39.7 % and 40.8 % for the radiographers and the radiologists, respectively. The most frequent reason for missed lesions in study III was SOS and showed a proportion at 32.3 % and 25.8 % (table 2, study III) for the radiographers and the radiologists, respectively. The difference between the two groups of readers is probably due to the longer clinical experience of the radiologists. In total, the majority (55.9 %, table 2, study III) of the false negative lesions in study III were located in the left colon (rectum, sigmoid colon, and the descending colon). These results are comparable to a study by Arnesen et al. [51] which showed the same tendency. The definition the false positive lesions were assessed by a third expert reader as earlier mentioned and the results from study III showed that 79.4 % of the false positive lesions ≥6 mm were due to residual stool for both groups of readers. These results are higher compared to the study by Arnesen et al. which showed that 42 % of the false positive lesions ≥5 mm were due to residual stool. The large amount of false positive lesions in study III probably demonstrates the efficacy of the insufficient preparation, and the results underscore the importance of optimal patient preparation and rigorous technique when performing CTC. Although the preparation at the two hospitals in study II and III was different, there was only a limited difference between the false positive rate for hospital A (16.7 %) and hospital B (13.5 %).
At hospital B, 36 patients were excluded and only 20 patients were included and at hospital B, 56 patients were included and only 3 patients were excluded (flow chart study III). The false positive rate would probably have shown a larger difference between the two hospitals, if the 36 patients had not been excluded at hospital B.

For every size category, the radiographers as a group had more false positive findings than the radiologists (table 3, study III). This result is similar to the study by Jensch et al. [28].

Limitations of the studies

There are several limitations of the studies in this PhD thesis. In study I the test of the radiographers only contained 20 test-cases including 27 polyps ≥6 mm. A larger number of cases in the test would probably show more accurate results. However the same number of cases was also used for testing the participants of the ACRIN study [18]. This means that they were probably more watchful for polyp detection compared to a screening setting with low disease prevalence. Furthermore we used clinical cases for CTC-training. A predefined set of CTC-examinations would assumably enhance the assessment of progress in
performance in the training. This was not possible in this study, as the last 20 training cases in study I, were presented with only two polyps.

Study II and III were performed at two centers and therefore the preparation protocols were not identical.

The diagnostic performance by the readers is probably impacted by these different conditions.

An example of the dissimilar preparation protocols was the lack of intra-venous injection of hyoscine butylbromide (Buscopan®, Boehringer Ingelheim, Germany) in study II and III.

Another factor that apparently had an unfavourable influence on the results in study II and III could be the insufficient stool and fluid tagging and residual material that could hamper the interpretation.

Particularly the sensitivity and specificity per-polyp showed low values probably due to the inadequate preparation in some cases.

In study II, two radiologists interpreted the CTCs, but they were not tested before the study started and the question could be if the radiologists were experienced enough and if other radiologists could have achieved better results.

Another limitation in study II and III is due to the fact that the preparation method was not evidence-based and has not been proven to be successful in other studies. Other preparation protocols could be recommended e.g. the preparation used in the study by Liedenbaum et al. [52,53].
Perspectives and Conclusion

The results achieved in study I confirms the efficacy of the training method used. Furthermore the results show that radiographers with sufficient training are able to achieve an adequate level of diagnostic performance in interpreting CTCs.

CTC has been shown to be sufficiently accurate in detecting colorectal neoplasia, and is now considered a valid alternative for CRC screening in the general population at average risk for Colorectal Cancer (CRC) [52]. Due to these results, widespread implementation of CTC in many radiology departments could be expected. However frequently efficient implementation is hampered because of time constraints with many radiologists lacking the time to learn the technique or even to interpret the cases during their busy daily practice. This would certainly be the case if the amount of CTCs would increase e.g. with screening for CRC. Hence, an alternative based on CTC interpretation by trained radiographers as reviewers under radiologist supervision may be deployed [23,25,28,54]. The results from study II suggest that dedicated radiographers trained in interpretation of CTCs can achieve accuracy comparable with that of experienced radiologists in the evaluation of CTC. The results in this study also show that the diagnostic performance can still be improved with further experience and better techniques. This finding is of particular interest in double-interpretation screening of CRC using
radiographers and radiologists. Screening for CRC in Denmark will start in 2014 and the results of this study could probably be a helpful tool for choosing a sufficient design for the screening model.

In conclusion, the results of study III showed that the main reason for the false positive findings was misinterpretation of stool. For the false negative lesions the most frequently reason was multiple lesions in one patient (SOS). These results are useful in the future planning of training readers in interpreting CTCs.

Study II and III included six incomplete OCs and two of them contained two stenosing tumors which were seen at CTC. Two other masses were detected by OC and CTC respectively, and these results confirms that CTC in the combination with OC is a very useful tool to detect CRC.
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen

Figures

Table 1

Per patient accuracy of CTC for lesions 10 mm or larger: meta-analyses and multicenter studies including asymptomatic patients.

<table>
<thead>
<tr>
<th>Author</th>
<th>Study type</th>
<th>Sensitivity (%) ≥10 mm</th>
<th>Specificity (%) ≥10 mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sosna [12] et al.</td>
<td>Meta-analysis</td>
<td>88 (CI 84-93)</td>
<td>95 (CI 94-97)</td>
</tr>
<tr>
<td>Mulhall [14] et al.</td>
<td>Meta-analysis</td>
<td>85 (CI 79-91)</td>
<td>97 (CI 96-97)</td>
</tr>
<tr>
<td>Rosman [15] et al.</td>
<td>Meta-analysis</td>
<td>82 (CI 76-88)</td>
<td>96 (CI 94-98)</td>
</tr>
<tr>
<td>Author</td>
<td>Study type</td>
<td>Sensitivity (%) ≥6 mm</td>
<td>Specificity (%) ≥6 mm</td>
</tr>
<tr>
<td>-----------------</td>
<td>------------</td>
<td>------------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Pickhardt [7] et al.</td>
<td>Multicenter</td>
<td>89 (CI 82.9-93.1)</td>
<td>80 (CI 77-82.0)</td>
</tr>
<tr>
<td>Johnson [18] et al.</td>
<td>Multicenter</td>
<td>78 (CI 71-85)</td>
<td>88 (CI 84-92)</td>
</tr>
<tr>
<td>Graser* [47] et al.</td>
<td>Multicenter</td>
<td>84 (CI 76-90.3)</td>
<td>47 (CI 40.2-54.7)</td>
</tr>
<tr>
<td>Rockey [17] et al.</td>
<td>Multicenter</td>
<td>55 (CI 47-63)</td>
<td>89 (CI 86-92)</td>
</tr>
<tr>
<td>Cotton [16] et al.</td>
<td>Multicenter</td>
<td>39 (CI 29.6-48.4)</td>
<td>91 (CI 89.9-93.1)</td>
</tr>
</tbody>
</table>

*Sensitivity and specificity for patients with polyps ≥9 mm
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen

Table 2. Studies in which radiographers interpreted CTCs

<table>
<thead>
<tr>
<th>Study</th>
<th>Number of Training cases</th>
<th>Number of Radiographers</th>
<th>Sensitivity per-polyp (%) for lesions ≥6 mm ≤9 mm</th>
<th>Sensitivity per-polyp (%) for lesions ≥1 mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jensch et al. [28]</td>
<td>20</td>
<td>2</td>
<td>65</td>
<td>66</td>
</tr>
<tr>
<td>Bodily et al. *[25]</td>
<td>50</td>
<td>2</td>
<td>45</td>
<td>74</td>
</tr>
<tr>
<td>Esgar study [23]</td>
<td>50</td>
<td>10**</td>
<td>*** 63.5</td>
<td></td>
</tr>
</tbody>
</table>

* Including 5 medical students and 2 radiographers

** From different centers

*** Sensitivity for all lesions ≥6 mm
Figure 1 – Flowchart – Showing the number patients included in the study

- Invited by mail: 350
  - 101 not included
  - 177 from Hospital A
  - 173 from Hospital B
  - 91 not included

- 158 patients included after telephone interview

- Written information sent by mail

- 16 excluded from hospital A
  - Recall of consent

- 16 excluded from hospital B
  - Recall of consent

- A total: 126 patients included in the study
Figure 2. Showing an example of the insufficient preparation. False positive polyp in the transverse colon. Misinterpreted as a sessile polyp at the size of 9.2 mm. Supine position.

a: ´intermediate´ window setting

b: ´abdominal´ window setting
Figure 3 – showing an example of the insufficient fecal tagging. False positive polyp in the descending colon. Misinterpreted as a sessile polyp at the size of 7.1 mm. Prone position.
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen

References


A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen


A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen


A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen


Summary in English

A comparison of the diagnostic performance in CT-Colonography interpreted by experienced radiologists and trained radiographers.

Introduction

Computed tomographic colonography (CTC) has been proposed as an alternative to optical colonoscopy (OC) for detecting colorectal polyps and cancer. CTC is appealing because it does not involve conscious sedation, and recovery time associated with OC. CTC represents a modified CT examination in a patient who has undergone bowel preparation and colonic distension, in which the images are then interpreted using advanced 2D and 3D display techniques.

Purpose

To investigate the diagnostic performance of four trained radiographers in comparison with that of two experienced radiologists in the evaluation of CTC-examinations with those of the reference standard (OC).
Materials and Methods

This PhD thesis is mainly based on three studies and to accomplish the purpose of the thesis it was essential to create an educational platform for the radiographers. The first study (study I) included an assessment of the diagnostic performance of four radiographers who underwent training in CTC using a tele-training program generally based on the interpretation of 75 training cases. Subsequently, the radiographers went through a test of 20 CTC-examinations.

In the second study (study II) 126 consecutive patients were examined at two hospitals and went through same-day CTC and OC.

The four trained radiographers and the two experienced radiologist interpreted the CTCs and were blinded to all clinical findings and each other’s findings. The OCs were performed by experienced gastroenterologists and the results were annotated in the study protocol. Sensitivity, specificity, and positive and negative predictive values from the CTC-interpretation were calculated per-polyp and per-patient including 95 % confidence interval.

The aim of the third study (study III) was to identify the pitfalls in CTC through analyses of false positive and false negative findings on CTC.

