UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Clinical evaluation of technical developments in CT colonography

de Vries, A.H.

Publication date
2009
Document Version
Final published version

Link to publication

Citation for published version (APA):
de Vries, A. H. (2009). *Clinical evaluation of technical developments in CT colonography*. [Thesis, fully internal, Universiteit van Amsterdam].

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

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

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

Download date:17 Nov 2026

Clinical Evaluation of Technical Developments in CT Colonography

Clinical Evaluation of Technical Developments in CT Colonography

Ayso de Vries

Ayso de Vries

This thesis was prepared at the Department of Radiology, Academic Medical Center, University of Amsterdam, the Netherlands.

Copyright © 2009, Ayso de Vries, Amsterdam, the Netherlands

For this research a grant was received from Philips Healthcare, Best, the Netherlands

The printing of this thesis was financially supported by

- the Department of Radiology, Academic Medical Center, Amsterdam, the Netherlands.
- J.E. Jurriaanse Stichting
- 'Stichting Nationaal Fonds tegen Kanker - voor onderzoek naar reguliere en aanvullende therapieën te Amsterdam'
- Bayer HealthCare AG.
- Guerbet Nederland B.V.
- Sectra Benelux
- Integraal Kankercentrum Amsterdam (IKA) / Comprehensive Cancer Center Amsterdam (CCCA)
- Philips Healthcare

The CCCA is one of eight Comprehensive Cancer Centers in the Netherlands. Its area is the north-western part of the Netherlands and involves 2,800,000 inhabitants, 16 general hospitals, two university hospitals and the Netherlands Cancer Institute.

The comprehensive cancer centres (CCC's) in the Netherlands have been founded to provide comprehensive and high-quality cancer care close to home for all cancer patients. The CCCA provides and coordinates a collaboration of all health care professionals and institutions involved in cancer and palliative care. The CCCA functions as a centre of knowledge and quality care that helps to improve cancer treatment, patient care and clinical research as well as prevention of cancer and decrease of cancer mortality.

Cover: AG Freelance-dtp
Layout: AG Freelance-dtp
Printed by : Scholma Druk BV

ISBN 978-90-9024632-1

CLINICAL EVALUATION OF TECHNICAL DEVELOPMENTS IN CT COLONOGRAPHY

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor

aan de Universiteit van Amsterdam

op gezag van de Rector Magnificus

prof. dr. D.C. van den Boom

ten overstaan van een door het college voor promoties ingestelde

commissie, in het openbaar te verdedigen in de Agnietenkapel

op vrijdag 16 oktober 2009, te 12.00 uur

door Ayso Harmen de Vries

geboren te Noordoostpolder

Promotiecommissie:

Promotor:	Prof. dr. J. Stoker
Co-promotores:	Dr. E. Dekker
	Dr. H.W. Venema
Overige leden:	Prof. dr. P. Fockens
	Prof. dr. W.M. Prokop
	Prof. dr. J.S. Laméris
	Prof. dr. L.J. van Vliet
	Dr. F.M. Vos

Faculteit der Geneeskunde

Inhoud

Chapter 1	**General Introduction**	7
Chapter 2	**Image Quality and Patient Acceptance of Four Regimens with Different Amounts of Mild Laxatives for CT Colonography**	17
Chapter 3	**Lesion Conspicuity and Efficiency of CT colonography with Electronic Cleansing Based on a Three-Material Transition Model**	35
Chapter 4	**Primary Uncleansed 2D Versus Primary Electronically Cleansed 3D in Limited Bowel Preparation CT Colonography. Is There a Difference for Novices and Experienced Readers?**	53
Chapter 5	**Influence of Tagged Fecal Material on Detectabilit of Colorectal Polyps at CT: Phantom Study**	69
Chapter 6	**Does a Computer-Aided Detection Algorithm in a Second Read Paradigm Enhance the Performance of Experienced Computed Tomography Colonography Readers in a Population of Increased Risk?**	85
Chapter 7	**Feasibility of Automated Matching of Supine and Prone CT Colonography Examinations**	101
Chapter 8	**Polyp Measurement Based on CT Colonography and Colonoscopy: Inter-observer Variability and Systematic Differences**	109
Chapter 9	**Summary and General Discussion**	123
Chapter 10	**Samenvatting en Algemene Discussie**	131
	List of Publications	139
	Dankwoord	141
	Curriculum Vitae	143

Voor mijn ouders en Jan

Chapter 1

General Introduction

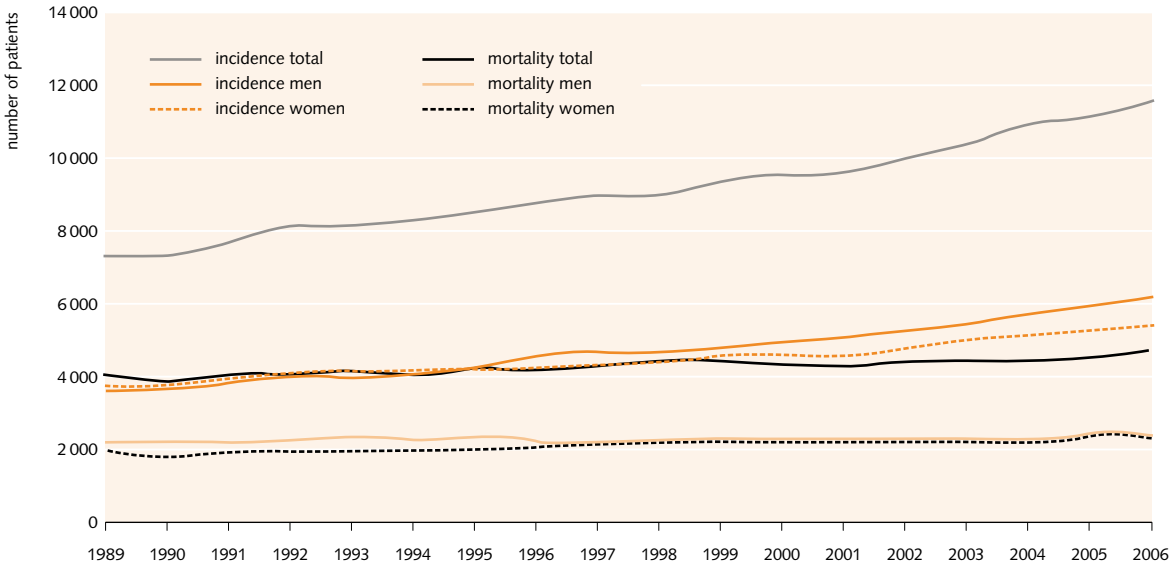
Ayso H. de Vries

Introduction

Colorectal cancer ranks third in incidence and second in cause of cancer related death for both men and women in the western world. In the Netherlands in 2006 colorectal cancer was diagnosed in 11,500 patients. In that year 4,700 patients died because of the disease [1]. These figures are increasing (Figure 1.1) and in 2015 13,800 patients are expected to be diagnosed with colorectal cancer if no other measures are taken [2].

At present 40 to 45% of the patients die of the disease within 5 years after diagnosis [3]. The high mortality rate is due to the fact that the disease is often diagnosed in a late stage. Early detection may reduce the high mortality rate. Colorectal cancer can be identified in a pre-symptomatic phase since most of the tumors are considered to arise from well detectable but asymptomatic benign colonic polyps (Figure 1.2). The vast majority of the cancers arise in benign (adenomatous) polyps that develop slowly. The time span of this so called adenoma-carcinoma sequence is approximately 10 to 15 years [4].

FIGURE 1.1
Incidence and mortality 1989-2005 in The Netherlands (source: ikcnet.nl).



Even though 20 to 30% of the population aged fifty years or older have these benign colonic polyps [5], only a small fraction of these polyps will actually develop into colorectal cancer [6]. Size is an important predictor for this development since the chance of an adenoma being malignant increases with its size [7-10]. Currently, polyps of 10 mm or larger are considered clinically relevant since approximately ten percent of these adenomas are malignant [7;11]. A debate is still going about polyps smaller than 10 mm; according to some investigators these polyps have little to no clinical relevance in intermediate and long-term follow-up [7;12], however others advocate removal of all polyps regardless of size [13].

Detection and removal of these benign precursors of colorectal cancer will decrease the incidence of colorectal cancer and its mortality [14]. Moreover, the detection of colorectal cancer in an early and localized stage also decreases the disease related mortality since the treatment for early stages of the disease is more effective [15-17].

Most newly diagnosed patients do not have a genetic predisposition for colorectal cancer [18] and are therefore not at increased risk for colorectal cancer. Evidence-based

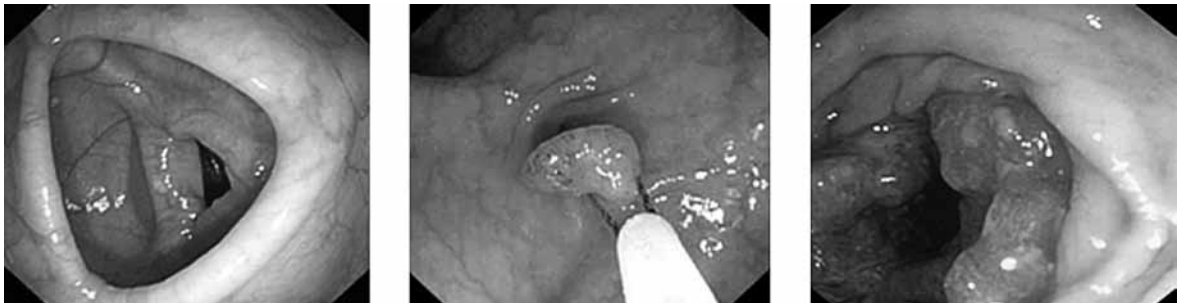


FIGURE 1.2
Normal colon mucosa in the transverse colon (left), adenomatous polyp of 10 mm just about to be removed (middle) and a colorectal tumor (right).

guidelines have recommended screening for a population aged fifty years or older that bear an average risk for developing colorectal cancer [19]. In the Netherlands a screening program is not routinely offered to the public yet, although research on this topic is done. For screening a range of options is available. These options can be divided in two general categories; stool tests that basically detect colorectal cancer and structural exams that detect both colorectal cancer and clinically relevant polyps. The stool tests include for occult blood (FOBT) or fecal DNA, the structural colonic exams include for sigmoidoscopy, colonoscopy, double-contrast barium enema (DCBE) or computed tomography (CT) colonography.

Surveillance guidelines advise patients with a personal or familial history of colorectal polyps or cancer a regular surveillance by colonoscopy [20-22]. In the Netherlands, increased-risk patients are in general under surveillance by a gastroenterologist. These patients undergo an optical colonoscopy once in 2 to 6 years for the detection and removal of adenomatous polyps and early carcinoma [23]. Colonoscopy is considered the reference standard since it has a very high accuracy for the detection of colorectal cancer and its precursors. Furthermore, colorectal polyps and early cancers can be removed during the procedure and a histopathological biopsy can be taken of larger, unresectable lesions.

An important disadvantage of the technique is that the colon needs to be cleansed before the colonic mucosa can be investigated. Therefore patients have to undergo an extensive bowel preparation. Patients consider this as the most burdensome aspect of the investigation [24] and it most likely reduces patient compliance [25-27]. Secondly, sedation is often needed to reduce pain, so patients have to be monitored for a while after the investigation. Thirdly, this invasive test has a small but not negligible risk of causing bleeding and perforation of the bowel. Low patient compliance may result in a low participation rate for surveillance and screening programs.
Currently, colonoscopy is the primary technique in symptomatic patients and it is the only recommended examination for patients under surveillance due to an increased risk for colorectal cancer.

In contrast to colonoscopy, the above mentioned stool tests and structural colonic exams are less burdensome. However, these tests have a low(er) detection rate for colorectal cancer and clinical relevant polyps. A good alternative for early polyp detection may be CT colonography.
CT colonography is a radiological technique to visualize the colon. After a bowel preparation, a rectal catheter is introduced and after administration of a muscle relaxant the colon is insufflated with 2-4 liter CO2. The patient is scanned in both prone and supine position, in order to obtain maximal bowel distension for the complete colon. Based on acquired cross sectional CT images the colon

can be assessed. During the procedure the patient is not sedated, so the patient does not need to stay in the hospital after completion of the examination.
The technique has shown comparable results to colonoscopy for the detection of colorectal cancer and clinically relevant polyps [28-30] i.e. polyps 10 mm or larger. Moreover, CT colonography goes with a relative limited patient burden[31;32]. Therefore, CT colonography could be an alternative for colonoscopy in polyp detection. The disadvantage of the technique is the exposure to ionizing radiation (effective dose of 5-10mSv) [33] that has a potential carcinogenic effect. However, because the dose is low the carcinogenic effect is limited. Still, the predicted mortality of cancer induced by CT may be less than the mortality that is related to complications of colonoscopy [34].
Recently, CT colonography has officially been introduced as one of the screening test for average risk patients in the United States [35]. In the Netherlands, CT colonography is not officially recommended as a test for screening or surveillance. Currently, the role of CT colonography in the Netherlands is as a diagnostic or adjuvant diagnostic test for patients with incomplete or inconclusive colonoscopy [36].

This thesis

Although a considerable body of evidence is present for CT colonography as a diagnostic test, a number of issues still need further research.
In a diagnostic test for the detection of colorectal cancer and clinically relevant lesions it is important to have a high yield of lesions to keep the number of false negative findings minimal. This can be done by choosing a test with a high sensitivity and a high participation rate. In this thesis we have focused on a number of subjects that aim to eventually increase the yield of CT colonography. We have focused (1) on reducing the patient burden of the test and (2) on increasing observer performance. Reducing the patient burden can be done by reducing bowel preparation, and increasing observer performance can be done by electronic cleansing (i.e. virtual removing fecal material from the CT data) and computer aided detection (i.e. computer support by reading CT colonography). Furthermore, we have focused on the effect of limited bowel preparation on the required radiation dose, matching of scans performed in prone and supine position and polyp measurement. The chapters of this thesis will be discussed briefly.

Bowel preparation

When CT colonography was introduced in 1994, the bowel preparation was similar to the extensive bowel preparation of colonoscopy. This is often described by patients as the most burdensome aspect of both colonoscopy and CT colonography [37-39]. This aspect may diminish the patients' willingness to undergo the examination [26;40;41]. In the last couple of years, efforts have been made to reduce the patient burden of CT colonography by reducing the bowel preparation in combination with fecal tagging (labeling of fecal material with a contrast agent) [42-46]. In literature various limited bowel preparation regimens have been described, although an optimal bowel preparation has still not been not found yet [47-49].
Our objective in **Chapter 2** was to determine the optimal dosage of laxatives with regard to diagnostic quality and patient acceptance in a limited bowel preparation regimen. Therefore, we compared diagnostic quality and patient acceptance between four regimens of limited bowel preparation.

Limited bowel preparation and review technique

Reading CT colonography studies can be done primary two-dimensionally (based on the original CT images) and primary three-dimensionally (based on a three-dimensional reconstruction of colonic surface using the original CT images). Until now, all CT colonography studies of patients that underwent a limited bowel preparation, have been

read using primary two-dimensional (2D) display methods [50-52]. The rationale for this approach is that the colon wall is always visible in a 2D display, even if it is submerged or covered by fecal material. This is not the case using primary 3D images.

Previous studies in patients undergoing an extensive bowel preparation (in which only a limited amount of fecal remains are present) have indicated that primary 3D reading may result in less missed findings compared to a primary 2D reading [53;54]. If a similar empty colon could be achieved in patients who have undergone limited bowel preparation, primary 3D reading could be the method of choice.

In a feasibility study in **Chapter 3** the effect of an electronic cleansing algorithm on lesion conspicuity, review time, assessment effort and observer confidence was evaluated.

In **Chapter 4** we assessed whether electronic cleansing actually increased the accuracy of CT colonography in patients that had undergone a limited bowel preparation. Therefore, we compared primary electronically cleansed 3D with the technique that is routinely used after a limited bowel preparation i.e. primary uncleansed 2D. Both techniques were evaluated in a consecutive group of fecal occult blood test positive patients.

Radiation dose

Computed tomography is an imaging technique that goes with exposure to ionizing radiation. The amount of ionizing radiation determines the amount of image noise. The less radiation, the more noise is present in the CT images. Image noise impairs the image quality. Because the contrast between the colon wall and air is very large, the dose can be very low with preservation of diagnostic quality [55]. However, if patients have undergone a limited bowel preparation, the contrast between the colon wall and the lumen is reduced by the presence of fecal remains, even if oral iodine contrast material has been added to the bowel preparation. This has consequences for the minimal radiation dose required for an adequate image quality.

In **Chapter 5** we performed a study to determine the effect of oral contrast material (i.e. reduced contrast) on the minimal radiation dose required for a qualitatively acceptable CT colonography examination. For this study we used a synthetic anthropomorphic phantom.

Computer aided detection

In the past years, efforts have been made to in order to increase CT colonography accuracy e.g. fecal tagging (as discussed above), automatic insufflation, training of CT colonography observers and improvement of workstations. Despite these efforts, detectable lesions are still missed, even by well-trained radiologists. Computer aided detection (CAD) is a promising technique [56-59] that could be helpful in detecting these missed colorectal lesions [60;61].

However, true positive CAD lesions initially missed by the observer could be discarded based on wrong considerations of the observer [62;63], in other words; true positive CAD suggestions could be wrongfully labeled as non-polyps. This stresses the complex interaction between CAD and the human observer.

In **Chapter 6** we determined whether CAD in a second read paradigm (CAD as backup reader of the human observer) could improve the performance of an observer in a practical setting. This was done in a prospective comparative study of CT colonography and colonoscopy in a consecutive series of 170 patients at increased risk for colorectal cancer.

Prone supine matching

State of art CT colonography is performed with the patient in both supine and prone position. Combining information of both scans will assist the reviewer in differentiating colorectal polyps from fecal material or folds. Verification of findings on supine and prone positions may be a time-consuming activity since reference-points (e.g. hepatic flexure) are often not fixed. Using an automated supine-prone matching algorithm may facilitate this process, and may lead to a more efficient interpretation of CT colonography.

In **Chapter 7** we have assessed the feasibility of automated matching of supine and prone CT colonography examinations.

Polyp measurement

Once a lesion is found with CT colonography, an assessment of the chance of malignancy by performing a biopsy for histology is not possible. As a surrogate for histopathology, polyp size is used for patient management strategies [35], therefore size is crucial for decision making in CT colonography.

As stated earlier in this chapter, the chance of a colorectal polyp being malignant increases with its size. Whereas polyps of 10mm or larger have a ten percent chance of being malignant [7;64], this chance is only 0.1percent for polyps smaller than 6mm [65]. Therefore, size thresholds are used for referral strategies for colonoscopy. Using these thresholds, variability of measurements within CT colonography and colonoscopy techniques and mean differences between these techniques should be minimal to avoid over-treatment or under-treatment.

In **Chapter 8** we evaluated the variability of measurements within CT colonography and colonoscopy techniques and the mean differences between these techniques on identical colorectal polyps.

Chapter 9 is a summary, general discussion and conclusions

Chapter 10 is a summary, general discussion and conclusions provided in Dutch.

References

1. Nederlands Kanker Registratie. 2008. 13-11-2008. Ref Type: Internet Communication
2. Signaleringscommissie Kanker van KWF Kankerbestrijding. Toekomstscenario's en aanbevelingen. Vroege opsporing van dikkedarmkanker. 2004; 91-96
3. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2008. CA Cancer J Clin 2008; 58:71-96
4. Winawer SJ, Fletcher RH, Miller L, et al. Colorectal cancer screening: clinical guidelines and rationale. Gastroenterology 1997; 112:594-642
5. Bond JH. Colon polyps and cancer. Endoscopy 2005; 37:208-212
6. Midgley R, Kerr D. Colorectal cancer. Lancet 1999; 353:391-399
7. Atkin WS, Morson BC, Cuzick J. Long-term risk of colorectal cancer after excision of rectosigmoid adenomas. N Engl J Med 1992; 326:658-662
8. O'Brien MJ, Winawer SJ, Zauber AG, et al. The National Polyp Study. Patient and polyp characteristics associated with high-grade dysplasia in colorectal adenomas. Gastroenterology 1990; 98:371-379
9. Simons BD, Morrison AS, Lev R, Verhoek-Oftedahl W. Relationship of polyps to cancer of the large intestine. J Natl Cancer Inst 1992; 84:962-966
10. Waye JD, Lewis BS, Frankel A, Geller SA. Small colon polyps. Am J Gastroenterol 1988; 83:120-122
11. Simons BD, Morrison AS, Lev R, Verhoek-Oftedahl W. Relationship of polyps to cancer of the large intestine. J Natl Cancer Inst 1992; 84:962-966
12. Spencer RJ, Melton LJ, III, Ready RL, Ilstrup DM. Treatment of small colorectal polyps: a population-based study of the risk of subsequent carcinoma. Mayo Clin Proc 1984; 59:305-310
13. Rex DK. PRO: Patients with polyps smaller than 1 cm on computed tomographic colonography should be offered colonoscopy and polypectomy. Am J Gastroenterol 2005; 100:1903-1905
14. Winawer SJ, Zauber AG, Ho MN, et al. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. N Engl J Med 1993; 329:1977-1981
15. Hardcastle JD, Chamberlain JO, Robinson MH, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet 1996; 348:1472-1477
16. Kronborg O, Fenger C, Olsen J, Jorgensen OD, Sondergaard O. Randomised study of screening for colorectal cancer with faecal-occult-blood test. Lancet 1996; 348:1467-1471
17. Mandel JS, Church TR, Bond JH, et al. The effect of fecal occult-blood screening on the incidence of colorectal cancer. N Engl J Med 2000; 343:1603-1607
18. Winawer SJ, Fletcher RH, Miller L, et al. Colorectal cancer screening: clinical guidelines and rationale. Gastroenterology 1997; 112:594-642
19. Levin B, Lieberman DA, McFarland B, et al. Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. Gastroenterology 2008; 134:1570-1595
20. Rex DK, Kahi CJ, Levin B, et al. Guidelines for colonoscopy surveillance after cancer resection: a consensus update by the American Cancer Society and the US Multi-Society Task Force on Colorectal Cancer. Gastroenterology 2006; 130:1865-1871
21. Winawer SJ, Zauber AG, Fletcher RH, et al. Guidelines for colonoscopy surveillance after polypectomy: a consensus update by the US Multi-Society Task Force on Colorectal Cancer and the American Cancer Society. Gastroenterology 2006; 130:1872-1885
22. Smith RA, Cokkinides V, von Eschenbach AC, et al. American Cancer Society guidelines for the early detection of cancer. CA Cancer J Clin 2002; 52:8-22
23. Vereniging Klinische Genetica Nederland (VKGN). Landelijke richtlijn erfelijke darmkanker; Start en interval colonoscopische surveillance. VIKC . 2009. 24-3-2009. Ref Type: Internet Communication

24. van Gelder RE, Birnie E, Florie J, et al. CT colonography and colonoscopy: assessment of patient preference in a 5-week follow-up study. Radiology 2004; 233:328-337
25. Gluecker TM, Johnson CD, Harmsen WS, et al. Colorectal cancer screening with CT colonography, colonoscopy, and double-contrast barium enema examination: prospective assessment of patient perceptions and preferences. Radiology 2003; 227:378-384
26. Rex DK. Current colorectal cancer screening strategies: overview and obstacles to implementation. Rev Gastroenterol Disord 2002; 2 Suppl 1:S2-11
27. Weitzman ER, Zapka J, Estabrook B, Goins KV. Risk and reluctance: understanding impediments to colorectal cancer screening. Prev Med 2001; 32:502-513
28. Kim DH, Pickhardt PJ, Taylor AJ, et al. CT colonography versus colonoscopy for the detection of advanced neoplasia. N Engl J Med 2007; 357:1403-1412
29. Pickhardt PJ, Choi JR, Hwang I, et al. Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. New England Journal of Medicine 2003; 349:2191-2200
30. Johnson CD, Chen MH, Toledano AY, et al. Accuracy of CT colonography for detection of large adenomas and cancers. N Engl J Med 2008; 359:1207-1217
31. Lefere PA, Gryspeerdt SS, Dewyspelaere J, Baekelandt M, Van Holsbeeck BG. Dietary fecal tagging as a cleansing method before CT colonography: initial results polyp detection and patient acceptance. Radiology 2002; 224:393-403
32. Zalis ME, Perumpillichira JJ, Magee C, Kohlberg G, Hahn PF. Tagging-based, electronically cleansed CT colonography: evaluation of patient comfort and image readability. Radiology 2006; 239:149-159
33. Liedenbaum MH, Venema HW, Stoker J. Radiation dose in CT colonography--trends in time and differences between daily practice and screening protocols. Eur Radiol 2008; 18:2222-2230
34. Hassan C, Pickhardt PJ, Laghi A, et al. Computed tomographic colonography to screen for colorectal cancer, extracolonic cancer, and aortic aneurysm: model simulation with cost-effectiveness analysis. Arch Intern Med 2008; 168:696-705
35. McFarland EG, Levin B, Lieberman DA, et al. Revised colorectal screening guidelines: joint effort of the American Cancer Society, U.S. Multisociety Task Force on Colorectal Cancer, and American College of Radiology. Radiology 2008; 248:717-720
36. Vereniging van Integrale Kankercentra. Primaire tumor en indien geen volledige colonoscopie. http://oncoline.nl/ . 2008. 14-11-2008. Ref Type: Internet Communication
37. Harewood GC, Wiersema MJ, Nelson H, et al. A prospective, blinded assessment of the impact of preoperative staging on the management of rectal cancer. Gastroenterology 2002; 123:24-32
38. Ristvedt SL, McFarland EG, Weinstock LB, Thyssen EP. Patient preferences for CT colonography, conventional colonoscopy, and bowel preparation. Am J Gastroenterol 2003; 98:578-585
39. van Gelder RE, Birnie E, Florie J, et al. CT colonography and colonoscopy: assessment of patient preference in a 5-week follow-up study. Radiology 2004; 233:328-337
40. Gluecker TM, Johnson CD, Harmsen WS, et al. Colorectal cancer screening with CT colonography, colonoscopy, and double-contrast barium enema examination: prospective assessment of patient perceptions and preferences. Radiology 2003; 227:378-384
41. Weitzman ER, Zapka J, Estabrook B, Goins KV. Risk and reluctance: understanding impediments to colorectal cancer screening. Prev Med 2001; 32:502-513
42. Iannaccone R, Laghi A, Catalano C, et al. Computed tomographic colonography without cathartic preparation for the detection of colorectal polyps. Gastroenterology 2004; 127:1300-1311
43. Bielen D, Thomeer M, Vanbeckevoort D, et al. Dry preparation for virtual CT colonography with fecal tagging using water-soluble contrast medium: initial results. Eur Radiol 2003; 13:453-458

44. Lefere P, Gryspeerdt S, Marrannes J, Baekelandt M, van HB. CT colonography after fecal tagging with a reduced cathartic cleansing and a reduced volume of barium. AJR Am J Roentgenol 2005; 184:1836-1842
45. Lefere PA, Gryspeerdt SS, Dewyspelaere J, Baekelandt M, Van Holsbeeck BG. Dietary fecal tagging as a cleansing method before CT colonography: initial results polyp detection and patient acceptance. Radiology 2002; 224:393-403
46. McFarland EG, Zalis ME. CT colonography: progress toward colorectal evaluation without catharsis. Gastroenterology 2004; 127:1623-1626
47. Iannaccone R, Laghi A, Catalano C, et al. Computed tomographic colonography without cathartic preparation for the detection of colorectal polyps. Gastroenterology 2004; 127:1300-1311
48. Jensch S, de Vries AH, Peringa J, et al. CT colonography with Limited Bowel Preparation: Performance Characteristics in an Increased-Risk Population. Radiology 2008;
49. Lefere P, Gryspeerdt S, Marrannes J, Baekelandt M, van HB. CT colonography after fecal tagging with a reduced cathartic cleansing and a reduced volume of barium. AJR Am J Roentgenol 2005; 184:1836-1842
50. Iannaccone R, Laghi A, Catalano C, et al. Computed tomographic colonography without cathartic preparation for the detection of colorectal polyps. Gastroenterology 2004; 127:1300-1311
51. Jensch S, de Vries AH, Peringa J, et al. CT colonography with limited bowel preparation: performance characteristics in an increased-risk population. Radiology 2008; 247:122-132
52. Lefere P, Gryspeerdt S, Marrannes J, Baekelandt M, van HB. CT colonography after fecal tagging with a reduced cathartic cleansing and a reduced volume of barium. AJR Am J Roentgenol 2005; 184:1836-1842
53. Pickhardt PJ, Lee AD, Taylor AJ, et al. Primary 2D versus primary 3D polyp detection at screening CT colonography. AJR Am J Roentgenol 2007; 189:1451-1456
54. van Gelder RE, Florie J, Nio CY, et al. A comparison of primary two- and three-dimensional methods to review CT colonography. Eur Radiol 2006; 17:1181-1192
55. van Gelder RE, Venema HW, Serlie IW, et al. CT colonography at different radiation dose levels: feasibility of dose reduction. Radiology 2002; 224:25-33
56. Summers RM, Yao J, Pickhardt PJ, et al. Computed tomographic virtual colonoscopy computer-aided polyp detection in a screening population. Gastroenterology 2005; 129:1832-1844
57. Yoshida H, Nappi J, MacEneaney P, Rubin DT, Dachman AH. Computer-aided diagnosis scheme for detection of polyps at CT colonography. Radiographics 2002; 22:963-979
58. Jerebko AK, Summers RM, Malley JD, Franaszek M, Johnson CD. Computer-assisted detection of colonic polyps with CT colonography using neural networks and binary classification trees. Med Phys 2003; 30:52-60
59. Mani A, Napel S, Paik DS, et al. Computed tomography colonography: feasibility of computer-aided polyp detection in a "first reader" paradigm. J Comput Assist Tomogr 2004; 28:318-326
60. Summers RM, Jerebko AK, Franaszek M, Malley JD, Johnson CD. Complementary role of computer-aided diagnosis for detecting colonic polyps with CT colonography. Radiology 2001; 120:391-399
61. Taylor SA, Halligan S, Burling D, et al. Computer-assisted reader software versus expert reviewers for polyp detection on CT colonography. AJR Am J Roentgenol 2006; 186:696-702
62. Halligan S, Altman DG, Mallett S, et al. Computed tomographic colonography: assessment of radiologist performance with and without computer-aided detection. Gastroenterology 2006; 131:1690-1699
63. Taylor SA, Charman SC, Lefere P, et al. CT colonography: investigation of the optimum reader paradigm by using computer-aided detection software. Radiology 2008; 246:463-471
64. Simons BD, Morrison AS, Lev R, Verhoek-Oftedahl W. Relationship of polyps to cancer of the large intestine. J Natl Cancer Inst 1992; 84:962-966
65. Waye JD, Lewis BS, Frankel A, Geller SA. Small colon polyps. Am J Gastroenterol 1988; 83:120-122

Chapter 2

Image Quality and Patient Acceptance of Four Regimens with Different Amounts of Mild Laxatives for CT Colonography

Sebastiaan Jensch
Ayso H. de Vries
Dennis Pot
Jan Peringa
Shandra Bipat
Jasper Florie
Rogier E. van Gelder
Jaap Stoker

American Journal of Roentgenology 2008; 191:158–167

Abstract

Objective

The purpose of our study was to prospectively evaluate image quality and patient acceptance of CT colonography (CTC) with fecal tagging using different levels of catharsis.

Material and methods

Forty consecutive increased-risk patients were randomized. Group 1 received orally 20 mg of bisacodyl, group 2 received 30 mg of bisacodyl, group 3 received 20 mg of bisacodyl and 8.2 g of magnesium citrate, and group 4 received 30 mg of bisacodyl and 16.4 g of magnesium citrate. All patients used a 2-day low-fiber diet and received diatrizoate meglumine and barium for fecal tagging. One reviewer blindly scored subjective image quality (fecal tagging, amount of residual feces [liquid or solid], luminal distention, and image readability) on a 5- to 6-point scale using a 2D review technique. The mean and SD of attenuation of tagging were measured as well as the relative SD as a measure of homogeneity. Furthermore, patient acceptance (burden related to diarrhea, abdominal pain, flatulence, and overall burden) was evaluated. Ordinal regression, generalized estimating equations, and parametric and nonparametric tests were used for analysis.

Results

Image readability was evaluated as good or excellent in all examinations except one in group 2 (non-diagnostic) and two in group 3 (moderate). Group 2 contained more feces than group 4 ($p = 0.04$). With regard to mean attenuation and homogeneity of tagging, no significant differences were observed between groups. Group 4 experienced more severe diarrhea than groups 1 and 2 and higher overall burden than groups 1 and 3 ($p < 0.042$).

Conclusion

The mildest preparation with 20 mg of bisacodyl provided good image quality of CTC images. Increasing the amount of laxatives did not improve image quality or tagging characteristics but was associated with a lower patient acceptance.

19.

Introduction

CT colonography (CTC) is being investigated as a possible screening technique for the detection of colorectal polyps and cancer [1–5]. However, the requisite cathartic bowel preparation, which is often described by patients as the most burdensome aspect of colonic examinations [6–8], might diminish patients' willingness to participate in a screening program [9–11]. Labeling of fecal material with a contrast agent (fecal tagging) has enabled the use of limited bowel preparation regimens containing no or only limited amounts of laxatives [12]. So far, feasibility studies investigating CTC with reduced catharsis have shown promising results with regard to image quality and patient acceptance [13–17]. In addition, one large accuracy study that used no cathartics has been published that reported excellent results with regard to polyp detection [18].

For labeling fecal material, three types of tagging agents are available: barium and nonionic and ionic iodinated contrast agents. Barium is traditionally used for solid fecal matter tagging and iodinated contrast agents are used to tag residual fluid. Furthermore, a variety of mild laxatives, for example, bisacodyl sodium or magnesium citrate [13, 14], can be added to reduce the amount of fecal matter in the colon. At present, no consensus exists about which contrast agent should be used and whether mild cathartics should be added to the tagging regimen, and if so, in what dose [19].

We hypothesized that a higher level of catharsis would improve image quality but reduce patient acceptance. Our objective was to determine the optimal dosage of laxatives for CTC with limited bowel preparation with regard to both image quality and patient acceptance. Therefore, we compared image quality and patient acceptance between four regimens with increasing levels of mild catharsis, using bisacodyl and magnesium citrate as laxative agents. Although some studies have compared image quality of CTC with and without laxatives [20, 21], to our knowledge, our study is the first that has investigated the effect of different amounts of mild laxatives in CTC.

Materials and Methods

Study population

Forty consecutive adult patients with an increased risk for colorectal cancer (i.e., a personal or family history of colorectal polyps or cancer) were included in this study between October 2004 and January 2005. All patients were scheduled to undergo a conventional colonoscopy for polyp detection at the endoscopic departments of the Academic Medical Center or the Onze Lieve Vrouwe Gasthuis. Exclusion criteria were a personal history of inflammatory bowel disease or familial adenomatous polyposis, prior allergic reaction to an iodine-containing contrast agent, colorectal polyps or cancer at prior endoscopy that were not removed, or participation in a research project that involved ionizing radiation within 12 months preceding the CTC examination. On the day of the CTC examination, patients were asked to indicate whether they had symptoms of colorectal disease (i.e., abdominal pain, hematochezia, or altered bowel habits) and if so, which symptoms were present. No formal power calculation was performed in this feasibility study and, for practical reasons, 40 patients were studied. The institutional review board of both hospitals gave approval for this feasibility study of 40 patients. All patients gave written informed consent.

Bowel Preparation

A low-fiber diet was prescribed for all patients for 2 days before the CTC examination. No specific meal kit was used. Through a customized computer program, patients were randomized into one of four groups using a randomization block design (Table 2.1).

Preparations contained solely orally administered bisacodyl or bisacodyl in combination with magnesium citrate (LoSo Prep, E-Z-EM) as laxative agents, and the extent of catharsis gradually increased from group 1 to group 4. We did not include a laxative-free regimen because such regimens have been shown to result in insufficient image quality [20, 21]. The most extensive preparation (group 4) contained a combination of bisacodyl and magnesium

TABLE 2.1 **Bowel preparation schemes for CT colonography (CTC)**

Group	Barium Sulphate[a] (mL) (40% w/v)	Diatrizoate Meglumine[b] (mL) (200 mgI/mL)	Bisacodyl[c] (mg)	Magnesium Citrate[d] (g)
I	80	110	20	-
II	80	110	30	-
III	80	110	20	8.2
IV	80	110	30	16.4

a Dosage: 2 days before CTC, 20 mL at dinner and 1 day before CTC, 20 mL at breakfast, lunch, and dinner.
b Dosage: 1 day before CTC, 10 mL at breakfast and 20 mL at lunch and dinner; on the day of CTC, 60 mL atbreakfast.
c Dosage: 16 hours before CTC, 10 mg of orally administered bisacodyl for groups 1 and 3 and 20 mg for groups 2 and 4; on the day of CTC, all groups received 10 mg of bisacodyl orally at breakfast.
d Dosage: 18 hours before CTC, groups 3 and 4 received magnesium citrate in the displayed dose.

citrate in a dosage that is commercially available as a preparation kit for CT colonography (LoSo Prep) and is considered a mild preparation [13, 14]. Tagging consisted of barium sulfate, 40% weight per volume (Tagitol V, E-Z-EM) and diatrizoate meglumine with an iodine concentration of 200 mg/mL. Patients were not informed of the content of the other regimens.

CT Colonography

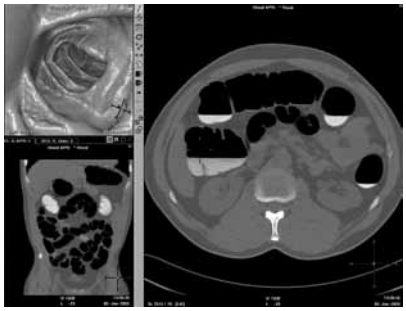
The CTC examination was performed in the supine and prone positions on an Mx8000 4-MDCT scanner (Philips Medical Systems) with the following parameters:120 kV; collimation, 4 × 2.5 mm; rotation time, 0.75 second; pitch, 1.25; slice thickness, 3.2 mm; and reconstruction interval, 1.6 mm. Patients with a circumference < 103 cm were scanned with 50 mAs (n = 15); patients with a circumference > 103 cm, with 70 mAs (n = 25). A total of 20 mg of butyl scopolamine bromide (Buscopan, Boehringer Ingelheim) (n = 22) or, if contraindicated, 1 mg of glucagon hydrochloride (GlucaGen, Novo Nordisk) (n = 17) was administered IV immediately before scanning. In one patient neither was administered because of contraindications. On average, 4 L of carbon dioxide (CO2) was insufflated via an automatic insufflator (PROTOCO2L, E-Z-EM) to achieve adequate distention. The time that patients spent in the CT room was recorded with a stopwatch.

Conventional Colonoscopy

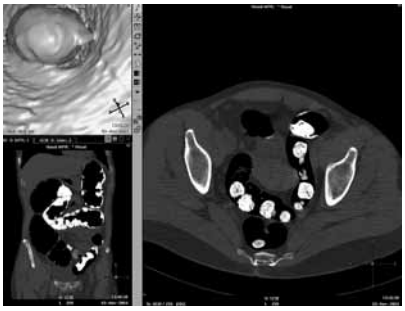
Colonoscopy was performed 4–30 days (mean, 17 days) after CTC. Bowel preparation for colonoscopy consisted of 3 (n = 2), 4 (n = 36), or 6 L (n = 2) of polyethylene glycol–electrolyte solution (Klean-Prep, Helsinn Birex Pharmaceuticals) administered the day before the examination. Patients received midazolam (n = 28), fentanyl (n = 20), or alfentanil (Rapifen, AstraZeneca) (n = 6) on request. Buscopan was administered IV to all patients. The colonoscopy was performed either by an experienced gastroenterologist or surgeon or by a resident under the direct supervision of a staff member. During the examination, a research nurse involved in our study was present.

Outcome Parameters

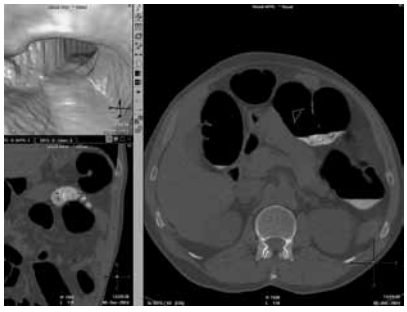
SUBJECTIVE IMAGE QUALITY – The CTC data were evaluated by a research fellow in CTC who was blinded to all clinical data, including the bowel preparation used. The reviewer had a prior experience of 150 non-tagged cathartic-prepared CTC examinations with colonoscopic verification. A primary 2D evaluation technique (axial views) was applied using multiplanar reformatted (MPR) images and 3D endoluminal views for problem solving (ViewForum 5.1, Philips Medical Systems). No electronic cleansing software was applied.



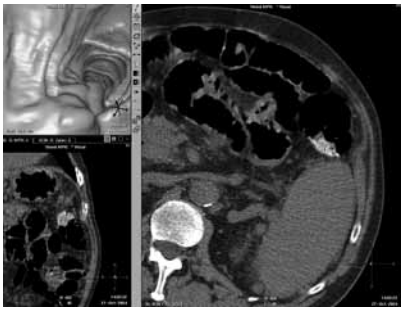
A In 50-year-old man with abdominal pain and family history of colorectal cancer, reviewer evaluated fecal tagging as excellent.



B In 51-year-old woman with history of colorectal polyps, reviewer evaluated fecal tagging as good.



C In 61-year-old man with history of colorectal cancer, reviewer evaluated fecal tagging as moderate.



D In 72-year-old man with history of colorectal cancer, reviewer evaluated fecal tagging as poor.

FIGURE 2.1
Axial multiplanar reformatted coronal and endoluminal 3D images show subjective evaluation in four different patients with various grades of fecal tagging.

The reviewer filled out a standardized questionnaire for every examination with regard to image quality. Table 2.2 displays the scoring system used by the reviewer. The image quality parameters of fecal tagging (Figure 2.1), luminal distention, amount of fecal residue (combined estimation for liquid and solid feces), and image readability were scored by the reviewer for the complete (supine and prone together) examination. This evaluation was considered the subjective image quality assessment on a per-patient basis.

The same image quality parameters were again evaluated on a per-segment basis. To this end, the colon was divided in six segments: cecum; ascending, transverse, descending,and sigmoid colon; and rectum. Fecal tagging

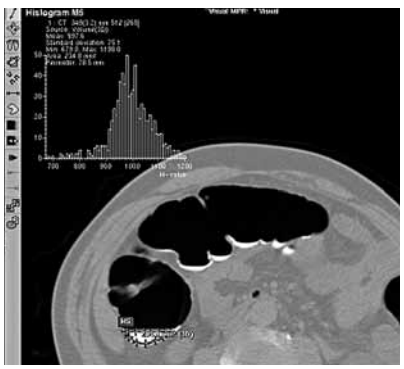


FIGURE 2.2
Axial multiplanar reformatted coronal and endoluminal 3D image shows mean attenuation and SD of fecal material in 63-year-old man with family history of colorectal cancer. Inset histogram shows attenuation values in Hounsfield units of tagged material (x-axis) and frequencies of measurement (y-axis). Observer placed region of interest with surface area of at least 60 mm2 in largest fluid level on randomized computer-picked slice, and CT colonography software subsequently provided mean attenuation and SD.

and the amount of fecal material were evaluated only with the patient in the supine position because we presumed these parameters would not be influenced by a position change. However, because alteration of the position of the patient might substantially influence distention of different colonic segments [22, 23], luminal distention and segmental image readability were again evaluated with the patient in the prone position. Furthermore, to examine the effect of increasing catharsis on the consistency of fecal material, the reviewer scored the ratio between feces with a solid consistency and feces with a liquid consistency. This was performed on a persegment basis with the patient in the supine position. A 5-point scale was applied: 1, 0–20% solid; 2,21–40% solid; 3, 41–60% solid; 4, 61–80% solid; and 5, 81–100% solid. If no stool was present in a segment, no evaluation of consistency was performed for that particular segment.

NUMERIC EVALUATION OF DEGREE AND HOMOGENEITY OF TAGGING – To evaluate the quality of fecal tagging, the mean and SD of attenuation of tagged material (in Hounsfield units) were measured on seven separate slices that were randomly selected by a computer program (Windows Excel 2002, Microsoft) for every patient. To this end, a resident in the fourth year of training, who was blinded as to the preparation used, placed a region of interest (ROI) of at least 60 mm^2 in the area. that contained the most fecal material on the randomly picked axial slices. The resident recorded whether the ROI was placed in liquid stool, stool adherent to the colon wall, or solid stool. The mean and SD of attenuation values within the ROIs were then calculated by our CTC software (ViewForum 5.1) (Figure 2.2). If not enough fecal material was present to draw an ROI, the randomization program provided another slice for evaluation. The SD of attenuation values of tagged material is a measure of homogeneity. However, because the mean attenuation of tagging can vary considerably among patients, we considered the relative SD (SD / mean) to be a better measure of homogeneity. After all, by adapting the window width when viewing images, tagging with a larger mean attenuation value and a large SD may have a similar visual appearance to tagging with a smaller attenuation value and a proportionally smaller SD. Furthermore, separate assessments were performed for solid, adherent, and liquid feces.

INTERPRETATION TIME – The time needed to interpret a complete examination for polyp detection, excluding report time and image quality evaluation, was recorded with a stopwatch by the reviewer. Because of the limited number of patients and the expected low prevalence of polyps in our cohort, performance characteristics were not part of our study design and are not discussed in this article.

PATIENT EXPERIENCE - Patients were asked to fill out a questionnaire on the day of the CTC examination with regard to the burden and side effects of the CTC preparation – that is, flatulence, abdominal pain, and diarrhea – as well as the discomfort caused by the intake of contrast material and laxative agents on a 5-point scale (Table 2.2). To understand if the everyday bowel habits of patients were of influence on image quality parameters, patients were asked to fill out their normal frequency of defecation: 1, more than one defecation per day; 2, once per day; 3, once per 2 days; 4, once in 3 days; 5, less than one defecation per 3 days; and 6, less than one defecation per 5 days. On the day of the colonoscopy, patients filled out a questionnaire with regard to the burden associated with the bowel preparation for colonoscopy (Table 2.2).

PATIENT PREFERENCE – Five weeks after colonoscopy, patients were sent a questionnaire in which they were asked which preparation had been most burdensome (CTC or conventional colonoscopy) and which examination (CTC or conventional colonoscopy) they would prefer in the future, and they were asked to indicate the most burdensome event of both examinations (Table 2.2). The patient preference and experience questionnaires were designed by the department of social medicine and had previously been used in two studies that were performed in our institution [7, 24].

TABLE 2.2 **Scales Used by Observer to Rate Image Quality and by Patients to Rate Burden and Preference**

Observer	Scale
Fecal tagging[a]	a = Poor, not interpretable b = Moderate, diagnostic with untagged feces < 10 mm c = Good, diagnostic with untagged feces < 6 mm d = Very good, diagnostic without limitations
Presence of feces[a]	a = Large amount of feces, segment fully filled b = Moderate amount of feces, ≈ 50% of lumen filled c = Small amount of feces d = Only layer on wall e = No feces at all
Luminal distention[a,b]	a = Collapsed b = Poorly distended c = Only moderately distended, but segment is distended over its full length d = Good e = Very good
Image readability[a,b]	a = Poor, not diagnostic b = Moderate, diagnostic for lesions ≥ 10 mm c = Good, diagnostic for lesions ≥ 6 mm d = Excellent, no limitations
Patients	**Scale**
Burden caused by intake of bisacodyl, magnesium citrate, barium sulfate,diatrizoate meglumine	a = None b = Mild c = Moderate d = Severe e = Extreme
Side effects: diarrhea, flatulence, abdominal pain, overall burden	a = None b = Mild c = Moderate d = Severe e = Extreme
Most burdensome preparation	CT colonography (CTC) or conventional colonoscopy
Most burdensome examination	CT colonography (CTC) or conventional colonoscopy
Most burdensome event	Bowel preparation prior to conventional colonoscopy Limited bowel preparation prior to CTC CTC examination Conventional colonoscopy examination
Preference for CTC or conventional colonoscopy examination in the future	Definitely CTC Probably CTC Possibly CTC Indifferent Possibly conventional colonoscopy Probably conventional colonoscopy Definitely conventional colonoscopy

a Subjective image quality parameters were scored per patient and per segment in the supine position.

b Luminal distention and image readability were again evaluated on a per-segment basis in the prone position.

Statistical Analysis

SUBJECTIVE IMAGE QUALITY – With regard to the per-patient analysis, possible differences in image quality parameters between the four preparations were assessed using ordinal regression analysis. In this analysis, first the preparation with the highest regression coefficient was determined; this was subsequently used as the reference group.

With regard to the per-segment analysis, ordinal regression analysis was applied using generalized estimating equations (GEE) to revise the data clustering and dependency [25]. This was done because more than one segment was obtained from each patient. The link function was set at log link, and an independent working correction matrix was used. Furthermore, associations between normal defecation frequency and subjective image quality parameters were tested with the chi-square test. If significant associations were present, additional ordinal regression analyses using GEE for revising the patient's normal defecation frequency were performed. Associations between the subjective image quality parameters (e.g., association between fecal tagging and the amount of residual feces) within each group were assessed using the chi-square test.

NUMERIC EVALUATION OF DEGREE AND HOMOGENEITY OF TAGGING – Because more than one measurement of Hounsfield units and SD was obtained from each patient, linear regression analysis was applied using GEE to revise the data clustering and dependency. Furthermore, ROIs were placed at random and were not evenly distributed among segments. This might have resulted in differences in measured homogeneity because of the physiologic variation of mean tagging attenuation within a given patient due to the normal dehydrating action of the colon. Therefore, the GEE analysis was adjusted to correct for this segmental distribution. Furthermore, because liquid material tends to be more homogeneously tagged than solid material, the GEE analysis was adjusted to correct for stool consistency (solid, adherent, or liquid feces). For each group, estimates of means with corresponding standard errors could be calculated from the results (intercept and slopes) obtained by the analysis.

PATIENT EXPERIENCE AND PREFERENCE – Differences in patient experience (burden) were analyzed using ordinal regression analysis and the preparation with the lowest regression coefficient (most burdensome) as the reference group. The chi-square test was used to test for significant differences in patients' indications of the most burdensome preparation and examination (CTC or conventional colonoscopy) between groups. Patient preference for either CTC or colonoscopy was tested using the chi-square test after the data were first dichotomized as preference for CTC versus preference for colonoscopy.

INTERPRETATION TIME – Differences in interpretation times for the different preparations were tested for significance using the independent samples Student's t test. Statistical analyses were performed with SPSS version 12.0.2 for Windows (SPSS) and SAS version 8.02 for Windows (SAS Institute). The Proc Genmod command was used to apply the generalized estimating equations.

A p-value of < 0.05 was considered to indicate a statistically significant difference.

Results

All patients included in this study accepted randomization. Eleven patients were randomized into group 1, 10 patients into group 2, 10 patients into group 3, and nine patients into group 4. Twenty-seven patients were men (67.5%) and 13 patients were women (32.5%), with an average age of 62 years (age range, 40–83 years). Fourteen (35%) patients had symptoms (abdominal pain, 6; hematochezia, 4; and altered bowel habits, 4), and 26 (65%) patients were asymptomatic. Thirty-three (83%) patients had a personal history of colorectal polyps (n = 21), cancer

(n = 6), or both (n = 6). A total of 11 patients had undergone surgery for a colorectal cancer: right-sided hemicolectomy (n = 1), transverse colectomy (n = 2), low anterior resection (n = 7), total mesorectal excision (n = 1). Thirty of 38 patients who filled out the questionnaire with regard to bowel frequency had a defecation frequency of one defecation or more per day, five patients had a defecation frequency of once in 2 days, two patients (groups 2 and 3) had a frequency of once in 3 days, and one patient had a bowel frequency of once in 3–5 days (group 2). Because one patient had undergone a hemicolectomy, a total of 476 bowel segments were available for analysis (per position: group 1, 66 segments; group 2, 58; group 3, 60; and group 4, 54 segments). The average room examination time for CTC was 15 minutes.

image quality results

SUBJECTIVE IMAGE QUALITY – On a per-patient basis (n = 40), no significant differences were found between groups with regard to image quality. All examinations were evaluated as good (diagnostic for lesions ≥ 6 mm) or excellent (diagnostic with no limitations) image readability. Only one examination (group 2) was of nondiagnostic value (because of poor distention) and two examinations (group 3) were of moderate image readability (because of moderate fecal tagging) (Table 2.3). Within each group, a higher degree of homogeneity was significantly associated with better image readability ($p < 0.006$). For group 2, poorer distention was associated with decreased image readability ($p = 0.002$).

On a per-segment basis, significantly more residual feces was present in group 2 in comparison with group 4 ($p = 0.04$). No other significant differences in image quality were found between groups (Figure 2.3). Fecal material was significantly more liquid in group 4 in comparison with groups 1, 2, and 3 ($p = 0.004$, $p = 0.002$, and $p < 0.001$) (Figure 2.4). A significant association was observed between daily bowel habits and the amount of residual feces in the colon for groups 2 and 3 ($p < 0.0001$). In other words, a lower defecation frequency in normal life was associated with a higher amount of residual feces in our study. Subsequently, when the normal defecation frequency of patients was included in the GEE analysis, the finding that group 2 contained more residual feces than group 4 was no longer significant ($p = 0.160$).

TABLE 2.3 **Overall Image Readability of 40 Complete (Supine and Prone) CT Colonography Examinations**

Group	Excellent; No Limitations	Good; Diagnostic for Lesions ≥ 6mm	Moderate; Diagnostic for Lesions ≥ 10mm	Poor; Not Diagnostic
1 (n=11)	5	6	0	0
2 (n=10)	3	6	0	1
3 (n=10)	4	4	2	0
4 (n=9)	5	4	0	0

Data are number of examinations that were subjectively evaluated by the reviewer.

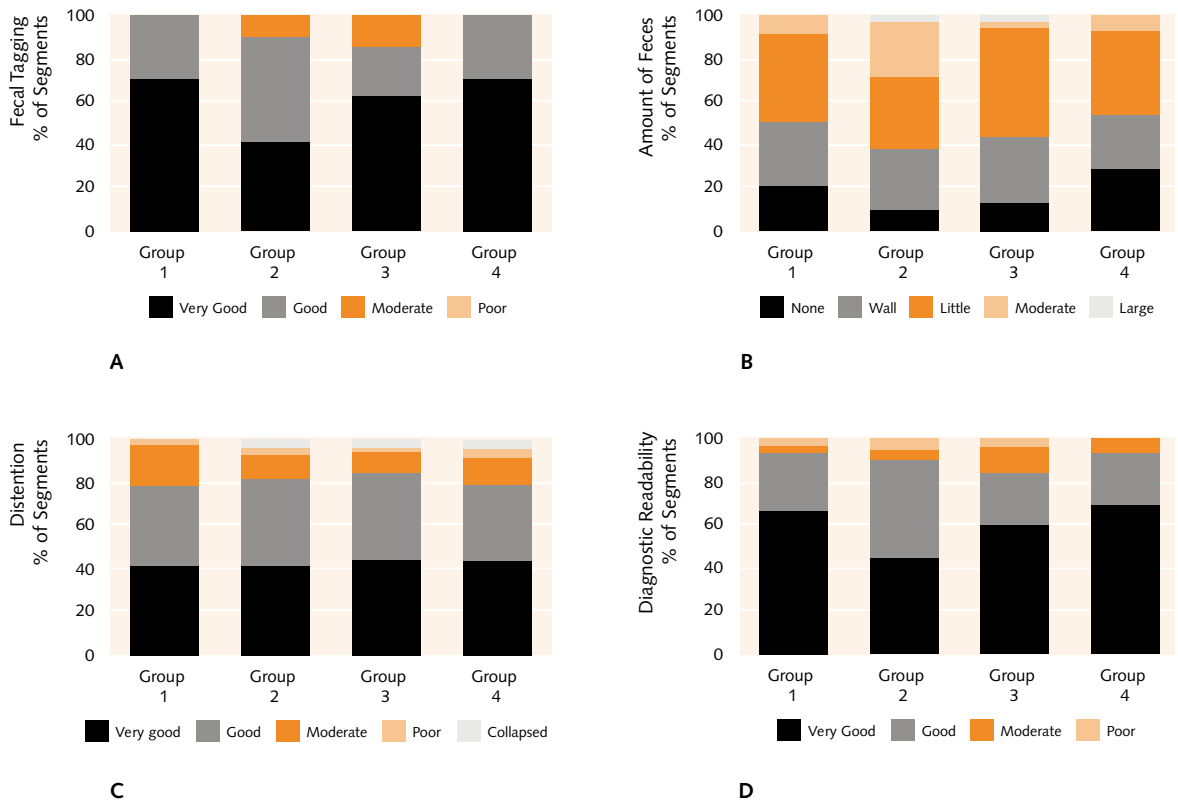


FIGURE 2.3
Graphs show subjective image quality scores on per-segment basis (A-D). Fecal tagging (A) and amount of feces (B) were scored on supine images (for group 1, 66 segments; group 2, 58 segments; group 3, 60 segments; and group 4, 54 segments). Distention (C) and diagnostic readability (D) were scored on supine and prone images (132, 120, 210, and 108 segments, respectively). Group 2 performed significantly poorer with regard to amount of residual feces in comparison with group 4 ($p = 0.04$).

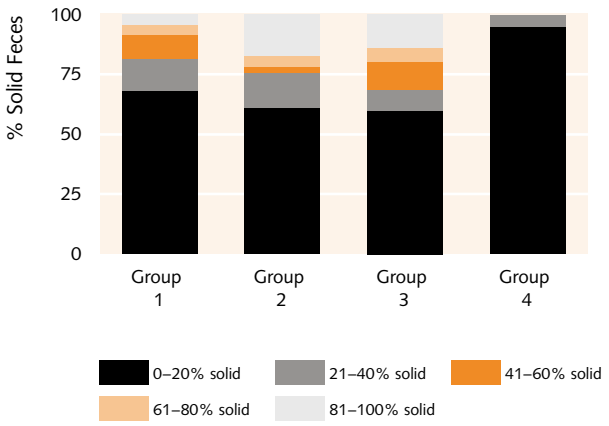


FIGURE 2.4
Graph shows proportion of solid feces in colon on per-segment basis (if feces was not solid it was considered liquid). Most segments contained only little feces with solid consistency (0–20%) in every preparation. However, group 4 contained significantly less solid feces compared with groups 1, 2, and 3 ($p = 0.004$, $p = 0.002$, $p < 0.001$).

TABLE 2.4 **Estimates of Mean Attenuation and Mean SD of Fecal Material**

		Mean attenuation (HU)	SD	Relative SD (SD / mean)
Preparation 1				
fecal material	(n=77)	720	106	0.15
liquid consistency	(71%)	683	98	0.14
adherent consistency	(19%)	803	115	0.16
solid consistency	(10%)	1030[a]	215[b]	0.18
Preparation 2				
fecal material	(n=66)	686	97	0.14
liquid consistency	(60%)	710	74	0.12
adherent consistency	(14%)	636	89	0.15
solid consistency	(26%)	634	140	0.21[c]
Preparation 3				
fecal material	(n=70)	654	80	0.14
liquid consistency	(60%)	684	73	0.11
adherent consistency	(29%)	677	103	0.18
solid consistency	(11%)	485	74	0.23[c]
Preparation 4				
fecal material	(n=60)	557	62	0.16
liquid consistency	(90%)	509	56	0.19
adherent consistency	(10%)	685	94	0.11
solid consistency	(0%)	n/a	n/a	n/a

Numbers in parentheses are the number of measurements. Percentages in parentheses refer to the percentage of measurements in the indicated form of fecal material.
NA = not applicable.

a Significantly higher mean value for solid consistency in comparison with liquid ($p = 0.003$) or adherent ($p=0.046$) consistency within preparation group 1.
b Significantly higher SD value for solid consistency in comparison with liquid ($p = 0.025$) or adherent ($p =.008$) consistency within preparation group 1.
c Significantly higher relative SD value for solid consistency in comparison with liquid consistency within preparation group 2 ($p = 0.006$) and within preparation group 3 ($p = 0.011$).

NUMERIC EVALUATION OF DEGREE AND HOMOGENEITY OF TAGGING – The degree of tagging, expressed as the mean attenuation of tagged material, decreased when more laxatives were used (Table 2.4); however, this decrease was not statistically significant (all $p \geq 0.253$). A gradual decrease in mean SD of pixel values was also found from group 1 to group 4.
This decline, however, was not significant between groups (all $p \geq 0.067$). Furthermore, the relative SD (SD / mean) did not differ significantly among groups (all $p \geq 0.157$), indicating no difference in homogeneity of tagging among groups. Regardless of the preparation, the relative SD was significantly higher for solid feces (0.22) versus adherent (0.16) or liquid (0.14) feces, showing that tagging was less homogeneous for solid feces than for adherent stool ($p = 0.008$) or liquid feces ($p = 0.001$) (Table 2.5).

TABLE 2.5 **Tagging Characteristics for Stool Consistency Regardless of the Preparation**

	Mean (HU)	SD	Relative SD (SD/mean)
Solid consistency (n=32)	737	143[a]	0.22[b]
Adherent consistency (n=50)	688	97	0.16
Liquid consistency (n=191)	631	79	0.16

Data in parentheses are number of measurements.

a SD was significantly higher for solid than for adherent ($p = 0.041$) or liquid ($p = 0.005$) feces.
b Relative SD was significantly higher for solid than for adherent ($p = 0.008$) or liquid ($p = 0.001$) feces.

INTERPRETATION TIME – The mean interpretation time for group 1 was 16 minutes, for group 2 was 24 minutes, for group 3 was 26 minutes, and for group 4 was 21 minutes (all $p \geq 0.107$).

Patient Experience

All patients experienced diarrhea except for three patients in group 2. The burden of diarrhea was evaluated as none or mild by six patients (55%) in group 1, six patients (60%) in group 2, four patients (40%) in group 3, and one patient (11%) in group 4. Diarrhea was considered significantly more burdensome in group 4 than in groups 1 ($p = 0.042$) and 2 ($p = 0.031$) but not compared with group 3 ($p = 0.179$). With regard to abdominal pain and flatulence, no significant differences were found among groups. With regard to abdominal pain (n = 37), one patient in group 4 had severe abdominal pain, and two patients (groups 2 and 4) had moderate abdominal pain. The remaining patients had little (n = 6) or no (n = 28) pain. With regard to flatulence (n = 37), one patient in group 4 experienced a severe burden, one patient in group 2 had a moderate burden, and the other patients experienced little (n = 5) or no (n = 30) burden.

The total burden of the bowel preparation was rated as none or mild by all patients (100%) in group 1, nine (90%) in group 2, eight (80%) in group 3, and five (55%) in group 4 – a significantly higher total burden in group 4 than in groups 1 ($p = 0.002$) and 3 ($p = 0.02$). No significant difference ($p = 0.082$) was found between groups 2 and 4, most likely because one patient in group 2 rated the CTC bowel preparation as extremely burdensome. This patient stated, before the CTC preparation, an explicit preference for the colonoscopy preparation because in his opinion that preparation was simpler and shorter. Excluding this patient, a significantly lower overall burden ($p = 0.04$) was found for group 2 in comparison with group 4. Most patients experienced no or only a mild burden with regard to the intake of barium (32/34), diatrizoate meglumine (36/40), bisacodyl (34/37), or magnesium citrate (9/15), with no significant differences between the groups.

Patient Preferences

Except four patients (one in group 2 and three in group 3), all patients found colonoscopy a more burdensome examination than CTC. If patients were asked what examination they would prefer in the future, most patients (30/40) indicated a preference for CTC, two patients were indifferent, and eight patients preferred conventional colonoscopy (Figure 2.5). The reasons patients preferred colonoscopy were: direct polypectomy (n = 4), shorter preparation time (n = 2), false-positive CTC (n = 1), and no particular reason (n = 1). No significant differences in preference were observed among the groups.

All (100%) patients in group 1, nine (90%) patients in group 2, eight (80%) in group 3, and seven (78%) patients in group 4 preferred CTC bowel preparation over the polyethylene glycol preparation, with no significant differences among groups ($p = 0.054$). In all groups, most patients evaluated the bowel preparation for colonoscopy (n = 30) or colonoscopy itself (n = 5) as the most

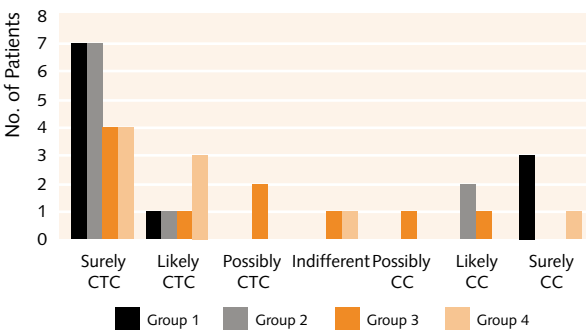


FIGURE 2.5
Graph shows patient preference 5 weeks after colonoscopy. Patients (n = 40) indicated whether they preferred CT colonography (CTC) or colonoscopy (CC) as colorectal examination in the future. Most patients (n = 30) preferred CTC.

TABLE 2.6 **Most Burdensome Event for 40 Patients**

Group	CTC Bowel Preparation	CTC Examination	Conventional Colonoscopy Bowel Preparation	Conventional Colonoscopy Examination
1	0	0	8	3
2	1	1	8	0
3	1	1	8	0
4	1	0	6	2

Appraisal by patients 5 weeks after CT colonography (CTC) and conventional colonoscopy. Most patients (n = 30) indicated the bowel preparation for colonoscopy as the most burdensome event regardless of the bowel preparation used for CTC.

burdensome event, with no significant differences among the groups (Table 2.6).

Discussion

In this feasibility study regarding image quality and patient acceptance, we investigated four different dosages of mild laxatives, consisting of bisacodyl (a bowel stimulant) and magnesium citrate (a hyperosmotic saline laxative), to prepare patients for CTC. Our results showed good to excellent image readability of CTC examinations (37/40) regardless of the preparation used. Increasing the amounts of laxatives did not lead to a higher attenuation of tagging or to more homogeneous tagging, and subjective image quality did not show significant improvement. A higher dosage of laxatives was significantly associated with a higher burden of diarrhea and a higher overall burden of the bowel preparation. Nevertheless, irrespective of the amount of laxatives for CTC, the majority of patients (35/40) preferred the bowel preparation for CTC over the cleansing laxative bowel preparation (polyethylene glycol) for colonoscopy.

In two previous studies, the same or a higher dose of magnesium citrate (without bisacodyl) as used in our group 4, was used; both reported significantly better image quality of CTC images compared with images made without the use of laxatives [20, 21]. Magnesium citrate liquefies residual feces and decreases the amount of remaining feces. This is associated with better readability [21] and can be expected to lead to increased homogeneity of tagged material. In concordance with these studies, we expected that image quality would gradually improve in our study and be best with group 4.

In our study, however, a higher level of catharsis resulted in neither better image quality nor improvement of tagging characteristics (subjective or numeric). Our data showed that residual feces in group 4 was more liquefied compared with the other groups (all $p \leq 0.004$), but the relative SD (SD / mean) did not decrease, indicating that homogeneity did not improve. A decline, although not significant ($p \geq 0.253$), of the mean attenuation of the tagged material from group 1 to group 4 was observed that we did not anticipate. The lower attenuation values in groups with more catharsis might be explained because lower concentrations of contrast agent were present as a result of dilution caused by the hyperosmotic effect of magnesium citrate.

Regardless of the preparation, solid feces was less homogeneously tagged than liquid stool. However, we

believe that homogeneity of solid feces was still satisfactory in all groups. Attenuation of solid feces was high (group 1, 1,030 H; group 2, 634 H; and group 3, 485 H) with relatively small SDs (215, 140, and 74, respectively) and nearly constant relative SDs (0.18, 0.21, and 0.23, respectively). As a result, the lowest attenuation value of tagged solid material was still well above soft-tissue density in all groups.

Most segments (88%) contained no or only small amounts of feces. Furthermore, no decrease in the amount of fecal material was observed using a higher dose of laxatives. This is in contrast with the results of Dachman et al. [21], who found no feces in only 65% of segments using magnesium citrate. That we used a combination of barium and diatrizoate meglumine for tagging purposes instead of only barium might explain this difference. Probably the laxative side effect of diatrizoate meglumine combined with (small amounts of) laxatives caused relatively clean colons with good homogeneity of tagged feces.

Increasing the amounts of laxatives did not improve subjective image quality but did increase the burden perceived by patients. In our study, most patients experienced diarrhea, regardless of the preparation used. These observations are not in line with the previously discussed study by Zalis et al. [20]. In that study, only one of 22 patients reported mild transient diarrhea after preparation with magnesium citrate. Two important differences between the studies may explain this discrepancy. First, we used ionic iodinated contrast material, which is known to have a strong osmotic effect, instead of a nonionic contrast agent. Second, in our study, all patients received bisacodyl, whereas in the Zalis et al. study, no bisacodyl was administered at all. It is likely that a combination of both factors led to some degree of diarrhea. Although diarrhea was present in all preparations, the burden of diarrhea and the overall burden of the preparation significantly increased with higher doses of laxatives. In fact, one could argue whether laxatives should be added at all to the preparation if ionic contrast agents are used in high concentrations. At the time of the writing of this article, we no longer add any form of laxatives to the preparation, and we believe this has not impaired the image quality of CTC.

An interval of 5 weeks was applied after colonoscopy to ask patients about their preference because the memory of adverse reactions may decrease over time [7, 8, 24]. The majority of patients (75%) preferred CTC as a colorectal examination and considered the polyethylene glycol–electrolyte solution for colonoscopy the most unpleasant factor (75%). In our institution, this bowel preparation is used by the gastroenterologists for all patients because it is safe to use and rigorously cleans the colon. Other bowel preparations for colonoscopy, such as an osmotic laxative ([sodium phosphates solution] Phospho-soda, Fleet), are considered more patient-friendly [26, 27]. Had these been used in our study, they might have caused a shift in patient preference toward colonoscopy. A disadvantage of Phospho-soda is that it can cause electrolyte imbalance and is contraindicated in patients with congestive heart failure or renal failure [28, 29].

Several limitations of our study must be considered. A limited cohort of 40 patients, randomized into four groups, participated in this study. Despite the relatively low numbers of patients per group, we did find a significantly higher patient burden with increasing laxative dosage. With regard to image quality, no trend of improvement in quality was observed with more catharsis. However, in group 2, segments contained significantly more feces compared with group 4. This was mainly attributed to the fact that two of the three patients in our study with a relatively low defecation frequency in normal life were placed in group 2. Despite more feces present in the colon, image readability of the examinations was not significantly affected. Although our data are limited, this might suggest that with adequate tagging, patients with bowel habits of one defecation in 3–5 days can be prepared with minimal

amounts of laxatives without affecting image quality. Furthermore, all patients without diarrhea (n = 3) were in group 2. Patients were not asked about compliance with the preparation. Because non-compliance could have contributed to more residual feces in group 2 but could also provide a possible explanation for the moderate image quality in two patients in group 3, we considered that a limitation of our study.

Furthermore, only one observer subjectively evaluated all data on image quality. It is possible that another observer would rate the data differently. Another possible qualifier is that we did not include a full catharsis regimen as a reference standard to which the studied preparations were compared. However, we believe that for this study a clinically relevant scoring system was constructed with regard to image quality of the examinations: using classifications as not diagnostic, diagnostic for all lesions, diagnostic for lesions ≥ 6 mm, or only diagnostic for lesions ≥ 10 mm. The experience of the reviewer before our study consisted of 150 cathartic CTC examinations, and consequently image quality of these examinations served as a reference standard for the reviewer. With regard to patient experience, the bowel preparation used for colonoscopy in our study was a full cathartic bowel preparation using a polyethylene glycol–electrolyte solution that can also be used for CTC. Most patients in all groups indicated the colonoscopy preparation as more burdensome when compared with the CTC bowel preparation.

We used an iodine-based contrast medium for fecal tagging in addition to barium. Some investigators prefer using only barium because of the side effects and possible adverse reactions to iodine-based contrast agents [10]. However, we believe that a combination of both contrast agents resulted in adequate tagging of both solid and liquid feces as stated in other studies [2, 29]. Although the laxative side effect of an iodine-based contrast agent will probably increase the patient burden in some way, an advantage is that CTC images are easier to interpret and a possible cleansing algorithm might be more effective [20, 30].

Finally, in our study we focused on image quality. Polyp conspicuity was not investigated. So far, one larger study without catharsis has reported an excellent sensitivity of 90% and specificity of 92% for patients with polyps of any size [18]. Although further research is warranted, these results underscore our findings that good image quality can be obtained with small amounts of laxatives.

We conclude that CTC with limited bowel preparation, using barium and ionic iodinated contrast agents for fecal tagging, requires only minimal doses of laxatives – in our study only 20 mg of bisacodyl – to obtain good image quality and minimize patient burden. This is important because a mild bowel preparation will undoubtedly increase patient willingness to participate in a screening program.

ACKNOWLEDGMENTS – We thank Karin Horsthuis for her critical review of this manuscript and Henk W. Venema for his help with the data analysis and critical review of this manuscript.