The readers and the study population were the same as mentioned in study II. The number, locations, and the reasons of the false positive and false negative findings were calculated and analyzed.
Results

In the training cases in study I the radiographers achieved an overall per-polyp sensitivity of 57 % (95 % CI 46.1-67.9) and 69.1 % (95 % CI 50.6-87.5) for lesions ≥6 mm and ≥10 mm respectively. Overall per-patient sensitivity, specificity and PPV were 86.4 % (95 % CI 76.7-96.1), 85.4 % (95 % CI 77-93.9) and 78.3 % (95 % CI 64.9-91.7), respectively.

In the test cases overall per-polyp sensitivity was 80.7 % (95 % CI 69.5-92) and 94.7 % (95 % CI 85.6-100*) for lesions ≥6 mm and ≥10 mm, respectively. Overall per-patient sensitivity, specificity and PPV were 92.9 % (95 % CI 83.1-100*), 64 % (95 % CI 13.1-100*) and 87.8 % (95 % CI 71.7-100*), respectively.

There was a statistically significant improvement in per-polyp sensitivity for lesions ≥6 mm in the test cases.

In study II, a total of 40 lesions were detected in 22 patients and the lesions included four masses and 36 polyps. Among the 40 lesions, 28 had a sessile morphology, 6 pedunculated, 2 flat and 4 were categorized as mass.

For the radiographers, overall per-polyp sensitivity using bootstrapping was 60.3 % (95 % CI 50.3-70.3) and 60.7 % (95 % CI 42.2-79.2) for polyps ≥6 mm and ≥10 mm, respectively.
For the radiologists, overall per-polyp sensitivity was 59.2 % (95 % CI 46.4-72.0) and 69.0 % (95 % CI 48.1-89.6) for polyps ≥6 mm and ≥10 mm, respectively.

Overall sensitivity per-patient with polyps ≥6 mm using bootstrapping was 76.2 % (95 % CI 61.4-91.0) and 76.2 % (95 % CI 61.7-90.6) for the radiographers and radiologists, respectively.

Overall specificity per-patient with polyps ≥6 mm using bootstrapping were 81.4 % (95 % CI 73.7-89.2) and 81.1 % (95 % CI 73.8-88.3) for the radiographers and the radiologists, respectively. There was no statistically significant difference in the overall per-patient sensitivity between the radiographers and the radiologists.

In study II and III, 39 patients (three from hospital A and 36 from hospital B) were excluded because of inadequate preparation or insufficient distension.

There were six incomplete OCs. In these cases, the CTCs were compared with the deficient OC-examination, but only with the colon segments that had been examined with both technologies. The six incomplete OCs included two cases with stenosing masses, one case with a polyp in the ascending colon and three cases which showed no polyps.

For the four radiographers and the two radiologists the false negative rate was 39.7 % versus 40.8 % and 39.3 % versus 31.0 % for polyps ≥6 mm and ≥10 mm, respectively.

For both groups of readers, the most frequent location of the false negative lesions was in the left hand side of the colon. The most frequent reason for the
false negative findings was categorized as multiple lesions in one patient for both the radiographers and the radiologists. The false positive rate per-patient was 18.6 % and 18.9 % for the radiographers and the radiologists, respectively. The most common reason for the false positive findings was mentioned as stool including 79.4 % of these lesions.

**Conclusion**

In conclusion, the results of this thesis suggest that dedicated radiographers trained in interpretation of CTC-examinations can achieve diagnostic accuracy comparable with that of experienced radiologists in the evaluation of CTC. The results in this study also show that the diagnostic performance can still be improved with further experience and better techniques. This finding is of particular interest in double-interpretation screening for CRC.
Summary in Danish

En sammenligning af den diagnostiske nøjagtighed for radiograf og radiolog bedømte CT-kolografier med anvendelse af koloskopier som referencestandard.

Indledning

Computer Tomografisk Kolografi (CT-K) er en billeddiagnostisk undersøgelse der simulerer en koloskopi, og den kan anvendes til påvisning af kolorektale polypper som kan være et forstadium til kolorektal cancer (KRC). CT-K kræver udrensning som ved en koloskopi, men scanningstiden er kortvarig og sedation er ikke nødvendig, og dermed er patientens ubehag ved CT-K mindre end ved koloskopier. Ved CT-K benyttes tre-dimensionelle intraluminale billeder af colon, rekonstrueret af to-dimensionelle billeder fra en Multislice CT-Scanner (MSCT).

Formål

At sammenligne den diagnostiske nøjagtighed ved henholdsvis radiograf og radiologbedømte CT-K undersøgelser med koloskopier som referencestandard.
Materiale og metoder

Denne PhD afhandling er baseret på tre studier, og i forhold til formålet for afhandlingen, var det afgørende at etablere en uddannelses platform for radiografer der udfører og beskriver CT-K undersøgelser. Det første studie (studie I) indeholdt en vurdering af resultaterne fra fem radiografer, som gennemgik en uddannelse i at beskrive CT-K undersøgelser. Uddannelsen indeholdt beskrivelse af 75 CT-K "trænings" undersøgelser med en efterfølgende test indeholdende 20 undersøgelser.

I det andet studie (studie II) blev i alt 126 patienter undersøgt ved anvendelse af både CT-K og koloskopier. Fire CT-K uddannede radiografer og to erfarne radiologer udførte beskrivelser af CT-K undersøgelserne, og de var alle blændet for hinandens resultater, samt for kliniske fund. Koloskopierne blev udført af gastroenterologer der noterede resultaterne i forsøgsprotokollen. Sensitivitet, specificitet og positive og negative prædiktive værdier blev beregnet per-patient og per-polyp med anvendelse af 95 % konfidensinterval. Formålet med det tredje studie (studie III) var at identificere og analysere falsk positive og falsk negative fund ved beskrivelser af CT-K-undersøgelser. Beskriverne og patienterne var de samme som i studie II. Antal, placering, og årsager af/til de falsk positive og falsk negative fund blev beregnet og analyseret.
**Resultater**

Ved de 75 "trænings-undersøgelser" i studie I opnåede radiograferne en samlet per-polyp sensitivitet på henholdsvis 57 % (95 % CI 46,1-67,9) og 69,1 % (95 % CI 50,6-87,5) for læsioner ≥6 mm og ≥10 mm. Den samlede per-patient sensitivitet, specificitet og PPV var henholdsvis 86,4 % (95 % CI 76,7-96,1), 85,4 % (95 % CI 77 til 93,9) og 78,3 % (95 % CI 64,9-91,7).

Ved testen opnåede radiograferne en samlet per-polyp sensitivitet på henholdsvis 80,7 % (95 % CI 69,5 til 92) og 94,7 % (95 % CI 85,6 til 100 *) for læsioner ≥6 mm og ≥10 mm. Den samlede per-patient sensitivitet, specificitet og PPV var henholdsvis 92,9 % (95 % CI 83,1 til 100 *), 64 % (95 % CI 13,1 til 100 *) og 87,8 % (95 % CI 71,7 til 100 *). Der var en statistisk signifikant forbedring i sensitivitet per-polyp for læsioner ≥6 mm ved testen i forhold til "trænings" undersøgelserne. I studie II indgik 22 patienter med ialt 40 læsioner indeholdende fire tumorer og 36 polypper. Blandt de 40 læsioner, var der 28 brede, 6 stilkede, 2 flade 2 og 4 tumorer. Den samlede per-polyp sensitivitet for radiograferne inklusive bootstrapping var henholdsvis 60,3 % (95 % CI 50,3-70,3) og 60,7 % (95 % CI 42,2-79,2) for polypper ≥6 mm og ≥10 mm. For radiologerne, var den samlede per-polyp sensitivitet henholdsvis 59,2 % (95 % CI 46,4-72,0) og 69,0 % (95 % CI 48,1-89,6) for polypper ≥6 mm og ≥10 mm. Sensitiviteten per-patient med polypper ≥6 mm var for henholdsvis radiografer og radiologer 76,2 % (95 % CI 61,4-91,0) og 76,2 % (95 % CI 61,7-90,6) inklusiv bootstrapping. Specificitet per-patient med
polypper ≥6 mm var for henholdsvis radiografer og radiologer 81,4 % (95 % CI 73,7-89,2) og 81,1 % (95 % CI 73,8-88,3) inklusiv bootstrapping. Der var ikke statistisk signifikant forskel i forhold til den samlede per-patient sensitivitet mellem radiografer og radiologer. I undersøgelse II og III blev 39 patienter (tre fra sygehus A og 36 fra hospital B) ekskluderet på grund af utilstrækkelig udrensning eller ikke optimal luft insufflation af colon. Der var seks imkomplette koloskopier. Disse undersøgelser blev sammenlignet med CK-K, men kun med de colon segmenter, der var blevet undersøgt med anvendelse af begge teknologier. De seks ufuldstændige koloskopier omfattede 2 stenoserende tumorer, en undersøgelse med en polyp i colon ascendens og tre undersøgelser, som ikke indeholdt nogen læsioner.
Falsk negativ rater for de fire radiografer og to radiologer var henholdsvis 39,7 % versus 40,8 % og 39,3 % versus 31,0 % for polypper ≥6 mm og ≥10 mm.
For både radiografer og radiologer, var den hyppigste placering af falsk negative læsioner i venstre side af colon. Den hyppigst årsag til falsk negative resultater blev kategoriseret som "flere læsioner hos en patient" for både radiografer og radiologer.
Falsk positiv rater per-patient for radiografer og radiologer var henholdsvis 18,6 % og 18,9 %. Den hyppigste (79,4 %) årsag til falske positive fund var resisterende afføring.
Konklusion

Resultaterne af denne PhD afhandling, tyder på at dedikerede radiografer der er uddannet til at beskrive CT-K undersøgelser, kan opnå en diagnostisk præcision der er sammenlignelig med erfarne radiologer. Resultaterne viser også, at kvaliteten af den diagnostiske nøjagtighed kan forbedres yderligere med anvendelse bedre teknikker i forhold til forberedelse af patienten. I forhold til screening for kolo-rectal cancer er anvendelse ´dobbelt-reader ´(radiografer/radiologer) af særlig interesse.
First Paper

Effect of a Tele-Training Program on Radiographers in the Interpretation of CT-Colonography.

By:

Carsten Lauridsen¹, Philippe Lefere², Oke Gerke³, Stefaan Gyspeerdt²

From:
¹Metropolitan University College, Radiography Education, Sigurdsgade 26, 2200 Kbh. N, Denmark.
²VCTC-Virtual Colonoscopy Teaching Centre, Akkerstraat 32c, B-8830 Hooglede, Belgium.
³Department of Nuclear Medicine, Odense University Hospital, Sdr. Boulevard 29, 5000 Odense C, Denmark.