References

1. Fenlon HM, Nunes DP, Schroy PC, et al. A comparison of virtual and conventional colonoscopy for the detection of colorectal polyps. *N Engl J Med* 1999; 341:1496–1503
2. Pickhardt PJ, Choi JR, Hwang I, et al. Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. *N Engl J Med* 2003; 349:2191–2200
3. Pineau BC, Paskett ED, Chen GJ, et al. Virtual colonoscopy using oral contrast compared with colonoscopy for the detection of patients with colorectal polyps. *Gastroenterology* 2003; 125: 304–310
4. Johnson CD, Harmsen WS, Wilson LA, et al.Prospective blinded evaluation of computed tomographic colonography for screen detection of colorectal polyps. *Gastroenterology* 2003; 125:311–319
5. Van Gelder RE, Nio CY, Florie, et al. Computed tomographic colonography compared with colonoscopy in patients at increased risk for colorectal cancer. *Gastroenterology* 2004; 127: 41–48
6. Harewood GC, Wiersema MJ, Melton LJ III. A prospective, controlled assessment of factors influencing acceptance of screening colonoscopy. *Am J Gastroenterol* 2002; 97:3186–3194
7. van Gelder RE, Birnie E, Florie J, et al. CT colonography and colonoscopy: assessment of patient preference in a 5-week follow-up study. *Radiology* 2004; 233:328–337
8. Ristvedt SL, McFarland EG, Weinstock LB, Thyssen EP. Patient preferences for CT colonography, conventional colonoscopy, and bowel preparation. *Am J Gastroenterol* 2003; 98:578–585
9. Weitzman ER, Zapka J, Estabrook B, Goins KV. Risk and reluctance: understanding impediments to colorectal cancer screening. *Prev Med* 2001; 32:502–513
10. Gluecker TM, Johnson CD, Harmsen WS, et al. Colorectal cancer screening with CT colonography, colonoscopy, and double-contrast barium enema examination: prospective assessment of patient perceptions and preference. *Radiology* 2003; 227:378–384
11. Rex DK. Current colorectal cancer screening strategies: overview and obstacles to implementation. *Rev Gastroenterol Disord* 2002; 2[suppl 1]:S2–S11
12. Callstrom MR, Johnson CD, Fletcher JG, et al. CT colonography without cathartic group: feasibility study. *Radiology* 2001; 219:693–698
13. Lefere P, Gryspeerdt S, Marrannes J, et al. CT colonography after fecal tagging with a reduced cathartic cleansing and a reduced volume of barium. *AJR* 2005; 184:1836–1842
14. Lefere PA, Gryspeerdt SS, Dewyspelaere J, et al. Dietary fecal tagging as a cleansing method before CT colonography: initial results polyp detection and patient acceptance. *Radiology* 2002; 224: 393–403
15. Thomeer M, Carbone I, Bosmans H, et al. Stool tagging applied in thin-slice multidetector computed tomography colonography. *J Comput Assist Tomogr* 2003; 27:132–139
16. Bielen D, Thomeer M, Vanbeckevoort D, et al. Dry group for virtual CT colonography with fecal tagging using water-soluble contrast medium: initial results. *Eur Radiol* 2003; 13:453–458
17. McFarland EG, Zalis ME. CT colonography: progress toward colorectal evaluation without catharsis. *Gastroenterology* 2004; 127: 1623–1626
18. Iannaccone R, Laghi A, Catalano C, et al. Computed tomographic colonography without cathartic group for the detection of colorectal polyps. *Gastroenterology* 2004; 127: 1300–1311
19. Taylor S, Laghi A, Lefere PA, Halligan S, Stoker J. European Society of Gastrointestinal and Abdominal Radiology (ESGAR): consensus statement on CT colonography. *Eur Radiol* 2007; 17: 575–579
20. Zalis ME, Perumpillichira JJ, Magee C, Kohlberg G, Hahn PF. Tagging-based, electronically cleansed CT colonography: evaluation of patient comfort and image readability. *Radiology* 2006; 239: 149–159
21. Dachman AH, Dawson DO, Lefere P, et al. Comparison of routine and unprepped CT colonography augmented by low fiber diet and stool tagging: a pilot study. *Abdom Imaging* 2007; 32: 96–104

22. Chen SC, Lu DS, Hecht JR, Kadell BM. CT colonography: value of scanning in both the supine and prone positions. *AJR* 1999; 172:595–599
23. Yee J, Kumar NN, Hung RK, Akerkar GA, Kumar PR, Wall SD. Comparison of supine and prone scanning separately and in combination at CT colonography. *Radiology* 2003; 226: 653–661
24. Florie J, Birnie E, van Gelder RE, et al. MR colonography with limited bowel preparation: patient acceptance compared with that of fullpreparation colonoscopy. *Radiology* 2007; 245: 150–159
25. Pepe MS. *The statistical evaluation of medical tests for classification and prediction: Oxford statistical science series 28*. New York, NY: Oxford University Press, 2003: 35–65
26. Ginnerup Pedersen B, Moller Christiansen TE, Viborg Mortensen F, Christensen H, Laurberg S. Bowel cleansing methods prior to CT colonography. *Acta Radiol* 2002; 43:306–311
27. Macari M, Lavelle M, Pedrosa I, et al. Effect of different bowel preparations on residual fluid at CT colonography. Radiology 2001; 218:274–277
28. Curran MP, Plosker GL. Oral sodium phosphate solution: a review of its use as a colorectal cleanser. Drugs 2004; 64: 1697–1714
29. Pickhardt PJ. CT colonography without catharsis: the ultimate study or useful additional option? Gastroenterology 2005; 128: 521–522
30. Pickhardt PJ, Choi JR. Electronic cleansing and stool tagging in CT colonography: advantages and pitfalls with primary three-dimensional evaluation. AJR 2003; 181: 799–805

Chapter 3

Lesion Conspicuity and Efficiency of CT colonography with Electronic Cleansing Based on a Three-Material Transition Model

Iwo W. O. Serlie
Ayso H. de Vries
Lucas J. van Vliet
Chung Y. Nio
Roel Truyen
Jaap Stoker
Franciscus M. Vos

American Journal of Roentgenology 2008; 191:1493–1502

Abstract

Objective

The purpose of this article is to report the effect on lesion conspicuity and the practical efficiency of electronic cleansing for CT colonography (CT colonography).

Material and methods

Patients were included from the Walter Reed Army Medical Center public database. All patients had undergone extensive bowel preparation with fecal tagging. A primary 3D display method was used. For study I, the data consisted of all patients with polyps ≥ 6 mm. Two experienced CT colonography observers (observer 1 and observer 2) scored the lesion conspicuity considering supine and prone positions separately.

For study II, data consisted of 19 randomly chosen patients from the database. The same observers evaluated the data before and after electronic cleansing. Evaluation time, assessment effort, and observer confidence were recorded.

Results

In study I, there were 59 lesions partly or completely covered by tagged material (to be uncovered by electronic cleansing) and 70 lesions surrounded by air (no electronic cleansing required). The conspicuity did not differ significantly between lesions that were uncovered by electronic cleansing and lesions surrounded by air (observer 1, $p < 0.5$; observer 2, $p < 0.6$).

In study II, the median evaluation time per patient after electronic cleansing was significantly shorter than for original data (observer 1, 20 reduced to 12 minutes; observer 2, 17 reduced to 12 minutes). Assessment effort was significantly smaller for both observers ($p < 0.0000001$), and observer confidence was significantly larger (observer 1, $p < 0.007$; observer 2, $p < 0.0002$) after electronic cleansing.

Conclusion

Lesions uncovered by electronic cleansing have comparable conspicuity with lesions surrounded by air. CT colonography with electronic cleansing sustains a shorter evaluation time, lower assessment effort, and larger observer confidence than without electronic cleansing.

Introduction

CT colonography is a minimally invasive procedure that is advocated for polyp screening [1]. Residual fecal material and fluid are well known to hinder CT colonography evaluation, especially in a 3D display mode. Fecal material may cover lesions (preventing detection) or, conversely, mimic polyps (necessitating superfluous 2D verification and possibly reducing specificity). Fecal tagging was introduced to discriminate between tissue and fecal material [2, 3]. It permits a limited bowel preparation, which may contribute to better patient compliance.

Electronic cleansing aims at replacing tagged material (i.e., fecal remains and fluid) with air. This is needed for primary 3D evaluation of data in which there is a large amount of fecal material. Moreover, it may aid 3D problem solving in a primary 2D reading under these conditions.

Several electronic cleansing algorithms based on the increased attenuation of tagged material have been described previously [4–6]. Incomplete processing is still reported to leave artifacts [7]. A specifically noticeable problem is posed by the distracting "ridges" emanating from locations in which air, soft tissue, and tagged material meet [7]. An electronic cleansing algorithm was devised to improve the accuracy at these three-material junctions [8, 9]. The method adapts to patientspecific conditions, such as the local density of the tagged material, and is automated.

The purpose of this study was to assess the electronic cleansing algorithm's effect on lesion conspicuity as well as its practical efficiency. We hypothesize that the studied electronic cleansing method does not affect the conspicuity of lesions and facilitates efficient evaluation gegarding time, assessment effort, and observer confidence.

Materials and Methods

Patient Population

Patients were included from the Walter Reed Army Medical Center (WRAMC) public database, made available by the National Cancer Institute [10, 11]. Study I explored the effect of electronic cleansing on lesion conspicuity. All patients were included whose results contained polyps ≥ 6 mm in the largest diameter (measured during colonoscopy), irrespective of shape or location (patient group 1).

Study II investigated the effect of electronic cleansing on efficiency. Ten randomly selected patients were taken from patient group 1 and nine patients were randomly selected from the public database who did not have polyps (patient group 2) (Table 3.1). No patients were excluded a priori (apart from those already excluded by WRAMC).

CT Colonography

Patients had undergone standard 24-hour colonic preparation with the oral administration of 90 mL of sodium phosphate (Fleet 1, Fleet Pharmaceuticals) and 10 mg of bisacodyl. Patients consumed 500 mL of barium (2.1% by weight) (Scan C, Lafayette Pharmaceuticals) and 120 mL of diatrizoate meglumine and diatrizoate sodium solution (Gastrografin, Bracco Diagnostics). Distention was achieved through patient-controlled insufflation of room air.

Colonoscopy

CT colonography findings were disclosed during colonoscopy using segmental unblinding (at WRAMC). The colonoscopy findings were available via a report that included photographs, size measurement, morphology (pedunculated, sessile, and flat), and location (cecum, ascending colon, hepatic flexure, transverse colon, splenic flexure, descending colon, sigmoid, and rectum).

TABLE 3.1 **Walter Reed Army Medical Center Cases Used to Measure Evaluation Efficiency With and Without Electronic Cleansing**

Patient No.	Colonoscopically Proven Polyps	Location	
		Supine	Prone
8	6 mm, sessile, rectum	A	A
9	None		
11	7 mm, pedunculated, hepatic flexure	A	A
19	10 mm, pedunculated, distal ascending colon	P	A
	7 mm, pedunculated, sigmoid	A	F
20	6 mm, pedunculated, sigmoid	P	A
32	None		
35	7 mm, pedunculated, sigmoid	P	P
38	10 mm, round, sigmoid	P	A
56	8 mm, pedunculated, proximal ascending colon	F	A
60	None		
64	None		
81	None		
138	None		
193	8 mm, pedunculated, hepatic flexure	A	P
	7 mm, sessile, sigmoid	A	P
180	None		
393	None		
408	12 mm, pedunculated, proximal ascending colon	A	A
416	None		
437	6 mm, sessile, rectum	A	A

A research fellow indicated whether a polyp resided in air (A), partly resided in fecal matter (P), or fully resided in fecal matter (F).

Electronic Cleansing

The electronic cleansing method assumes that the measured density in a voxel arises because of a combination of three materials: soft tissue, air, and tagged material. Initially, the percentage of materials in each voxel is determined. Subsequently, the partial volume of tagged material is replaced by air and the new density is calculated. Finally, a 3D method visualizes the colon from an endoluminal perspective as if there were no fecal remains. Appendix 1 contains an intuitive description of the algorithm; an exact (mathematic) explanation has been previously published [8, 9].

Observers

An abdominal radiologist (observer 1) and a research fellow (observer 2) were involved in reading all data for both studies. Observer 1 had previous experience of approximately 1,200 colonoscopy-verified CT colonography examinations at the start of the study. Observer 2 had previous experience of 350 such examinations. The interval between the two studies was 4 months, during which both observers evaluated approximately 100 additional patients using CT colonography. This setup was chosen to avoid observer bias (the positive patients from study II are also in study I). Both observers were familiar with pitfalls using primary 3D evaluation with stool tagging and electronic cleansing as described previously [7].

Study I: Image Review

The conspicuity study focused on all polyps from patient group 1 (study I). The prone and supine positions were considered as separate findings. Each finding was assessed by a researcher who had a background in designing electronic cleansing algorithms. This researcher determined whether the finding completely resided in air, was partly covered with tagged material, or was fully covered with tagged material.

Each polyp was marked and presented to the observers on a 3D display. An enhanced 3D display (unfolded cube display; ViewForum, Philips Healthcare) was chosen to measure the conspicuity of a polyp without having to turn the camera [12]. The polyps partly or fully covered by fecal material were presented in 3D only after processing by electronic cleansing. The polyps residing in air (not requiring electronic cleansing) were presented directly as such. The cases were presented in random order. The observers were not informed whether the polyp had been uncovered by electronic cleansing or not.

The viewing position along the colon's centreline was controlled by the observer. Two-dimensional reformatted views of the original CT data were initialized by clicking on a position in the 3D endoluminal images. The window width and level default of the reformats was width, 1,250 HU; level, –50 HU; however, this could be freely adapted by the observers.

Study I: Assessments

Initially, the observers indicated on a 5-point Likert scale for each case whether it was suspected that the polyp completely resided in air or was uncovered by electronic cleansing. On this scale, 1 corresponded with a polyp definitely considered not uncovered by electronic cleansing and 5 with a polyp strongly suspected to be uncovered. The observers performed this rating based on the 3D display only.

Subsequently, each polyp was scored with respect to conspicuity on a 5-point Likert scale: 1 indicated inadequate, the lesion is not visible and will not prompt a 2D inspection; 2 indicated moderate to questionable, the lesion is hardly visible and the location might lead to a 2D inspection; 3 indicated average, lesion is expected to be detected and the location will prompt a 2D inspection; 4 indicated good, the lesion is well visible and the location should lead to a 2D inspection; and 5 indicated excellent, the lesion is very well visible and the location will certainly lead to a 2D inspection.

The observers performed the conspicuity rating based on the 3D display only. After having provided the conspicuity ratings, each observer had to indicate for the cases scored as "inadequate" or "moderate" whether the electronic cleansing algorithm caused the difficulty or if there was another cause: inappropriate electronic cleansing or some other reason (e.g., CT artifact or a flat lesion). The observers were aware that electronic cleansing algorithms may leave artifacts, specifically clouds of debris and lines of ridges at junctions. The first item was selected if such an electronic cleansing artifact was particularly disturbing, for example, a smoothened polyp surface because of image processing or visible ridges emanating from junctions. In the latter case, the observer was asked o indicate the reason for the difficulty. Both the enhanced 3D display with electronic

cleansing and 2D reformats of original data were available to assess the cause of low conspicuity.

Study I: Outcome Parameters and Statistical Analysis

UNCOVERED BY ELECTRONIC CLEANSING ASSESSMENT - The polyps residing in air were compared with those partly or completely covered by tagged material. The difference in the consideration whether a polyp was "uncovered" was tested by means of the Mann-Whitney test (Wilcoxon's rank sum test). A p value ≤ 0.05 was considered to indicate a statistically significant difference.

CONSPICUITY - Differences in conspicuity between polyps were also statistically tested by means of the Mann-Whitney test. The outcomes were stratified by observer, polyp size, and polyp environment (i.e., residing in air or partly or fully covered by tagged material). Two-sided tests were used. Again, a p value ≤ 0.05 corresponded to a significant difference.

Study II: Image Review

The data not processed by electronic cleansing from patient group 2 (study II) were reviewed as follows. In part 1, 3D cine loops (using the unfolded cube display, ViewForum) were examined in the prone and supine positions [12]. The two positions were electronically linked. After clicking in one unfolded cube image, the corresponding unfolded cube image of the other position was displayed [13]. The frame rate was controlled by the observers. Two-dimensional reformatted views of the original CT data became accessible by clicking on a position in the 3D endoluminal images. Lesion size was measured by means of electronic calipers on the 3D view. In part 2, after reviewing the unfolded cube images, the original axial CT slices were inspected to verify whether surface parts were obscured by tagged material. The window width and level setting of the reformats and axial CT slices was by default width, 1,250 HU; level, –50 HU; however, this could be freely adapted by the observers. The data processed by electronic cleansing were reviewed similarly.

For part 1, the electronically cleansed data were reviewed using the same approach as the data that were not processed by electronic cleansing. However, the 2D reformatted views (obtained after clicking in a 3D image) were based on the original data previously proposed [7] to avoid pitfalls caused by electronic cleansing in both 3D and 2D. In part 2, after reviewing the 3D images, the original axial CT slices were inspected as previously described [7]. Both the evaluation of the original data and the data processed by electronic cleansing include a 2D review of the original axial CT slices (i.e., data not processed by electronic cleansing). Figure 3.1 illustrates the interface used to read the data.

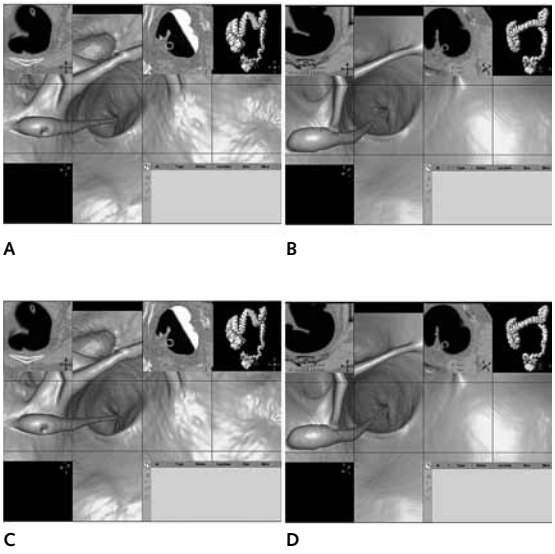


FIGURE 3.1
Interface for CT colonography evaluation in 57-year-old man (WRAMC 393). **A** and **B**, Supine (**A**) and prone (**B**) enhanced 3D display images from part 1, in which 3D cine loops (using unfolded cube display; ViewForum, Philips Healthcare) were examined. Note that views are linked. **C** and **D**, Axial CT images from part 2, in which original axial slices were examined to verify whether surface parts were obscured by tagged material while tracking colon's centerline.

Study II: Assessments.

Observer 1 and observer 2 independently evaluated the original data and the electronically cleansed data from patient group 2. To maintain objectivity, both were blinded to findings during colonoscopy, findings by themselves for the same patient (without and with electronic cleansing), and each other's findings. In addition, they were unaware of the prevalence of polyps. The evaluation is using the original data preceded those with the electronically cleansed data. The interval period between evaluating original and electronically cleansed data of the same patient was at least 4 weeks to avoid biased results. The cases were presented in random order. Each surmised polyp was scored with respect to size, morphology, location (colon segment), and lesion confidence. The lesion confidence was qualified on a 5-point Likert scale: zero, not a lesion; and 4, absolute confidence of a lesion. The zero score was used to annotate a potential polyp initially detected through the 3D display but subsequently discarded as "not a lesion" after consulting the 2D reformats or axial CT slices.

The observers were aware that in a clinical setting a confidence of 2 or more would indicate a relevant lesion. The observers rated the assessment effort per colon segment on a 4-point Likert scale: 1, extremely easy; 2, good; 3, difficult; 4, extremely difficult. In addition, they rated their confidence in the reading per patient on a 3-point Likert scale: 1, confident; 2, in doubt; 3, extremely doubtful. The assessment effort and confidence in the reading were recorded after complete evaluation of a patient.

Study II: Reference Standard

Lesions detected during CT colonography evaluation were matched with the findings in the colonoscopy report (as previously discussed). This was done by a research nurse (with experience of > 500 such matchings) under the supervision of a research fellow (with experience of > 350 CT colonography examinations). Neither was involved in reading the CT colonography studies.

A lesion detected during CT colonography in this study was considered true-positive if it matched a colonoscopy finding with regard to size, morphology, and location (i.e., same or adjacent segment). A deviation in size by at most 5 mm was accepted to accommodate the inherent inaccuracy of colonoscopic size measurement. It should be noted that in this study the matching strategy from a previous study 14] is dopted, which differs slightly from some others [10, 11]. A false-negative finding was defined as a polyp detected during colonoscopy that did not match any CT colonography finding. A CT colonography finding with a lesion confidence of at least 2, a diameter ≥ 6 mm, and not matching a polyp detected with colonoscopy was considered to be false- positive. All detections with a diameter < 6 mm or scored with an observer confidence lower than 2 were not considered relevant and discarded.

Study II: Outcome Parameters and Statistical Analysis

PROCESSED VOLUME FRACTION AND PROCESSED SURFACE FRACTION - In this study, the terms "processed volume fraction" and "processed surface fraction" are introduced to assess the extent to which electronic cleansing had an effect. The colon volume not processed by electronic cleansing was identified by thresholding the original data (before electronic cleansing) at –650 HU. All voxels with a density below –650 HU after electronic cleansing (notice that electronic cleansing replaced tagged material with air) approximate the colon volume after processing. The "processed volume" is constituted by all those voxels that are part of the colon volume after electronic cleansing and not part of the colon volume before electronic cleansing. The mean processed volume fraction is defined as the average processed volume divided by the average colon volume after electronic cleansing. It is taken to represent the fraction of additionally exposed volume.

The colon surface voxels after electronic cleansing are those voxels that are immediately adjacent to the colonic volume

after processing. The processed surface voxels are those voxels that are directly proximate to the colon volume after electronic cleansing and not proximate to the colon volume before processing. Effectively, the latter condition aims to select voxels at the softtissue–tagged-material transition and exclude the voxels at the soft-tissue–air transition. The mean processed surface fraction is defined as the average number of processed surface voxels divided by the average number of colonic surface voxels after electronic cleansing. It is considered to reflect the fraction of colon surface added by the electronic cleansing algorithm.

EVALUATION TIME - For both observers, the evaluation time per patient - including the supine and prone acquisitions - was measured for the original and the electronically cleansed data. The time for evaluating the 3D cine loop (part 1) and the axial slices (part 2) was recorded separately. The measured times included annotation of the lesions - size measurement and indications of location, morphology, and lesion confidence. All evaluation times for the original data were compared with the corresponding times for the electronically cleansed data using the Wilcoxon's matched pairs signed rank test. Differences were considered significant at $p \leq 0.05$. The initialization time (e.g., loading the data) was disregarded in the analysis.

ASSESSMENT EFFORT AND OBSERVER CONFIDENCE - Differences in assessment effort and observer confidence between the original data and the electronically cleansed data were statistically tested per segment by means of the Wilcoxon's matched pairs signed rank test. Two-tailed p values were used and $p \leq 0.05$ was considered statistically significant.

SENSITIVITY AND SPECIFICITY - We do not consider a statistical comparison of sensitivity and specificity between the two methods on the basis of the set of 19 patients possible. Consequently, the measures of sensitivity and specificity by themselves merely serve as descriptive statistics for polyps ≥ 6 mm. The sensitivity of both display methods was determined on a per-polyp basis. The specificity of both display methods was determined by the false-positive rate.

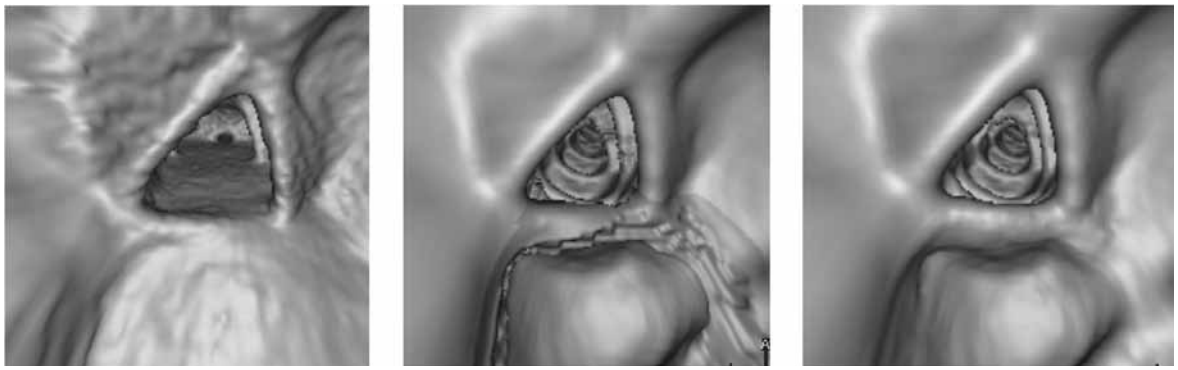
Results

Study I

Four patients from the WRAMC database (WRAMC 26, 98, 220, and 256) were excluded because of incomplete data. In the remaining data, there were 36 polyps between 6 and 10 mm in largest diameter in 32 patients and 29 polyps equal to or larger than 10 mm in 24 patients, mounting to 65 polyps in total. There were 129 polyp findings (regarding supine and prone as separate cases); one polyp remained retrospectively invisible in one position (WRAMC 213p). Eighteen of 129 findings (14%) concerned polyps completely covered and 41 of 129 findings (32%) related to polyps partly covered by tagged material (all uncovered by electronic cleansing); 70 of 129 findings (54%) referred to polyps completely surrounded by air (no electronic cleansing action needed). The median age of the patients was 57 years (age range, 48–76 years); 36 were men, and 15 were women.

UNCOVERED BY ELECTRONIC CLEANSING ASSESSMENT - Both observers rated polyps residing in air significantly lower on the scale reflecting whether they suspected that a polyp partly or fully resided in tagged material ($p < 0.05$). The readers reported a slight noise suppression by electronic cleansing that was apparent by a slightly smoothened surface (Figure 3.2).

CONSPICUITY - Both observers rated the conspicuity of all polyps residing in air as not significantly different from all those partly or fully residing in tagged material ($p = 0.5$ for observer 1 and $p = 0.6$ for observer 2). None of the substratifications yielded a significant difference. The numeric data, that is, the distribution of the ratings on



A B C

FIGURE 3.2
Walter Reed patient with fluid surface bordering on soft tissue. **A**, Image shows 3D display rendered without electronic cleansing. **B**, Image shows result as electronic cleansing is applied only to "pure" materials and two material transitions. **C**, Image shows result after applying electronic cleansing to three-material voxels as well (indicated by shaded band (arrow) in image **B**).

conspicuity, are shown in the graphs in Figures 3.3 and 3.4 These figures contain the normalized histograms of the conspicuity stratified by polyp size and polyp environment, respectively. Figure 3.5 shows examples of 3D views. Observer 1 scored four supine lesions, one prone lesion, and one supine and prone lesion (WRAMC 35s, 159s, 177p, 213s, 185s, and 311sp) as "inadequate" or "moderate to questionable" conspicuity. Observer 2 scored two supine lesions and one supine and prone lesion (WRAMC 35s, 185s, 311sp) as "inadequate" or "moderate to questionable" conspicuity. All these cases were attributed to the complex shape of the polyps (i.e., "other" and not "inappropriate electronic cleansing"). The polyp findings in two patients (WRAMC 35s, 177p) were at a fold. Polyps of two other patients (WRAMC 159, 213) were flat lesions. The other polyps were at the ileocecal valve. The polyp findings involving electronic cleansing were of two supine and one prone patient (WRAMC 159s, 177p, 213s). Images of all polyps after electronic cleansing are accessible via the Internet [15].

Study II

One patient (WRAMC 32) was excluded from the study because of inadequate diagnostic quality: both observers considered the colonic distention in the prone and supine positions insufficient for evaluation. In the remaining data, there were nine polyps between 6 and 10 mm in diameter in seven of 19 patients and three polyps equal to or greater than 10 mm in three of 19 patients, amounting to12 polyps in total. The median age of the patients was 55 years (age range, 50–78 years); 13 were men, and six were women.

PROCESSED SURFACE FRACTION AND PROCESSED VOLUME FRACTION – The mean processed surface fraction was 269,640 of 975,378 voxels = 0.27 (range, 0.16–0.46) and the mean processed volume fraction was 0.36 of 1.84 L = 0.20 (range, 0.13–0.45). Figure 3.6 illustrates the effect of electronic cleansing for the cases with the largest and smallest processed volume fractions, respectively.

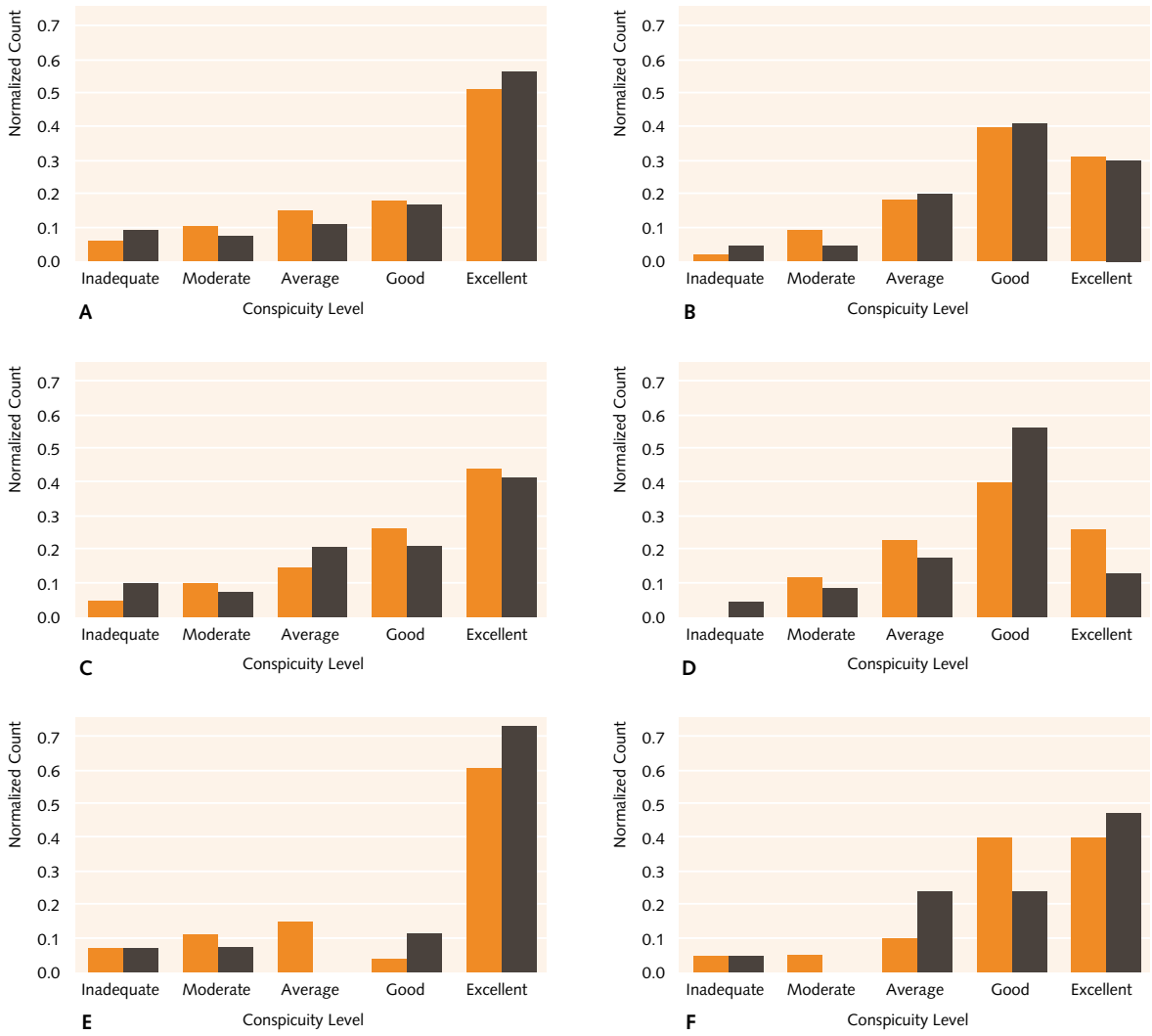


FIGURE 3.3
Histograms of polyp conspicuity for both observers stratified by polyp size. Light gray indicates air and dark gray indicates tagged material. **A–F**, Histograms for observer 1, all data (**A**); observer 2, all data (**B**); observer 1, size < 10 mm (**C**); observer 2, size < 10 mm (**D**); observer 1, size ≥ 10 mm (**E**); and observer 2, size ≥ 10 mm (**F**).

EVALUATION TIME - The median total evaluation time per patient using the original data (observer 1,20 minutes 45 seconds; observer 2, 17 minutes 14 seconds) was significantly larger than that using the electronically cleansed data (observer 1, 12 minutes 25 seconds; observer 2, 11 minutes 8 seconds) for both observers (observer 1, $p < 0.001$; observer 2, $p < 0.004$). The median evaluation time per patient for inspecting the 3D images (with the added surface by electronic cleansing) without including the 2D time (i.e., part 1) differed significantly between the original data and the electronically cleansed data for observer 1 (p = 0.037) and not for observer 2 (p = 0.396)

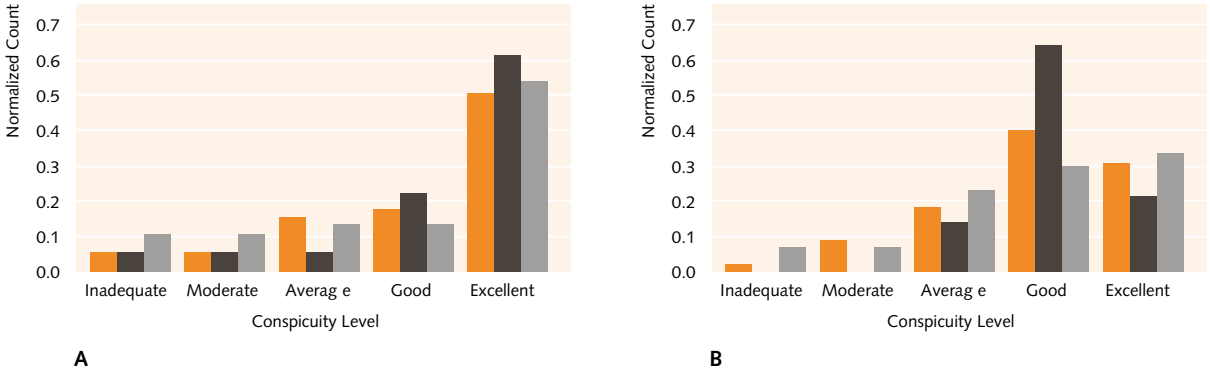


FIGURE 3.4
Histograms of polyp conspicuity for both observers stratified by polyp environment (completely surrounded by air [grey], fully submerged [black], or partly submerged [orange]). A and B, Histograms for observer 1, all data (A) and observer 2, all data (B).

(Table 3.2). The median time per patient for reviewing the axial 2D slices (i.e., part 2 after the 3D inspection) was significantly larger for the original data than for the electronically cleansed data for both observers ($p < 0.001$). The measured evaluation times take into account both the supine and prone positions (hence, time per patient).

ASSESSMENT EFFORT AND OBSERVER CONFIDENCE
Both observers rated the assessment effort for inspecting the original data significantly larger than for the electronically cleansed data in each segment. The sum of the signed ranks for observer 1 was 4,095 and for observer 2 was 2,045.5 (for both $p < 0.0000001$) (Figure 3.7). The observer confidence over all patients was rated significantly smaller for the original data than for the electronically cleansed data. The sum of the signed ranks in this respect for observer 1 was 91 ($p < 0.007$) and for observer 2 was 120 ($p < 0.0002$) (Figure. 3.7). Both reviewers reported a "cloud of debris" artifact in one location emanating from acquisition artifacts. It was not considered to preclude polyp detection. No distracting lines or ridges at junctions were encountered.

SENSITIVITY AND SPECIFICITY - The sensitivity per polyp was identical for the original and the electronically cleansed data for both observers: 10 of 12 polyps = 0.83.

TABLE 3.2 **Median Evaluation Time per Patient Using Original and Electronically Cleansed Data**

Evaluation Part	Original Data		Electronically Cleansed Data (min:s)		Paired Difference (min:s)	
	Observer 1	Observer 2	Observer 1	Observer 2	Observer 1	Observer 2
Unfolded cubes (part 1)	14:46 (6:44)	11:47 (4:38)	11:29 (5:12)	10:41 (5:30)	2:54 (6:29)	1:16 (8:24)
Axial CT slices (part 2)	5:6 (2:22)	4:17 (1:46)	0:43 (0:34)	0:0 (0:00)	4:9 (2:4)	3:53 (1:34)
Total	20:45 (6:48)	17:14 (5:49)	12:25 (5:42)	11:8 (5:34)	7:51 (7:52)	5:58 (7:49)

Data are stratified by the primary 3D part (unfolded cubes [part 1]) and the 2D part (axial CT slices [part 2]) of the evaluation.
Data in parentheses are interquartile range.

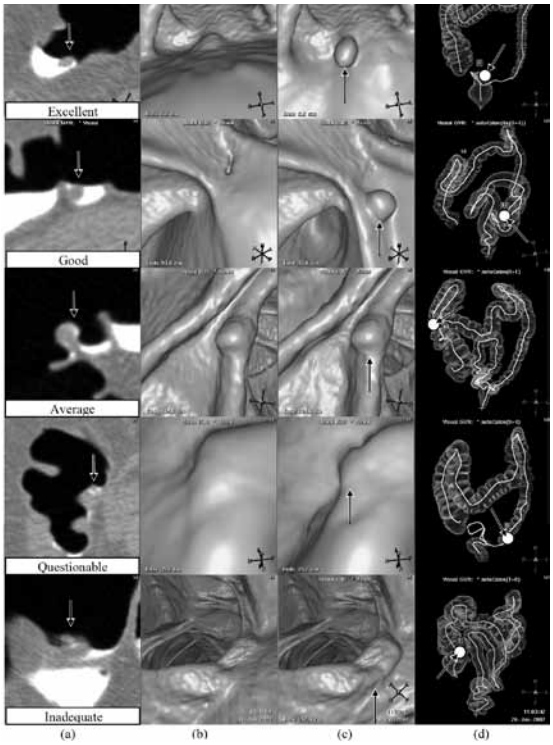


FIGURE 3.5
Examples of polyps (arrows) with varying conspicuity before (first two columns and last column) and after (third column) electronic cleansing from 52-year-old man with rating of excellent conspicuity (top row), 58-year-old man with rating of good conspicuity (second row), 53-year-old woman with rating of average conspicuity (third row), 53-year-old woman with rating of questionable conspicuity (fourth row), 54-year-old woman with rating of inadequate conspicuity (bottom row). Rating of conspicuity was given by both observers to electronically cleansed data. Circles indicate approximate polyp location. Asterisks and calipers indicate orientation of volume.

The same polyp was missed by both observers, in addition to a different polyp for each observer in the original data. The observers missed the same polyps in the data after electronic cleansing. The polyps were missed because of perceptive errors.