Corresponding author:
Carsten Lauridsen, Lecturer, Metropolitan University College, Sigurdsgade 26, DK-2200 Copenhagen
Tel.: +45 7228 2833.
E-mail addresses: Cala@phmetropol.dk (C. Lauridsen)
Introduction

CT colonography (CTC) has now reached a high performance level in detecting tumoral lesions in the colon. Indeed, several large trials could demonstrate very good results of polyp detection [1-4]. These good results are only obtained if state-of-the-art technique and interpretation are applied by experienced teams being adequately trained in CTC. All experts [5-7] agree that this adequacy level is hard to reach, as CTC has a long and steep learning curve. This was demonstrated in the large trial performed by Rockey et al. [8]. Indeed, a review of this trial showed that technical failure and observer perceptual errors were by far the major cause of false negative results [9].

Whereas training in the practical aspects seems straightforward, only needing some short theoretical training and practical demonstration, acquiring sufficient experience and confidence in CTC interpretation is a rather arduous task. To date numerous efforts are being made to build up an efficient CTC training programme.

Furthermore, despite good sensitivity for polyp detection using state-of-the-art CTC technique, the recent ACRIN 6664 trial [3] was inflicted with positive predictive values as low as 40% and 23% for polyps $> 6$ mm and $> 10$ mm, respectively, indicative of a high false positive rate. Hence, improving CTC technique and interpretation through adequate training is deemed mandatory.
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen

The combination of interpreting 50-75 cases validated by optical colonoscopy (OC) with the participation in a hands-on CTC workshop is considered a solid basis for starting to read CTC exams[10,11]

It is generally accepted [12] that the training process can be improved by augmenting the interaction with and feedback from the teachers. This type of mentored training has been proposed in the ACR Colon Cancer Committee’s white paper [10].

Tele-training or tele-medicine, which is now widely available and easy to apply, could be a promising companion in achieving these goals [13].

Likewise, this premise makes the establishment of a structured CTC education a priority and yet induces additional problems.

Population-based training for colorectal cancer would encompass a large number of radiologists participating in the screening process possibly increasing their workload to unacceptable levels.

It may be supposed that a reduction of radiologist hours and costs could probably be accomplished by working with a team of radiographers supervised by one radiologist. With a sufficient level of experience for the radiographers, the interpretation time for the radiologist can be reduced significantly with the ultimate goal being the radiologist only checking the findings of the radiographers. This would allow supervision and validation of a substantial daily number of examinations per radiologist, reducing the total number of staff radiologists needed for population-based screening.
Keeping this in mind the purpose of this study was to assess the efficacy of radiographers interpreting CTC after a tele-radiology based training.

**Materials and methods**

The study was started in May 2008 and ended in October 2009. An IRB approval was not needed according to the Committees on Biomedical Research Ethics in Denmark.

The study was granted by Metropolitan University College (DK), University College Nordjylland (DK), Odense University Hospital (DK), Copenhagen University Hospital Herlev (DK) and the Danish Association of Radiographers.

Radiographers and tutors

Five radiographers (1 male, 4 females) participated in this study on a voluntary basis. Their age varied between 30 and 59 (mean age 50 years). They had no experience with CTC, and only very basic experience with colonic anatomy and pathology. They had practical experience with numerous abdominal CT and Barium enema examinations. CTC training of the radiographers was performed by two expert radiologists with an experience of > 6000 CTC (with > 800 validated by optical colonoscopy) and with great experience with the organisation of CTC-workshops. This training exclusively focused on the colon and did not consider the extracolonic structures.
Training

The training programme consisted of different stages:

Workshop

A 3-day workshop was organized in the teaching centre introducing the radiographers to CTC and to the normal and pathological conditions in the colon. The workshop consisted of a mix of theoretical presentations dealing with all technical and interpretational aspects of CTC and hands-on sessions. In total, 50 CTC cases were interpreted.

Reading Material

After the workshop, a book on CTC was administered to the radiographers to refine the knowledge and experience obtained during the workshop. A CTC course was also available on the website of the teaching centre. This web-course focussed on CTC imaging characteristics from normal to pitfalls. The web-course was concluded by a large Q/A section.
Training cases

75 randomly selected cases performed after incomplete colonoscopy in the local department were used for further training. The cases were interpreted by the radiographers at a rate of four cases/two weeks in a period of 34 weeks (holidays excluded) between July 2008 and June 2009. The use of CAD (Computer Aided Detection) was not allowed. Colorectal polyps ≥ 6 mm were reported and classified in two size categories (≥ 6 mm and ≥ 10 mm). Tumours were included in the calculations and analysed as polyps, but were described separately as well. The C-RADs classification was used [14]. All observers read the examinations independently and were blinded to all clinical findings, the colonoscopic results and each other’s findings.

Image processing and interpretation in the novice department were performed with the use of a CT-workstation (Extended Brilliance workspace 3.5, Philips, The Netherlands) provided with dedicated CTC software and allowing primary two-dimensional and three dimensional reading of the colon.

During this tele-training, all observers continued their normal professional activity and interpreted the examinations during their spare time.

Due to technical limitations of the workstation, simultaneous projection of the supine and prone acquisition, allowing fast comparison between both acquisitions, was impossible.
Polyps were measured with electronic calipers on the two dimensional view and recorded according to the segment (cecum, ascending colon, transverse colon, descending colon, sigmoid colon or rectum)

Per polyp detected, the radiographers annotated the segmental location, the size, the attenuation, the slice numbers per acquisition, and the distance to the anal margin of the polyp in a report including a snap shot per polyp.

The anonymized examinations were sent in DICOM format to the teaching centre using secured ultrafast lines for interpretation by both experts. The experts interpreted the cases on a Vitrea workstation (Vital Images, Minnetonka, USA) and made a report based upon consensus reading. The results from the two experts were considered as the gold standard except for parts of the colon where optical colonoscopy (OC) was performed.

All CTCs were performed after incomplete OC. The reports from the radiographers were corrected by one of the experts and returned with comments. There was an active communication via e-mail for answering questions concerning CTC and the examinations at regular intervals, group discussions were performed via tele-conference. The use of CAD was not allowed.

Test cases

Finally, the radiographers underwent a test of 20 cases validated by optical colonoscopy. This test was composed by the teaching centre. These cases were
sent in DICOM format to the novice department using secured ultrafast lines allowing for interpretation of the cases on the local workstation.

The test consisted of five normal cases and 15 cases with colonic polyps. There were a total of 27 polyps $\geq 6$ mm with 12 and 15 polyps 6-9 mm and $\geq 10$ mm, respectively. According to the C-RADs classification, 10 polyps had a sessile morphology, 11 were pedunculated, three were flat polyps, and three were masses with malignant characteristics. There was one lipoma.

The cases were interpreted by the radiographers at a rate of four cases/two weeks in a period of 10 weeks (holidays excluded) between August 2009 and October 2009. The use of CAD was not allowed.

The interpretation and reporting was performed the same way as the training cases mentioned above.

The outcome measure of the test was to achieve a per-polyp sensitivity per radiographer at 80% for polyps $\geq 6$mm.

Statistical analysis

Sensitivity, specificity, and PPV were evaluated by means of point estimates and respective 95% confidence intervals (95% CI). The results were given both on patient basis and on polyp basis, stratified according to the respective size categories, i.e. polyps $\geq 6$ mm and $\geq 10$ mm, respectively. Assuming a prevalence of 33% and a true (but unknown) sensitivity on per patient-basis of 0.85, 75 patients included in the study were sufficient for an expected width of
a 95% Wilson-score confidence interval of 0.27. The success criterion for the single reader is an estimated per patient sensitivity of at least 0.8. For per polyp-based analyses, bootstrapping [15] was applied as supplementary sensitivity analysis (results not shown here). For polyp-based analyses as well as for average reader analyses on patient-basis, linear regression models were used with the constant term as only explanatory variable and clustered sandwich estimators of variance to allow for intra group correlation (due to several polyps in the same patient) Confidence intervals which emanate from these linear regression models are Wald-type confidence intervals. These may exceed the boundaries of 0% or 100%, and were therefore truncated and indexed with an asterisk where required. Patient-based analyses by reader were carried out by means of 95% CI based on the Wilson-score method [16]. Group comparisons were performed by comparing the respective 95% confidence intervals, hence, significance level was 5%.

All Results were kept on a worksheet (Microsoft Excel version 2007, Microsoft Corporation, Redmond, Wash., USA) and analyzed by using Stata 11 (StataCorp, Texas 77845 USA).

**Results**

Training cases

The training cases presented 69 polyps \( \geq 6 \) mm, with 47 and 22 polyps 6-9 mm and \( \geq 10 \) mm, respectively. Nine cases were excluded and considered as
inadequate exams due to insufficient distension or incomplete preparation. The polyps were detected in 25 of 66 patients (39.4%).

Two colorectal carcinomas were detected and were categorized as polyps $\geq 10$ mm. Overall per-polyp sensitivity was 57.0% (95% CI 46.1-67.9) and 69.1% (95% CI 50.6 – 87.5) for polyps $\geq 6$ mm and $\geq 10$ mm, respectively.

Individual per-polyp sensitivity with 95% CI is shown in figure 1 and ranged between 54.7% and 61.7% and between 61.1% and 78.9% for polyps $\geq 6$ mm and $\geq 10$ mm, respectively.

The overall per-polyp sensitivity analysis for polyps $\geq 6$mm and $\geq 10$ mm using bootstrapping was 56.3% (95% CI 37.2 – 76.1) and 69.1% (95% CI 33.3 – 100), respectively.

Overall per-patient sensitivity was 86.4% (95% CI 76.7 – 96.1) for polyps $\geq 6$mm. Individual per patient sensitivity with 95% CI is shown in figure 2a and ranged between 76.0% and 92.0% for polyps $\geq 6$mm.

Overall per-patient specificity was 85.4% (95% CI 77.0 – 93.9) for polyps $\geq 6$mm. Individual specificity with 95% CI is shown in figure 2b and ranged between 75.6% and 90.2%.

Overall per-patient PPV was 78.3% (95% CI 64.9 - 91.7) for polyps $\geq 6$mm. Individual PPV with 95% CI is shown in figure 3 and ranged between 68.8% and 84.6%.
Test cases

Overall per-polyp sensitivity was 80.7% (95% CI 69.5-92.0) and 94.7% (95% CI 85.6 -100*) for polyps $> 6$ mm and $> 10$ mm, respectively. Individual per-polyp sensitivity with 95% CI is shown in figure 4 and ranged between 77.8% and 85.2% and between 93.3% and 100% for polyps $> 6$ mm and $> 10$ mm, respectively.

Compared to the training cases, there was a statistically significant improvement of sensitivity for polyps $> 6$ mm in the test cases, since the respective confidence intervals did not overlap (figure 1 and 4). The overall per-polyp sensitivity analysis for polyps $> 6$ mm and $> 10$ mm using bootstrapping was 86.4% (95% CI 73.5 – 96.9) and 95.8% (95% CI 87.0-100) respectively. The bootstrapping analysis of the data from the training cases for per-polyp sensitivity did not show any important difference.

Overall per-patient sensitivity was 92.9% (95% CI 83.1-100*) for polyps $> 6$mm. Individual per-patient sensitivity with 95% CI is shown in figure 5a and ranged between 92.9% and 100%.

No significant difference in sensitivity was observed on a per-patient basis between the training cases and the test cases (figure 2a and 5a).

Overall specificity was 64.0% (95% CI 13.1-100*) for polyps $> 6$mm. Individual specificity with 95% CI is shown in figure 5b and ranged between 40% and 100%.
The overall specificity was higher in the training cases compared to the test cases (figure 2b and 5b).

This difference of specificity was probably due to the fact that the radiographers were more focused on finding polyps in the test cases, and due the low number of negative cases in the test.

Overall per-patient PPV was 87.8% (95% CI 71.7 – 100*) for polyps ≥ 6mm. Individual per-patient PPV with 95% CI for polyps is shown in figure 6 and ranged between 81.3% and 100%.

This high level of PPV on a per patient basis indicated a low number of false positives in the test cases compared to the training cases, however the difference is not statistically significant (figure 3 and 6)

All readers detected the three colorectal cancers.

Discussion

Using a tele-training programme, five radiographers were successfully educated in CTC. After initial training, good sensitivity and PPV for polyps ≥ 6mm were obtained in a test with statistically significant improvement for sensitivity compared to the training cases. The training focused on both FN and FP findings. This was achieved by combining a workshop, reading material and interpretation of 75 cases. During this training period the radiographers could rely on the mentorship of two experienced CTC readers.
Compared to other studies this training method obtained good results. Jensch et al. [17] trained two radiographers with 20 cases with feedback. In 145 patients per-polyp sensitivity was 65% and 66% for polyps \( \geq 6 \) mm and \( \geq 10 \) mm, respectively. PPV was 37.5% and 59.5% for polyps \( \geq 6 \) mm and \( \geq 10 \) mm, respectively. In the ESGAR CTC study [18], 10 radiographers without any previous experience were trained in 50 cases. In a test of 40 cases, per-polyp sensitivity was 63.5% after exclusion of six cases, where it was difficult to detect lesions. In another study, Burling et al. [19] obtained a per-polyp sensitivity of 76% for polyps \( \geq 6 \) mm with five radiographers. PPV for cancer was 74%. Bodily et al. [20] trained two radiographers with teaching files. One file provided the basic imaging characteristics of colonic lesions and imaging pitfalls. The second file consisted of interpreting 50 CTC cases. At the end they were tested on 56 cases. This test was repeated after six weeks of performing second reads. After this period of additional training, results improved from 61% to 79.5% for polyps \( \geq 5 \) mm. These results are comparable to the results of the present study and probably confirm the efficacy of this education method.

Training in detecting polyps in a well distended colon could appear straightforward. Indeed, once a basic level of experience is reached and sufficient knowledge concerning pathologic conditions in the colon has been acquired, one would expect an adequate observer performance. However, several studies have shown that a basic experience does not guarantee performance in polyp detection [21,22]. Firstly, numerous errors of
interpretation causing FN are made [23,24]. Secondly, erroneous characterization of luminal defects may result in disappointing PPV. This was demonstrated in the ACRIN trial [3] with a PPV of 40% and 23% for polyps $> 6$ mm and $> 10$ mm, respectively. Interpretation errors are not only caused by lacking experience in interpretation. Frequently, technical inadequacy is a major source of error as was the case in the Rockey trial with 26% of important lesions missed due to technical shortcomings in colonic preparation and distension [8]. From this it has been learned that state-of-the-art application of CTC is mandatory [25]. This concept of good CTC technique was learned when interpreting the 75 training cases. We preferred not to use a predefined teaching file of 75 cases, as was done by Dachman et al [6], but deliberately chose cases coming out of clinical practice performed in the novice centre. With nine out of 75 examinations considered inadequate for interpretation, we were able to show that state-of-the-art application of CTC technique needs special attention during CTC implementation in a novice center. This approach also added a real time aspect to the training as the teachers were faced with unexpected problems enabling the novices to assess how they solved particular problems.

This training method based on mentored supervision with tele-training had some advantages. It allowed for continuous guidance during the training period and could be considered a virtual fellowship. The radiographers could compare their findings with the reports of the experienced CTC-readers. Feedback via e-mail and/or tele-conferences provided a continuous source of information. In
that way, the radiographers not only learned the importance of good CTC technique, they were most of all faced with their own interpretational errors. During the group meetings each case was reviewed and an explanation was provided when a lesion was overlooked or incorrectly interpreted. This type of mentored supervision has been considered important for CTC-training in the White Paper on CTC published by the ACR colon cancer committee [10]. Tele-training also allowed the radiographers to integrate the training in normal daily activity and learning the technique on their own workstation. Finally, it can be expected that this mentored training would also be helpful for radiologists.

To our knowledge this is the first study assessing the effect of a mentored training performed with tele-education over several months. This method could be improved by organizing webinars for learning the basic CTC principles. The use of tele-radiology also allowed for fast and efficient communication with the experts and allowed the radiographers to learn the technique on their own workstation.

What was the rationale for educating radiographers? A team of radiographers under supervision of one radiologist could be helpful when screening for colorectal cancer, as this approach would reduce radiologist time and the procedural costs. Furthermore, it would also allow for double reading which has been proven to increase performance [10].

This training method had limitations. Firstly, the radiographers were tested on only 20 cases with a total of 27 polyps ≥ 6 mm. However the same number of cases was also used for testing the participants at the ACRIN trial [26].
This means that they were more alert for polyp detection compared to a screening setting with low disease prevalence. Together with a low number of normal cases, this probably explains the low specificity. Secondly, the gold standard for the training cases was the report after consensus reading by two CTC experts. This could have resulted in some false negative and false positive findings. However, this avoided the necessity of validating each CTC by optical colonoscopy and hence examining patients twice. Furthermore, the test cases had optical colonoscopy as ground truth. As mentioned above, we used clinical cases for CTC-training. A predefined set of CTC examinations would enable improved monitoring with assessment of progress in performance. This was not possible in this study, as the last 20 cases were presented with only two polyps. Finally, the training was performed over a considerable time span of 34 weeks. This long period is probably not favourable for a focused training and could have got a detrimental effect on the final results.

These results cannot be generalized because of the low number of radiographers educated. Further studies are required to prove this.

In conclusion, this training in CTC based on tele-training proved successful and could be a useful method in training radiographers in CTC.
Figure 1

Per Polyp Sensitivity for polyps ≥ 6 mm/Training cases

Per Polyp Sensitivity for polyps ≥ 10 mm/Training cases
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen

Figur 2

Per patient sensitivity for polyps ≥ 6 mm/Training cases

<table>
<thead>
<tr>
<th>Percentage score</th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>R4</th>
<th>R5</th>
<th>R1 - R5</th>
</tr>
</thead>
<tbody>
<tr>
<td>100,00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95,00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90,00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>85,00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80,00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>75,00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70,00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65,00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60,00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55,00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50,00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Per patient specificity for polyps ≥ 6 mm/Training cases

<table>
<thead>
<tr>
<th>Percentage score</th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>R4</th>
<th>R5</th>
<th>R1-R5</th>
</tr>
</thead>
<tbody>
<tr>
<td>100,00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95,00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90,00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>85,00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80,00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>75,00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70,00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65,00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60,00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55,00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen

**Figure 3**

**PPV Per Patient for polyps ≥ 6 mm/Training cases**

<table>
<thead>
<tr>
<th>Percentage score</th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>R4</th>
<th>R5</th>
<th>R1-R5</th>
</tr>
</thead>
<tbody>
<tr>
<td>100.00</td>
<td>93.9</td>
<td>90.8</td>
<td>92.1</td>
<td>89.8</td>
<td>91.7</td>
<td></td>
</tr>
<tr>
<td>95.00</td>
<td>84.6</td>
<td>79.2</td>
<td>82.1</td>
<td>82.0</td>
<td>78.3</td>
<td></td>
</tr>
<tr>
<td>90.00</td>
<td>66.5</td>
<td>59.5</td>
<td>64.4</td>
<td>60.5</td>
<td>64.9</td>
<td></td>
</tr>
<tr>
<td>85.00</td>
<td>59.9</td>
<td>51.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R1 - R5
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen

Figure 4

Per Polyp Sensitivity for polyps ≥ 6 mm/Testcases

Per Polyp Sensitivity for polyps ≥ 10 mm/Testcases
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen

**Figure 5**

**Per patient Sensitivity for polyps ≥ 6 mm/Testcases**

- R1: 98.7%
- R2: 98.7%
- R3: 100.0%
- R4: 98.7%
- R5: 100.0%

**Per patient Specificity for polyps ≥ 6 mm/Testcases**

- R1: 68.5%
- R2: 68.5%
- R3: 68.5%
- R4: 68.5%
- R5: 60.1%

R1 – R5
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen

Figure 6

Per patient PPV for polyps ≥ 6 mm/Testcases

Percentage score

R1

R2

R3

R4

R5

R1 – R5
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen

References


A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen


Second Paper

Comparison of the Diagnostic Performance of CT Colonography

Interpreted by Radiologists and Radiographers.

By:

Carsten Lauridsen¹, Philippe Lefere², Oke Gerke³, Stefaan Gryspeerdt², Steven Hageman², Jens Karstoft⁴

From:

¹Metropolitan University College, Radiography Education, Sigurdsgade 26, 2200 Kbh. N, Denmark.

²VCTC-Virtual Colonoscopy Teaching Centre, Akkerstraat 32c, B-8830 Hooglede, Belgium.

³Department of Nuclear Medicine, Odense University Hospital, Sdr. Boulevard 29, 5000 Odense C, Denmark.

⁴Department of radiology, Odense University Hospital, Sdr. Boulevard 29, 5000 Odense C, Denmark.