Observer 1 had, in total, five false-positive detections in the original data and four falsepositive detections in the electronically cleansed data. Observer 2 had, in total, two false-positive detections in the original data and four in the electronically cleansed data.

Discussion

In study I, no significant difference was found in the conspicuity of polyps that were uncovered by electronic cleansing and polyps surrounded by air. Study II showed that electronic cleansing enabled comprehensive visibility (mean processed surface fraction, 27%) and time-efficient inspection (approximately 12 minutes per patient, saving 40% in inspection time). Moreover, the observers indicated a low assessment effort and a high confidence while reading the electronically cleansed data. For polyps ≥ 6 mm, the sensitivity was high (10/12) and the false-positive rate was low: four false-positive findings, in total, in 18 patients.

Several previous articles have described technical innovations regarding electronic cleansing [4-7]. Zalis et al. [4] indicated that it is essential to ascertain how artifacts still present in successfully tagged electronically cleansed data affect observer performance. Pickhardt and Choi [7] identified a reading pitfall caused by an artifact at junctions, that is, locations where air–fluid levels interface with the colon wall. The proposed electronic cleansing method aims at improving the accuracy at such locations. We found that 41 of 129 findings related to polyps partly covered by fecal material, signifying the importance of accurate electronic cleansing in these circumstances. We hypothesize that the insignificant difference of polyp conspicuity (study I), short average 3D evaluation time, high confidence, low assessment effort, and high specificity (study II) indicate that electronic cleansing did not result in artifacts that complicated the reading. The reviewers encountered one artifact in study II, but that was not caused by electronic cleansing.

In study I, it was found that the reviewers could identify whether a lesion was uncovered by electronic cleansing through a slight smoothing of the colon surface. Although smoothing could lead to missing subtle (flat) lesions, we did not experience this in the study. The conspicuity of uncovered polyps did not differ significantly from polyps completely surrounded by air.

It is a complex problem to measure the performance of an electronic cleansing algorithm. Effectively, one should

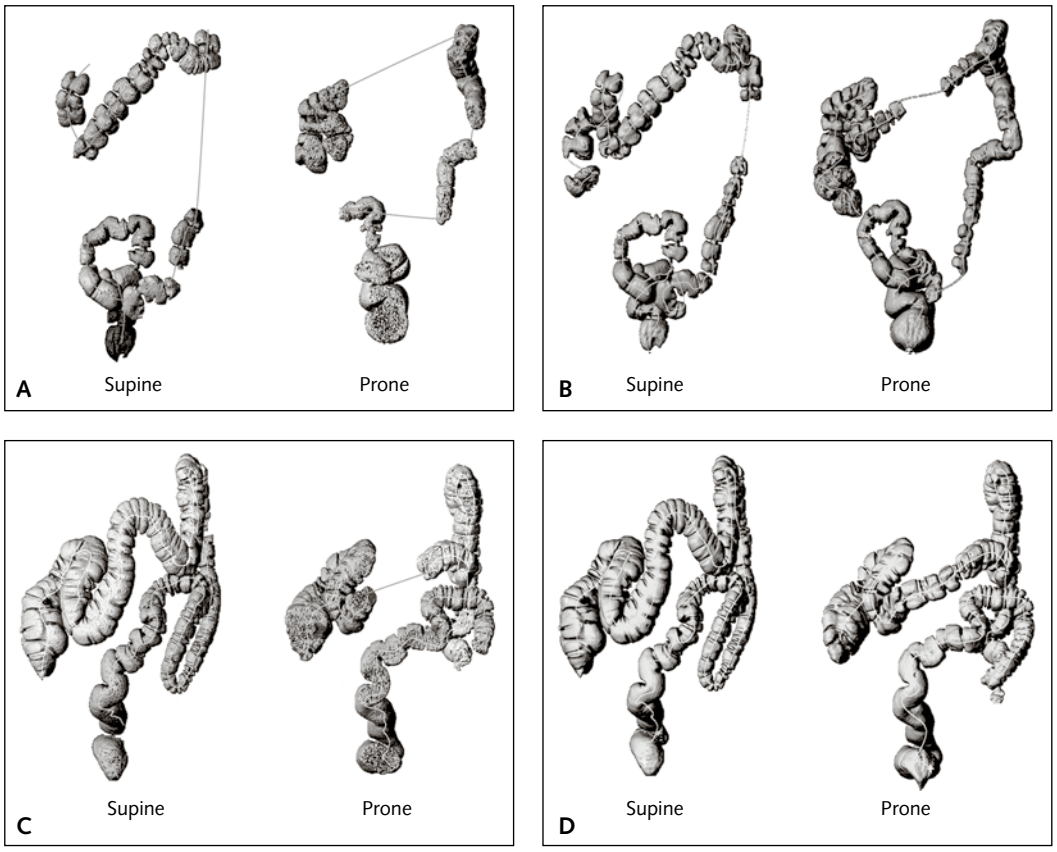


FIGURE 3.6
Visualization of the segmented colonic volume before and after electronic cleansing. **A** and **B**, Images from patient with highest processed volume fraction in both scans (WRAMC 9) show segmented colonic volume before (**A**) and after (**B**) electronic cleansing (processed volume fraction: supine, 0.44 L / 0.98 L = 0.45; prone, 0.44 L / 1.33 L = 0.33). **C** and **D**, Images from patient with lowest processed volume fraction in both scans (WRAMC 408) show segmented colonic volume before (**C**) and after (**D**) electronic cleansing (processed volume fraction: supine, 0.31 L / 2.33 L = 0.13; prone, 0.30 L / 2.33 L = 0.13).

determine the extent to which the algorithm modifies the polyp shape so that it is no longer detected. Study I compared the conspicuity of electronically cleansed polyps and those surrounded by air (not affected by the algorithm). If the electronic cleansing algorithm would change the polyp shape in a destructive manner, this might lead to different distributions of conspicuity rating. In other words, electronically cleansed polyps could become less 'conspicuous'. We did not find an indication of such an effect.

We did not find previous results on the effectiveness of electronic cleansing regarding processed surface fraction or processed volume fraction, as in study II. A related finding on surface visibility is that, on average, 99.5% of the colon

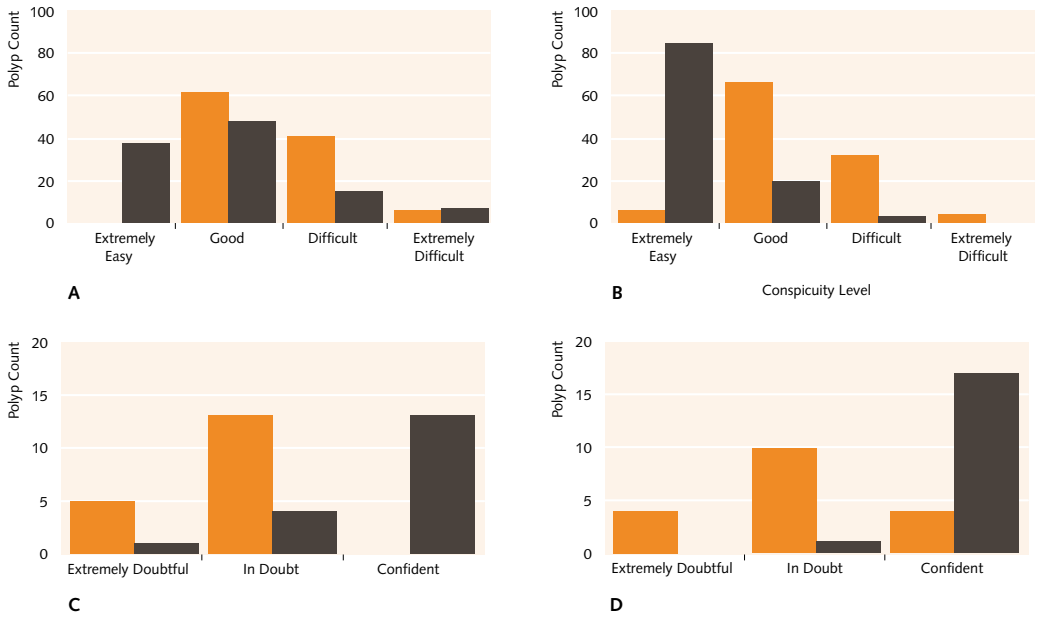


FIGURE 3.7
Histograms of observer effort and confidence level. Original data = light gray, electronically cleansed data = dark gray. **A** and **B**, Histograms of assessment effort of CT colonography inspection for observer 1 (**A**) and observer 2 (**B**) over all segments and patients.**C** and **D**, Histograms of observer confidence in reading using original and electronically cleansed data for observer 1 (**C**) and observer 2 (**D**) over all patients.

surface becomes visible using the unfolded cube imaging sequences and 95% using forward- and backward-looking imaging sequences [12]. The average processed surface fraction (0.27) by far exceeds the missed parts in those mentioned display modes. Both the display mode and electronic cleansing increase surface visibility such that electronic cleansing should be used with an enhanced 3D display to best benefit from the added surface by electronic cleansing.

The measured reading time per patient (12 minutes) in study II slightly improves on previous work [14]. In that study, a reading time of approximately 14 minutes was measured, with an enhanced 3D display technique. However, fecal tagging was not applied, with the result that additional reading of axial 2D slices for verification of obscured parts was not possible. Previous articles applying 3D displays report evaluation times of 12 [16] and 16 minutes [17].

A limitation of the entire study is that results may not be extrapolated to patients undergoing a limited bowel preparation. It might be expected that a rigorous preparation simplifies electronic cleansing somewhat because there may be less tagged material than in a limited purgation scheme. Moreover, such remains may have a more homogeneous appearance. Still, the complex nature of the problem prompted us to initially study the algorithm's performance regarding extensively prepared patients. It should be noted that we found a considerable processed volume fraction (0.20) despite the extensive preparation. The processed volume and surface fractions might depend on the nature of the preparation (wet or dry). Probably any electronic cleansing algorithm will

perform optimally with homogeneously tagged stool [18]. Accordingly, we expect the algorithm to be better suited for removing pools of fluid compared with fecal residue because the former will present a more homogeneous density.

A limitation of study II is the small study population. The sample size of the patient population was calculated to be sufficiently large to meet the primary aim of comparing the original data reading versus electronically cleansed data reading regarding time efficiency, surface visibility, assessment effort, and observer confidence. The reported sensitivity is in the same range as reported in another study [14] in which a similar display method was used. As indicated, the findings on sensitivity and specificity should merely be regarded as descriptive statistics because of the small study population.

We conclude that lesions uncovered by electronic cleansing had similar conspicuity in the 3D display as lesions already exposed to air in extensively prepared patients. In such a population, the electronic cleansing algorithm led to a shorter evaluation time, lower assessment effort, and greater observer confidence than CT colonography without electronic cleansing. The algorithm may contribute to a practical evaluation strategy involving either a primary 3D display or 3D problem solving in a primary 2D display.

References

1. Vining DJ, Gelfand DW, Bechtold RE, Scharling ES, Grishaw EK, Shifrin RY. Technical feasibility of colon imaging with helical CT and virtual reality. (abstr) AJR 1994; 162[suppl]:104
2. Callstrom MR, Johnson CD, Fletcher JG, et al. CT colonography without cathartic preparation: feasibility study. Radiology 2001; 219:693–698
3. Lefere PA, Gryspeerdt SS, Dewyspelaere J, Baekelandt M, Van Holsbeeck BG. Dietary fecal tagging as a cleansing method before CT colonography: initial results - polyp detection and patient acceptance. Radiology 2002; 224:393–403
4. Zalis ME, Perumpillichira JJ, Hahn PF. Digital subtraction bowel cleansing for CT colonography using morphological and linear filtration methods. IEEE Trans Med Imaging 2004; 23:1335–1343
5. Chen D, Liang Z, Wax MR, Li L, Li B, Kaufman AE. A novel approach to extract colon lumen from CT images for virtual colonoscopy. IEEE Trans Med Imaging 2000; 19:1220–1226
6. Lakare S, Wan M, Sata M, Kaufman A. 3D digital cleansing using segmentation rays. (abstr) In: Ertl T, Hamann B, Varshney A, eds. IEEE Visualization 2000 (Proceedings of the 2002 IEEE International Conference). New York, NY: ACM Press, 2002:37–44
7. Pickhardt PJ, Choi JHR. Electronic cleansing and stool tagging in CT colonography: advantages and pitfalls with primary three-dimensional evaluation. AJR 2003; 181:799–805
8. Serlie IWO, Truyen R, Florie J, Post FH, van Vliet LJ, Vos FM. Computed cleansing for virtual colonoscopy using a three-material transition model, MICCAI 2003. Lecture Notes in Computer Science, vol. 2879:175–183
9. Serlie IWO, Vos FM, Truyen R, Post FH, van Vliet LJ, Classifying CT image data into material fractions by a scale and rotation invariant edge model. IEEE Trans Image Process 2007; 16:2891–2904
10. Pickhardt PJ, Choi JR, Hwang I, et al. Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. N Engl J Med 2003; 349:2191–2200
11. Walter Reed Army Medical Center, Virtual Colonoscopy Center Website. CT colonography in early detection of colorectal cancer. wramc.vcscreen. com/training/index.html. Accessed March 1, 2006
12. Vos FM, van Gelder RE, Serlie IWO, et al. Threedimensional display modes for CT colonography: conventional 3D virtual colonoscopy versus unfolded cube projection. Radiology 2003; 228:878–885
13. de Vries AH, Truyen R, van der Peijl J, et al. Feasibility of automated matching of supine and prone CT colonography examinations. Br J Radiol 2006; 79: 740–744
14. van Gelder RE, Nio CY, Florie J, et al. Computed tomographic colonography compared with colonoscopy in patients at increased risk for colorectal cancer. Gastroenterology 2004; 127:41–48
15. Serlie IWO. WRAMC polyps. Delft Technical University Website. www.ist.tudelft.nl/qi. Accessed July 11, 2008
16. Macari M, Milano A, Lavelle M, Berman P, Megibow AJ. Comparison of time-efficient CT colonography with two- and three-dimensional colonic evaluation for detecting colorectal polyps. AJR 2000; 174:1543–1549
17. Neri E, Vannozzi F, Vagli P, Bardine A, Bartolozzi C. Time efficiency of CT colonography: 2D vs 3D visualization. Comp Med Imag Graph 2006; 30: 175–180
18. Zalis ME, Perumpillichira JJ, Magee C, Kohlberg G, Hahn PF. Tagging-based, electronically cleansed CT colonography: evaluation of patient comfort and image readability. Radiology 2006; 239:149–159

Appendix

An Intuitive Description of the Electronic Cleansing Algorithm

The electronic cleansing algorithm uses the environment of CT densities around each voxel to solve the partial volume effect. Effectively, three types of environments are modeled: pure materials, transitions between two materials, and junctions in which three materials meet. Initially, it is assessed whether the local environment matches "pure" soft tissue, tagged material, or air. If so, the voxel under investigation is labelled to contain 100% of that material. Alternatively, the local environment is compared with models of two-material transitions. This procedure involves selecting the type of transition (e.g., air or tagged material) and, subsequently, estimating the percentages of the (two) constituting materials. Effectively, the voxel is assigned material percentages depending on its location with respect to the transition. Simply applying this algorithm to a sample data set leads to the 3D view of Figure 3.2B. The well-visualized artifact emanates from the band-shaped area where three materials meet (also described in the study by Pickhardt and Choi [7]).

Separate models were created for potential junctions, such as a thin layer of tagged material and tagged material attached to the colon wall at varying angles. The type of junction is identified by the model that best fits the local configuration of CT values and the material percentages derived from the considered voxel's location with respect to the junction. Figure 3.2C shows the same rendering after also processing the three-material transitions. Notice that the ridges at the junction that were visible in Figure 3.2B (in the shaded band) are now properly handled by the electronic cleansing algorithm.The algorithm takes into account that the image values of the materials may vary (this specifically goes for tagged material).

We implemented the electronic cleansing method on a proprietary, experimental version of the ViewForum workstation (Philips Healthcare).

Chapter 4

Primary Uncleansed 2D Versus Primary Electronically Cleansed 3D in Limited Bowel Preparation CT colonography. Is There a Difference For Novices and Experienced Readers?

Ayso H. de Vries
Marjolein H. Liedenbaum
Shandra Bipat
Roel Truyen
Iwo W. O. Serlie
Rutger H. Cohen
Saskia G. C. van Elderen
Anneke Heutinck
Oskar Kesselring
Wouter de Monyé
Lambertus te Strake
Tjeerd Wiersma
Jaap Stoker

European Radiology 2009; 19: 1939-1950

Abstract

Objective

The purpose of this study was to compare a primary uncleansed 2D and a primary electronically cleansed 3D reading strategy in CTC in limited prepped patients.

Material and methods

Seventytwo patients received a low-fibre diet with oral iodine before CTcolonography. Six novices and two experienced observers reviewed both cleansed and uncleansed examinations in randomized order. Mean per-polyp sensitivity was compared between the methods by using generalized estimating equations. Mean per-patient sensitivity, and specificity were compared using the McNemar test. Results were stratified for experience (experienced observers versus novice observers).

Results

Mean per-polyp sensitivity for polyps 6 mm or larger was significantly higher for 30 novices using cleansed 3D (65%; 95% CI 57–73%) compared with uncleansed 2D (51%; 95%CI 44–59%). For experienced observers there was no significant difference. Mean per-patient sensitivity for polyps 6 mm or larger was significantly higher for novices as well: respectively 75% (95%CI 70–80%) versus 64% (95%CI 59–70%).

Conclusion

For experienced observers there was no statistically significant difference. Specificity for both novices and experienced observers was not significantly different. For novices primary electronically cleansed 3D is better for polyp detection than primary uncleansed 2D.

55.

Introduction

CT-colonography (CTC) has consistently been shown to have a high accuracy for the detection of colorectal neoplasia, and has recently been included in the official guidelines for colorectal cancer screening [1]. An important disadvantage of the technique is that many patients find the bowel preparation burdensome [2]. Therefore efforts have been made to prepare patients for CTC with a less extensive bowel preparation [3–7]. Minimizing bowel preparation may increase patient compliance [8–10], but will result in larger amounts of residual faeces in the colon. A prerequisite is that faecal material is labelled with oral contrast (i.e. faecal tagging) in order to differentiate faecal material from colonic structures.

To our knowledge, all limited prepared CTC studies have been performed using primary two-dimensional (2D) display methods [3–5]. The rationale for this approach is that submerged segments can be better assessed in 2D. Previous studies in patients with extensive bowel preparation have indicated that primary three-dimensional (3D) reading may result in less false negative findings compared with primary 2D reading [11, 12]. If a similar empty endoluminal view could be achieved by electronic removal of tagged material ("electronic cleansing") in patients who have undergone limited bowel preparation, primary 3D could be a method of choice.

However, specific artefacts of electronic cleansing were described that potentially reduced the accuracy of CTC [13]. This may be the reason for the paucity of papers on the use of electronic cleansing. A specifically noticeable problem is posed by the distracting 'ridges' or 'pseudopolyps' emanating from locations where air, soft tissue and tagged material meet.

We hypothesized that especially for inexperienced observers a primary electronically cleansed 3D (PEC3D) method may have advantages for evaluation of the colon: polyps are visible for longer than in a 2D method and PEC3D provides a more intuitive reproduction of reality. In this study we used a cleansing algorithm that was devised to improve 3D image quality at the junctions of air, soft tissue and tagged material.

Therefore the purpose of this study was to assess whether there was any difference in accuracy between two different reading strategies for the detection of colorectal polyps in a patient population that had undergone a reduced bowel preparation. The results of primary uncleansed 2D (PU2D) and primary electronically cleansed 3D (PEC3D) were stratified for reader experience.

Materials and methods

Study population

The institutional review board of our hospital approved the study. All patients gave written informed consent.

The CT datasets used in this study were a consecutive series of FOBT (faecal occult blood test) positive patients that were included in the framework of a previous comparative study of two different faecal occult blood tests [14].

Bowel preparation

Bowel preparation started 2 days before CTC and consisted of seven 50-ml aliquots of meglumine ioxithalamate (Telebrix Gastro 300 mg I/ml; Guerbet, Cedex, France) administered orally (undiluted) with each meal (breakfast, lunch and dinner). The use of oral contrast was combined with a low-fibre diet. The evening and morning before the CTC examination no solid foods were allowed. Explicit instructions about fluid intake were not given. No laxatives were used in order to minimize patient discomfort.

Scan parameters

CTC was executed according to state-of-the-art techniques. Patients were examined in prone and supine position after the intravenous administration of bowel relaxants (Buscopan; Boehringer Ingelheim, Germany or, if contraindicated, Glucagon; Novo ordisk, Bagsvaerd, Denmark). CO2 was automatically insufflated (PROTOCO2L, EZ-E-M). Patients were not

TABLE 4.1 **Individual Observer Performance in PU2D and PEC3D**

Observer	Experience	Number of primary 2D CTC's verified by colonoscopy	Number of primary 3D CTC's verified by colonoscopy	Review times (s)		Mean interval (days) between both review methods (SD)
				2D	3D	
Observer 2	Experienced observer	500	100	574	593	32 (15)
Observer 6	Experienced observer	300	50	432	579	66 (38)
Mean experienced observer				**503***	**586***	
Observer 1	Novice	40	10	765	870	11 (4)
Observer 3	Novice	40	10	756	964	36 (19)
Observer 4	Novice	40	10	404	376	7 (3)
Observer 5	Novice	40	10	259	419	60 (38)
Observer 7	Novice	40	10	634	795	30 (23)
Observer 8	Novice	40	10	403	447	20 (9)
Mean novice observers				**537***	**645***	

SD standard deviation

* Denotes statistically significant difference between the review methods of both groups.

given intravenous contrast medium. Examinations were performed on a 64-slice multidetector CT system (Brilliance, Philips Medical Systems, Best, the Netherlands) with a reference mA s value of 40 mA s (z-axis tube modulation and automatic current selection was used). Collimation was 64×0.625 mm, pitch 1.2, slice thickness 0.9 mm, rotation time 0.4 s and tube voltage 120 kV.

Reading methods

The examinations were read in random order i.e. the PU2D and PEC3D datasets were interspersed. The observers were blinded to the results of the reference standard. To avoid recall bias, we aimed to maximize the interval between the PU2D and PEC3D reading of the same patient. This period varied per observer (mean 33 days, range 7–66) (Table 4.1). All detected lesions were recorded in a digital database. The method of detection (PU2D or PEC3D), colon segment and size of the lesion (as measured in the primary review method) of all findings were documented.

PRIMARY UNCLEANSED 2D (PU2D) – The PU2D interface is illustrated in Figure 4.1 (ViewForum 6.1, PhilipsMedical Systems), using axial CT images (W, 250, L 50). Observers were free to adjust the window setting when appropriate. To further elucidate suspected findings on 2D, an uncleansed 3D display and a 3D colon overview could be viewed for problem solving. The dual-screen interface simultaneously displayed both supine and prone scan positions.

PRIMARY ELECTRONICALLY CLEANSED 3D (PEC3D) – The PEC3D interface is illustrated in Figure 4.2. The unfolded 56 cube review method (ViewForum, Philips Medical Systems) was developed to maximize the area of

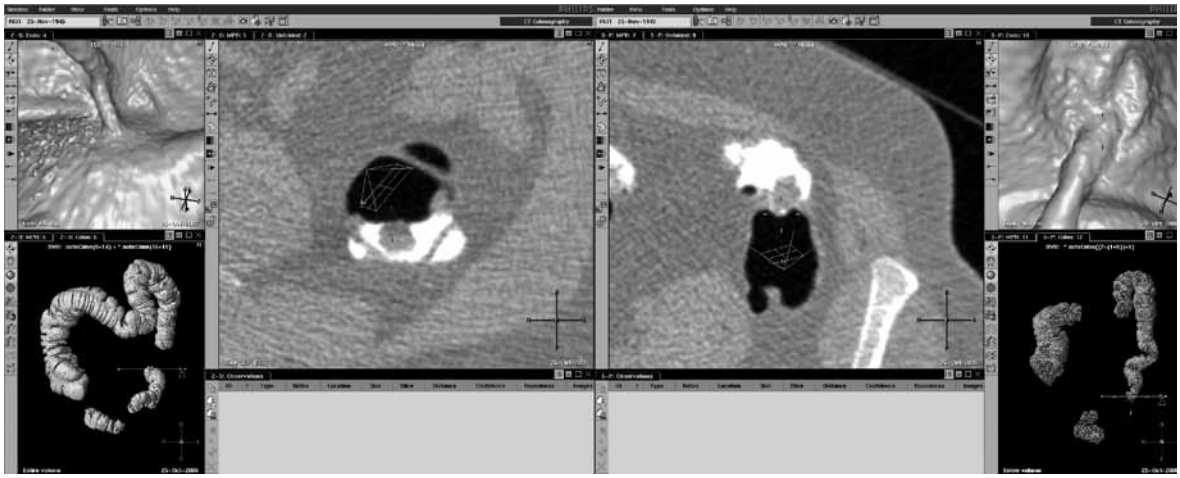


FIGURE 4.1
Interface of the PU2D method. Uncleansed 3D images were used for problem solving. In the correspondingscan positionsa large stalked polyp can be seen in the descending colon

visible colon surface and was previously validated [15]. The unfolded cube display method was simultaneously displayed with corresponding original – uncleansed – 2D multiplanar reformatted images and a 3D colon overview for problem solving. The dual-screen interface simultaneously displayed both supine and prone examinations. At the time of study the cleansing algorithm (View-Forum, Philips Medical Systems, Best, the Netherlands) [16] was not yet commercially available. In short, the algorithm assumes that the measured density in a voxel arises due to a

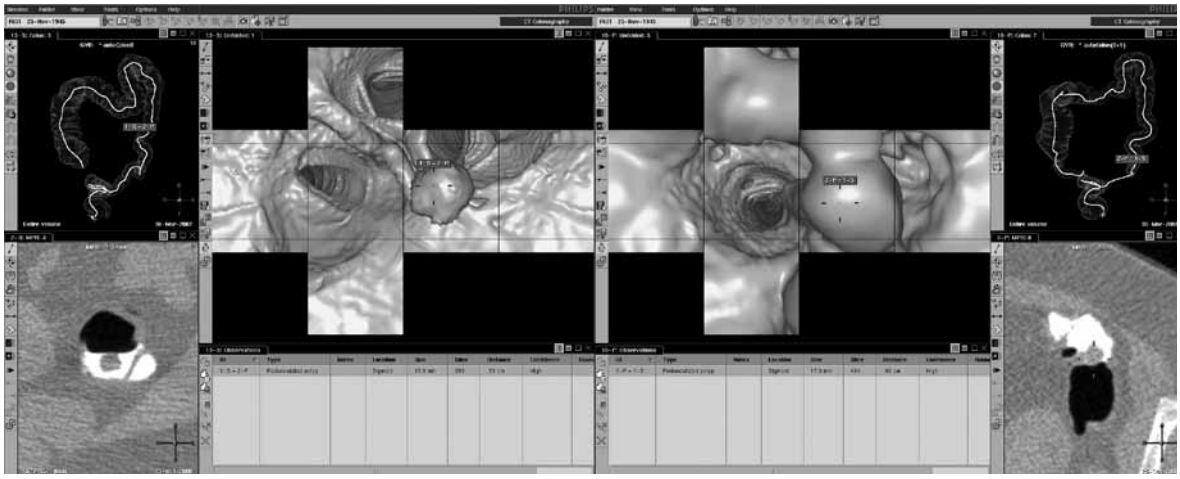


FIGURE 4.2
Interface of the PEC3D method. Uncleansed 2D images were used for problem solving. In the corresponding scan positions the same large stalked polyp can be seen as in Figure 4.1. This submerged polyp has become visible in 3D after electronic cleansing

combination of three materials: soft tissue, air and tagged material. Initially, the percentage of materials in each voxel is determined. Subsequently, the partial volume of tagged material is replaced by air and the new density is calculated. Finally, a 3D method visualizes the colon from an endoluminal perspective as if there were no faecal remains. During the study the algorithm was not yet integrated into the system and was therefore processed on a separate computer (Precision 690, Dell, Round Rock, USA). Afterwards, the cleansed data were reloaded on the workstation.

Observers

Eight physicians participated in this study: two experienced observers (two research physicians working full time on CTC) and six novice observers (five radiologists, one recently qualified physician).

The two experienced observers had seen over 350 CTC examinations verified by colonoscopy (among these the 75 patients included in this study). All novice observers, without any notable prior experience with CTC, had undergone the same learning curve. The learning curve consisted of 50 selected CTC examinations from a publicly available database [17]. Forty examinations were read using a primary 2D method and 0 using a primary 3D method. The ratio for this distribution was the assumption that detecting lesions in 2D is more difficult than in 3D, although the assessment of a suspected lesion is similar in both review methods. Personal feedback on false negative and false positive findings for polyps 6 mm or larger was provided by a radiology research fellow (more than 500 CTC examinations verified by colonoscopy).

The observers were aware of the fact that the actual study was done on FOBT positive patients.

Reference standard

The reference standard was based on the findings of the initial CTC reading, which was double read by two experienced observers (more than 200 CTC) and verified by colonoscopy with segmental unblinding. Colonoscopy was performed within 2 weeks after CTC.

A polyp seen during CTC was considered true positive if (1) its appearance resembled the corresponding adenomatous or nonadenomatous polyp at colonoscopy, (2) its segment or adjacent segment corresponded with the segment of the reference standard and (3) the polyp size as estimated by the endoscopist corresponded with the CTC size, considering a margin of error of 50%. Since the colonoscopy measurement is subject to inaccuracy [18, 19] this criterion could be overruled by the first two criteria. All other annotations were considered false positives.

The relation of polyps to faecal material

To illustrate the influence of faecal material on the visibility of polyps, a research fellow determined whether each polyp was completely covered by faecal material (i.e. completely submerged in both positions), partially covered by faecal material (i.e. not covered in both scan positions but at least partially covered in at least one position) or not covered by faecal material at all.

Power analysis

A power calculation was performed based on an assumed 15% difference between the methods in sensitivity for polyps 6 mm or larger (i.e. 70% versus 85%). The number of visible polyps required to detect a statistically significant difference by using the McNemar test was 75 (p=0.05). As we expected some of the patients to be excluded due to insufficient diagnostic quality, we included a total of 75 examinations of FOBT positive patients that had 84 visible polyps.

Performance

PER POLYP – To investigate differences between PU2D and PEC3D, we calculated the mean per-polyp sensitivity for both experienced and novice readers. Statistical differences between the review methods were assessed by using generalized estimating equations (GEE) (SPSS, 5.0, Statistics, Chicago, USA) to revise for data clustering and dependency.

In this GEE method, regression analyses were performed to compare the mean sensitivity values of the two methods [20]. Since the per-polyp specificity cannot be calculated because it is a nonexisting entity, we confined the per-polyp results to the number of false positive findings.

PER PATIENT – Other main outcome per-patient parameters were per-patient sensitivity and specificity. Statistical differences in mean per-patient sensitivity and mean specificity measures were assessed with the McNemar test.

Both mean per-polyp and per-patient outcome measures were analysed according to cutoff values of 6 and 10 mm. P values < 0.05 were considered statistically significant.

Review time

The review time, defined as the time measured with a stopwatch to review a complete examination, was compared for both methods. The review time did not include the time required for processing the images. These procedures are highly dependent on calculation power, are performed semiautomatically and require no reviewer interaction. Differences in mean review time of experienced and novice observers were assessed with a paired Student's t test. P values<0.05 were considered statistically significant.

Image quality

The image quality of the examinations was rated (after reading) on a four-point Likert scale: diagnostic without artefacts; diagnostic with a small number of artefacts, polyps 6 mm or larger cannot be missed; diagnostic with many artefacts, polyps 6-9 mm can be missed; not diagnostic, polyps 10 mm or larger can be missed.
Firstly, if at least four observers rated the examination as "not diagnostic" the patient was excluded.
Secondly, we determined the percentage of PU2D and PEC3D examinations in the various rating categories. The percentage represented the mean rating of all eight observers.

Thirdly, we assessed per-observer differences in quality between PU2D and PEC3D using ordinal regression analysis with PU2D as reference standard. A lower relative diagnostic odds ratio (RDOR) implies inferior image quality of the PEC3D compared with the PU2D. Confidence intervals not reaching indicate significant inferiority.
Fourthly, each observer determined the presence of artefacts per examination. The impact of each artefact on the "readability" of the examination was assessed on a four-point Likert scale: not disturbing; disturbing, but cannot hinder detection of polyps 6 mm or larger; disturbing, can cover polyps 6–9 mm; and disturbing, can cover polyps 10 mm or larger. We report the number of patients with artefacts observed by at least four observers and the number of patients with artefacts classified by at least four observers as "disturbing, can cover polyps". These were analysed according to cutoff values of 6 and 10 mm.

Results

All 75 patients were scanned between October 2006 and January 2007 and underwent colonoscopy within 9.4 days (SD 6.6 days). We excluded three of the 75 patients because more than four observers rated the diagnostic quality of three PEC3D examinations "not diagnostic". In 2D these examinations were rated by one to three observers as "not diagnostic". The remaining 72 patients consisted of 38 men and 34 women (mean age 59.5 years, SD 6.4 years, range 50–73). Bleeding during polypectomy was reported for three patients, none of whom required reintervention. No adverse events were reported to occur in any of the CT examinations.
A total of 90 polyps 6 mm or larger were detected: 17 polyps were 20 mm or larger (median size 25 mm, range 20–40 mm), 36 polyps were 10–9 mm (median size mm, range 10–7 mm) and 37 polyps were 6–9 mm (median size 7 mm, range 6–9 mm). The histology revealed 82 adenomatous polyps, 4 nonadenomatous polyps and 4 colorectal carcinomas.

TABLE 4.2 **The Relation of Polyps to Faecal Material.**

	6-9 mm frequency (%)		≥ 10mm frequency n (%)	
	n=	%	n=	%
Polyps completely covered by faecal material in both scan positions	1	3	2	4
Polyps partially covered by faecal material in one or both scan positions or completely covered in one position	13	35	28	53
Polyps not covered by faecal material at all in both scan positions	18	49	22	42
In retrospect not visible in both scan positions	5	14	1	2
Total number of polyps	37	100	53	100

Table shows the number and percentage of visible polyps either completely covered by faecal material, partially covered by faecal material or not covered by faecal material at all in two scan positions. Polyps that are not visible at all are reported as well.

Polyps 6 mm or larger were found in 50 out of 72 patients (69%) and polyps 10 mm or larger were found in 34 out of 72 patients (47%). As shown by Table 4.2 a substantial part of most polyps was covered by faecal material.

Performance

PER POLYP – Mean per-polyp sensitivity for novices and experienced observers for both review methods are listed in Table 4.3 Novice observers had a significantly higher mean sensitivity when using PEC3D for polyps 6 mm or larger (+14%, p<0.001), 6–9 mm (+19%, p<0.001) and 10 mm or larger (+9%, p<0.001). Within the group of novice observers there was a considerable difference in sensitivity for both review methods (Tables 4.3 and 4.4), despite the fact that all observers had undergone a similar training protocol. For experienced observers there was no significant difference in mean sensitivity between both methods for polyps 6 mm or larger (p=0.755), 6–9 mm (p=0.70) and 10 mm or larger (p=0.207).

PER PATIENT – The per-patient performance characteristics for PU2D and PEC3D are shown in Table 4.4. Novice observers had a significantly higher mean sensitivity when using PEC3D for polyps 6 mm or larger (+11%, p<0.001) and 10 mm or larger (+6%, p=0.033) compared with PU2D. For experienced observers there was no significant difference in mean sensitivity between both methods for polyps 6 mm or larger (p=0.549) and 10 mm or larger (p=0.125).

Specificity for novice observers when using PEC3D was not significantly lower for polyps 6 mm or larger (p=0.057) and 10 mm or larger (p=0. 36) compared with PU2D. For experienced observers there was no significant difference for polyps 6 mm or larger (p=0.5) and 10 mm or larger (p=.0) as well. Thus, specificity did not significantly differ between both methods in any size category for both experienced and novice observers.

Review time

For novice observers mean review time for PU2D was 8.9min (range 4.3–12.8min) compared with 10.8min (range 6.3–6.1 min) for PEC3D (p<0.00). For experienced observers the review times were respectively 8.4 min (range 7.2–9.6 min) and 9.8 min (range 9.6–9.9 min) (p<0.00).

TABLE 4.3 **Per-polyp sensitivity and false positives rate of experienced observers and novices**

Observer	Experience	Per polyp sensitivity						Number of false positives ≥ 6mm		Number of false positives ≥ 10mm	
		2D			3D			2D	3D	2D	3D
		≥6 mm (95% CI)	6-9mm (95% CI)	≥10mm (95% CI)	≥6mm (95% CI)	6-9mm (95% CI)	≥10mm (95% CI)				
Observer 2	Experienced observer	80% (70-87)	65% (47-79)	91% (79-96)	82% (74-88)	76% (61-86)	87% (77-93)	11	7	2	0
Observer 6	Experienced observer	78% (69-85)	57% (43-70)	92% (79-98)	78% (67-86)	65% (48-79)	87% (73-94)	9	12	1	2
Mean experienced observers		**79% (70-86)**	**61% (46-74)**	**92% (79-97)**	**80% (72-86)**	**70% (57-81)**	**87% (75-93)**	**10**	**10**	**2**	**1**
Observer 1	Novice	63% (53-71)	42% (27-58)	77% (66-85)	77% (67-85)	67% (49-81)	85% (74-91)	2	4	1	1
Observer 3	Novice	59% (50-67)	38% (24-54)	74% (59-84)	63% (54-72)	46% (31-62)	75% (61-86)	7	8	1	3
Observer 4	Novice	56% (45-66)	27% (14-45)	75% (59-87)	66% (55-75)	41% (23-61)	83% (68-92)	8	10	3	4
Observer 5	Novice	53% (44-62)	24% (12-42)	74% (61-83)	68% (58-76)	46% (29-64)	83% (74-90)	2	9	1	3
Observer 7	Novice	29% (21-39)	8% (3-23)	45% (31-57)	53% (43-64)	32% (19-50)	68% (49-82)	4	13	4	7
Observer 8	Novice	49% (35-64)	22% (9-43)	70% (54-82)	64% (52-75)	49% (32-65)	76% (61-86)	14	18	7	2
Mean novice observers		**51%* (44-59)**	**27%* (17-39)**	**69%* (59-77)**	**65%* (57-73)**	**46%* (33-60)**	**78%* (68-86)**	**6**	**10**	**3**	**3**

Table displays individual per-polyp sensitivity stratified for polyp size (95% confidence interval between brackets). The mean sensitivity values are given for experienced readers and novices. Mean sensitivities marked with * have a statistically significant difference between the review methods. In addition, the number of false positive findings is shown.

For most observers PU2D was faster. One observer evaluated the examinations faster in primary cleansed 3D method (Table 4.1).

Diagnostic quality

The mean rating of the diagnostic quality is displayed in Table 4.5. Figure 4.3 shows that all observers rated the image quality of PEC3D significantly lower than PU2D, as the confidence intervals did not reach 1.

Floating debris (Figure 4.4) and holes in the colon wall (Figure 4.5) were important causes of artefacts in PEC3D (Table 4.6). According to the observers, floating debris in particular may hinder the diagnostic accuracy by covering polyps 6 mm or larger. In PU2D virtually no disturbing artefacts were reported.

TABLE 4.4 **Per-patient sensitivity and specificity of experienced observers and novices**

Observer	Experience	Per patient sensitivity				Specificity			
		2D		3D		2D		3D	
		≥6mm (95% CI)	≥10mm (95% CI)	≥6mm (95% CI)	≥10mm (95% CI)	≥6mm (95% CI)	≥10mm (95% CI)	≥6mm (95% CI)	≥10mm (95% CI)
Observer 2	Experienced observer	88% (79-97)	97% (91-100)	86% (76-96)	91% (81-100)	95% (87-100)	97% (92-100)	91% (79-100)	100% (100-100)
Observer 6	Experienced observer	84% (74-94)	97% (91-100)	80% (69-91)	91% (82-100)	95% (87-100)	100% (100-100)	91% (79-100)	95% (88-100)
Mean values experienced observers		**86% (79-93)**	**97% (93-100)**	**83% (76-90)**	**91% (84-98)**	**95% (89-100)**	**99% (96-100)**	**91% (82-99)**	**97% (94-100)**
Observer 1	Novice observer	71% (58-84)	82% (69-95)	85% (75-95)	91% (81-100)	100% (100-100)	97% (92-100)	91% (79-100)	97% (92-100)
Observer 3	Novice observer	74% (62-86)	85% (73-97)	72% (60-84)	79% (66-93)	91% (79-100)	100% (100-100)	95% (87-100)	95% (88-100)
Observer 4	Novice observer	68% (55-81)	85% (73-97)	74% (62-86)	91% (82-100)	95% (87-100)	95% (88-100)	100% (100-100)	95% (88-100)
Observer 5	Novice observer	70% (57-83)	88% (77-99)	78% (67-89)	91% (82-100)	100% (100-100)	97% (92-100)	91% (79-100)	97% (92-100)
Observer 7	Novice observer	42% (28-56)	59% (42-75)	72% (60-84)	85% (73-97)	100% (100-100)	97% (92-100)	91% (79-100)	95% (88-100)
Observer 8	Novice observer	62% (48-76)	81% (67-95)	68% (55-81)	81% (67-95)	95% (87-100)	95% (88-100)	77% (60-95)	95% (88-100)
Mean values novice observers		**64% (59-70)**	**80% (74-86)**	**75%* (70-80)**	**86%* (82-91)**	**97% (94-100)**	**97% (95-99)**	**91% (86—96)**	**96% (93-98)**

Table displays individual per-patient sensitivity and specificity stratified for polyp size (95% confidence interval between brackets). The mean sensitivity and specificity values are given for experienced readers and novices. Mean values marked with * have a statistically significant difference between the review methods.

Discussion

This study shows that novice observers (compared with experienced observers) have a higher sensitivity in limited prepared patients when using PEC3D compared with PU2D. The higher sensitivity comes without a statistically significant lower specificity. On average more review time was needed for PEC3D. Paradoxically, despite its superior performance in polyp detection, the image quality of PEC3D was rated significantly less than for PU2D. Recently, two comparative studies of primary 2D and primary 3D review methods have addressed the merit of both techniques [11,12]. As in this study, more polyps of 6 mm or larger were detected using primary 3D, although in only one study the difference was statistically significant. The reason for this superior sensitivity may be that abnormalities are visible to the observer for longer. Secondly, polyps that may have a similar appearance to folds on 2D are easier to distinguish from folds in 3D. However, the fact that primary 3D is often not used in CTC is probably based on practical grounds such as long review time and high computer requirements associated with this review technique in the past.

TABLE 4.5 **Mean Rating of Diagnostic Quality.**

Image quality of the examinations	2D	3D
	% of the examinations	% of the examinations
Not diagnostic, polyps ≥ 10mm can be missed.	1%	4%
Diagnostic with many artefacts, polyps 6-9mm can be missed	4%	18%
Diagnostic with a small number of artefacts, polyps ≥ 6mm cannot be missed	20%	53%
Diagnostic without artefacts	69%	19%
Unknown	5%	5%

Table displays the percentage of PU2D and PEC3D examinations in the various rating categories.
The percentage represented the mean rating of all eight observers in 72 patients.

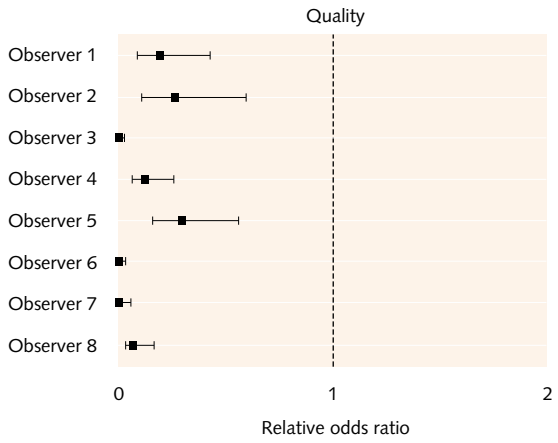


FIGURE 4.3
Image quality assessment showing regression coefficients of primary electronically cleansed 3D (PEC3D). It estimates the change in the log transformed diagnostic odds ratio (DOR) compared with primary uncleansed 2D (PU2D). A lower relative diagnostic odds ratio (RDOR) implies inferior image quality. Confidence intervals not reaching 1 indicate significant inferiority. Thus, this figure shows that all observers rated the image quality of PEC3D significantly less qualitative than PU2D

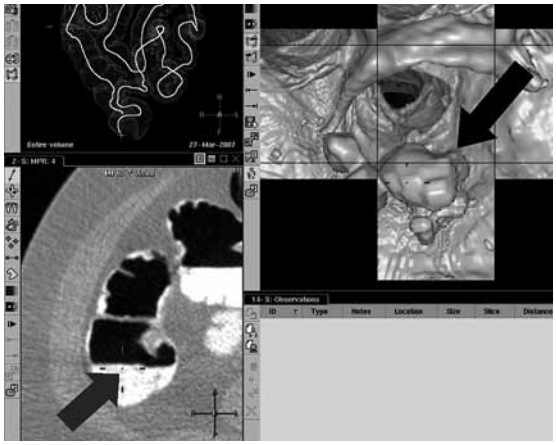


FIGURE 4.4
Floating debris (black arrow) as a result of incomplete electronic cleansing of inhomogeneously tagged faecal material (grey arrow on 2D image)

TABLE 4.6 **Number of Patients with Artefacts.**

Artefacts	Number of patients	Number of patients having artefacts classified as "disturbing, can cover polyps of ≥ 6mm"	Number of patients having artefacts classified as "disturbing, can cover polyps of ≥ 10mm"
		Primary electronically cleansed 3D	
Floating debris	55 (76%)	18 (25%)	1 (1%)
Holes in the colon wall	39 (54%)	2 (3%)	0 (0%)
Windmill artefacts	1 (1%)	0 (0%)	0 (0%)
Ridges	0 (0%)	0 (0%)	0 (0%)
		Primary uncleansed 2D	
Windmill artefacts	2 (3%)	0 (0%)	0 (0%)
Insufficient faecal tagging	2 (3%)	1 (1%)	0 (0%)

Table displays the number of patients with artefacts observed by at least 4 observers, the number of patients having artefacts classified by at least 4 observers as "disturbing, can cover polyps of ≥ 6mm" and the number of patients having artefacts classified by at least 4 observers as "disturbing, can cover polyps of ≥ 10mm".

FIGURE 4.5
Hole in the wall on the 3D image (black arrow). The white arrow indicates the corresponding colon wall in 2D. After electronic cleansing, the colon wall has become so thin that virtual holes appear between two air-containing structures

Compared with the abovementioned previous studies, in this study patients underwent a limited bowel preparation. There is a risk that the reported superiority of 3D in polyp detection would be neutralized by the reduction of visible colonic surface. In this study we report a large number of at least partially submerged polyps. This increases the risk of overlooking polyps [21, 22]. This is the reason why electronic cleansing was used in the 3D examinations. Electronic cleansing has been subject to study for several years now [16, 23–27]. Recently two comparative studies of electronic cleansing were published. In both studies electronic cleansing had an additional value in terms of sensitivity for some observers [23, 24]. In this study we assumed it had an additional value as well: nearly half of the polyps were in at least one position at least partially covered by faecal material (Table 4.2).

Specific artefacts of electronic cleansing are described in the literature [13] e.g. ridges, pseudopolyps due to partial volume effect and floating debris due to untagged faecal material. These may be the reason that electronic cleansing for primary CTC evaluation has not often been used. The electronic cleansing algorithm we used in this study was specially designed to overcome artefacts of distracting 'ridges' emanating from locations where air, soft tissue and tagged material meet [16]. These ridges were in fact noted by none of the observers in this study.

Floating debris, though, was detected in the majority of patients examined in PEC3D (Table 4.6). An important causeof debris is noise due to heterogeneously or insufficiently tagged stool. The three patients were excluded because of these artefacts. These artefacts stress the fact that more than just a good cleansing algorithm is important in order to achieve good 3D image quality i.e. a good tagging regimen, good patient compliance and good CT parameters. However, the three excluded patients were rated in PU2D by one to three observers "not diagnostic" as well. This stresses the fact that although 3D is more susceptible to tagging artefacts, 2D suffers from heterogeneous or insufficient tagging as well.

Artefacts were seen in the majority of patients reviewed in PEC3D. To be able to easily distinguish artefacts from polyps, it is important to be able to correlate electronically cleansed 3D images with complementary original uncleansed 2D images. This combination has limited the number of false positive findings when using PEC3D (Table 4.3). Using PEC3D did not statistically decrease specificity for any observer group at any size per-patient threshold i.e. true negative patients were not erroneously classified using PEC3D.

We have not studied a primary 3D reading method without electronic cleansing or primary 2D with electronic cleansing. Although interesting from a methodological point of view, we think that these approaches are not meaningful; the former is not since a large number of (partially) submerged polyps are prevented from being detected because they are otherwise covered by faecal material; the second approach is not since there is no need to electronically remove faecal material that can already be distinguished from colonic structures. However, one study [23] demonstrated an additional value in terms of polyp detection of cleansing in a 2D approach. An important difference compared with our study is the nature of the preparation: barium instead of iodine and no low-fibre diet. This results in more adherent and solid stool that is "mentally" more difficult to read than the quiet homogeneous fluid levels seen in our patient population. Therefore, electronic cleansing may prevent reader fatigue in this patient population.

The mean difference between PU2D and PEC3D was in accordance with the expected difference between both techniques. However, the expected baseline sensitivity for polyps 6 mm or larger was higher (70%) than the actual measured sensitivity (51%). This had consequences for the statistical power of the comparison; however, statistical significance was still reached for the group of inexperienced readers.

The higher per-polyp sensitivity of PEC3D mainly concerned polyps 6–9 mm (Table 4.3). The prevalence of adenomas with advanced features (i.e. villous components or high-grade dysplasia) in this size category tends to be low [28]. The joint guideline of the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer and the American College of Radiology recommends colonoscopy and polypectomy for polyps 6 mm or larger [1]. Thus, polyps in this size category may not be neglected.

The novice observers were trained according to the recommendations of the American College of Radiology and the European Society of Gastrointestinal and Abdominal Radiology [29, 30]; 50 CTC studies with 20–50% prevalence with personal feedback on all false positive and negative findings for polyps 6 mm or larger [29]. The response to training, though, is unpredictable and competence cannot be assumed after 50 cases [31]. In this study the two experienced observers (350 CTCs or more) outperformed the six novice observers. Thus, it is likely that the optimum number of training cases is more than 50, as suggested earlier [32]. However, the difference between these levels of experience in PEC3D is less than PU2D. So, in the phase of familiarization with CTC primary cleansed 3D is advantageous.

The group of experienced readers consisted of two observers compared with the group of inexperience readers that consisted of six observers. In general, experienced observers show less difference in polyp detection between review methods [33] and will perform better when

compared with inexperienced readers. Therefore, we expected the differences between both review methods to be insignificant in a very experienced observer group. This was confirmed by the two experienced reviewers in this study who performed very well using both methods. Thus, we think that although interesting from a methodological point of view adding four extra reviewers would not change the conclusion of this study.

A computer-aided detection algorithm has not been used in this study. Its effect will be twofold when applied to electronically cleansed data: the number of detectable polyps will increase as well as the number of detected artefacts. Since these effects are not yet balanced, this will be subject to further research.

This study has limitations. First the prevalence of polyps in this FOBT positive population was fairly high compared with an asymptomatic screening population. This may limit the generalization to an average risk screening population. Secondly, in this patient population we used a low-dose scan protocol combined with sub-millimetre slices. Noise in the images may have limited the quality of the cleansing since faecal material appeared less homogeneous. Still, the quality of the images was rated diagnostic in the vast majority of the cases i.e. only three cases were excluded. Thirdly, all patients had been prepared with oral iodine tagging resulting in a fairly homogeneously tagged colon content. Probably any electronic cleansing algorithm will perform optimal with homogeneously tagged stool [27]. Accordingly, we expect the algorithm to be better suited for removing pools of (iodine) fluid compared with adherent heterogeneous faecal residue encountered in barium tagging.

Fourthly, before this study the experienced observers had evaluated all patients in the framework of a comparative study of colonoscopy and CTC (data published elsewhere). These patients were evaluated atleast 1 year before this study with a primary 2D review method. In the period in between both studies atleast 100 other CTC examinations were read. So, it is not likely that this has influenced the performance characteristics of the experienced observers. Fifthly, in this study we have used an enhanced 3D display i.e. the unfolded cube display. The advantage of this technique is that it covers nearly all colonic mucosa without image distortion in a single fly-through [15], compared with the conventional 'endoscopic' view that needs a bidirectional fly-through to cover nearly all colonic mucosa. Therefore, our approach is a more time efficient method than a conventional 3D technique [15]. This may limit the generalizability of difference in review time, however not in accuracy.

In summary, we conclude that novice observers have a significantly higher sensitivity for the detection of clinically relevant polyps when using primary electronically cleansed 3D compared with primary 2D. For experienced observers, who performed better overall, there is no difference between both methods. Specificity is not affected when using primary electronically cleansed 3D. Therefore we recommend primary electronically cleansed 3D for novice observers in evaluating CTC in patients that have undergone limited bowel preparation.

References

1. Levin B, Lieberman DA, McFarland B, et al. Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. Gastroenterology 2008; 134:1570-1595
2. van Gelder RE, Birnie E, Florie J, et al. CT colonography and colonoscopy: assessment of patient preference in a 5-week follow-up study. Radiology 2004; 233:328-337
3. Iannaccone R, Laghi A, Catalano C, et al. Computed tomographic colonography without cathartic preparation for the detection of colorectal polyps. Gastroenterology 2004; 127:1300-1311
4. Jensch S, de Vries AH, Peringa J, et al. CT colonography with limited bowel preparation: performance characteristics in an increased-risk population. Radiology 2008; 247:122-132
5. Lefere P, Gryspeerdt S, Marrannes J, Baekelandt M, van HB. CT colonography after fecal tagging with a reduced cathartic cleansing and a reduced volume of barium. AJR Am J Roentgenol 2005; 184:1836-1842
6. Iafrate F, Hassan C, Zullo A, et al. CT colonography with reduced bowel preparation after incomplete colonoscopy in the elderly. Eur Radiol 2008; 18:1385-1395
7. Gryspeerdt S, Lefere P, Herman M, et al. CT colonography with fecal tagging after incomplete colonoscopy. Eur Radiol 2005; 15:1192-1202
8. Gluecker TM, Johnson CD, Harmsen WS, et al. Colorectal cancer screening with CT colonography, colonoscopy, and double-contrast barium enema examination: prospective assessment of patient perceptions and preferences. Radiology 2003; 227:378-384
9. Rex DK. Current colorectal cancer screening strategies: overview and obstacles to implementation. Rev Gastroenterol Disord 2002; 2 Suppl 1:S2-11
10. Weitzman ER, Zapka J, Estabrook B, Goins KV. Risk and reluctance: understanding impediments to colorectal cancer screening. Prev Med 2001; 32:502-513
11. Pickhardt PJ, Lee AD, Taylor AJ, et al. Primary 2D versus primary 3D polyp detection at screening CT colonography. AJR Am J Roentgenol 2007; 189:1451-1456
12. van Gelder RE, Florie J, Nio CY, et al. A comparison of primary two- and three-dimensional methods to review CT colonography. Eur Radiol 2006; 17:1181-1192
13. Pickhardt PJ, Choi JH. Electronic cleansing and stool tagging in CT colonography: advantages and pitfalls with primary three-dimensional evaluation. AJR Am J Roentgenol 2003; 181:799-805
14. van Rossum L, van Rijn A, Laheij R et al (2008) Random Comparison of Guaiac and immunochemical fecal occult blood tests for colorectal cancer in a screening population. Gastroenterology 135(1):82-90
15. Vos FM, van Gelder RE, Serlie IW, et al. Three-dimensional display modes for CT colonography: conventional 3D virtual colonoscopy versus unfolded cube projection. Radiology 2003; 228:878-885
16. Serlie IW, Vos FM, Truyen R, Post FH, van Vliet LJ. Classifying CT image data into material fractions by a scale and rotation invariant edge model. IEEE Trans Image Process 2007; 16:2891-2904
17. Walter Reed Army Medical Center. Walter Reed Army Medical Center. http://wramc.vcscreen.com/training/index.html. date accessed: 23-11-2005 Ref Type: Internet Communication
18. Fennerty MB, Davidson J, Emerson SS, Sampliner RE, Hixson LJ, Garewal HS. Are endoscopic measurements of colonic polyps reliable? Am J Gastroenterol 1993; 88:496-500
19. Gopalswamy N, Shenoy VN, Choudhry U, et al. Is in vivo measurement of size of polyps during colonoscopy accurate? Gastrointest Endosc 1997; 46:497-502
20. Pepe MS. The statistical evaluation of medical tests for classification and prediction. New York, NY: Oxford University Press, 2003; 35-65
21. Doshi T, Rusinak D, Halvorsen RA, Rockey DC, Suzuki K, Dachman AH. CT colonography: false-negative interpretations. Radiology 2007; 244:165-173
22. Gluecker TM, Fletcher JG, Welch TJ, et al. Characterization of lesions missed on interpretation of CT colonography using a 2D search method. AJR Am J Roentgenol 2004; 182:881-889

23. Johnson CD, Manduca A, Fletcher JG, et al. Noncathartic CT colonography with stool tagging: performance with and without electronic stool subtraction. AJR Am J Roentgenol 2008; 190:361-366
24. Juchems MS, Ernst A, Johnson P, Virmani S, Brambs HJ, Aschoff AJ. Electronic colon-cleansing for CT colonography: diagnostic performance. Abdom Imaging 2008; [Epub ahead of print]
25. Serlie IW, de Vries AH, Vos FM, et al. Lesion Conspicuity and Efficiency of CT Colonography with Electronic Cleansing Based on a Three-material Transition Model. AJR Am J Roentgenol 2008; 191:1493-502.
26. Wang Z, Liang Z, Li X, et al. An improved electronic colon cleansing method for detection of colonic polyps by virtual colonoscopy. IEEE Trans Biomed Eng 2006; 53:1635-1646
27. Zalis ME, Perumpillichira JJ, Magee C, Kohlberg G, Hahn PF. Tagging-based, electronically cleansed CT colonography: evaluation of patient comfort and image readability. Radiology 2006; 239:149-159
28. Kim DH, Pickhardt PJ, Taylor AJ. Characteristics of advanced adenomas detected at CT colonographic screening: implications for appropriate polyp size thresholds for polypectomy versus surveillance. AJR Am J Roentgenol 2007; 188:940-944
29. ACR practice guideline for the performance of Computed Tomography (CT) colonography in adults. http://www.acr.org/SecondaryMainMenuCategories/quality_safety/guidelines/dx/gastro/ct_colonography.aspx . date accessed 8-7-2008. Ref Type: Internet Communication
30. Taylor SA, Laghi A, Lefere P, Halligan S, Stoker J. European Society of Gastrointestinal and Abdominal Radiology (ESGAR): consensus statement on CT colonography. Eur Radiol 2007; 17:575-579
31. Taylor SA, Halligan S, Burling D, et al. CT colonography: effect of experience and training on reader performance. Eur Radiol 2004; 14:1025-1033
32. Spinzi G, Belloni G, Martegani A, Sangiovanni A, Del FC, Minoli G. Computed tomographic colonography and conventional colonoscopy for colon diseases: a prospective, blinded study. Am J Gastroenterol 2001; 96:394-400
33. Johnson CD, Chen MH, Toledano AY, et al. Accuracy of CT colonography for detection of large adenomas and cancers. N Engl J Med 2008; 359:1207-1217

Chapter 5

Influence of Tagged Fecal Material on Detectabilit of Colorectal Polyps at CT: Phantom Study

Ayso H. de Vries
Henk W. Venema
Jasper Florie
Chung Y. Nio
Jaap Stoker

American Journal of Roentgenology 2008; 191: W181–W189

Abstract

Objective

The purpose of this study was to determine the influence of tagged material on the minimal radiation dose needed to detect colorectal polyps at CT.

Materials and methods

The study was conducted in two phases. In the first,three experienced observers determined the visibility of sessile polyps (6 mm) at five contrast levels (300, 480, 790, and 1,040 HU and air) and five tube charge levels (10, 14, 20, 28, and 40 mAs) in an anthropomorphic phantom. Each polyp was present in one of eight possiblelocations. The mean tube charge threshold for 90% correct responses was determined for each contrast level. Blinded observers performed independent 2D readings. In the second phase of the study, three 150-cm virtual colons were evaluated at two contrast levels (300 and 480 HU) and at five tube charge levels between 20 and 80 mAs. The three colons contained 18 randomly located polyps. The mean tube charge threshold for 90% sensitivity was determined for each contrast level.

Results

In the first phase of the study, the estimated tube charge thresholds for 300, 480, and 790 HU were 24.0, 16.3, and 6.2 mAs. At 1,040 HU and in air, all polyps were detected atthe lowest tube charge setting (10 mAs). In the second phase, the tube charge thresholds for 90% sensitivity at 300 and 480 HU were 70 and 35 mAs, respectively.

Conclusion

If polyps are covered by fecal material, a considerably higher tube charge setting is needed for adequate visualization than is needed for polyps in a completely cleansed colon, especially when the density of the tagged residue is low.

Introduction

CT colonography (CTC) is a promising technique of screening for colorectal cancer and colonic polyps. An important disadvantage of the technique is that many patients consider the bowel preparation burdensome [1]. Labeling fecal material with an oral contrast agent (fecal tagging) enables patients to prepare for CTC with less extensive bowel preparation [2–4]. Minimizing bowel preparation, however, may increase patient compliance [5–7] but results in a higher amount of residual feces in the colon [8].

Because of reduced contrast (i.e., reduced difference in attenuation) between the polyp and the luminal contents, polyps covered by fecal material are less conspicuous than polyps surrounded by air, even if oral contrast material has been administered. Figure 5.1 shows that the submerged polyp is less visible than the haustral fold in air at 40 mAs. Another factor that influences polyp conspicuity is radiation dose. Especially in low- dose scan protocols, the noise of CT images is prominent. In Figure 5.1, this phenomenon is illustrated by the submerged polyp visible at 40 mAs but hardly visible at 5 mAs.

With regard to using CTC for screening, the goal should be the lowest possible patient burden and radiation dose. To our knowledge, the influence of the use of oral contrast material on the minimal radiation dose for qualitatively acceptable CTC has not been determined. We performed a phantom study to determine the magnitude of this effect.

Materials and Methods

Study Design

This phantom study consisted of two parts. The aim of the first phase was to determine the detectability of clinically relevant polyps (≥ 6 mm) in tagged material of various densities or in air as a function of the tube charge - that is, the product of tube current and exposure time divided by the pitch - in an idealized setup. We performed the second phase to translate the results of the first phase into clinically applicable performance values, such as sensitivity and number of falsepositive findings.

Phantom

A cylindric water-filled polyethylene drum with a diameter of 34 cm was used to mimic the abdomen of a fairly obese patient (Figure 5.2). We devised a synthetic colonic segment by placing a cylindric polymethyl methacrylate (PMMA) tube in the center of this drum.

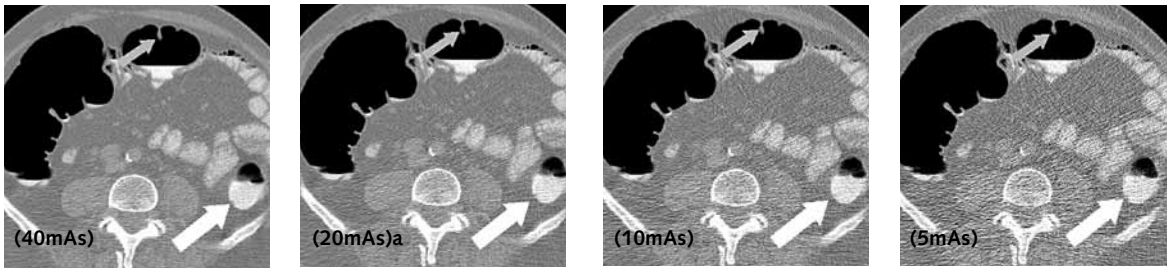


FIGURE 5.1
Influence of tagged material on visibility. A, 75-year-old man with colonic polyp. CT scan (40 mAs; window width, 1,250 HU; level, –50 HU) shows colonic wall and polyp (white arrow). Gray arrow indicates haustral fold in transverse colon surrounded by air. B–D, Simulated CT scans at tube charges lower than in A: 20 mAs (B), 10 mAs (C), and 5 mAs (D). White arrow indicates 6-mm polyp submerged in fecal material in descending colon proven at colonoscopy. Polyp is highly visible in A but hardly visible in D. In comparison, haustral fold in transverse colon surrounded by air (gray arrow) remains highly visible in D.

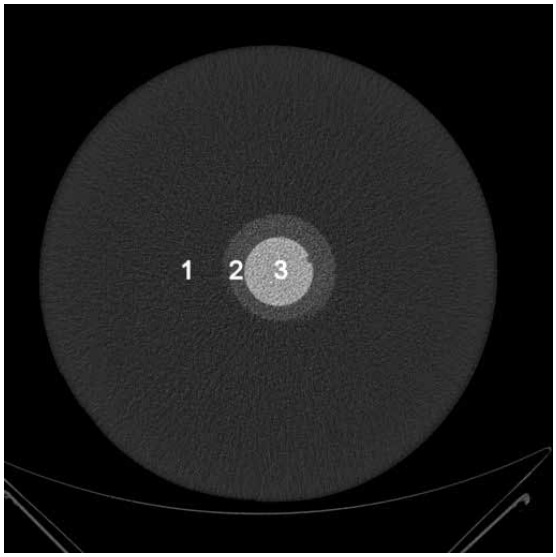


FIGURE 5.2
Phantom (width, 1,250 HU; level, –50 HU). CT scan shows water-filled drum (1) with centrally placed polymethyl methacrylate cylinder (2), representing colon. Cylinder is filled with contrast material (3). Contrast between lumen and border is 480 HU. Filling defect represents 6-mm polyp.

We positioned this phantom with the long axis of the tube parallel to the long axis of the drum. The tube contained eight PMMA rings with a thickness of 11 mm and an inner diameter of 50 mm in the order of the diameter of a distended colon segment. Each ring could contain a PMMA hemisphere with a diameter of 6 mm, mimicking a 6-mm sessile polyp. We chose 6 mm because it often is considered the smallest clinically relevant size for polyps [9]. All rings were separated by a thin PMMA ring (inner diameter, 40 mm; thickness, 2 mm), representing an artificial haustral fold.

We use the terms 'contrast' and 'contrast level' to indicate the difference in attenuation between polyps and contrast material or air in the lumen. The attenuation of PMMA is approximately 120 HU. We used PMMA because it is readily available and its attenuation is reasonably close to that of soft tissue. The tube was completely filled with either contrast material or air for each scan. The contrast material was a mixture of iodine (sodium amidotrizoate, 300 mg I/mL, hospital pharmacy) and water. We chose this concentration of contrast material so that the differences in attenuation between the polyp and the lumen would be approximately 300, 500, 800, and 1,000 HU. These attenuation values encompassed the range of contrast measured in tagged CTC examinations performed at our hospital with a regimen of 4 L of polyethylene glycol and 200 mL of 300 mg I/mL of meglumine ioxithalamate (Telebrix, Guerbet). Other regimens with similar attenuation have been described [10, 11].

CT Data and Preprocessing

We obtained scans of this phantom with a 64-MDCT scanner (Brilliance 64, Philips Healthcare) with the following parameters: 120 kV; collimation, 64; pitch, 1.2. Effective tube charge settings were used in steps of the square root of two from the lowest tube charge value available: 10, 14, 20, 28, 40, 56, and 80 mAs. In this way, the noise in the images decreased by a constant factor between successive values. For the first phase of the study, scans of the phantom filled with the four iodine mixtures and air were obtained with tube charge values increasing from 10 mAs until the value was reached at which all polyps were visible for all observers. Because we expected that the task for the observers would be more difficult for the second phase of the study, higher tube charge values (20–80 mAs) were used in this phase. In the second phase, only the two lowest concentrations were used. For each contrast level, a scan at 300 mAs was made for reference purposes. All images were reconstructed at a slice thickness of 0.9 mm and an increment of 0.45 mm with reconstruction filter C (sharper standard filter). The pixel size was 0.6 mm.

Measurement of Contrast and Noise

Because it is difficult to exactly obtain the intended contrast levels, for each intended contrast level we measured the actual contrast level, that is, the difference in attenuation

between the wall of the colonic phantom and its content. We also measured the SD as a measure of noise in the images at the location of the polyps. This measurement of noise also was performed at the same position in a cross section of the drum that contained water only. This approach made it possible to determine the influence of the tagged material on noise. The measuring procedure is described in the Appendix .

Observer Studies

Scans of the phantom were used to construct virtual colons for the observer studies. These colons consisted of a large number of rings separated by thin PMMA rings mimicking haustral folds. Before images of the virtual colons were presented to the observers, a research fellow assessed the visibility of the polyps for all contrast levels at the lowest tube charge setting (10 mAs). If the polyps could absolutely not be missed, the particular contrast level was not further evaluated by the observers. This measure was taken because a visibility threshold cannot be determined if all polyps are highly visible. For all other contrast levels, the visibility of polyps was determined in the first phase of the study. Three observers, an experienced abdominal radiologist, a resident in radiology, and a research fellow reviewed all images. All observers had read more than 250 colonoscopically verified CTC examinations before this study. The observers reviewed the images using a 2D review method (ViewForum, Philips Healthcare) and indicated the location of the polyp with a cursor. All annotations were digitally stored. Because window and level settings can influence the conspicuity of submerged polyps [12], the window width and level were preset by two researchers at a setting deemed optimal for polyp detection at the given contrast level. These settings varied from 1,000/200 HU for contrast of 300 HU to 2,400/550 HU for contrast of 1,000 HU. The observers, however, were free to adjust the window setting to their preference. This procedure is in accordance with that used in the clinical evaluation of CTC scans.

First Phase of Study

In the first phase of the study, we determined the visibility of the polyps in an eight-alternative forced-choice paradigm. In this paradigm a polyp always was present in each ring at one of eight possible locations (Figure 5.3), and the observers had to choose one of the eight locations. With this paradigm we could determine the visibility of polyps in a time-efficient way.

For every contrast level, a virtual colon was composed of 80 rings. Each virtual colon contained rings from scans made at five tube charge levels (10, 14, 20, 28, and 40 mAs), so each observer had to locate 16 polyps at each combination of tube charge and contrast. Each ring was rotated so that the polyp was present in one of eight possible clock-face positions on the phantom wall. The choice of location was made at random (Figure 5.3).

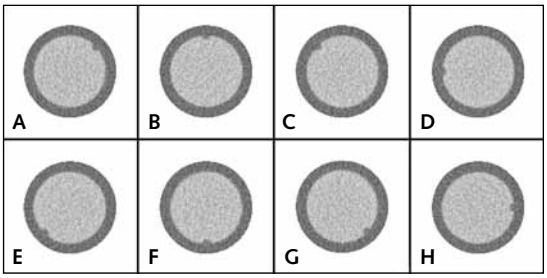


FIGURE 5.3
Colonic phantom in first phase of study. CT scans (300 mAs; 300 HU; width, 1,000 HU; level, 200 HU) show polyp in one of eight possible locations.

A Half past one-o'clock position.
B Twelve-o'clock position.
C Half past ten-o'clock position.
D Nine-o'clockw position.
E Half past seven-o'clock position.
F Six-o'clock position.
G Half past four-o'clock position.
H Three-o'clock position.

After the reading by the three observers, a researcher not involved in reading the data assessed the correctness of the annotations. On the basis of these annotations, the number and average percentages of correct responses were determined for every combination of tube charge level and contrast level. Psychometric curves with constant steepness were fitted to this average percentage of correct responses as a function of tube charge [13]. For the fitting we applied a maximum likelihood procedure [14] for every contrast level. At each contrast level we determined the tube charge for 90% correct responses and the SD of this estimate [15]. We also determined whether the number of correctly detected polyps increased significantly with the density of the contrast material. We counted the correctly detected polyps for each virtual colon for each observer and for the three observers combined. A chi-square test (SPSS 15.0, SPSS) was used for the comparison of proportions in two independent groups.

Contrast-to-Noise Ratio

The conspicuity of a polyp of a certain size depends both on the contrast of the polyp to its surroundings and on the noise in the image, quantified by the contrast-to-noise ratio (CNR) [16, 17]. For different contrast levels, the conspicuity at the tube charge for 90% correct responses is by definition the same. The CNRs at these tube charge levels should therefore also be equal. We checked whether this was indeed the case to cross-validate our results.

The SD of the attenuation is used as a measure of noise. The CNR at 90% correct responses was determined by dividing each contrast value by the SD for that contrast value at the tube charge for 90% correct responses. The SD can be considered inversely proportional to the square root of the tube charge [18]. Therefore, the SD at the tube charge for 90% correct responses can be derived from the SD in the 300 mAs reference scan as follows:

$$SD_{mAs@90\%} = SD_{300mAs}\sqrt{\frac{300}{mAs@90\%}}$$

where mAs@90% is the tube charge value for 90% correct responses.

Second Phase of Study

Although in an eight-alternative forced-choice design, the tube charge value for 90% correct responses can be determined efficiently, no clinically applicable measures such as sensitivity and number of false-positive findings can be determined. In the second phase of the study, these performance values were measured, and the threshold tube charge value for a sensitivity of at least 90% was determined.

The second phase was performed at the two lowest contrast levels, 300 and 480 HU, and scans were obtained at five tube charge levels between 20 and 80 mAs. For each contrast and tube charge level, three virtual colons were composed, each consisting of 120 rings (Figure 5.4). Five percent (18 of 360) of these rings contained a polyp. Rings containing polyps were randomly distributed among the three colons. All rings were randomly rotated. Thus the virtual colons in the second phase of the study differed in three important aspects from those of the first phase: The prevalence of polyps was only 5%; in each ring the polyps were situated at any angle along the phantom wall; and colons were composed of rings scanned at the same tube charge levels.

The same three observers as in the first phase evaluated the virtual colons starting at the level nearest the tube charge value for 90% correct responses determined in the first phase. The tube charge was increased until the level was reached at which the average per-polyp sensitivity of the observers was greater than 90%.