Corresponding author:

Carsten Lauridsen, Lecturer, Metropolitan University College, Sigurdsgade 26, DK-2200 Copenhagen
Tel.: +45 7228 2833.
E-mail addresses: Cala@phmetropol.dk (C. Lauridsen)
Introduction

CT-colonography (CTC), also referred to as virtual colonoscopy (VC), has attracted multidisciplinary attention as a minimally invasive and structural evaluation of the entire colon and rectum for the detection of colorectal neoplasia. Introduced for the first time in 1994 by Vining [1], rapid advancements in technology improved visualization of the colon. Multidetector CT (MDCT) now permits image acquisition of thin 1-2 mm slices of the entire large intestine well within breath-hold imaging times. Computer imaging graphics constantly refine three-dimensional (3D) visualization with endoscopic fly-throughs of the colon with simultaneous interactive depiction of multi-planar two-dimensional (2D) images. This integrated use of the 3D and 2D techniques improves polyp detection. CTC has been shown to be sufficiently accurate in detecting colorectal neoplasia, and is now considered a valid alternative for colorectal cancer (CRC) screening in the general population at average risk of getting Colorectal Cancer (CRC) [2]. Due to these results, widespread implementation of CTC in many radiology departments could be expected. However, frequently efficient implementation is hampered because of time constraints with many radiologists lacking the time to learn the technique or even to interpret the cases during their busy daily schedule. This would certainly be the case if the amount of CTC would increase e.g. with screening. Hence, an alternative based on CTC-interpretation by trained radiographers as reviewers under radiologist supervision may be deployed [3-
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen

6]. The aim of this study was to investigate the reviewer performance of four trained radiographers in comparison with that of two experienced radiologists in the evaluation of CTC-examinations of 90 patients by comparing the sensitivity, specificity, PPV, and NPV of CTC in polyp detection with the reference standard, optical colonoscopy (OC).

**Materials and Methods**

Study Design

The prospective study started in September 2008 and ended in November 2010 and the study protocol was approved by the Institutional Review Board (Videnskabsetisk Komité) in accordance with the Declaration of Helsinki. All patients provided written informed consent before participation in the study and signed an informed consent before the examination. The study was granted by Metropolitan University College (DK), University College Nordjylland (DK), Odense University Hospital (DK), Copenhagen University Hospital Herlev (DK) and the Danish Association of Radiographers. Four radiographers trained in CTC and two radiologists with an experience of more than 200 CTCs interpreted the 90 CTC-examinations. The radiographers were trained and tested in CTC previously in their competence of CTC-interpretation [7]. One of the radiologists was trained at a 2-day ESGAR-
workshop (European Society of Gastrointestinal Radiology) and had clinical experience of more than 200 CTCs. The other radiologist had the same training as the radiographers.

Study Population

A total of 126 consecutive symptomatic outpatients examined in two hospitals (70 from hospital A and 56 from hospital B) underwent same-day CTC and OC (74 men and 52 women, 35–90 years of age, mean (SD) 63 (11.0) years). CTC was performed on the patients immediately prior to OC.

Inclusion and Exclusion Criteria

Inclusion criteria were referral for OC, age ≥18 years, and the ability to give written and orally informed consent. Patients were excluded in case of inflammatory bowel disease, colostomy after colorectal surgery, colorectal biopsy performed within 72 hours, and/or polypectomy within two weeks prior to CTC, and/or known allergy with Buscopan® and pregnancy.
Examination Technique

All patients underwent a colonic preparation using a low-fibre diet, 2 L of polyethylene glycol electrolyte solution (Moviprep®; Norgine Limited, Mid Glamorgan, UK) and faecal tagging. In 70 patients (hospital A), faecal tagging was obtained with 100 ml of ionic iodinated contrast (Gastrografin® 370mgI/ml, Bracco Diagnostics, Princeton, USA) soluted in 400 ml of water and administered the day before their CTC. In 56 patients (hospital B), faecal tagging was obtained with 20ml of non-ionic iodinated contrast (Iomeron® 300 mgI/ml, Bracco Diagnostics, Princeton, USA) soluted in 200 ml of water and administered in the late afternoon the day before the examination.

In 70 patients (hospital A), 20 mg i.v. hyoscine butylbromide (Buscopan®, Boehringer Ingelheim, Germany) was used for bowel relaxation [8]. All patients underwent colonic insufflation with carbon dioxide using a CO2 injector (PROTOCO® 2 L, Bracco, Princeton, USA). At hospital B, there was no use of medicine for bowel relaxation.

All the examinations were performed using a 64-channel multislice CT scanner (Hospital A, Brilliance Philips Medical Systems, The Netherlands; Hospital B Lightspeed, General Electric Medical Systems, France).
Interpretation

All readers read the examinations independently and were blinded to all clinical findings, the results from OC, and each other’s findings. Image processing and interpretation in the novice department were performed with the use of a CT-workstation (Extended Brilliance workspace 3.5, Philips, The Netherlands) provided with dedicated CTC-software and allowing 2D and 3D reading of the colon. This system was used by the radiographers and by one radiologist. Due to local technical limitations of the workstation, simultaneous projection of the supine and prone acquisition, allowing fast comparison between both acquisitions, was impossible. The other radiologist interpreted the examinations on a Vitrea workstation (Vital Images, Minnetonka, USA).

Per-polyp detected, the readers annotated the segmental location, the size, the attenuation, the slice numbers per acquisition, and the distance to the anal margin of the polyp in a report including a screen dump of the polyp. Tumors were included in the calculations and analyzed as polyps but were described separately as well.

There were six incomplete OCs (6.7 %). In these cases the CTC-examination was compared with the deficient OC-examination, but only with the colon.
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers  
PhD Thesis by Carsten Lauridsen

segments that had been examined with both technologies. The extra-colonic findings were not considered in this study.

Colonoscopy Protocol

OCs were performed by an experienced staff member (gastroenterologist or gastrointestinal surgeon) or by a gastroenterology fellow under direct supervision of experienced staff using 165 cm colonoscopes (Olympus CF-Q1; 160DL; Olympus Europe Ltd., Hamburg, Germany). While performing the OCs, the endoscopist was unaware of the CTC-findings. Patients received 2.5-7.5 mg of midazolam (Dormicum®, Roche, Basel, Schweiz) and 0.05-0.1 mg of fentanyl (Fentanyl®-Janssen, Janssen Pharmaceuticals, Titusville New Jersey, USA) on request. The size, morphologic features, segmental location and the distance from the anal margin of the polyps were documented on a case record form by the endoscopist who performed the examination and by the attending research fellow. Polyp size was measured at endoscopy using open biopsy forceps.

According to the adopted segmental checking procedure, a lesion found at CTC was matched to a corresponding one found at OC if it was located in the same or adjacent colon segment and when its size differed by no more than 50 % [9]. Discrepancies in the results of the lesion-matching were adjudicated by a third expert reader.
Statistical Analyses

Sensitivity, specificity and positive and negative predictive values (PPV and NPV respectively) were assessed by means of point estimates and respective 95 % confidence intervals (95 % CI) on a per-patient basis. Moreover, sensitivity was analyzed on a per-polyp basis and stratified according to the respective size categories (polyps ≥6 mm as well as polyps ≥10 mm). Patient-based analyses per-reader were carried out by means of 95 % CI based on the Wilson-score method [10]. For average reader-analyses as well as for polyp-based analyses, linear regression models were used with the constant term as only explanatory variable and clustered sandwich estimators of variance to allow intra-group correlation. Bootstrapping [11] was applied in order to account for the correlated nature of the data when computing point estimates and 95 % CI. Group comparisons were performed by comparing the respective 95 % confidence intervals using a significance level of 5 %. Inter-reader agreement was assessed for both radiologists and radiographers using Cohen’s kappa [12] and Fleiss’ kappa [13], respectively. Supplementary 95 % CI were calculated using bootstrapping.

Assuming a prevalence of patients with colorectal neoplasia of 33 % and a true (but unknown) sensitivity on a per-patient-basis of 0.85, including 90 patients in the study was sufficient for an expected width of a 95 % Wilson-score confidence interval of 0.25. This precision was deemed appropriate for this exploratory study. All results were kept in a worksheet (Microsoft Excel version
2007, Microsoft Corporation, Redmond, Wash., USA), and analyzed by using Stata/MP 11.1 (StataCorp, Texas 77845 USA).

Results

There were a total of 40 polyps ≥6 mm with 24 and 16 polyps measuring 6–9 mm and ≥10 mm, respectively. The polyps were detected in 22 of 87 patients (25 %). A total of 39 patients (3 from hospital A and 36 from hospital B) were excluded by one of the expert readers because of inadequate preparation or insufficient distension. Four masses and 36 polyps were detected with 28, six and two having a sessile, pedunculated and flat morphology. The four masses included a 2.5 cm and a 3 cm lesion in the rectum, a malignant extracolonic lymphoma (82*87 mm) 70 cm from the anal margin obstructing the sigmoid and a tumor (9 mm) at the ileo-cecal valve (figure 7).

Sensitivity

The radiographers obtained an overall per-patient sensitivity (using bootstrapping) of 76.2 % (95 % CI 61.4–91.0) for patients with polyps ≥6 mm. Individual per-patient sensitivity with 95 % CI is shown in Fig. 3 and ranged between 71.4 % and 85.7 % for polyps ≥6 mm. The radiologists achieved an overall per-patient sensitivity (using bootstrapping) at 76.2 % (95
% CI 61.7-90.6) for patients with polyps ≥6 mm. Individual per-patient sensitivity with 95 % CI is shown in Fig. 3 and ranged between 66.7 % and 85.7 % for polyps ≥6 mm. The bootstrapping analysis of the data for the overall per-patient sensitivity for both the radiographers and the radiologists demonstrated no difference between the two groups. Figure 8 and 9 shows some examples of the insufficient preparation in some cases. The overall sensitivity per-patient inter-reader agreement between radiologists and radiographers separately showed moderate and good (Altman et al. [14]) kappa values at 0.42 (95 % CI: 0.23-0.60) and 0.69 (95 % CI 0.58-0.80) respectively.

The radiographers achieved an overall per-polyp sensitivity (using bootstrapping) at 60.3 % (95 % CI 50.3-70.3) and 60.7 % (95 % CI 42.2-79.2) for polyps ≥6 mm and ≥10 mm, respectively. Individual per-polyp sensitivity (using bootstrapping) with 95 % CI is shown in Fig. 1-2 and ranged between 53.8 % and 71.8 % and between 50.0 % and 71.4 % for polyps ≥6 mm and ≥10 mm, respectively.

The radiologists obtained an overall per-polyp sensitivity (using bootstrapping) of 59.2 % (95 % CI 46.4-72.0) and 69.0 % (95 % CI 48.1-89.6) for polyps ≥6 mm and ≥10 mm, respectively. Individual per-polyp sensitivity (using bootstrapping) with 95 % CI is shown in Fig. 1-2 and ranged between 51.3 % and 67.6 % and between 66.7 % and 71.4 % for polyps ≥6 mm and ≥10 mm, respectively.
There was no statistically significant difference in per-polyp sensitivity between the radiographers as a group and the radiologists as a group. For polyps ≥10 mm there was a larger difference compared to polyps ≥6 mm (Fig. 1-2).

Specificity

Overall per-patient specificity (using bootstrapping) for the radiographers was 81.4 % (95 % CI 73.7 – 89.2) for patients with polyps ≥6 mm. Individual specificity with 95 % CI is shown in Fig. 4 and ranged between 78.8 % and 83.3 %. The radiologists obtained an overall per-patient specificity using bootstrapping at 81.1 % (95 % CI 73.8 – 88.3) for patients with polyps ≥6 mm. Individual specificity with 95 % CI is shown in Fig. 4 and ranged between 74.2 % and 87.9 %.

Positive Predictive Value

The radiographers achieved an overall per-patient PPV of 56.6 % (95 % CI 40.1–73.2) for patients with polyps ≥6 mm. Individual PPV with 95 % CI is shown in Fig. 5 and ranged between 51.7 % and 60.0 %.

Overall per-patient PPV for the radiologists was 56.1 % (95 % CI 40.0–72.3) for polyps ≥6 mm. Individual PPV with 95 % CI is shown in Fig. 5 and ranged between 51.4 % and 63.6 %.
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen

Negative Predictive Value

Overall per-patient NPV for the radiographers was 91.5 % (95 % CI 85.2-97.8) for polyps ≥6 mm. Individual NPV with 95 % CI is shown in Fig. 6 and ranged between 89.7 % and 94.7 %. The radiologists obtained an overall per-patient NPV of 91.5 % (95 % CI 85.4 – 97.5) for polyps ≥6 mm. Individual NPV with 95 % CI is shown in Fig. 6 and ranged between 89.2 % and 94.2 %.

Discussion

CTC is now established as a reliable radiological test for detecting colonic neoplasia [9]. In some countries, radiographers are likely to play a useful role in its wider dissemination including conversion of current services from barium enema by decreasing the radiologists’ workload. Like other recent prospective CTC-studies [9,15-17], our study focused on polyps measuring 6 mm or more, since the prevalence of advanced histologic features in small polyps (i.e., <6 mm) is reportedly low [18]. As for breast cancer screening, double-reading of CTC by two radiologists has shown to be an effective solution to improve inter-reader variability [15]. This solution, however, is costly and logistically impractical in daily clinical practice and legitimates exploration of alternative solutions. As an alternative radiographers trained in CTC could be considered. If sufficient experience of these radiographers is obtained and validated, their interpretation of CTC under supervision of a radiologist could be considered. In
the present study, the diagnostic performance of trained radiographers was comparable to that of experienced radiologists interpreting CTC-examinations. We investigated the performance characteristics of CTC by trained radiographers and experienced radiologists in 87 consecutively enrolled symptomatic outpatients. No statistically significant differences were found in detection rates between radiologists and radiographers. We found detection rates for radiographers similar to those of experienced radiologists. The overall sensitivity (Fig. 3) per-patient with polyps ≥6 mm was 76.2 % for both radiographers and radiologists. The overall specificity per-patient with polyps ≥6 mm for radiographers and radiologists was 81.1 % and 81.4 % (Fig. 4) respectively. The overall per-patient sensitivity with polyps ≥6 mm in this study is lower than results reported in three large studies (table 1) on average risk individuals. A probable reason for these results could be the failed preparation in some cases (figure 8 and 9).

Compared to other studies including trained radiographers and radiologists interpreting CTC this present study demonstrated similar results. Bodily et al. [3] found that in a selected data set of 56 cases, two trained radiographers and 15 radiologists achieved a sensitivity and specificity per-patient with polyps ≥5 mm at 70 % versus 84 % and 80 % versus 74 % respectively. Per-polyp sensitivity for radiographers and radiologists was 79.5 % versus 71 % for polyps ≥5 mm and polyps ≥10 mm respectively. In the study by Jensch et al. [4] two trained radiographers and two radiologists (one experienced and one in training) evaluated 145 cases with a sensitivity and
specificity per-patient with polyps ≥6 mm at 87 % versus 81 % and 67 % versus 71 %, respectively. Per-polyp sensitivity for radiographers and radiologists was 65 % versus 71 % and 66 % versus 69 % for polyps ≥6 mm and ≥10 mm respectively. In our study, the overall per-polyp sensitivity for radiographers and radiologists was 60.3 % versus 59.2 % and 60.7 % versus 69 % for polyps ≥6 mm and ≥10 mm, respectively (Fig. 1). In the two mentioned studies by Jensch and Bodilly, there was no statistically significant difference in the diagnostic performance between radiographers and radiologists. Compared to our study there was a larger difference in terms of per-patient sensitivity between the two groups of readers. Our study showed exactly the same per-patient sensitivity at 76.2 % for both groups. This probably confirms the efficacy of the training of the radiographers in our study.

In our study, there was a difference between patient sensitivity and per-polyp sensitivity (≥ 6 mm) at approximately 16 % for both the radiographers and the radiologists. This difference is very variable as seen in the Acrin study [19] and the study by Graser et al.[20] showing a difference between the per-patient and per-polyp (≥ 6 mm) sensitivity at 8 % and 32.7 % respectively.

For calculation of inter-reader variability, the kappa value is the accepted statistic method, and in our study we have calculated lower values of sensitivity per-patient inter-reader agreement between the two experienced radiologists (0.42) compared to the four radiographers (0.69).
The reason for the difference between the two groups of readers could be that the radiologists had different training in CTC-reading, compared to the radiographers who all went through the aforementioned training. This result concurs with another study by Burling et al. [21], which showed an agreement between the reference standard (consensus between expert radiological review, colonoscopy data, and clinical follow-up) and Computer Aided Detection(CAD)-assisted radiographers demonstrating the kappa value at 0.72 (95 %CI 0.65-0.78). Some of the CTC-examinations performed by the radiographers were also evaluated by those radiographers. The radiographers, however, had no knowledge of previous findings, and because there was considerable time between the examination and the evaluation, we do not believe this factor influenced their results.

There were limitations to our study, though. The first limitation is that, because this was a two center study, CTC and preparation protocols were not uniform across participating centers. The performance characteristics found in our study are probably affected by these variable conditions. Since segmental unblinding was not used in defining the reference standard in our study it is likely that some reader false-positive findings are actually true-positive findings. A third factor that probably had a negative influence on our results could be the lack of sufficient stool and fluid tagging, poor patient compliance and residual material that could make interpretation difficult. Another explanation for the rather low results particularly the sensitivity and specificity
per-polyp could be that the preparation was not sufficient for some patients. This could have compromised the interpretation in some cases.

For a further evaluation of the results, an analysis of the pitfalls made from the radiographers and the radiologists in this study could be of great interest. However the results imply that deployment of radiographers as reviewers in CTC is acceptable. In screening, however, the huge number of patients markedly increases the workload of radiologists, and double interpretation by radiologists is probably not cost-effective in that situation. One could certainly ask the question if the participating radiologists were experienced enough, and if other radiologists could have achieved a better level of diagnostic performance.

In conclusion, the results of this study suggest that dedicated radiographers trained in interpretation of CTC-examinations can achieve accuracy comparable to that of experienced radiologists in the evaluation of CTC. The results in this study also show that the diagnostic performance can still be improved with further experience and better techniques. This finding is of particular interest in double-interpretation screening.
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers

PhD Thesis by Carsten Lauridsen

Fig 1 Sensitivity Per-Polyp ≥6 mm

Fig 2 Sensitivity Per-Polyp ≥10 mm
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers

PhD Thesis by Carsten Lauridsen

Fig 3 Sensitivity Per-Patient with polyps ≥6 mm

Fig 4 Specificity Per-Patient with polyps ≥6 mm
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers

PhD Thesis by Carsten Lauridsen

Fig 5 PPV Per-Patient with polyps ≥6 mm

Fig 6 NPV Per-Patient with polyps ≥6 mm
Figure 7. Tumor at the Ileo-cecal valve with the size of 9 mm.
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen

Figure 8 – showing an example of the insufficient preparation. False positive polyp in the transverse colon. Misinterpreted as a sessile polyp at the size of 9.2 mm. Supine position.

a: intermediate window setting

b: Abdominal window
Figure 9 – showing an example of the insufficient fecal tagging. False positive polyp in the descending colon. Misinterpreted as a sessile polyp at the size of 7.1 mm. Prone position.
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen

References


A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen


A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen


A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen

Third Paper

Analysis of the False Negative and False Positive Findings in CTC Interpreted by Radiographers and Radiologists.

By:
Carsten Lauridsen¹, Philippe Lefere², Oke Gerke³, Stefaan Gryspeerdt², Jens Karstoft⁴

From:
¹Metropolitan University College, Radiography Education, Sigurdsgade 26, 2200 Kbh. N, Denmark.
²VCTC-Virtual Colonoscopy Teaching Centre, Akkerstraat 32c, B-8830 Hooglede, Belgium.
³Department of Nuclear Medicine, Odense University Hospital, Sdr. Boulevard 29, 5000 Odense C, Denmark.
⁴Department of radiology, Odense University Hospital, Sdr. Boulevard 29, 5000 Odense C, Denmark.