The observers knew that the polyps, if present, were of the same size and shape as in the first phase and were in random locations. The observers did not, however, know the polyp prevalence. After the reading, a researcher not involved in interpretation of the data assessed the correctness of the scores by comparing the true polyp locations with the annotations of the observers. For each

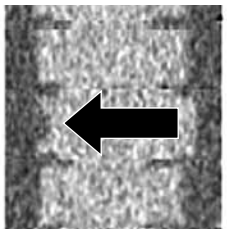


FIGURE 5.4
Coronal section of part of virtual colon (24 of 120 rings) composed of scans at 80 mAs (width, 1,000 HU; level, 200 HU) for second phase of study. Contrast between colon and its contents is 300 HU. Arrow indicates 6-mm sessile polyp.

contrast level, per-polyp sensitivity and the number of falsepositive findings were determined for every tube charge value used. The tube charge value for 90% sensitivity was determined by linear interpolation.
Also in the second phase, we determined whether the number of correctly detected polyps significantly increased with the density of the tagging material (300 and 480 HU). In this case, we compared the total numbers of correctly detected polyps at 28 and 40 mAs for which data were available for both densities. This procedure was performed for each observer and for the three observers combined. The statistical procedure was the same as that used in the first phase of the study.

Ratio of Thresholds of the First and Second Phases of the Study

The eight-alternative forced-choice task of the first phase of the study was relatively easy. The second phase of the study was performed to obtain results closer to the reality of clinical practice, but the task in this phase was more difficult for the observers. The ratio of the tube charge value for 90% sensitivity (second phase) to the tube charge value for 90% correct responses (first phase) was determined as an indicator of the relative difficulty of detection of polyps in the second versus the first phase of the study. This ratio can be used to translate the thresholds to extrapolate the results of the first phase of the study to those of the more clinically orientated second phase. The significance of this ratio is elucidated in the Discussion section.

Results

Measurement of Contrast and Noise

The actual contrast values were close to the intended values. The actual values of contrast and SD are listed in Table 5.1. Table 5.1 shows that the SD, which is a measure of the noise at the location of the polyps, increases as he contrast level increases. For the highest contrast level (1,040 HU), the noise is approximately twice as high as for a phantom filled with air (SD, 70 and 33 HU, respectively).

TABLE 5.1 **Contrast Values and the Standard Deviation (SD) at the Location of the Polyps**

Nominal contrast [HU]	Actual contrast [HU]	SD [HU] at site polyps at 300 mAs
–1120*	–970	33
300	300	51
500	480	58
800	790	62
1000	1040	70

*1,000 HU (air) – 120 HU (polymethyl methacrylate).

Contrast is difference between attenuation of material in colonic lumen and that of polyps. Actual contrast of 480 HU, for example, corresponds to differences of approximately 600 HU (lumen) and 120 HU (polyp) in the phantom. The contrast and SD at the location of the polyps were measured as explained in the Appendix. In the drum filled with water only, the SD at the same location as the polyps was 40 HU.

Observer Study

The polyps in the air-filled colon were clearly visible at 10 mAs (Figure.5.5). Consequently, this contrast level was not further evaluated

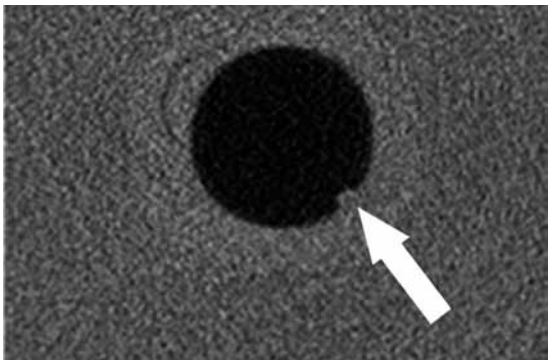


FIGURE 5.5
Polymethyl methacrylate colon phantom in center of 34-cm-diameter water-filled drum. CT scan (10 mAs; width, 2,000 HU; level, 0 HU) shows 6-mm polyp (arrow) surrounded by air. Polyp was considered visible beyond doubt.

First Phase of Study

With increasing tube charge levels, the number of errors decreased. At contrast levels of 300, 480, and 790 HU, the observers correctly located all polyps at, respectively, 40, 28, and 20 mAs. The tube charges for 90% correct responses for the three observers combined were 24.0, 16.3, and 6.2 mAs for these three contrast levels (Figure.5.6). At 790 HU, nearly all polyps were located correctly. The fitting procedure remained valid under these circumstances [14, 15], although the precision of the estimate was impaired. This phenomenon was evidenced by the much higher SD of the estimated threshold (Table 5.2). At 1,040 HU, all polyps were correctly located at all tube charge levels. Therefore, the tube charge value for 90% correct responses could not be determined for this contrast level.

For the 300, 480, and 790 HU contrast levels, the numbers of correctly detected polyps in each virtual colon (i.e., for the five tube charge levels combined) are listed in Table 5.3. Nearly all differences between contrast levels were significant.

The CNRs at the tube charge value for 90% correct responses varied between 1.7 and 1.9, with an average of 1.8 (Table 5.2).

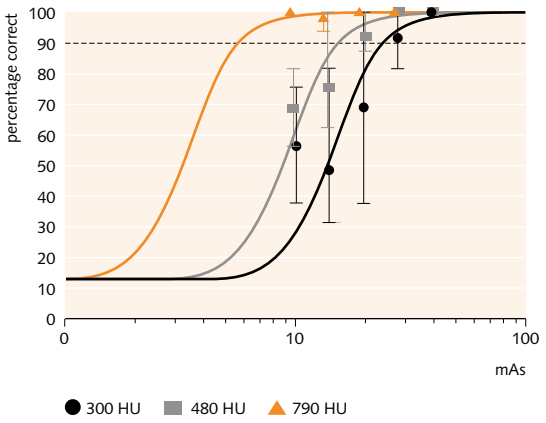


FIGURE 5.6
Graph shows percentages of correct responses for each tube charge level for three lowest contrast levels. Data points are mean values for three observers. Ranges from lowest to highest scores are indicated. Curves are psychometric curves fitted to these data points. Dashed line indicates score of 90% correct.

TABLE 5.2 **Tube Charges for 90% Correct Responses Determined in First Phase of Study**

Contrast (HU)	Tube Charge in mAs for 90% Correct Responses (SD of Tube Charge in %)	Contrast-to-Noise Ratio at Tube Charge for 90% Correct Responses
300	24.0 (5%)	1.7
480	16.3 (6%)	1.9
790	6.2 (27%)	1.8
Average		1.8

The tube charge for 90% correct responses and corresponding contrast-to-noise ratio for polyps at the highest contrast level (1,040 HU) and for polyps in air could not be determined because all polyps were visible at the lowest available tube charge (10 mAs).

TABLE 5.3 **Number of Polyps Correctly Detected in First Phase of Study**

Contrast (HU)	Observer 1(n = 80)	Observer 2 (n = 80)	Observer 3 (n = 80)	All Observers (n = 240)
300 HU	49	73	53	175
480 HU	67	77	65	209
790 HU	80	80	79	239
1040 HU	80	80	80	240
300-480 HU	p<0.001	n.s.	p=0.031	p<0.001
480-790 HU	p<0.001	n.s.	p<0.001	p<0.001
300-790 HU	p<0.001	p=0.014	p<0.001	p<0.001

Significance of differences between the three lowest contrast levels and 1,040 HU were not determined because at 790 HU all but one polyp were detected.
NS = not significant.

Second Phase of Study

Sensitivities and numbers of false-positive findings by the three observers combined at contrast levels of 300 and 480 HU are shown in Figure 5.7. Sensitivity increased with tube charge level and with the density of the contrast material. The sensitivity at a contrast of 300 HU exceeded 90% at 80 mAs and at a contrast of 480 HU exceeded 90% at 40 mAs. Using linear interpolation, we obtained tube charges for 90% sensitivity of 70 mAs and 35 mAs for 300 HU and 480 HU, respectively. At these tube charge levels, each observer had a maximum of one false-positive finding. For 300 and 480 HU, the total numbers of correctly detected polyps per observer at 28 and 40 mAs are listed in Table 5.4. All observers detected more polyps in the colon filled with the densest contrast material. This difference was statistically significant for two of the three observers and for all observers combined.

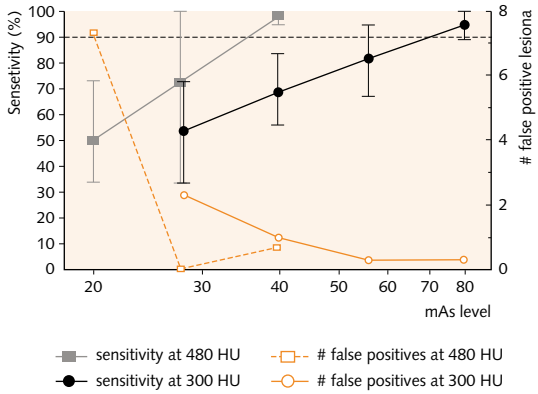


FIGURE 5.7
Graph shows sensitivity and number of false-positive findings at contrast levels of 300 HU (28–80 mAs) and 480 HU (20–40 mAs). Data points are average values for three observers. Sensitivity ranges from lowest to highest value are indicated. Horizontal dashed line represents 90% sensitivity.

TABLE 5.4 **Total Number of Correctly Detected Polyps at 28 and 40 mAs Combined in Second Phase of Study**

Attenuation (HU)	Observer 1 (n = 36)	Observer 2 (n = 36)	Observer 3 (n = 36)	All Observers (n = 108)
300	16	25	25	66
480	23	33	35	91
300-480	NS	p=0.017	p=0.002	p<0.001

Data at 28 and 40 mAs were chosen because at the other tube charge, only one contrast level was evaluated. NS = not significant.

Ratio

The ratios of the tube charge for 90% sensitivity (second phase of the study) and the tube charge for 90% correct responses (first phase of the study) were 2.9 for 300 HU and 2.2 for 480 HU, with an average value of 2.5.

Discussion

The results of this CTC study show that the minimal radiation dose needed to visualize a 6-mm polyp increases considerably if the contrast between the polyp and its surroundings is reduced. Although all polyps in air were well visible at 10 mAs, approximately 70 mAs was needed to achieve 90% sensitivity at a contrast level of 300 HU.

We determined the minimal tube charge level to visualize a clinically relevant polyp in two study phases. The first phase was performed to obtain the threshold visibility in a simple and time-efficient way. It showed that if contrast was increased from 300 to 790 HU, the tube charge for 90% correct responses decreased from 24 to 6.2 mAs. The CNRs at these tube charge levels were nearly constant. This finding can be expected because the detectability of structures with relatively low contrast, such as polyps, is primarily limited by noise on the images [18]. With a 1.8 mean CNR at the tube charge for 90% correct responses, the tube charge level for 90% correct responses at the highest contrast (1,040 HU) and for air can be estimated to be 4 and 1 mAs, respectively.

We realized that extrapolation to a clinical setting is a problem with phantom studies. Therefore, we performed the second phase of the study to obtain results closer to the reality of clinical practice. Important points of the setup of the second phase are that only a few polyps were present, at uncertain locations, and that measures of practical clinical relevance could be determined. We expected that in this setup, the task for the observers in the second phase of the study would be more difficult than in the first phase. Our expectation turned out to be correct. In the second phase of the study, a sensitivity of 90% at 300 and 480 HU was obtained at, respectively, 70 and 35 mAs, with less than one false-positive finding per observer per tube charge level. Thus, this sensitivity was obtained at tube charge levels approximately 2.5 times higher than the corresponding levels in the first phase of the study.

This ratio depends on factors such as the number of locations in which a polyp can be present (only eight for the first phase of the study and at a large number of locations in the second phase), the certainty or lack of certainty that a polyp is present, and the possibility of choice (forced in the first phase of the study, free in the second phase). Thus this ratio reflects the relative difficulty of the tasks in the two phases, and the ratios can therefore be expected to be similar for each contrast level.

The independence of contrast level was also found by Burgess and Ghandeharian [19], who compared the percentage of correct responses in a two-alternative forced-choice study with those of alternative forced-choice studies with up to 1,800 possible locations. Therefore, we could have estimated this ratio using only one of the contrast levels in the second phase of the study. For cross check and to improve accuracy and precision, we decided to use two contrast levels. We could have performed the second phase of the study with a third contrast level (790 HU). However, we did not do so since nearly all polyps were visible at this contrast level in the first phase of the study. As a consequence, the precision of the threshold for the first phase was rather poor (Table 5.3). Using the ratio of 2.5, we can estimate that for the second phase of the study, the tube charges for 90% sensitivity should be on the order of 15, 10, and 3 mAs for contrast of 790 and 1,040 HU and for air.

Our results are for data from all three observers combined. Both observer studies showed interobserver variability. This finding is evident from the error bars in Figures 5.6 and 5.7. The results in Tables 5.3 and 5.4 show that systematic differences were present between observers. Yet the general trend is confirmed by the results for the individual

observers, which consistently showed an increase in polyp detectability with an increase in the density of the tagging material. In most cases this increase was significant. We conclude that although individual thresholds may vary, the influence of tagged material on visibility is important.

A number of phantom studies and simulation studies have been conducted to determine optimal scan parameters for CTC [20– 27]. In all of those studies, the visibility of polyps was investigated in air only. Because of differences in study design and scan parameters, a comparison of these studies is difficult. The main finding is that all polyps 6 mm in diameter or larger are detected, even at the lowest tube charge levels (5 or 10 mAs). These studies are thus in line with our findings on polyps in air.

For this study we used the range of contrast levels measured in tagged CTC examinations at our hospital. Polyp detectability increases with the density of the tagging material. Aiming at even higher densities than used in this study may seem advantageous. It appears, however, that as contrast level increases, image noise increases as well (Table 5.1). This phenomenon is a consequence of the increased attenuation of x-rays in dense material. The increase in noise, also noted by Zalis et al. [28], counteracts the improved visibility owing to the higher contrast and may eventually lead to streak artifacts in the directions of the highest attenuation of the x-rays. For these reasons, striving for higher densities of tagging material than the ones used in this study may be of limited value.

Many factors contribute to the quality of CT images. In this study we varied the density of the contrast agent and the tube charge level. Other image-quality factors such as tube voltage, pitch, and slice thickness were left constant and were not subject to study. The conspicuity of polyps is also considerably influenced by the size of the patient. In this study we used a water-filled drum with a diameter of 34 cm, mimicking a rather corpulent patient. If we had used a smaller human phantom, the noise would have been lower, and therefore the same visibility of the polyps would have been obtained with lower tube charge values. In state-of-the-art scanners, dose modulation can be used in which the tube current can be automatically adapted to the size of a patient. The results of this study can be used to determine the reference setting for dose modulation.

With regard to ability to generalize our results to other CT scanners, it is important to realize that CT scanners differ in effective dose and image quality when scans are obtained with identical parameter settings. These differences have to be taken into account in translation of our threshold values to other scanners. The conclusion of our study, however, applies to any CT scanner.

The main factor in visibility of a polyp is its size. The size of a polyp correlates with its malignant potential [29–31]. Polyps measuring 6–9 mm in diameter, especially polyps 10 mm in diameter and larger, are associated with increased risk of abnormal histologic findings. Polyps smaller than 6 mm can be disregarded because of minimal risk. This policy is endorsed by a CTC consensus proposal by Zalis et al. [9], which defines a normal colon as a colon without polyps larger than 5 mm, and a European Society of Gastrointestinal and Abdominal Radiology consensus statement [32] that a reasonable minimum size for reported polyps is 5 or 6 mm, depending on local preference.

We found that the visibility of polyps is substantially reduced when the polyp is submerged in tagged material with relatively low density. In practice, however, the overall per-polyp sensitivity would be affected less because not all segments are completely filled with tagged material, even if a regimen of reduced bowel preparation is undertaken. Owing to acquisition of prone and supine images, a polyp may not be surrounded by tagged material in both scan positions. We believe, however, that one should aim to visualize all relevant polyps and thus choose the scan parameters to reach this goal. Doing so means that the dose of the scan has to be increased in comparison with that of a scan in which the colon is completely cleansed. The higher dose level also may influence the visibility of extracolonic findings.

This study had limitations. Although the setup in the second phase was chosen to more closely resemble a clinical situation than did the first phase, it still is an idealization of real life. Polyp detection in patients can be more difficult than it was in the second phase of the study. Thus tube charge values needed in practice to adequately visualize 6-mm polyps may be somewhat higher than in the study.

A second limitation was that the colon phantom was made of PMMA. This material has an attenuation of approximately 120 HU, slightly higher than the 40-HU attenuation of colonic wall and polyps in vivo [33]. We reported polyp conspicuity as a function of the contrast, that is, the difference between the attenuation of polyps and that of the luminal content. Our results therefore also apply to these contrast levels in vivo in that each contrast level corresponds to slightly lower attenuation of the tagging material. Because the noise at the location of the polyp depends on the attenuation of the tagging material (Table 5.1), the noise would also be somewhat lower. One can calculate from the data in Table 5.1 that this difference is slight (≈ 4%) and that the main outcome of our study remains unaffected.

A third limitation was that we considered detection of polyps only in a colon completely filled with contrast material. The increase in noise due to increased attenuation in the tagged material would be lower in a smaller lumen or a lumen only partly filled with tagged material. Therefore, in these situations, the visibility of the polyps for the same tube charge level would be somewhat better.

We used a homogeneous iodine solution to mimic colonic content because it completely covered the inner surface of the phantom. It is known, however, that different types of preparation produce tagging with varying degrees of heterogeneity [11]. Therefore, a fourth limitation was that the influence of heterogeneity of fecal material was not evaluated. Heterogeneity may have a negative effect on the visibility of colonic lesions. With inhomogeneous tagging, it can be expected that the tube charge values needed to adequately visualize a 6-mm polyp will have to be somewhat increased.

This phantom study is, to our knowledge, the first to show the negative influence of tagging on the detectability of polyps covered by tagged material. For adequate visualization of these polyps, a considerably higher tube charge level is needed than for polyps in a cleansed colon, especially when the density of the tagging material is low.

References

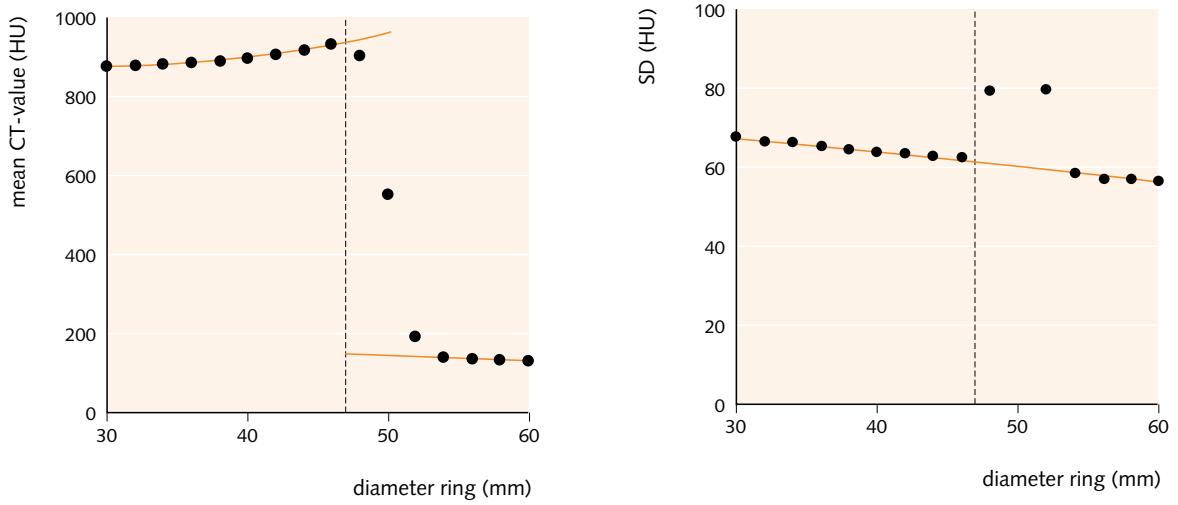
1. van Gelder RE, Birnie E, Florie J, et al. CT colonography and colonoscopy: assessment of patient preference in a 5-week follow-up study. Radiology 2004; 233:328–337
2. Iannaccone R, Laghi A, Catalano C, et al. Computed tomographic colonography without cathartic preparation for the detection of colorectal polyps. Gastroenterology 2004; 127:1300–1311
3. Lefere P, Gryspeerdt S, Marrannes J, Baekelandt M, Van Holsbeeck B. CT colonography after fecal tagging with a reduced cathartic cleansing and a reduced volume of barium. AJR 2005; 184:1836– 1842
5. Gluecker TM, Johnson CD, Harmsen WS, et al. Colorectal cancer screening with CT colonography, colonoscopy, and double-contrast barium enema examination: prospective assessment of patient perceptions and preferences. Radiology 2003; 227:378–384
6. Rex DK. Current colorectal cancer screening strategies: overview and obstacles to implementation. Rev Gastroenterol Disord 2002; 2[suppl 1]:S2–S11
7. Weitzman ER, Zapka J, Estabrook B, Goins KV. Risk and reluctance: understanding impediments to colorectal cancer screening. Prev Med 2001; 32:502–513
8. Dachman AH, Dawson DO, Lefere P, et al. Comparison of routine and unprepped CT colonography augmented by low fiber diet and stool tagging: a pilot study. Abdom Imaging 2007; 32:96–104
9. Zalis ME, Barish MA, Choi JR, et al. CT colonography reporting and data system: a consensus proposal. Radiology 2005; 236:3–9
10. Johnson KT, Carston MJ, Wentz RJ, Manduca A, Anderson SM, Johnson CD. Development of a cathartic-free colorectal cancer screening test using virtual colonoscopy: a feasibility study. AJR 2007; 188:145; [web]W29–W36
11. Zalis ME, Perumpillichira JJ, Magee C, Kohlberg G, Hahn PF. Tagging-based, electronically cleansed CT colonography: evaluation of patient comfort and image readability. Radiology 2006; 239:149–159
12. Slater A, Taylor SA, Burling D, Gartner L, Scarth J, Halligan S. Colonic polyps: effect of attenuation of tagged fluid and viewing window on conspicuity and measurement: in vitro experiment with porcine colonic specimen. Radiology 2006; 240:101–109
13. Veldkamp WJ, Thijssen MA, Karssemeijer N. The value of scatter removal by a grid in full field digital mammography. Med Phys 2003; 30:1712–1718
14. Harvey LO Jr. Efficient estimation of sensory thresholds with ML-PEST. Spat Vis 1997; 11:121–128
15. Coppens JE, Franssen L, van den Berg TJ. Reliability of the compensation comparison method for measuring retinal stray light studied using Monte-Carlo simulations. J Biomed Opt 2006; 11: 054010
16. Nakayama Y, Awai K, Funama Y, et al. Lower tube voltage reduces contrast material and radiation doses on 16-MDCT aortography. AJR 2006; 187: 1266; [web]W490–W497 17. Doyle P, Martin CJ, Gentle D. Application of contrast-to-noise ratio in optimizing beam quality for digital chest radiography: comparison of experimental measurements and theoretical simulations. Phys Med Biol 2006; 51:2953–2970
18. Kalender W. Computed tomography. Erlangen, Germany: Publicis, 2005:98–153
19. Burgess AE, Ghandeharian H. Visual signal detection. Part 2. Signal-location identification. J Opt Soc Am A 1984; 1:906–910
20. Embleton KV, Nicholson DA, Hufton AP, Jackson A. Optimization of scanning parameters for multi-slice CT colonography: experiments with synthetic and animal phantoms. Clin Radiol 2003; 58:955–963
21. Johnson KT, Johnson CD, Anderson SM, Bruesewitz MR, Mccollough CH. CT colonography: determination of optimal CT technique using a novel colon phantom. Abdom Imaging 2004; 29: 173–176
22. Ozgun A, Rollven E, Blomqvist L, Bremmer S, Odh R, Fransson A. Polyp detection with MDCT: a phantom-based evaluation of the impact of dose and spatial resolution. AJR 2005; 184:1181–1188
23. Taylor SA, Halligan S, Bartram CI, et al. Multidetector row CT colonography: effect of collimation, pitch, and orientation on polyp detection in a human colectomy specimen. Radiology 2003; 229:109–118

24. Wessling J, Fischbach R, Meier N, et al. CT colonography: protocol optimization with multidetector row CT - study in an anthropomorphic colon phantom. Radiology 2003; 228:753–759
25. Won HJ, Choi BI, Kim SH, Kim YI, Youn BJ, Han JK. Protocol optimization of multidetector computed tomography colonography using pig colonic phantoms. Invest Radiol 2005; 40:27–32
26. van Gelder RE, Venema HW, Florie J, et al. CT colonography: feasibility of substantial dose reduction - comparison of medium to very low doses in identical patients. Radiology 2004; 232:611–620
27. Wessling J, Fischbach R, Borchert A, et al. Detection of colorectal polyps: comparison of multidetector row CT and MR colonography in a colon phantom. Radiology 2006; 241:125–131
28. Zalis ME, Perumpillichira JJ, Kim JY, Del FC, Magee C, Hahn PF. Polyp size at CT colonography after electronic subtraction cleansing in an anthropomorphic colon phantom. Radiology 2005; 236:118–124
29. Atkin WS, Morson BC, Cuzick J. Long-term risk of colorectal cancer after excision of rectosigmoid adenomas. N Engl J Med 1992; 326:658–662
30. O'Brien MJ, Winawer SJ, Zauber AG, et al. The National Polyp Study: patient and polyp characteristics associated with high-grade dysplasia in colorectal adenomas. Gastroenterology 1990; 98: 371–379
31. Simons BD, Morrison AS, Lev R, Verhoek-Oftedahl W. Relationship of polyps to cancer of the large intestine. J Natl Cancer Inst 1992; 84:962–966
32. Taylor SA, Laghi A, Lefere P, Halligan S, Stoker J. European Society of Gastrointestinal and Abdominal Radiology (ESGAR): consensus statement on CT colonography. Eur Radiol 2007; 17: 575–579
33. Neri E, Vagli P, Picchietti S, et al. CT colonography: contrast enhancement of benign and malignant colorectal lesions versus fecal residuals. Abdom Imaging 2005; 30:694–697

Appendix

Procedure for Measuring Contrast Level and Noise

For each filling of the phantom, contrast between the polyp and its surroundings and the SD was measured, both at the location of the polyps. Because the polyps were located at a high-contrast interface that prohibits straightforward measurements, the following procedure was used. The attenuation of the iodine contrast agent, air, and polymethyl methacrylate (PMMA) and the SD of the attenuation appeared to vary systematically between the center and the periphery of the phantom. Therefore, both the average and the SD of attenuation were determined in regions of interest consisting of contiguous concentric rings around the center of the phantom. The rings were 2 mm wide and had diameters of 30–46 mm within the lumen and 54–60 mm within the PMMA border. Linear or quadratic curves were fitted to these values. Contrast was taken as the difference between the attenuation of the material within the lumen (iodine contrast agent or air) and that of PMMA, both extrapolated to a ring with a diameter of 47 mm. At the same position, the SD was obtained by interpolation with the fitted values at 30–46 and 54–60 mm (Figure 5.8).



A B

FIGURE 5.8

Measuring contrast level and noise. A and B, Graphs show mean (A) and SD (B) of attenuation in regions of interest with shape of concentric rings within lumen of phantom and within polymethyl methacrylate border. Circles indicate measured values; lines, fitted values. Value of contrast and SD at site of polyp was taken as difference between fitted values of contrast and SD for ring with diameter of 47 mm (indicated by dashed vertical line), on which each polyp is centered. SDs for regions of interest at 48 and 52 mm are high because these regions of interest contain pixels of both lumen and border. SD at 50 mm is higher than 100 HU and therefore not shown.

Chapter 6

Does a Computer-Aided Detection Algorithm in a Second Read Paradigm Enhance The Performance of Experienced Computed Tomography Colonography Readers in a Population of Increased Risk?

Ayso H. de Vries
Sebastiaan Jensch
Marjolein H. Liedenbaum
Jasper Florie
Chung Y. Nio
Roel Truyen
Shandra Bipat
Evelien Dekker
Paul Fockens
Lubbertus C. Baak
Jaap Stoker

European Radiology 2009; 19: 941–950

Abstract

Objective

We prospectively determined whether computer-aided detection (CAD) could improve the performance characteristics of computed tomography colonography (CTC) in a population of increased risk for colorectal cancer.

Materials and methods

We included 170 consecutive patients that underwent both CTC and colonoscopy. All findings ≥ 6 mm were evaluated at colonoscopy by segmental unblinding. We determined perpatient sensitivity and specificity for polyps ≥ 6 mm and ≥ 10 mm without and with computer-aided detection (CAD). The McNemar test was used for comparison the results without and with CAD.

Results

Unblinded colonoscopy detected 50 patients with lesions ≥ 6 mm and 25 patients with lesions ≥ 10 mm. Sensitivity of CTC without CAD for these size categories was 80% (40/50, 95% CI: 69–81%) and 64% (16/25, 95% CI: 45–83%), respectively. CTC with CAD detected one additional patient with a lesion ≥ 6 mm and two with a lesion ≥ 10 mm, resulting in a sensitivity of 82% (41/50, 95% CI: 71–93%) (p=0.50) and 72% (18/25, 95% CI: 54–90%) (p=1.0), respectively. Specificity without CAD for polyps ≥ 6 mm and ≥ 10 mm was 84% (101/120, 95% CI: 78–91%) and 94% (136/145, 95% CI: 90–98%), respectively. With CAD, the specificity remained (nearly) unchanged: 83% (99/120, 95% CI: 76– 89%) and 94% (136/145, 95% CI: 90– 98%), respectively.

Conclusion

Although CTC with CAD detected a few more patients than CTC without CAD, it had no statistically significant positive influence on CTC performance.

Introduction

Computed tomography colonography (CTC) has consistently shown to have a high accuracy for colorectal neoplasia, and has recently been included in the official guidelines for colorectal cancer screening [1].

In the past years, efforts have been made to in order to increase its accuracy, e.g., labeling fecal material with a contrast agent (fecal tagging), automatic insufflation and improvement of workstations. Despite these efforts, visible lesions are still missed, even by well-trained radiologists. Computer-aided detection (CAD) is a promising technique [2–5] that could be helpful in reducing these falsenegative findings [6, 7]. However, even if the CAD performance would be excellent, it does not automatically translate into equivalent reader performance [8, 9], i.e., CAD hits can be disregarded by the observer. This stresses the complex interaction between CAD and the observers.

Recent studies concluded that in a selected population CAD significantly improved per-polyp sensitivity for less experienced observers [10–13]. Though experienced observers benefited proportionally less from CAD [14, 15]. Therefore, the potential increase by CAD in accuracy for experienced observers is still controversial.

The additional value of CAD was tested in a selected and polyp-enriched population only. This may have a positive effect on the observer performance since observers may be more easily triggered to detect polyps. Secondly, the a priori chance that a finding is indeed a polyp has increased. Therefore, the additional value of CAD (that will have a similar detection pattern irrespective of the population) may be larger.

To our knowledge, the effect of CAD on the performance of observers has not been prospectively evaluated in an unselected patient population of increased risk for colorectal cancer. Therefore, the purpose of this study was to determine whether CAD in a second read paradigm could improve the performance characteristics in a practical setting. Based on indirect comparison of two experienced observers and CAD [16], we hypothesized that CAD could still improve experienced observer performance.

Materials and methods

The institutional review board of both hospitals approved the study. All patients gave written informed consent.

Study population

Consecutive patients with a personal or family history of colorectal polyps or cancer were invited to participate from February 2006 until July 2007. All patients were scheduled to undergo a routine colonoscopy at one or other of the two participating hospitals. Exclusion criteria were: age under 18 years, pregnancy, personal history of inflammatory bowel disease, familial adenomatous polyposis, Peutz- Jeghers syndrome, hereditary non-polyposis colorectal cancer, prior allergic reaction to iodine contrast, untreated hyperthyroidism, known colorectal polyps that were not removed at an earlier endoscopy.

Patients ingested 4 l polyethylene glycol electrolyte solution (KleanPrep; Helsinn Birex Pharmaceuticals, Dublin, Ireland) on the day before and the day of the examinations. If contraindicated, other regimes were used. Patients ingested 50 ml oral iodine contrast (ioxithalamate, 300 mg ml–1) (Telebrix, Guerbet, Roissy, France) with each liter of polyethylene glycol electrolyte solution.

All CT-examinations were performed on two different CT systems. The CT parameters for the four-slice CT were 120 kV, 50 mAs (abdominal circumference ≤ 103 cm) or 70 mAs (>103 cm), effective slice thickness 3.2 mm, pitch1.25 and reconstruction interval 1.6 mm. The CTparameters for the 64-slice CT were 120 kV, 58 mAs (abdominal circumference ≤ 103 cm) or 82 mAs (>103 cm), effective slice thickness 0.9 mm, pitch 0.984 and reconstruction interval 0.7 mm. Procedural details and baseline characteristics are listed in Table 6.1.

TABLE 6.1 **Baseline Patient Characteristics and Procedural Details of CTC (n=170)**

Male/Female	**98/72**
Age in years: mean ± SD	57 ± 12
Symptomatic (according to referring physician)	45
Abdominal pain	12
Weight loss	2
Changed bowel habits	25
Rectal bleeding	16
Hospital	
(blinded for review)	137
(blinded for review)	33
History of colonoscopy or sigmoidoscopy 10 years prior to CTC	
Yes	107
No	53
Unknown	10
Patients	
Without polyps ≥ 6mm	120
With polyps ≥ 6mm	50
With a polyp ≥ 10mm	25
Bowel preparation	
PEG	169
Sodium phosphate	1
CT-system used	
4-slice[a]	7
64-slice[b]	163
Spasmolytics	
20 mg Butyl scopolamine bromide[c]	136
1 mg Glucagon[d]	28
No spasmolytics	6
Abdominal circumference	
< 103cm	82
≥ 103cm	88
Mean volume of insufflated CO_2[e]	4.5 l (range: 2.2-7.6)
Mean scanner room examination time	21 minutes (range 12-35)
Number of complications	0

a Mx 8000, Philips Medical Systems, Best, The Netherlands
b Brilliance 64, Philips Medical Systems, Best, The Netherlands
c Boehringer Ingelheim, Ingelheim, Germany
d Glucagon; Novo Nordisk A/S, Bagsvaerd, Denmark
e ProtoCO2l, E-Z-EM, Lake Success, N.Y., USA

Observers

The CTC examinations were evaluated by one observer of a group of five observers; one board certified abdominal radiologist, two radiology residents (2nd and 4th year) and two radiology research fellows. The observers read the CTC examinations in a quiet environment not pressured to provide rapid reports, although they knew that the colonoscopy would be performed within 3 h. They were blinded to clinical data.

Although their experience varied, all had seen at least 100 CTC examinations verified by colonoscopy, often combined with additional examinations without direct feedback (Table 6.2).

Just prior to the study, all had passed a test of 25 selected CTC examinations [17] by scoring above a predefined perpolyp sensitivity threshold of 90%. In 12 of these 25 patients, 19 polyps ≥ 6 mm (one flat lesion) and 10 polyps larger than 10 mm could be detected.

CTC image analysis

The observers were blinded to the CAD results during the initial reading. All patients were evaluated with a primary three-dimensional (3D) method (Endo 3D Unfolded, ViewForum, Philips Medical Systems, Best, The Netherlands). This validated method [18] was used to increase surface visibility and reduce reading time. Additional two-dimensional (2D) displays with instant on-screen correlation were used for problem solving. Stool subtraction software was not used. The observers digitally recorded size (mm), morphology (pedunculated, sessile, flat) and colon segment (cecum, ascending colon, transverse colon, descending colon, sigmoid colon or rectum).

After their unassisted reading they were able to access the CAD results. Readers were permitted to discard unassisted findings after CAD application. The incorporated commercially available CAD algorithm (ColonCAD, Philips Medical Systems, Best, The Netherlands) had a fixed sensitivity threshold that was not changed during the study.

TABLE 6.2 **Number of Read CTC Cases per Observer and Observer Experience with and without Colonoscopic Verification**

	Number of read cases	Nr. of endoscopic examinations *with* colonoscopic verfication	Nr. of endoscopic examinations *without* colonoscopic verfication[b]
Observer 1	31/170	500 CTC	225 CTC
Observer 2	36/170	300 CTC and 300 MRC	225 CTC
Observer 3	29/170	300 CTC	75 CTC
Observer 4	36/170	230 CTC and 30 MRC[a]	100 CTC & 240 MRC
Observer 5	51/170	100 CTC	

a Including: Matching polyps in 200 CTC studies of patients of increased risk.

b Including: 25 test patients

The CAD algorithm was trained on annotated polyp data from 13 patients from a comparative study of 249 patients [19]. These datasets had been verified by colonoscopy and contained a total of 80 polyps ≥ 5 mm. In this study, by mouse-clicking a listed candidate, corresponding 3D, 2D axial and 2D MPR views were shown with a mark on the polyp candidate (Figure 6.1).

If the observers identified CAD lesions that were not detected in the unassisted evaluation, these could be added to the initial list of findings.

Interpretation time and image quality

Interpretation times for the unassisted read and for the evaluation of CAD candidates were recorded with a stopwatch for both the prone and supine positions. When the reading was completed, the observer assessed the degree of colonic distension and quality of the fecal tagging on a four-point Likert-scale (good, sufficient, moderate, poor). The overall quality of the examination was assessed as 'diagnostic' or 'non-diagnostic'. If the quality was assessed as 'non-diagnostic' by the examining physician, the patient was excluded.