Corresponding author:
Carsten Lauridsen, Lecturer, Metropolitan University College, Sigurdsgade 26, DK-2200 Copenhagen
Tel.: +45 7228 2833.
E-mail addresses: Cala@phmetropol.dk (C. Lauridsen)
Introduction

Computed tomographic colonography (CTC) has been proposed as an alternative to optical colonoscopy (OC) for detecting colorectal polyps and cancer [1]. CTC is appealing because it does not involve conscious sedation, and recovery time associated with OC. Until now published results have shown per-patient sensitivity and specificity of CTC for colorectal cancer (CRC) and its precursor the adenomatous polyp ≥6 mm in a range from 39-85% [2-6]. CTC is now a better examination than double contrast barium enema (DCBE) with results comparable to OC in major studies [4,5,7,8]. Good results depend on adequate bowel cleansing, with fecal tagging, colonic distension and interpretation by experienced readers. Technical adequacy and extensive training in CTC-reading have been shown to be crucial in obtaining good results [9]. By achieving state-of-the-art CTC-technique and interpretation by experienced CTC-readers many pitfalls in interpretation are avoided. Identification of pitfalls is an important key to enhance the validation of and to improve training in this field [10]. The aim of this study was to identify the pitfalls in CTC through analyses of false positive and false negative findings on CTC interpreted by four trained radiographers and two experienced radiologists.
Materials and Methods

Study Design

The study was started in September 2008 and ended in November 2010, and the study protocol was approved by the Institutional Review Board (Videnskabsetisk Komité) in accordance with the Declaration of Helsinki. All participants provided written informed consent before participation in the study, and signed an informed consent before the examination. The study was granted by Metropolitan University College (DK), University College Nordjylland (DK), Odense University Hospital (DK), Copenhagen University Hospital Herlev (DK) and the Danish Association of Radiographers. Four educated radiographers and two experienced radiologists with an experience of more than 200 CTCs interpreted the 90 CTC-examinations. The radiographers were educated and tested in CTC previously in their competence of CTC-interpretation by two expert radiologists with an experience of > 6000 CTC (with > 800 validated by OC).
Study Population

A total of 126 consecutive symptomatic outpatients examined in two hospitals (70 from hospital A and 56 from hospital B) underwent same-day CTC and OC (74 men and 52 women, 35–90 years of age, mean (SD) 63 (11.0) years). CTC was performed immediately prior to OC.

Inclusion and Exclusion Criteria

Inclusion criteria were: referral for OC, age ≥18 years, and the ability to give written and orally informed consent.

Patients were excluded if they had: inflammatory bowel disease, colostomy after colorectal surgery, colorectal biopsy performed within 72 hours, -and/or polypectomy within two weeks prior to CTC, and/or known pregnancy.

Diagnostic Procedures

Examination Technique

All patients underwent a colonic preparation using a low-fibre diet, 2 L of polyethylene glycol electrolyte solution (Moviprep®; Norgine Limited, Mid Glamorgan, UK) and faecal tagging.
In 70 patients (hospital A), fecal tagging was obtained with 100 ml of ionic iodinated contrast (Gastrografin® 370mgI/ml, Bracco Diagnostics, Princeton, USA) soluted in 400 ml of water and administered the day before CTC.

In 56 patients (hospital B), fecal tagging was obtained with 20ml of non-ionic iodinated contrast (Iomeron® 300 mgI/ml, Bracco Diagnostics, Princeton, USA) soluted in 200 ml of water and administered in the late afternoon the day before the examination.

In 70 patients (hospital A), 20mg i.v. hyoscine butylbromide (Buscopan®,, Boehringer Ingelheim, Germany) was used for bowel relaxation [11]. All patients underwent colonic insufflation with carbon dioxide using a CO₂ injector (PROTOCO® 2 L, Bracco, Princeton, USA). At hospital B, there was no use of medicine for bowel relaxation.

All the examinations were performed using a 64-channel multislice CT-scanner (Hospital A, Brilliance Philips Medical Systems, The Netherlands; Hospital B Lightspeed, General Electric Medical Systems, France).

Interpretation

All observers read the examinations independently and were blinded to all clinical findings, the OC results, and each other’s findings.

Scans were obtained at 50 mAs (Hospital A) and 40 mAs (Hospital B) with 120 kV. Patients were examined in supine and prone positions with identical
scanning parameters for both positions: collimation 64x0.625 - slice thickness: 1 mm – increment: 1 mm - rotation time 0.5 seconds.

Image processing and interpretation in the novice department were performed with the use of a CT-workstation (Extended Brilliance workspace 3.5, Philips, The Netherlands) provided with dedicated CTC-software allowing 2D and 3D reading of the colon. This system was used by the radiographers and by one radiologist. Due to local technical limitations of the workstation, simultaneous projection of the supine, and prone acquisition allowing fast comparison between both acquisitions was impossible. The other radiologist interpreted the examinations on a Vitrea workstation (Vital Images, Minnetonka, USA).

For each polyp detected, the readers annotated the segmental location, the size, the attenuation, the slice numbers per acquisition, and the distance to the anal margin of the polyp in a report including a screendump of the polyp. Tumors were included in the calculations and analyzed as polyps, but were described separately as well. The extra-colonic findings were not considered in this study.

Colonoscopy Protocol

OCs were performed by an experienced staff member (gastroenterologist or gastrointestinal surgeon) or by a gastroenterology fellow under direct supervision of experienced staff using 165 cm colonoscopes (Olympus® CF – Q1; 160DL; Olympus Europe Ltd., Hamburg, Germany). While performing the
OCs, the endoscopist was unaware of the CTC-findings. Patients received 2.5-7.5 mg of midazolam (Dormicum®, Roche, Basel, Schweiz) and 0.05-0.1 mg of fentanyl (Fentanyl®-Janssen, Janssen Pharmaceuticals, Titusville New Jersey, USA) on request. The size, morphologic features, segmental location, and the distance from the anal margin of the polyps were documented on a case record form by the endoscopist who performed the examination and by the attending research fellow. Polyp size was measured at endoscopy using open biopsy forceps. According to the adopted segmental checking procedure, a lesion found at a CTC was matched to a corresponding one found at a OC when it was located in the same or adjacent colon segment and when its size differed by no more than 50% [4].

Statistical Analyses

All results were kept in a worksheet (Microsoft Excel version 2007, Microsoft Corporation, Redmond, Wash., USA). A polyp was considered to be a false negative (FN) if the polyp in the respective size category was not detected by one of the readers. A polyp was considered to be a false positive (FP) if the polyp in the respective size category was wrongly identified by one of the readers.

FN-rate and FP-rate per-polyp and per-patient were assessed by using the formulas (1-sensitivity) and (1-specificity), respectively.
A comparison of the Diagnostic Performance in CT-Colonography interpreted by experienced Radiologists and trained Radiographers
PhD Thesis by Carsten Lauridsen

The sensitivity and specificity per-patient and per-polyp were calculated in a previous study by Lauridsen et. al [12].

**Results**

39 patients (three from hospital A and 36 from hospital B) were excluded by one of the expert readers because of inadequate preparation or insufficient distension.

40 lesions were detected in 22 of 87 patients (25%). There were four masses and 36 polyps with 28, 6, and 2 having a sessile, pedunculated and flat morphology, respectively (fig. 1). There were six incomplete OCs (6.7%). In these cases the CTCs were compared to OCs for the colon segments examined by both technologies.

Incomplete OCs were normal in three cases, and showed a stenosing mass in two cases and a polyp in the ascending colon in one case.

Among the six incomplete OCs, the two stenosing masses were detected by all the readers. One of these masses was an extracolonic lymphoma (82 * 87 mm) 70 cm from the anal margin obstructing the sigmoid colon. The second was a mass with the size of 25 mm in rectum located 12 cm from the anal margin.

This lesion was seen by all readers.

One 17 mm rectal mass located 7.5 cm from the anal margin was initially missed when using OC. This mass was detected by five out of six CTC-readers.
Review OC confirmed the lesion with histology revealing an adenocarcinoma.

**False Negative Findings**

The number of false negative findings and the false negative rate are displayed in table 1.

For polyps $\geq 6$ mm, there was only a marginal difference between the two groups of readers. The false negative rate for polyps $\geq 10$ mm was 26.8% higher for the radiographers compared to the radiologists.

One of the missed lesions was a sessile polyp on a fold (6 mm) in the sigmoid colon (figure 2).

This lesion was not seen by any of the readers.

All the readers missed a 9 mm neuroendocrine tumor located at the ileo-cecal valve. One sessile polyp (7 mm) in the ascending colon 130 cm from the anal margin was missed when using OC, but detected by four of the readers (three radiographers and one radiologist). The patient was recalled for an additional OC six months later for polypectomy. The histology showed a hyperplastic polyp. Table 2 shows the number of false-negative findings $\geq 6$ mm stratified according to location and reasons. For the radiographers and the radiologists the false negative findings $\geq 6$ mm were most frequently located in the first three segments of the colon with 54.8% (34/62) and 58.1% (18/31), respectively.
In total, 55.9% (52/93) lesions of the false negative findings ≥6 mm were located in the left hand side of the colon (the rectum, sigmoid, and descending colon) (table 2). For the radiographers, the most regular reasons for missed polyps ≥6 mm in descending order were multiple lesions in one patient, polyps on a fold, flat polyps, polyps at the ileo-caecal valve and polyps hidden by the rectal tube (table 2).

The radiologists most frequently missed polyps ≥6 mm because of multiple lesions in one patient and polyps on a fold, flat lesions, and polyps in the cecal area (table 2).

False positive findings

Table 3 shows the number of false positive findings stratified according to lesion size. The false positive rate per-patient was 18.6% and 18.9% for the radiographers and the radiologists, respectively. In total, 131 false positive findings were made by both groups of readers and two of them was misinterpreted as pedunculated polyps in the ascending colon (figure 3). The reasons for the false positive findings were divided into three categories (table 4). Totally, 79.4% of all the false positive findings were categorized as stool (table 4). One of these are shown in (Figure 4) and was misinterpreted as a sessile polyp (9.2 mm) in the transverse colon.
Discussion

Altogether 126 patients examined in two hospitals underwent same-day CTC and OC and 39 were excluded due to poor preparation, poor distension, and collapsed segments. The two hospitals had different procedures for preparation. At hospital B, they did not use suppressant medicine for bowel relaxation and a lesser amount of faecal tagging material, and this method can probably explain the exclusion of 36 out of 56 patients from hospital B. The false negative rate per-polyp (table 1) in this study showed favourable results compared to a study by Doshi et al [9] and the study by Rockey at al. [13], which demonstrated a false negative rate per-polyp at 48.7% and 51% for lesions ≥6 mm. In the present study, the false negative rate for polyps ≥6 mm was 39.7% and 40.8% for the radiographers and the radiologists respectively. Small polyps ≤ 10 mm are more likely to be missed than larger lesions are at CTC [9,14,15]. Analogous to that our study showed a higher false negative rate (40.25%) (table 1) for missed lesions ≥6 mm than for lesions ≥10 mm (35.15%), which is similar to two other studies [9,14]. The overall morphology of the ileo-cecal valve is much easier to appreciate on the 3D endoluminal view [16] compared to 2D projections where it is much more difficult to exclude a superimposed polyp. A mass involving or replacing the valve itself may represent a more challenging problem [17]. In our study a mass at the IC valve was missed by all the readers (figure 5). The successful detection of a target in a radiological search can reduce the detectability of a
second target, a phenomenon termed Satisfaction of Search (SOS). SOS occurs when an abnormality is missed because another abnormality has been detected. SOS has been used as an explanation for false-negative findings or ‘under-interpreting’ (a source of diminished accuracy) in chest and abdominal radiography, and could probably be a reason for misinterpretation of CTC-examinations as well [18]. In our study, we assume that SOS was the reason for 32.3% and 25.8% (table 2) of the false negative findings for the radiographers and radiologists, respectively. The difference between the two groups is probably due to the longer clinical experience of the radiologists.