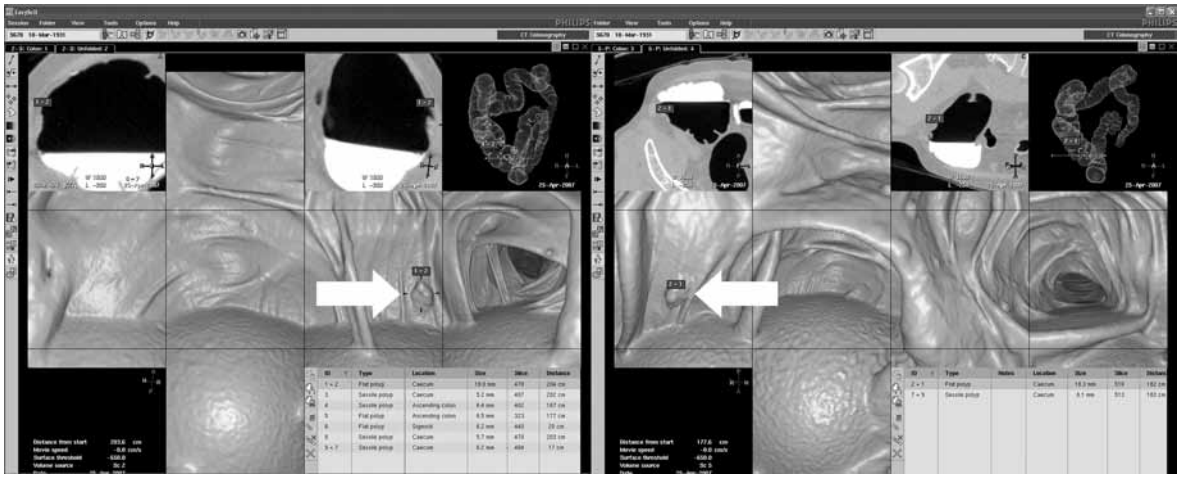


FIGURE 6.1
Screenshot of both monitors of the workstation displaying the patient in the supine (left) and prone positions (right). The white arrows mark a 16 mm sessile polyp that was detected by both CAD and the observer

Colonoscopy

All patients underwent colonoscopy within 3 h after CTC. A gastroenterologist (>200 colonoscopies), or fellow or nurse under direct supervision of a staff member performed the colonoscopy with a standard colonoscope (CF-140L; Olympus, Tokyo, Japan). Chromoendoscopy or narrowband imaging to improve flat polyp detection was not performed. Patients received on request midazolam (Dormicum, Roche, Basel, Switzerland) and fentanyl (Hameln Pharmaceuticals, Hameln, Germany) or propofol (Fresenius Kabi, Uppsala, Sweden) and fentanyl. The examination was digitally recorded.

Polyp characteristics (size, morphology and segmental location) were documented on a case record form by an attending research nurse. Polyp size was measured with open biopsy forceps (8 mm). The determination of the morphology of polyps was done by the gastroenterologist based on the endoscopic classification of superficial neoplastic lesions [20]. In this classification, flat polyps were defined as lesions with a maximum height of 2.5 mm (closed cups of biopsy forceps). Segmental unblinding was performed for CT lesions 6 mm or larger. Histology was obtained at colonoscopy, except in those cases in which polyp removal was technically impossible or when material was lost during the procedure.

Determination of lesion status

Observers were instructed that only hyperplastic, adenomatous (advanced and not-advanced) and potentially malignant lesions were considered true-positive lesions. This qualification was based on the histology report or – if histology was not acquired – based on the endoscopic report.

For CTC, a polyp was considered true-positive, if: (1) its appearance resembled the corresponding polyp at colonoscopy, (2) its segment or adjacent segment corresponded with the reference standard segment and (3) the polyp size as estimated by the endoscopist corresponded with size as measured on CTC, considering a margin of error of 50%. Since the colonoscopy measurement is subject to inaccuracy [21, 22] this criterion could be overruled by the first two criteria.

Polyps ≥ 6 mm at colonoscopy that were not identified by the observer without or with CAD, were re-evaluated with knowledge of the colonoscopic findings by a research fellow with experience of more than 300 CTC examinations verified by colonoscopy. In this re-evaluation the nature of all detection errors ≥ 6 mm (false-negative findings) was assessed and differentiated between perception errors (visible in retrospect) and non-perception errors (lesions not visible in retrospect).

Lesions not confirmed by colonoscopy ≥ 6 mm (false positives) were assessed by consensus by two experienced research fellows (300 colonoscopy verified CTC). The consensus panel determined whether the finding was related to bowel preparation.

Power calculation

Based on a prior feasibility study [16], we expected a 15% increase in sensitivity. In order to determine a statistically significant increase of 15% for polyps ≥ 6 mm, at least 39 lesions were required. For this approach, a McNemar test with continuity correction and a p value of 0.05 to indicate statistical significance was used. Based on prior studies in this patient population, we assumed that the prevalence of patients with polyps ≥ 6 mm would be 25% [23]. We therefore required a minimal number of 39/0.25= 156 patients. The total number of patients determined was 170.

Outcome parameters per patient

Sensitivity, specificity, positive- and negative-predictive values of CTC without and with CAD were calculated. Sensitivity and number of false-positive findings were calculated for CAD without interaction of the observers (stand alone). Furthermore, sensitivity was calculated for unblinded colonoscopy. The outcome parameters were determined for polyps ≥ 6 mm and ≥ 10 mm.

A patient was considered true-positive if CTC detected

at least one polyp seen at colonoscopy, based on the matching criteria described previously. A patient was categorized as false negative if CTC detected no polyps (although present at the reference standard) or only those of a lower size category in comparison to the reference standard.

We used the McNemar test to compare per-patient sensitivity and specificity values between CTC without and with CAD.

Outcome parameters per polyp

We calculated the per-polyp sensitivity for CTC without and with CAD, CAD (stand alone) and blinded colonoscopy for lesions 6–9 mm and ≥ 10 mm. In this study, more than one polyp was detected in some patients. Therefore, generalized estimating equations (GEE) (SPSS, 15.0, Statistics, Chicago, USA) was used to revise the data clustering and dependency. In the GEE, the adjusted confidence intervals with regard to per-polyp sensitivity for CTC and blinded colonoscopy (i.e., before unblinding of CTC results) were assessed for CTC without and with CAD. In this same GEE method, regression analyses were done to compare the sensitivity values.

Outcome time parameters

The median interpretation time of the CTC reading without CAD and the median time to evaluate all CAD results were calculated.

Prevalence of flat polyps stratified for endoscopic colon examination

Because a relatively high number of flat polyps were detected in this population we retrospectively determined whether a colon examination 10 years or less prior to the CTC in the patient's history could effect the prevalence of these polyps. The rationale for this retrospective study was an article published by MacCarty et al. [24] that suggested a higher prevalence of polyps in patients who had undergone a previous endoscopic colon examination.

We did not specify the type of colon examination in colonoscopy, sigmoidoscopy or proctoscopy because it was not always clear which part of the colon was examined. The arbitrary period of 10 years was chosen since we assumed this would be the period from a polyp to grow into a tumor and a colon examination executed earlier may have effect the prevalence of flat lesions at CTC.

The prevalence of polyps in the group that had undergone a colon examination and the group that had not were compared with the McNemar test and stratified for size.

Results

Of 448 eligible patients that were scheduled to undergo optical colonoscopy during the inclusion period, 170 'diagnostic' examinations were included in this study (Figure 6.2). The baseline characteristics and procedural details are listed in Table 6.1.

The degree of bowel distention was assessed as 'good' or 'average' in 161 patients (95%), 'moderate' in eight patients (5%) and 'poor' in one patient (1%).

Fecal tagging was assessed as 'good' or 'average' in 144 patients (85%), 'moderate' in 24 (14%) and 'poor' in two patients (1%).

Reference standard

Unblinded colonoscopy revealed that 50 out of 170 patients (29%) harbored one or more polyps ≥ 6 mm and 25 of 170 patients (15%) one or more polyps ≥ 10 mm. Table 6.3 displays the histological and morphological characteristics. One colorectal carcinoma (50 mm) was found.

Per-patient analysis

The per-patient sensitivity and specificity is displayed in Table 6.4. CAD did not significantly alter per-patient sensitivity and specificity for lesions ≥ 6 mm and ≥ 10 mm. Assisted by CAD, the observers detected one additional patient with a lesion ≥ 6 mm and two additional patients with a lesions ≥ 10 mm, resulting in a sensitivity of 82%

FIGURE 6.2
The flowchart of this study

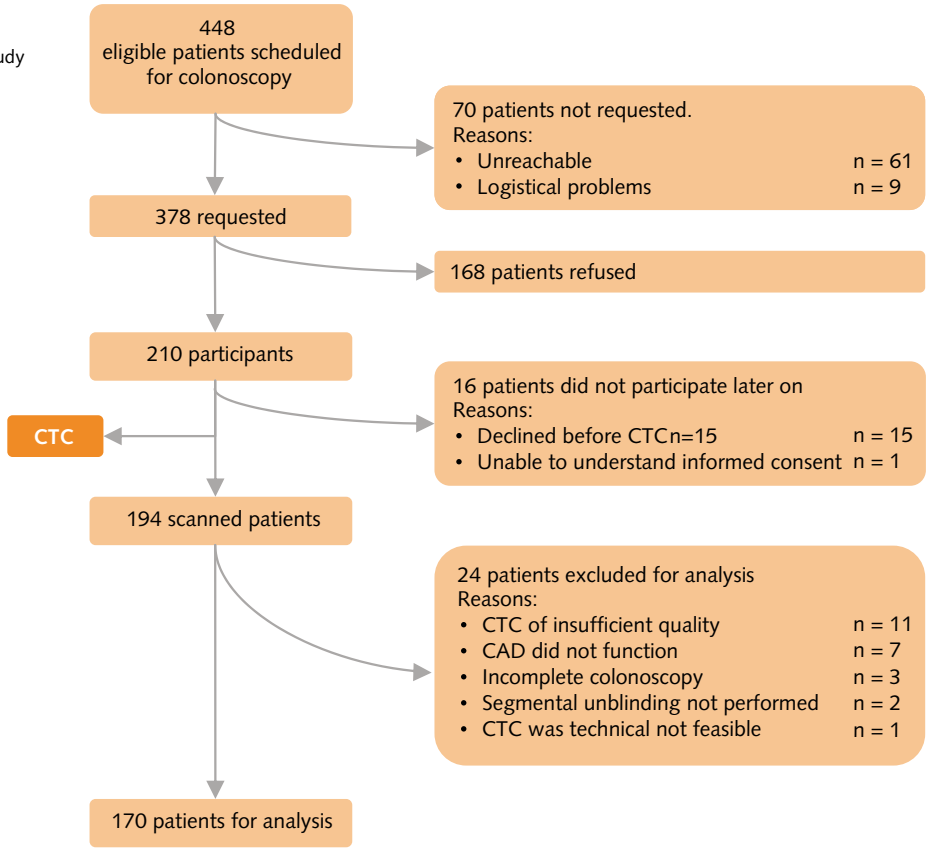


TABLE 6.3 **The Histology and Morphology of Polyps Seen and Removed During Colonoscopy.**

	6-9mm (n=58)	≥ 10mm (n=30)
Non-adenomatous lesions	24	4
Adenomatous lesions classified as 'advanced'	4	20
Adenomatous lesions classified as 'not advanced'	23	0
Unknown histology	7	4
CRC	0	1
Sessile	40	7
Pedunculated	5	9
Flat	13	13
Tumor	0	1

(p=1.0) and 72% (p=0.5), respectively.

Two patients were erroneously classified as having a lesion ≥ 6 mm after accepting a CAD hit and no patients without a lesion ≥ 10 mm were wrongly added to the list.

CAD on a stand-alone basis detected 74% (37/50) and 64% (16/25) of the patients with lesions ≥ 6 mm and ≥ 10 mm. There was no statistically significant difference between the observers and CAD in the respective size categories (p=0.375 and p=1.0).

CAD had a median number of nine hits per-patient (25–75% quartiles: 5–15). Blinded colonoscopy detected 96% (48/50) and 100% (25/25) of the patients in the respective size categories. As displayed in Table 6.4, both the positive- and negativepredictive values of the observers with and without CAD were nearly unchanged.

Per-polyp analysis

Per-polyp sensitivity for polyps of 6–9 mm and ≥ 10 mm are displayed in Table 6.4. CAD detected one lesion 6–9 mm and three polyps ≥ 10 mm initially missed by the observer, but it did not significantly increase sensitivity of the observer for the respective size categories (p=0.31 and p=0.08). No true-positive CAD hits were erroneously dismissed by observers. Per-polyp sensitivity was better for polyps 6–9 mm than for polyps ≥ 10 mm, without CAD as well as with CAD. To a large extent this can be explained by the difficulty of the observers and CAD in detecting the relative prevalent number of undetected flat lesions ≥ 10 mm; 23% (3/13) of the flat lesions ≥ 10 mm were detected without CAD and 31% (4/13) with CAD (Figure 6.3). Figure 6.4 shows that the largest part (6/10) of the missed flat large lesions (either without or with CAD) were not visible in retrospect (non-perception errors). Since these polyps cannot be detected, it is difficult to assess exactly why these non-perception errors were missed.

The three polypoid perception errors 6–9 mm (Figure 6.4) were missed because they were situated on a fold (n=1), were clearly smaller than 6 mm when measured on CT (n= 1) or could be defined as flat on CT (n=1). The two polypoid perception errors ≥ 10 mm were missed because they were situated on a fold (n=1) or because of unclear reasons (n=1).

Of the 53 false positive lesions detected by the reader without CAD, 23 (43%) findings were according to consensus related to bowel preparation. None of the four false-positive findings suggested by CAD and incorporated in the final list by the observer were related to bowel preparation.

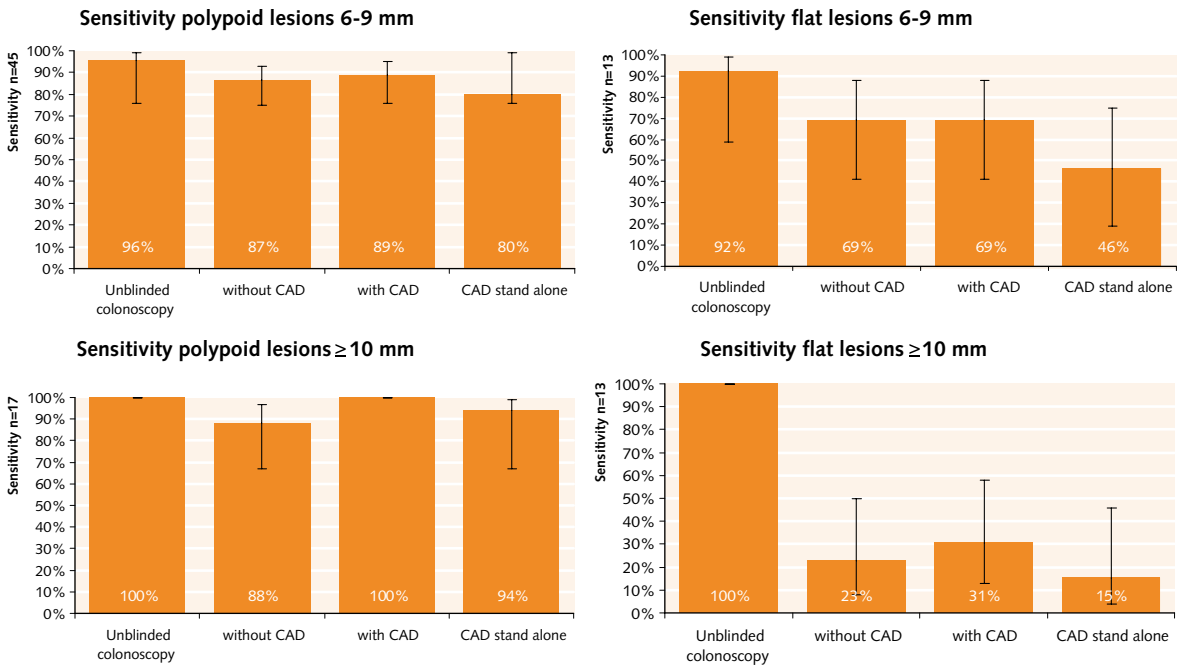


FIGURE 6.3

The left histograms show the sensitivity of polypoid lesions of 6–9 mm (n=45) and ≥ 10 mm (n=17), the right histograms show the sensitivity of lesions lesions of 6–9 mm (n=13) and ≥ 10 mm (n=13)with a flat morphology. The difference in sensitivity between polypoid and flat lesions is more striking at lesions ≥ 10 mm

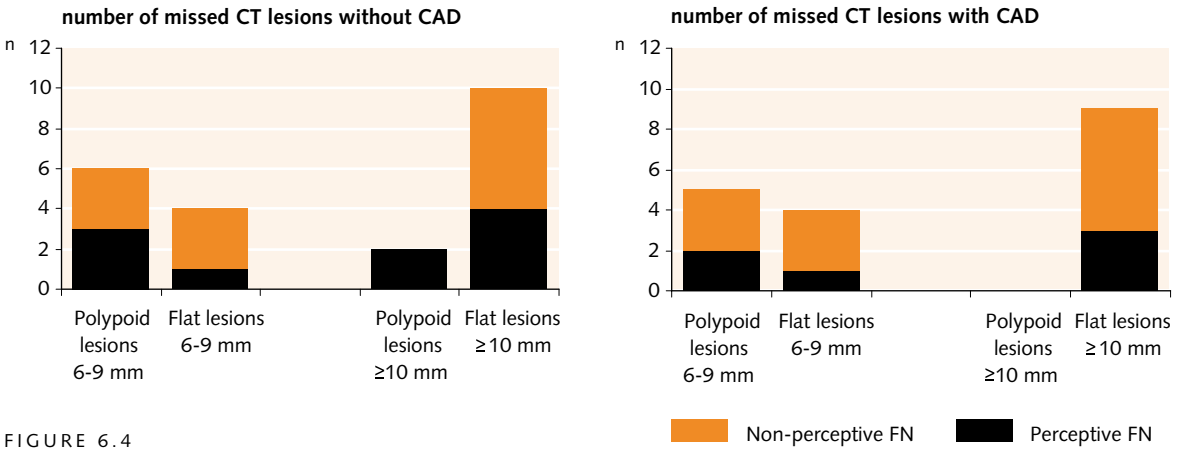


FIGURE 6.4
The number of false-negative findings and distribution of perceptive and non-perceptive errors among flat and non-flat lesions

TABLE 6.4 **Results of CAD Without and With CAD on a Per-patient and Per-poly Basis (%) (95% CI in parentheses)**

			CTC without CAD	CTC with CAD
PER PATIENT	**Sensitivity**	≥ 6mm	80% (69-91) 40/50	82% (71-93) 41/50
			64% (45-83) 16/25	72% (54-90) 18/25
	Specificity	≥ 6mm	84% (78-91) 101/120	83% (76-89) 99/120
		≥ 10mm	94% (90-98) 136/145	94% (90-98) 136/145
	Positive predictive value	≥ 6mm	68% (56-80) 40/59	66% (54-78) 41/62
		≥ 10mm	64% (45-83) 16/25	67% (48-84) 18/27
	Negative predictive value	≥ 6mm	91% (86-96) 101/111	92% (86-97) 99/108
		≥ 10mm	94% (90-98) 136/145	95% (92-98) 136/143
PER POLYP	**Sensitivity**	6-9mm	83% (72-90) 48/58	84% (73-92) 49/58
		≥ 10mm	60% (43-75) 18/30	70% (51-84) 21/30
	Total number of false positive lesions	6-9mm	42	45
		≥ 10mm	11	12

Although CAD did not significantly increase sensitivity, it did not significantly alter specificity either: three extra false-positive lesions 6–9 mm and one extra lesion ≥ 10 mm in 170 patients were added to the list of the observer (Table 6.4).

CAD on a stand-alone basis detected 72% (42/58) of the polyps 6–9 mm and 60% (18/30) of the polyps ≥ 10 mm. Blinded colonoscopy detected 95% (55/58) of the polyps 6–9 mm and 100% (30/30) of the polyps ≥ 10 mm.

Interpretation time

The observers had an a median interpretation time of 16 min 00 s (25–75% quartiles: 11 min 35 s-23 min 6 s) to complete the examination and a median time of 1 min 26 s (25–75% quartiles: 28 s-2 min 46 s) to evaluate all CAD results after the initial reading.

Prevalence of flat polyps stratified for endoscopic colon examination

Sixty-three percent (107/170) of the patients had undergone an endoscopic colon examination prior to CTC, 31% (53/170) had not. For ten patients, the history could not be retrieved. Of the polyps 6–9 mm, in patients with a history of endoscopy 27% (12/44) were flat, in contrast to 8% (1/12) of the polyps in patients without a history of endoscopy (p=0.259). Of the polyps ≥ 10 mm, in patients with a history of endoscopy 60% (9/15) were flat, in contrast to 33% (4/12) of polyps in patients without a history of endoscopy (p=0.168). Thus, the prevalence of flat polyps in both size categories was higher in the group that had undergone colon examination, though statistical significance was not reached.

Discussion

Although CAD in a second-read paradigm detected one additional patient with a lesion ≥ 6 mm and two patients with a lesion ≥ 10 mm, it did not significantly improve per-patient sensitivity in this increased risk patient population. Several CTC studies in which the additional value of CAD was evaluated (after the interaction with the observer) have reported good results in terms of polyp detection [12, 25–27]. All concluded that the observers detected statistically more polyps with CAD.

In contrast to these studies, we did not find a significant additional value for CAD. The study design of these studies differs from our study in a number of aspects: patient selection, inclusion and exclusion criteria and reference standard. However, we think that the most important difference between the aforementioned studies and our study is the fact that the observers in our study were more experienced, i.e., more than 100 CTC cases verified by colonoscopy. Since there is good evidence that experience in CTC results in less false-negative findings [28, 29], it is logical that it is more difficult to substantially increase the sensitivity of the observer with CAD. In studies that report data about the additional value of CAD on experienced observers [30, 31], experienced observers benefited proportionately less from CAD when compared with the inexperienced readers. This finding is supported by the results of this study.

Although a CAD algorithm has the potential to decrease the number of perceptual errors by exposing the observer to candidate lesions, it cannot account for interpretative errors. In the above-mentioned papers, a significant increase of false positives have been reported. Though the specificity in this study was not significantly increased, there were only two patients erroneously classified as having a lesion. Both false-positive lesions measured 6– 9 mm, none was larger than 10 mm.

Even though the sensitivity of the observers was low (i.e., 72% for lesions ≥ 10 mm), CAD was not able to increase their performance. In our opinion, the reported sensitivity requires looking for causes in the population itself. All nine polyps ≥ 10 mm missed by the observer with CAD had a flat morphology (Figure 6.4). These flat polyps are an important cause of false-negative findings [32, 33]. In this population, 13 of the 30 lesions ≥ 10 mm were flat and therefore an important explanation of the moderate

sensitivity, not only for the observer but for CAD as well. The unexpectedly high number of flat lesions may be related to the history of patients; MacCarty et al. [34] reported in a prospective study of 75 consecutive patients that more than 50% of the false negatives missed by experienced readers were not even visible in retrospect in a population that had been screened by colonoscopy 5 years prior to CTC. Nearly all these polyps were flat. In this study population, the prevalence of flat lesions was higher (although not statistically significant) in the group of patients that had undergone a colonoscopy or sigmoidoscopy prior to CTC as well (Table 6.1).

We concur with MacCarty and coworkers that previous screening could aversively affect CTC sensitivity in two ways: first, it is likely that many easy-to-see polyps would be detected and removed at the initial screening, and fewer hard-to-see polyps would be detected and removed; second, endoscopic polypectomy may have been incomplete. Remnants of polyps are flatter than the original intact lesions, and would, therefore, be more difficult to detect. So we think that the selection of patients has an important influence on the test characteristics.

The type of bowel preparation, i.e., extensive or reduced, with or without oral contrast (iodine and/or barium) may influence the performance in terms of polyp detection and number of false-positive findings of CAD; polyps can be covered by fecal material or fecal remains may simulate polyps. In this study, we used PEG as an extensive bowel preparation for colonoscopy, combined with Telebrix that has a laxative effect as well. Although we have not evaluated the nature of all CAD candidates, the additional value of CAD did not seem to be impaired by this type of bowel preparation used in this study, since none of the false-positive lesions incorporated in the final list of the observers were prep-related and only two of 22 falsenegative findings were covered by fecal material (although we think this is not the reason why they were missed).

This study has limitations; due to the small time-frame between the CTC and colonoscopy, the patients could be read by only one observer out of a group of five different observers. Each observer had a different level of experience. Therefore, although no statistically significant difference in sensitivity was measured between the observers, and none of the observers had a significant improvement in performance after CAD (data not shown), the bestperforming observers could have leveled out the sensitivity of the least performing observers. Still, the situation as described in this paper is similar to the practical setting of many hospitals; each examination will not be read by five different radiologists but by only one of a pool of experts.

Secondly, the level of experience of the group of five readers was relatively high. It is likely that the additional value of CAD would be larger if the data were read by a relatively inexperienced reader group [35, 36]. Therefore, our conclusionmay not apply to relatively untrained readers.

Thirdly, we evaluated CAD using a primary 3D reading paradigm. Although the discussion as to whether to read the data in 2D or 3D is still not settled, a (slight) superiority in terms of polyp detection with 3D is reported by some studies [37, 38]. Since the reported sensitivities of a primary 2D paradigm tend to be lower, CAD may have a larger additional value when used in a 2D reading protocol.

Fourthly, the relatively large number of flat polyps may limit the generalization of the results of this study. However, if we leave out all 26 flat polyps, and we only consider the remaining 62 polypoid lesions, we still cannot demonstrate a significant contribution of CAD to the sensitivity of the observer. This number was still more than the 39 polyps that were needed to demonstrate a 15% sensitivity difference according to our power analysis.

In conclusion, although CTC with CAD in a second read paradigm detected a few more lesions than CTC without CAD, CAD has no statistically significant positive influence on CTC performance in an increased-risk population when used by a relative experienced group of observers.

ACKNOWLEDGEMENTS – We thank Prof. M.J. van de Vijver, MD, PhD for assessing the histology of the polyps. We thank Mrs. A.C. Dobben, MSc, PhD for her statistical support.

References

1. Levin B, Lieberman DA, McFarland B et al (2008) Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. Gastroenterology 134:1570–1595
2. Summers RM, Yao J, Pickhardt PJ et al (2005) Computed tomographic virtual colonoscopy computer-aided polyp detection in a screening population. Gastroenterology 129:1832–1844
3. Yoshida H, Nappi J, MacEneaney P, Rubin DT, Dachman AH (2002) Computer-aided diagnosis scheme for detection of polyps at CT colonography. Radiographics 22:963–979
4. Jerebko AK, Summers RM, Malley JD, Franaszek M, Johnson CD (2003) Computer-assisted detection of colonic polyps with CT colonography using neural networks and binary classification trees. Med Phys 30:52–60
5. Mani A, Napel S, Paik DS et al (2004) Computed tomography colonography: feasibility of computer-aided polyp detection in a 'first reader' paradigm. J Comput Assist Tomogr 28:318–326
6. Summers RM, Jerebko AK, Franaszek M, Malley JD, Johnson CD (2001) Complementary role of computer-aided diagnosis for detecting colonic polyps with CT colonography. Radiology 120:391–399
7. Taylor SA, Halligan S, Burling D et al (2006) Computer-assisted reader software versus expert reviewers for polyp detection on CT colonography. AJR Am J Roentgenol 186:696–702
8. Halligan S, Altman DG, Mallett S et al (2006) Computed tomographic colonography: assessment of radiologist performance with and without computer-aided detection. Gastroenterology 131:1690–1699
9. Taylor SA, Charman SC, Lefere P et al (2008) CT colonography: investigation of the optimum reader paradigm by using computer-aided detection software. Radiology 246:463–471
10. Halligan S, Altman DG, Mallett S et al (2006) Computed tomographic colonography: assessment of radiologist performance with and without computer- aided detection. Gastroenterology 131:1690–1699 949
11. Baker ME, Bogoni L, Obuchowski NA et al (2007) Computer-aided detection of colorectal polyps: can it improve sensitivity of less-experienced readers? Preliminary findings. Radiology 245:140–149
12. Nicholas Petrick P, Maruf Haider M, Ronald M, Summers MP et al (2008) CT colonography with computer-aided detection as a second reader: observer performance study. Radiology 246:148–156
13. Mang T, Peloschek P, Plank C et al (2007) Effect of computer-aided detection as a second reader in multidetectorrow CT colonography. Eur Radiol 17:2598–2607
14. Mang T, Peloschek P, Plank C et al (2007) Effect of computer-aided detection as a second reader in multidetectorrow CT colonography. Eur Radiol 17:2598–2607
15. Taylor SA, Charman SC, Lefere P et al (2008) CT colonography: investigation of the optimum reader paradigm by using computer-aided detection software. Radiology 246:463–471
16. de Vries AH, Peters J, Grigorescu S, Truyen R, Gerritsen FA, Stoker J (2005) The performance of CAD versus the performance of an expert in CT colonography. Proc Sixth Int Symp Virtual Colonoscopy, Boston, p 189
17. Pickhardt PJ, Choi JR, Hwang I et al (2003) Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. N Engl J Med 349:2191–2200
18. Vos FM, van Gelder RE, Serlie IW et al (2003) Three-dimensional display modes for CT colonography: conventional 3D virtual colonoscopy versus unfolded cube projection. Radiology 228:878–885
19. van Gelder RE, Nio CY, Florie J et al (2004) Computed tomographic colonography compared with colonoscopy in patients at increased risk for colorectal cancer. Gastroenterology 127:41–48 20. The Paris endoscopic classification of superficial neoplastic lesions: esophagus, stomach, and colon: November 30 to December 1, 2002 (2003). Gastrointest Endosc 58:S3–43
21. Fennerty MB, Davidson J, Emerson SS, Sampliner RE, Hixson LJ, Garewal HS (1993) Are endoscopic measurements of colonic polyps reliable? Am J Gastroenterol 88:496–500
22. Gopalswamy N, Shenoy VN, Choudhry U et al (1997) Is in vivo measurement of size of polyps during colonoscopy accurate? Gastrointest Endosc 46:497–502

23. Jensch S, de Vries AH, Peringa J et al (2008) CT colonography with limited bowel preparation: performance characteristics in an increased-risk population. Radiology 247:122–132
24. MacCarty RL, Johnson CD, Fletcher JG, Wilson LA (2006) Occult colorectal polyps on CT colonography: implications for surveillance. AJR Am J Roentgenol 186:1380–1383
25. Baker ME, Bogoni L, Obuchowski NA et al (2007) Computer-aided detection of colorectal polyps: can it improve sensitivity of less-experienced readers? Preliminary findings. Radiology 245:140–149
26. Halligan S, Altman DG, Mallett S et al (2006) Computed tomographic colonography: assessment of radiologist performance with and without computer- aided detection. Gastroenterology 131:1690–1699
27. Mang T, Peloschek P, Plank C et al (2007) Effect of computer-aided detection as a second reader in multidetectorrow CT colonography. Eur Radiol 17:2598–2607
28. Burling D, Halligan S, Altman DG et al (2006) CT colonography interpretation times: effect of reader experience, fatigue, and scan findings in a multicentre setting. Eur Radiol 16:1745– 1749
29. Johnson CD, Toledano AY, Herman BA et al (2003) Computerized tomographic colonography: Performance evaluation in a retrospective multicenter setting. Gastroenterology 125:688–695
30. Mang T, Peloschek P, Plank C et al (2007) Effect of computer-aided detection as a second reader in multidetectorrow CT colonography. Eur Radiol 17:2598–2607
31. Taylor SA, Charman SC, Lefere P et al (2008) CT colonography: investigation of the optimum reader paradigm by using computer-aided detection software. Radiology 246:463–471
32. Fidler JL, Fletcher JG, Johnson CD et al (2004) Understanding interpretive errors in radiologists learning computed tomography colonography. Acad Radiol 11:750–756
33. Gluecker TM, Fletcher JG, Welch TJ et al (2004) Characterization of lesions missed on interpretation of CT colonography using a 2D search method. AJR Am J Roentgenol 182:881–889
34. MacCarty RL, Johnson CD, Fletcher JG, Wilson LA (2006) Occult colorectal polyps on CT colonography: implications for surveillance. AJR Am J Roentgenol 186:1380–1383
35. Mang T, Peloschek P, Plank C et al (2007) Effect of computer-aided detection as a second reader in multidetectorrow CT colonography. Eur Radiol 17:2598–2607
36. Taylor SA, Charman SC, Lefere P et al (2008) CT colonography: investigation of the optimum reader paradigm by using computer-aided detection software. Radiology 246:463–471
37. van Gelder RE, Florie J, Nio CY et al (2007) A comparison of primary twoand three-dimensional methods to review CT colonography. Eur Radiol 17:1181–1192
38. Pickhardt PJ, Lee AD, Taylor AJ et al (2007) Primary 2D versus primary 3D polyp detection at screening CT colonography. AJR Am J Roentgenol 189:1451–1456 950

Chapter 7

Feasibility of Automated Matching of Supine and Prone CT Colonography Examinations

Ayso H. de Vries
Roel Truyen
Jorien van der Peijl
Jasper Florie
Rogier E. van Gelder
Frans Gerritsen
Jaap Stoker

The British Journal of Radiology, 2006; 79: 740–744

Abstract.

Objective

Matching of prone and supine positions in CT colonography may improve accuracy of polyp detection. The purpose of this study was to investigate the feasibility of automatic prone-supine matching in CT colonography using proven polyps as fixed points of reference.

Material and Methods

The method is based on similarities in the direction of centre-lines and allows for compression and extraction of the centre-lines in both positions. To illustrate the impact of the match error of the new method in practice, the visibility of the matched polyps in a primary three-dimensional unfolded cube setting was determined as well. The method was compared with a method that relies on the normalized distance along the centre-line (NDAC method).

Results

The median absolute match error was 14 mm (range 0–59 mm, average 20 mm) either proximal or distal from the actual polyp in prone position. In the observer study, 70% (26/37) of the polyps were directly visible in prone view. The overall difference in median absolute match error between both methods was small (2 mm), although half way along the centre-line there were polyps with substantial differences in match error (larger with NDAC).

Conclusion

We concluded that automated prone-supine matching of CT colonography studies is feasible and has a low match error. The difference with the NDAC method was small and not significant, although half way along the centre-line some differences were seen.

Introduction

Colorectal cancer is the second leading cause of cancer-related mortality in the western world. CT colonography is considered as a potential screening tool for colorectal cancer. To improve the accuracy of polyp detection, patients are both scanned and examined in the prone and supine position. Combining information from these scans will assist the reviewer in evaluating colon segments and differentiating polyps from faeces or folds. As most faecal material is subject to gravity, combining both scans may be able to increase specificity of CT colonography.

Since reference-points (e.g. hepatic flexure) are often not fixed, manual verification of findings on supine and prone positions may be a time-consuming activity. Using an automated supine-prone matching algorithm may facilitate this process, and may lead to a more efficient interpretation of CT colonography.

The method evaluated in this article is based on similarities in direction of the centre-line and allows for compression and extraction of the centre-lines in both positions.

The first aim of the study was to assess the feasibility of automated matching of supine and prone CT colonography examinations with colonoscopically proven polyps as fixed points of reference. Therefore the match error was calculated. The visibility of matched polyps was determined as well. This was carried out in a three-dimensional setting after a match with this new method was calculated. Third, the method was compared with a method that relies on the relative location between start and end of both paths.

Materials and methods

To evaluate the algorithm, CT colonography examinations with colonoscopically proven polyps (>5 mm) were used. These examinations formed part of a comparative study of CT colonography and colonoscopy in a surveillance population [1]. All polyps could be identified three dimensionally in both prone and supine positions. The polyps in both positions were linked, based on the three criteria of size, location and morphology. Polyps were excluded in cases where there could be doubt about the accuracy of these links.

After manual insufflation of colon and rectum and intravenous administration of either 20 mg butylscopolamine bromide (Buscopan; Boehringer-Ingelheim, Ingelheim, Germany) or 1 mg glucagon hydrochloride (Glucagen; Novo Nordisk A/S, Bagsvaerd, Denmark), patients were scanned with a four-slice CT scanner (Mx8000; Philips Medical Systems, Best, The Netherlands) in supine and prone positions. Scanparameters were as follows: 120 kV, collimation 4 mm 62.5 mm, rotation-time 0.75 s, pitch 1.25, slice-thickness 3.2 mm, reconstruction interval 1.6 mm and reconstruction filter C. The tube-current varied between 25 mAs and 70 mAs depending on the waist circumference of the patient.

Following scanning, centre-lines were calculated semiautomatically; a "seed" was placed in the proximal colon and discontinuities of the colon were bridged manually.

For calculation of the centre-line, software was used to reconstruct the colon three-dimensionally (Easy Vision; Philips Medical Systems, Best, The Netherlands). The matching of prone and supine was performed based on two principles:

1 maximal alignment of centre-line-directions of both prone and supine positions.
2 limited compression or expansion of the centrelines to make the maximal alignment of corresponding segments possible.

In order to reach maximal alignment, the outcome of the following function needs to be minimum:

$$f_{i,j} = \sqrt{(xdir_i - xdir_j)^2 + (ydir_i - ydir_j)^2 + (zdir_i - zdir_j)^2}$$

Low values of fi,j imply a small difference in centre-line direction (dir) between sample points i (prone) and j (supine) in the three dimensions (x, y and z), indicating a good match. In this step, the points with the same direction in both positions are therefore matched as well as possible. Second, since corresponding directions of both centrelines

are often not situated in exactly the same part of the colon, expansion and compression of the paths was applied. To avoid matching of the hepatic flexure in supine position to a curvature in the sigmoid colon in prone position, a penalty-value for expansion and compression of the paths was applied. This penalty value was proportional to the expansion or compression used. Therefore, the more the centre-line was manipulated in order to match parts with a similar direction, the higher the penalty value was.