Sessile polyps are quite easy to detect if they are located between folds, but if they are located on a haustral fold they could be difficult to detect particularly on axial 3D images [19]. In the present study, 25.8% (or second most frequent) of the missed lesions ≥6 mm were sessile polyps on a fold for both groups of readers (table 2). This result is comparable to the study by Arnesen [20] where the most frequent reason for false negative lesions was slightly elevated lesions and irregularly shaped folds.

Flat lesions are defined as lesions with a height < 3 mm [21]. Similar to their appearance at OC, flat lesions are also less conspicuous at CTC compared to polypoid lesions [22]. In the study by Jensch et. al. [23] including radiographers and radiologists, 75% of the missed findings were flat polyps ≥6 mm. In our study the amount of flat lesions of the false negative findings for the radiographers and the radiologists was 19.4% and 22.6%, respectively (table 2). The present study included only two flat lesions, and this is probably
the reason for the large difference between the results compared to the results by Jensch et. al.

The locations of false negative findings has been analyzed in a study by Arnesen et al. [15] including 100 patients with 90 polyps in 41 patients. We found a total of 52/93 (55.9%) (table 2) of the findings for polyps ≥6 mm in the left colon compared to 76% in the study by Arnesen et al.

Residual stool is one of the most common causes for both false positive and false negative findings on CTC [16,24,25], and these findings may be misdiagnosed as a polyp particularly if homogeneous and adherent to the colon wall. Stool can also obscure polyps and cancer.

False positive findings are mainly related to small fecal residue sticking to the colonic wall as larger residues are subject to positional shifts. Therefore, there is a need for fecal tagging as fecal tagging reduces false positive diagnoses [26]. A properly cleansed colon maximizes the ability to differentiate polyps from folds and residual stool, which will minimize false-positive results.

In our study, 79.4% (table 3) of the false positive lesions ≥6 mm were due to residual stool for both the radiographers and the radiologists. The study by Arnesen et al. [15] showed that 42% of the false positive lesions ≥5 mm were due to retained stool. The large amount of false positive lesions in our study probably demonstrates the efficacy of the insufficient preparation and the results underscore the importance of optimal patient preparation and rigorous technique when performing CTC.
Thickened or complex folds are a common cause of false positives on CTC, particularly on 2D. 3D can frequently be helpful to delineate whether its morphology is linear (interhaustral fold) or polypoid (stool or polyp) [16,25,27]. In the present study, 16% (table 3) of the false positive lesions occurred due to misinterpreting thickened or complex folds as lesions. In a study by Fidler et al. [28] including a test of 15 radiologists after a training program in interpreting CTC, they had 73% of the false positive findings due to misinterpreting folds as polyps. The large difference between the two studies could be due to the fact that the patients from our study are out of clinical practice including outpatients referred for OC. The study by Fidler et al. encompassed particular selected training cases for interpretation.

To our knowledge this study is the first systematic evaluation of false negative and false positive findings between radiologist and radiographers interpreting CTC. For every size category, the radiographers as a group had more false positive findings than the radiologists (table 3). This result is similar to the study by Jensch at al. [23].

Although the preparation at the two hospitals was different, there was only a small difference between the false positive rate for hospital A (16.7%) and hospital B (13.5%). At hospital B, only 20 patients were included and this probably explains the lower value compared to hospital A with 65 patients included.

There were some limitations to this study, though. The main limitation is that because this was a two center study the preparation protocol and experience in
performing CTCs were not uniform at the two centers. Another limitation was that the effect of the insufficient preparation particularly at hospital B thereby resulted in exclusion of 36 patients.

At hospital B the examination technique did not include i.v. injection of hyoscine butylbromide (Buscopan®, Boehringer Ingelheim, Germany) and this could probably be another reason for exclusion of 36 patients.

Our study included six incomplete OCs and two of them contained two stenosing tumors which were seen at CTC. Two other masses were detected by OC and CTC, respectively, and these results confirm that CTC in combination with OC is a very useful tool for detecting CRC.

In conclusion, the results of our study show that the main reason for the false positive findings was misinterpretation of stool. For the false negative lesions the most frequent reason was multiple lesions in one patient (Satisfaction of Search).
Figure 1. Flowchart showing the number of the included patients, number of polyps and number of incomplete OCs.
Figure 2. Missed lesion. Small sessile polyp on a fold with the size of 6 mm in the sigmoid colon. Missed by all the readers.
Figure 3.
Two false positive findings. Pedunculated polyps in the ascending colon with the size 7.2 mm (upper red arrow) and 6.4 mm (lower red arrow). Prone position. The polyps have air inclusions and should have been categorized as stool.
3a: Abdominal window

3b: 3D
Figure 4. Showing an example of the insufficient preparation. False positive polyp in the transverse colon. Misinterpreted as a sessile polyp at the size of 9.2 mm. Supine position.

a: ‘intermediate’ window setting

b: ‘abdominal’ window setting
Figure 5. Tumor at the Ileo-cecal valve with the size of 9 mm. Missed by all the readers.
Table 1

<table>
<thead>
<tr>
<th>Size</th>
<th>Radiographers*</th>
<th></th>
<th>Radiologists**</th>
<th></th>
<th>Total</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>rate</td>
<td>n</td>
<td>rate</td>
<td>n</td>
<td>rate</td>
</tr>
<tr>
<td>≥ 6 mm</td>
<td>40</td>
<td>39.7</td>
<td>22</td>
<td>40.80</td>
<td>62</td>
<td>40.25</td>
</tr>
<tr>
<td>&gt;10 mm</td>
<td>22</td>
<td>39.3</td>
<td>9</td>
<td>31.00</td>
<td>31</td>
<td>35.15</td>
</tr>
</tbody>
</table>

* The digits (n) are the total number of the false negative findings for four radiographers in the particular size category.

** The digits (n) are the total number of the false negative findings for two radiologists in the particular size category.
Table 2

<table>
<thead>
<tr>
<th>Location of the missed lesions</th>
<th>Radiographers*</th>
<th>Radiologists**</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Rectum</td>
<td>19</td>
<td>30.6</td>
<td>11</td>
</tr>
<tr>
<td>Sigmoid colon</td>
<td>13</td>
<td>21.0</td>
<td>7</td>
</tr>
<tr>
<td>Descending colon</td>
<td>2</td>
<td>3.2</td>
<td>0</td>
</tr>
<tr>
<td>Transverse colon</td>
<td>0</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Ascending colon</td>
<td>15</td>
<td>24.2</td>
<td>6</td>
</tr>
<tr>
<td>Cecum</td>
<td>13</td>
<td>21.0</td>
<td>6</td>
</tr>
</tbody>
</table>

**Reasons for the missed lesions**

<table>
<thead>
<tr>
<th>Reason</th>
<th>Radiographers*</th>
<th>Radiologists**</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satisfaction of Search (Multiple lesions in one patient)</td>
<td>20</td>
<td>32.3</td>
<td>8</td>
</tr>
<tr>
<td>Small Sessile polyps on a fold</td>
<td>16</td>
<td>25.8</td>
<td>8</td>
</tr>
<tr>
<td>Flat Lesions</td>
<td>12</td>
<td>19.4</td>
<td>7</td>
</tr>
<tr>
<td>Polyps at the ileo - caecal valve</td>
<td>10</td>
<td>16.1</td>
<td>7</td>
</tr>
<tr>
<td>Polyps hidden by the rectal tube</td>
<td>4</td>
<td>6.5</td>
<td>1</td>
</tr>
</tbody>
</table>

* The digits (n) are the number of the false negative findings for the particular segment or reason divided by the total number of false negative lesions for four radiographers.

** The digits (n) are the number of the false negative findings for the particular segment or reason divided by the total number of false negative lesions for two radiologists.
Table 3

<table>
<thead>
<tr>
<th>Size</th>
<th>Radiographers</th>
<th>Radiologists</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>n</td>
<td>n</td>
<td></td>
</tr>
<tr>
<td>≥ 6 mm ≤ 9 mm</td>
<td>67</td>
<td>41</td>
<td>108</td>
</tr>
<tr>
<td>&gt;10 mm</td>
<td>12</td>
<td>11</td>
<td>23</td>
</tr>
<tr>
<td>All sizes</td>
<td>79</td>
<td>52</td>
<td>131</td>
</tr>
</tbody>
</table>

Table 4

<table>
<thead>
<tr>
<th>Reason</th>
<th>Radiographers*</th>
<th>Radiologists**</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reason</td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Stool</td>
<td>63/79</td>
<td>79.7</td>
<td>41/52</td>
</tr>
<tr>
<td>Thickened fold</td>
<td>13/79</td>
<td>16.5</td>
<td>8/52</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>3/79</td>
<td>3.8</td>
<td>3/52</td>
</tr>
</tbody>
</table>

* The digits (n) are the number of the false positive findings for the particular reason divided by the total number of false positive lesions for four radiographers.

** The digits (n) are the number of the false positive findings for the particular reason divided by the total number of false positive lesions for two radiologists.
References


A comparison of the diagnostic performance in CT-Colonography interpreted by experienced radiologists and trained radiographers.

Ph.D. thesis by Carsten Lauridsen


[15] Arnesen RB, Adamsen S, Svendsen LB, Raaschou HO, von Benzon E, Hansen OH. Missed lesions and false-positive findings on computed-
A comparison of the diagnostic performance in CT-Colonography interpreted by experienced radiologists and trained radiographers.

Ph.D. thesis by Carsten Lauridsen


A comparison of the diagnostic performance in CT-Colonography interpreted by experienced radiologists and trained radiographers.

Ph.D. thesis by Carsten Lauridsen