The sum of the outcome of the function of maximal alignment and the penalty value of the expansion/ compression formed the match cost.

With the use of dynamic programming [2], the match cost was computed for each possible combination of points on the centre line in both positions. The combination of points was made such that the total cost (i.e. the sum of all individual costs) was minimal.

The performance of the algorithm was measured using the absolute match error (in millimetres). This match error was calculated by measuring the distance from the actual polyp location to the calculated (expected) polyp position, [A] (Figure 7.1). This was done along the central axis in prone position after the match was made.

To illustrate the impact of the match error of the new method in practice, the visibility of the matched polyps in a primary three-dimensional unfolded cube setting was determined as well. This was done by a research fellow (JF, medical doctor), who had evaluated over 100 primary three-dimensional

CT-colonographies (all colonoscopically verified). The reviewer was presented a polyp in supine position that was to be indicated in the adjacent prone position, after the match was calculated (Figure 7.2).

The visibility was measured on a four point scale:

I. The corresponding polyp was clearly visible without scrolling along the central axis.
II. As I, but to be certain the reviewer scrolled along the central axis.
III. The corresponding polyp could not be seen instantly, but was traceable after scrolling along the central axis within a margin of 150 mm from the calculated location.
IV. The corresponding polyp could not be found by the reviewer within a margin of 150 mm from the initial spot of the virtual camera.

The match was considered successful if the matched polyp could be classified under category 1 or 2. This method was compared with a method based on the relative location between start and endpoint of this centre line (Normalized Distance Along the Centre-line, NDAC) [3]. Here, the beginning of the centre-line, the caecum, is located at index 0 on the path and the end, the rectum, is defined as index 1. All other points will have indices between 0 and 1. The NDAC absolute match error is then defined as:

Absolute match error=
[NDACpolypsupine-NDACpolypprone] × Length Central Axis prone

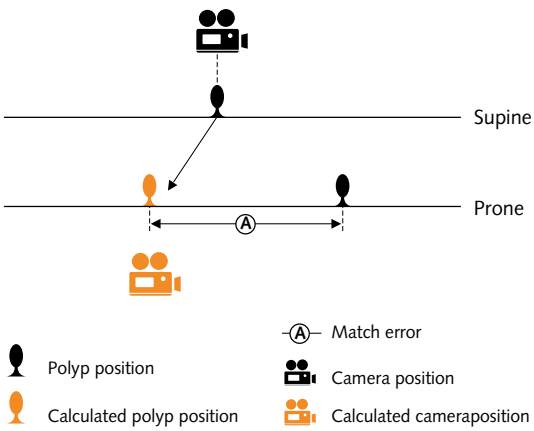


FIGURE 7.1
The performance of the algorithm was calculated by measuring the distance from the actual polyp location in supine position to the calculated (expected) polyp position in prone position after the match was made. (A, match error).

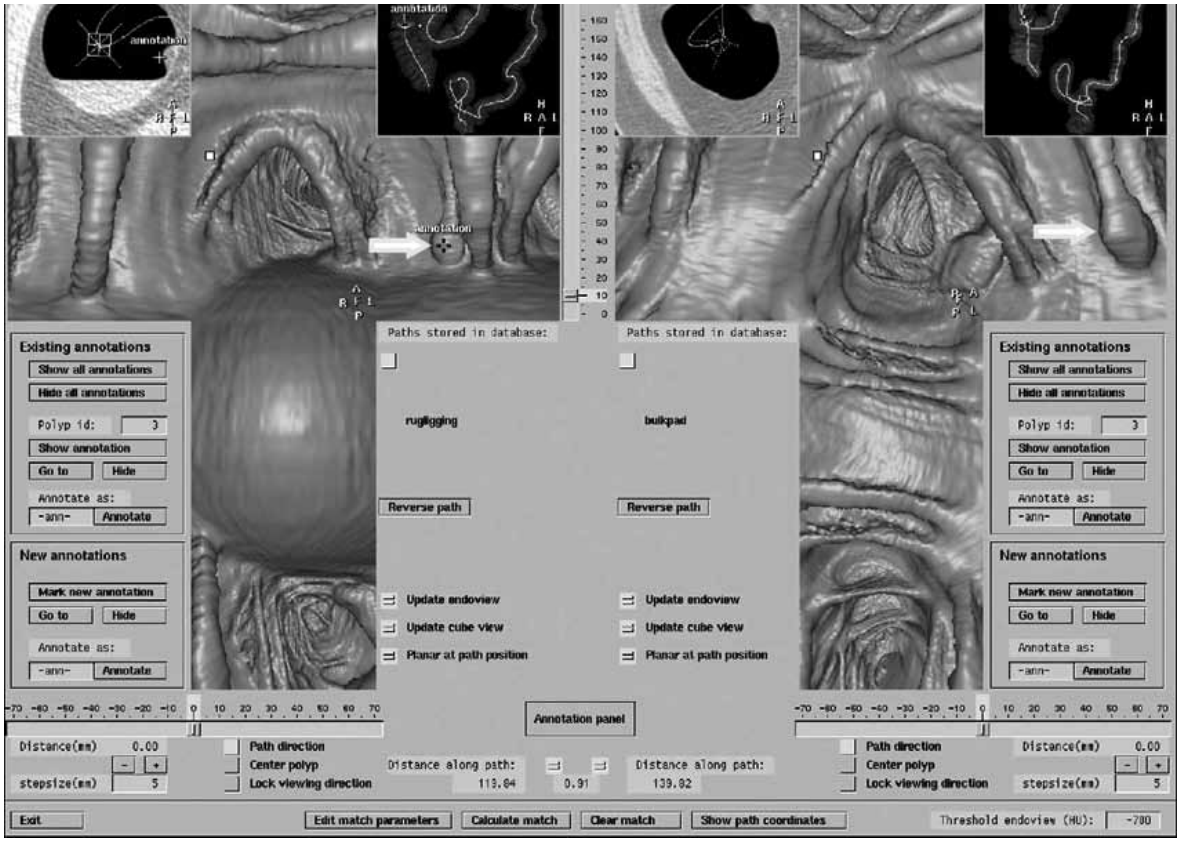


FIGURE 7.2
User interface panel shows both the supine and prone dataset in an unfolded three-dimensional rendering, in a twodimensional axial rendering and in a volume rendered overview of the colon. On the left the supine position was displayed, on the right the prone position. Interactive navigation along the pre-calculated centre-line was possible. This panel shows a calculated match. In this example a match error of 27 mm was measured and the corresponding polyp was immediately visible (category I). This match error must not be confused with the difference in "distance along path". The matching was evaluated with the use of a colonoscopically proven polyp (arrow) (Prototype based on EasyVision; Philips Medical Systems, Best, The Netherlands).

The Wilcoxon-test was performed to test for differences in absolute match error in both methods. The null hypothesis was that there was no difference between both methods. With a p-value of less than 0.05, this hypothesis could be rejected.

Results

32 of the 249 CT colonography studies from the former study included one or more polyps >5 mm that were visible in both supine and prone examinations. These 32 examinations included 53 polyps (5 mm or larger) visible in both positions. One polyp in two patients was excluded because of invisibility on the three dimensional display. These polyps in the rectum were both hidden behind the balloon of the inserted catheter and could only be seen in a two-dimensional read.

One patient was excluded because 12 of the 14 polyps were situated in the rectum and sigmoid. Here there could be doubt about the correctness of the exact linking of the polyps seen in colonoscopy and colonography since many morphologically less specific polyps were seen.

In total, 16 polyps in three patients were excluded. Of the remaining 37 polyps, 26 (70%) were sessile, 8 (22%) were pedunculated and 3 (8%) were flat. 24 polyps were 5–9 mm, 11 polyps were 10–14 mm and 2 polyps were larger than 15 mm. These were both carcinomas (5%). Of the remaining polyps 14 (38%) were adenomas, 10 (27%) were non-adenomas and of the remaining polyps histology was not obtained.

Two polyps were situated in the descending colon and four polyps in the rectum. In the remaining four segments, 7 to 9 polyps were situated. Four of the 29 patients had undergone a hemicolectomy.

The median absolute match error was 14 mm (range 0– 59 mm, average 20 mm) either proximal or distal from the actual polyp in prone position. In the observer study, 70% (26/37) of the polyps was directly visible in prone view. Of these directly visible polyps, 20 polyps (54%) were seen instantly without flying though the colon (category I), the remaining 6 polyps (16%) were also instantly visible, but the observer chose to move the virtual camera over a small distance to verify the polyp (category II). The remaining 11 polyps (30%) could not be seen instantly but all were found after a flight through the colon within 150 mm of the matching position (category III).

When the NDAC method was applied to our CT colonography examinations, this resulted in a median absolute match error of 16 mm (range 0.5–105 mm, average 26 mm), slightly larger than the other method. There was no significant difference (p50.502) in match error between both methods.

Half way along the centre-line three polyps were visible, with a substantial difference in match error in favour of the method of maximal alignment of centrelines (Figure 7.3). These polyps were all situated in the transverse colon. None of these three patients had undergone a hemicolectomy.

Discussion

This study shows that the matching of prone and supine CT colonography examinations is feasible; the median absolute match error was 14 mm and 70% of all polyps were visible after the match was made.

The overall difference in median absolute match error between both methods was small (2 mm), although half way along the centre-line there were polyps with substantial differences in match error (larger with NDAC). All these polyps were situated in the transverse colon. The transverse colon is the largest intraperitoneally situated colonic segment [4] that can move relatively freely in the

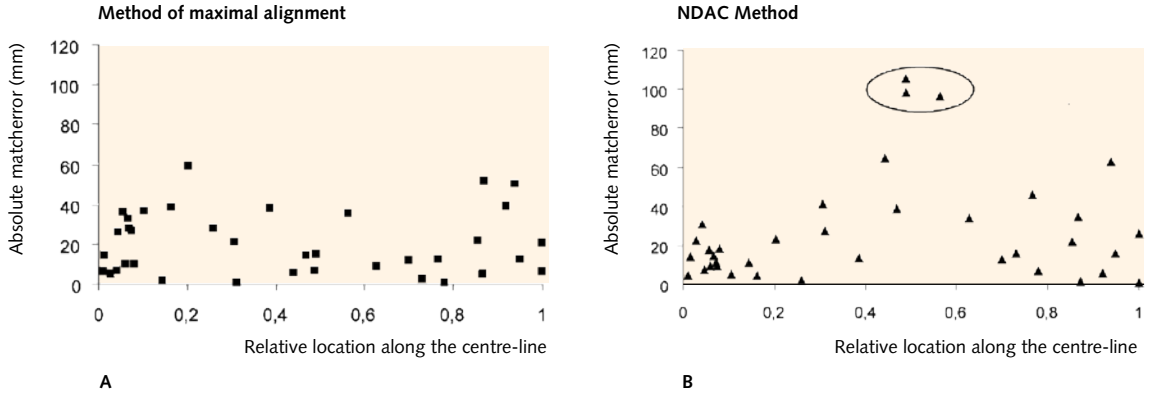


FIGURE 7.3

A The absolute match error of the method based on the maximal alignment of the centre-line plotted against the relative location of a polyp along the centre-line in supine position. The 0 on the x-axis corresponds to caecum and 1 to rectum.

B The absolute match error of the NDAC method plotted against the relative location of a polyp along the centre-line in supine position. The 0 on the x-axis corresponds to caecum and 1 to rectum. The three arrows indicate polyps from three different patients with a relatively large absolute match error.

abdominal cavity. We assume that this is the most mobile part of the colon, based on our observations of the three-dimensional overviews in both positions. The fact that polyps with a substantial difference in match error are situated in this segment might suggest that the method that relies on maximal alignment of centre-lines has an advantage in this mobile region.

The methods of matching prone and supine positions were evaluated based on the assumption that the included polyps were (immobile) true polyps and, second, that the actual link between these polyps in both positions was correct. In case there was doubt about the status of the polyp and linking, the polyp was excluded. With these exclusion criteria a very robust reference standard was created. Consequently, using this reference standard we were able to determine most precisely the matching quality of the methods.

Matching is important in determining whether a suspected lesion is a polyp or faecal material. A rule of thumb is that a suspected lesion with an unchanged position in prone and supine is most likely a polyp, while a change of position is related to stool. An accurate matching algorithm is therefore helpful to reduce the number of false positive findings in an efficient way.

Pitfalls in diagnosis can be caused by adherent stool mimicking a polyp, in case faecal tagging is absent or inadequate. On the other hand, pedunculated polyps (large stalk) or polyps situated in mobile segments may move (or seem to move) when both positions are compared. Therefore, when assessing whether a suspicious finding is a polyp, it is important not only to take the location into account, but also homogeneity and morphology of the lesion.

Incorporated in a CAD algorithm, this matching tool could also be used to reduce the number of false positive findings. Na¨ppi et al [5] used a region-growing scheme with distance calculations to divide the colonic lumen into overlapping segments that match in the supine and prone data sets. Polyp candidates detected by means of a CAD scheme were eliminated if they could be seen in only one of the two corresponding segments.

A limiting factor for generalizing the visibility of the matched polyps is the use of a 3D unfolded cube method for reviewing the data. Using this method, 70% of the polyps could be seen instantly when the match was made. This review method may have had a positive influence on the percentage of polyps that were seen instantly, since the unfolded cube display method creates a 360°view without major distortion of the image.

Although evaluated here in a 3D-setting, the method of matching prone and supine datasets can be applied in a 2Dsetting as well. This can be done on condition that a centreline is calculated through both corresponding colons. Theoretically, the matching quality of the method of maximal alignment is not altered by absence of a part of the beginning or ending of a centre-line in one position (e.g. caecum or rectum). In other words, the absence of baseline adjustment, a condition in other clinically evaluated methods [3, 6, 7], will not have a large influence on the match error in this new method. This baseline adjustment may not be possible in patients with an inadequately distended or insufficiently cleansed caecum or rectum in one position. Although this condition was not present in any of the data sets used, this proposition was confirmed in a test case. Here we removed 50 cm of the beginning of the centre-line without large consequences for the match error in the remaining colon segments.

The design of this feasibility study, with selected polyps as fixed reference points, precludes an assessment of the clinical value of the matching tool. Further study should assess matching quality and time efficiency of the matching tool in a series of unselected CT colonography examinations.

From this study we conclude that automated pronesupine matching of CT colonography studies is feasible and the difference in median absolute match error from theNDACmethodwas small and not significant, although half way along the centre-line (transverse colon) more pronounced differences were seen in some cases.

References

1. van Gelder RE, Nio CY, Florie J, Bartelsman JF, Snel P, De Jager SW, et al. Computed tomographic colonography compared with colonoscopy in patients at increased risk for colorectal cancer. Gastroenterology 2004;127:41–8.
2. Rouchka E. Dynamic programming. Stockholm Bioinformatics Center 2004 June 1. Available from: http:// www.sbc.su.se/,per/molbioinfo2001/dynprog/ dynamic. html [Accessed 23 March 2006].
3. Iordanescu G, Summers RM. Automated centerline for computed tomography colonography. Acad Radiol 2003;10:1291–301.
4. Saunders BP, Fukumoto M, Halligan S, Jobling C, Moussa ME, Bartram CI, et al. Why is colonoscopy more difficult in women? Gastrointest Endosc 1996;43(2 Pt 1):124–6.
5. Nappi J, Okamura A, Frimmel H, Dachman A, Yoshida H. Region-based supine-prone correspondence for the reduction of false-positive CAD polyp candidates in CT colonography. Acad Radiol 2005;12:695–707.
6. Acar B, Napel S, Paik DS, Li P, Yee J, Beaulieu CF. Registration of supine and prone CT colonography data: Method and evaluation. Radiology 2001;221:332.
7. Li P, Napel S, Acar B, Paik DS, Jeffrey RB, Beaulieu CF. Registration of central paths and colonic polyps between supine and prone scans in computed tomography colonography: pilot study. Med Phys 2004;31:2912–23.

Chapter 8

Polyp Measurement Based on CT Colonography and Colonoscopy: Inter-Observer Variability and Systematic Differences

Ayso H. de Vries
Shandra Bipat
Evelien Dekker
Marjolein H. Liedenbaum
Jasper Florie
Paul Fockens
Roel van der Kraan
Elizabeth M. Mathus-Vliegen
Johannes B. Reitsma
Roel Truyen
Frans M. Vos
Aeilko H. Zwinderman
Jaap Stoker

Accepted for publication European Radiology

Abstract

Objective

To assess variability and systematic differences of polyp measurements of colonoscopy and CT colonography.

Materials and Methods

Gastroenterologists measured 51 polyps by visual estimation, forceps comparison and linear probe. CT colonography observers randomly assessed polyp size of these polyps two-dimensionally (abdominal and lung window) and three-dimensionally (manually and semi-automatically). We used linear mixed models to assess variability and systematic differences between CT colonography and colonoscopy techniques.

Results

The variability of forceps and linear probe measurements were comparable and showed both less variability than measurement by visual assessment. Measurements by linear probe were 0.7mm smaller than measurements by visual assessment or by forceps. Variability of all CT colonography techniques was smaller compared to measurements by forceps or visual assessment and sometimes smaller (only 2D lung window and manual 3D) compared to measurements by linear probe. All CT colonography measurements measured polyps larger than colonoscopy varying from 0.7 mm to 2.3 mm.

Conclusion

A linear probe does not reduce the measurement variability of endoscopists compared to the forceps. Measurement differences between observers in CT colonography were most often smaller compared to measurements by optical colonoscopy. Polyps appear larger when using various CT colonography techniques than when measured during colonoscopy.

Introduction

CT colonography has consistently shown to have a high accuracy for the detection of colorectal neoplasia [1]. The method is less invasive and less burdensome than optical colonoscopy [2]. However, assessment of malignancy by obtaining tissue samples for histological analysis is not possible with this technique. As a surrogate for histopathology, polyp size is used for patient management strategies [3].

Currently, according to the American screening guidelines of CT colonography [4], all patients with a large polyp (10mm or larger) or medium sized polyp (6-9 mm) should be referred for colonoscopy. However, whether these medium sized polyps are an indication for colonoscopy is still under debate since the prevalence of advanced features was reported to be low[5;6]. Surveillance for growth with CT colonography has been suggested as a safe alternative [7]. Small (<6mm) polyps may be safely left in situ because of a negligible risk of malignant transformation.

So, size is crucial for decision making in CT colonography. Differences in polyp size measurements between CT colonography and colonoscopy should be minimal to avoid difficulties in the choice of management.

Two pivotal in-vitro studies report an underestimation of polyp size by the endoscopists and an accurate or slight overestimation of CT colonography[8;9] compared with colonoscopy measurements. However, factors that may influence polyp measurement such as difficult viewing angles or bowel motility are not considered in these analyses. Several in-vivo studies about differences in polyp size measurements and variability between CT colonography and optical colonoscopy have been published [10-13] as well. In these studies different CT colonography review modes, window settings and automatic measurement tools were compared to one of the various colonoscopy reference standards e.g. measurement by linear probe or forceps. These studies did have contradictory results i.e. some studies reported an underestimation [10;11] of polyp size whereas others reported an overestimation [12] of polyp size by 2D and 3D CT colonography measurements.

A comparison of the various CT colonography and colonoscopy measurement techniques within one study most likely will give more insight into the level of agreement between these measurement techniques. Therefore the purpose of our study was to assess the variability in size measurements within CT colonography and colonoscopy techniques. The variability can be considered a result of an inherent difference of techniques and/or differences of observers using the technique. Secondly, the purpose was to assess systematic differences in polyp size measurements between CT colonography and optical colonoscopy techniques.

Material and Methods

The study was approved by the institutional review board of our institute, and all patients provided written informed consent for participation in this study.

We compared 2D (abdominal and lung window) and 3D (manual and semi-automatic) CT colonography techniques with optical colonoscopy measurement techniques i.e. measurement by visual estimation, by comparison with a forceps and comparison by a calibrated linear probe. The respective measurements were done by three experienced CT colonography readers and three experienced colonoscopists on identical colorectal polyps. Patients were enrolled from two comparative studies of CT colonography and optical colonoscopy [14;15]. Polyps detected in these patients were used in a later phase for this measurement study.

Inclusion and exclusion criteria measurement study

Patients were excluded from the measurement study if the colonoscopy examination was not digitally stored, measuring devices were not available during the examination, or there was too much time pressure to properly execute the three different measurements during

the colonoscopy. Patients without polyps of 4mm or larger and patients with too many polyps that therefore could not be properly matched, were excluded as well. Polyps were included in the measurement study if they (1) were seen with both modalities, (2) measured by all techniques during the recorded colonoscopy, (3) were estimated at least 4mm (to be sure that all medium sized polyps would be included) and not larger than the linear probe (20mm) based on initial visual assessment by the executing endoscopist, (4) were not (partially) covered by fecal material in both positions during CT colonography (which would require electronic cleansing to sustain 3D measurement) and (5) were unambiguously matched.

A polyp detected on both CT colonography and optical colonoscopy was matched based on two criteria; (1) its appearance visually resembled the corresponding polyp at the colonoscopy movie and (2) its segment or adjacent segment corresponded with one of the six the reference segments. The polyps were matched to the colonoscopy findings by a research fellow with an experience of at least 150 CT colonography examinations verified by colonoscopy. A maximum of three polyps per patient were included to prevent substantial lengthening of the colonoscopy procedure.

CT colonography

Patients were scanned in supine and prone position after intravenous administration of bowel relaxants (20 mg butylscopolamine; Buscopan; Boehringer Ingelheim, Germany or, if contraindicated, 1 mg glucagon hydrochloride; Glucagon; Novo-Nordisk, Bagsvaerd, Denmark). CO2 was automatically insufflated up to maximum patient tolerance (PROTOCO2L, EZ-E-M, Lake Success, USA). Intravenous contrast was not administered. For the exact preparation schemes we refer to the comparative studies from which the polyps for this study were enrolled[14;15]. In short, oral iodine contrast was added to the patient's low fiber diet at least one day before the examination. In study I [14] this preparation scheme was combined with 4 l polyethylene glycol electrolyte solution (KleanPrep; Helsinn Birex Pharmaceuticals, Dublin, Ireland).

Examinations were performed on a 64-slice CT-scanner (Brilliance, Philips Healthcare, Best, the Netherlands). The collimation was 64*0.625 mm, pitch 1.2, slice thickness 0.9 mm, rotation time 0.4 s and tube voltage 120 kV. In Study I [14] the tube current was either 58 mAs or 82 mAs depending on the abdominal circumference (respectively less than 102,5cm or larger). In Study II [15] the tube current was modulated automatically (reference 40 mAs).

Colonoscopy

Within two weeks after CT colonography, optical colonoscopy was performed by an experienced staff member or a gastroenterology fellow under direct supervision of the attending gastroenterologist. All patients had undergone an extensive bowel preparation consisting of 4 liter of polyethylene glycol and a low-fiber diet. The procedure was done with a standard colonoscope (CF-140L; Olympus, Tokyo, Japan). The colonoscopy was videotaped and subsequently digitally stored. Segmental unblinding was performed according to the findings of CT colonography.

During colonoscopy a polyp was measured by three different techniques in succession; by (1) visual assessment, (2) comparison with an opened forceps with a size of 8 mm and (3) a calibrated linear probe (Figure 8.1). The linear probe (Olympus, Tokyo, Japan) had ten markings with a spacing of 2 mm on the distal flexible tip. During the procedure polyps were recorded with respect to size (these measurements were not used in this study), morphology and segment for patient management purposes and polyp matching purposes. Digital movie extracts of the polyp measurements were later presented to three experienced endoscopists in the framework of this polyp measurement study.

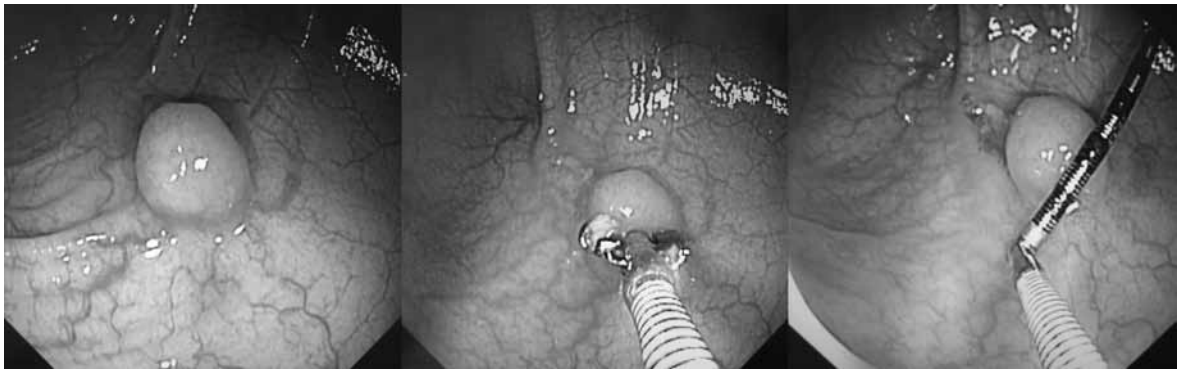


FIGURE 8.1
The three different measurement techniques of colonoscopy; visual assessment (left), measurement by forceps (middle) and measurement by calibrated linear probe (right).

Assessment of polyp size at CT colonography

The measurements in the framework of the polyp measurement study were done by three observers who had read respectively 200, 500 and 500 CT colonography studies verified by colonoscopy. The measurement of each polyp was done on a tailor made computer program based on the ViewForum workstation.

Using this program measurements were done 2D and 3D. In 2D, each polyp was measured in a reformatted cross sectional plane through the polyp. The plane could be rotated to identify the longest object dimension.

A polyp was measured by placing electronic calipers along the largest diameter in the reformatted image. Measurements were done in a preset window width and level setting of 1500/-50 HU (lung window) and 400/40 HU (abdominal window). The 3D measurements were done both manually and semiautomatically. The observers were instructed to measure the maximum diameter by electronic calipers in the endoluminal display. The semi-automatic measurements were based on automatic measurements using a prototype algorithm that will be integrated in future ViewForum software. The observers either accepted these measurements of the algorithm or modified the measurements manually by repositioning the two software cursors. The measurements were done on the default surface-rendered colonic wall threshold setting of ViewForum of -650 HU. The observers were explicitly instructed to carefully place the cursors at the edge of the polyp and to not let the cursors 'fall off' the edges of the polyp. Like in our clinical practice, we used the larger of the two measurements performed in prone and supine position to avoid under treatment. To avoid recall bias, polyp size (mm) was not displayed after the measurement was performed i.e. the CT colonography observers were blinded to their own size measurements. Each 2D measurement using the abdominal window setting was directly followed by a 3D semiautomatic measurement. The 2D measurement in lung window was directly followed by a manual 3D measurement. These paired measurements were performed in random order. There was an interval of at least one day between these clustered measurements.

Assessment of polyp size at colonoscopy

The three experienced gastroenterologists had performed more than 1,500 and 5,000 and 2500 colonoscopies respectively. For each polyp three separate movie extracts were made, each corresponding to one measurement method (visual assessment, forceps or linear probe).

All the extracts were presented in random order to the

gastroenterologists. This was done on a laptop computer. Each movie extract was rated by each colonoscopist as 'good', 'sufficient' and 'insufficient' depending on the possibility to properly assess the polyp's largest diameter. Each gastroenterologist was blinded for the measurements by the other gastroenterologists as well as for the measurements that were done during CT colonography and the initial colonoscopy.

Statistical analysis

The difference in measured sizes between the mostly used colonoscopy measurement techniques (i.e. forceps and visual assessment) and the various CT colonography measurements were illustrated by Bland- Altman plots. In a Bland-Altman plot the mean value of two measurements is plotted against their difference [16]. The data consist of measurements on the same polyp by three experienced CT colonography readers and three experienced endoscopists with different techniques. For this study we used a linear mixed (regression) model. The model contained the following parameters:

1. fixed parameters to estimate the systematic differences in size measurement between techniques (technique as independent variable); i.e. to estimate whether measurement technique A measured polyps systematically larger or smaller than measurement technique B;
2. heterogeneous error variance structure to estimate the residual variance (variability) for each technique. We assumed that the variability of measurements within a technique is an indication for the precision of the measurement technique: the smaller the difference between the observers, the better the technique [17]. To test whether the residual variance (obtained by the model above) was statistically different between two techniques we fitted the same model but then assuming that residual variances of these techniques were similar. The Akaike's information criterion (AIC) values of the models were compared using a chi2 distribution with one degree of freedom. If the models did not differ (i.e. the models had a comparable AIC), the variances in size measurements between the techniques were considered similar. A poorer fit (higher AIC) of the latter model would indicate that the residual variances (variability) in measurement were different[18].
3. fixed effect parameters to estimate the systematic differences in size measurement between observers within a technique (observers within a technique as independent variable) i.e. to estimate whether observer A measured polyps consequently larger or smaller than observer B when using the same technique.

To illustrate the variability within the measurements of observers per technique and the variability within the measurements between techniques, we used Bland-Altman plots. In these plots we have corrected for systematic differences between observers since we assumed that these differences could be corrected by calibrating the measurements.

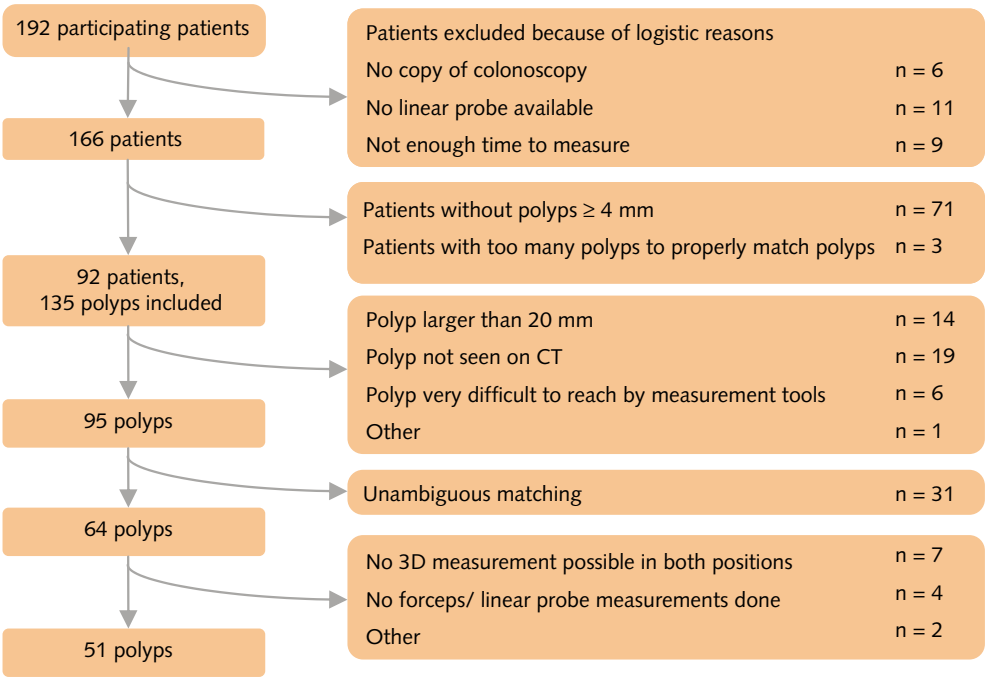
All analyses were performed using the linear mixed model procedure of a commercially available statistical software program (proc mixed, SAS Institute Inc 9.2, Cary, USA). P-values of < 0.05 were considered statistically significant.

Results

Polyp characteristics: In this study 192 patients were included (Figure 8.2). Fifty-one polyps in 44 patients fulfilled the selection criteria. Fifteen polyps of Study I [14] and 36 polyps of study II [15] were included.

The mean age of the participating patients was 61 years (SD 7.2) and consisted of 27 (66%) men. Thirty-seven polyps revealed an adenomatous histology, seven a hyperplastic histology and of six polyps the histology could not be retrieved. One carcinoma was included. Twenty-four polyps were classified by the executing endoscopist as sessile, six polyps as flat (according to the Paris criteria [19]) and 21 polyps as pedunculated. Ninety percent of the visual assessment movie extracts, 85% of the forceps measurement movie extracts and 74% of the linear probe

FIGURE 8.2
Flowchart



measurements movie extracts were rated at least 'sufficient' by the gastroenterologists. The median duration of the extracts was 33 seconds (P25-P75: 21-52s). The average polyp size (as measured by the different measurement techniques) is displayed in Table 8.1. The table shows that measurements by CT colonography tend to measure polyp size larger than measurements by colonoscopy, especially when using 2D lung window and semi-automatic 3D measurements.

The difference of size measurement between techniques between the mostly used colonoscopy measurement techniques (i.e. forceps and visual assessment) and the various CT colonography measurements are illustrated by the Bland-Altman plots in Figure 8.3 and Figure 8.4.The figures show that CT measures polyp size larger than forceps (Figure 8.3) and visual assessment (Figure 8.4). Using the linear regression model we determined the systematic differences in size measurement between techniques and the variability between techniques. The results are presented in Table 8.2.

TABLE 8.1 **Average Polyp Size (SD)**

Colonoscopy	
Visual assessment	8,1(7,0-9,3)
Forceps	8,1(7,1-9,0)
Linear probe	7,4(6,5-8,4)
CT colonography	
2D abdominal window	8,8 (7,9-9,7)
2D lung window	9,7 (8,8-10,7)
3D manual	9,3 (8,4-10,2)
3D semi-automatic	9,7 (8,7-10,7)

Systemic difference between techniques

Table 8.2 shows that measurements by visual assessment and forceps are systematically 0.69 mm larger than measurements by linear probe. Visual assessment and measurement by forceps do not differ. All four CT colonography techniques measure polyp size larger than

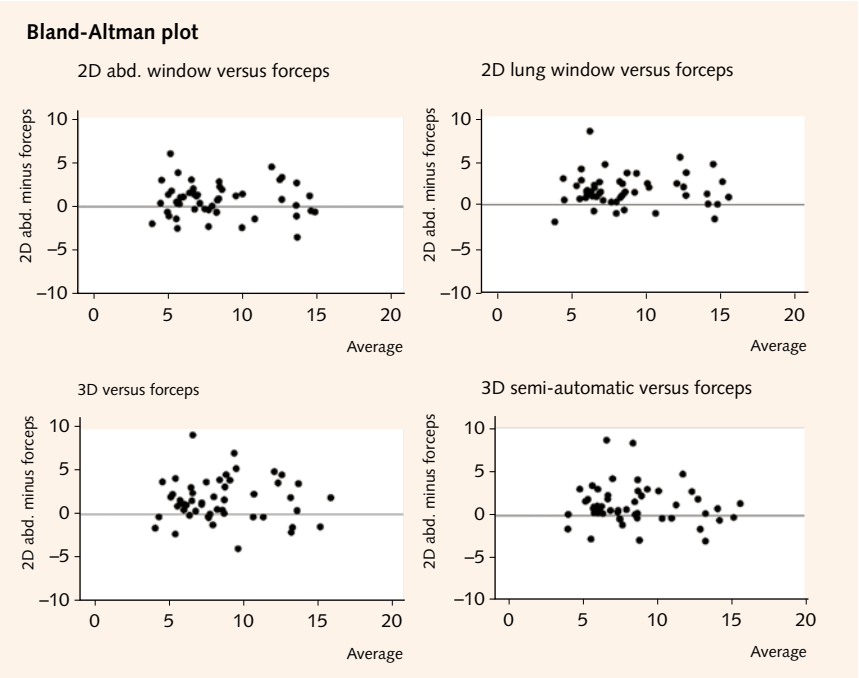


FIGURE 8.3
Most of the dots of this Bland-Altman plot are situated above the horizontal axis. This illustrates that CT measures polyp size most often larger than measurements by forceps.

TABLE 8.2 **Systematic Differences in Size Measurements Between Techniques and Estimates of Inter-observer Variability Within Techniques as Estimated by Linear Mixed Models.**

Technique 1	Technique 2	Mean systematic difference in size measurement between techniques (SD)	p	variability of measurements within technique 1	variability of measurements within technique 2	diference variance	p
visual assessment	forceps	0.03 (-0.53; 0.58)	0.92	7.53	4.42	3.11	0.0052
visual assessment	linear probe	0.71 (0.17;1.26)	0.01	7.53	3.94	3.59	0.0005
forceps	linear probe	0.69 (0.22; 1.15)	0.004	4.42	3.94	0.48	0.24
visual assessment	2D abdominal window	-0.65 (-1.17; -0.14)	0.01	7.53	2.87	4.65	< 0.0001
visual assessment	2D lung window	-1.61 (-2.10; -1.12)	< 0.0001	7.53	1.97	5.56	< 0.0001
visual assessment	3D	-1.19 (-1.68; -0.70)	< 0.0001	7.53	2.02	5.51	< 0.0001
visual assessment	3D semi automatic	-1.59 (2.10; -1.08)	< 0.0001	7.53	2.68	4.85	< 0.0001
forceps	2D abdominal window	-0.68 (-1.11; -0.25)	0.002	4.42	2.87	1.55	0.0455
forceps	2D lung window	-1.63 (-2.04; -1.23)	< 0.0001	4.42	1.97	2.45	< 0.0001
forceps	3D	-1.21 (-1.62; -0.81)	< 0.0001	4.42	2.02	2.40	0.0002
forceps	3D semi automatic	-1.62 (-2.05; -1.19)	< 0.0001	4.42	2.68	1.74	0.027
linear probe	2D abdominal window	-1.37 (-1.79; -0.95)	< 0.0001	3.94	2.87	1.06	0.29
linear probe	2D lung window	-2.32 (-2.71; -1.93)	< 0.0001	3.94	1.97	1.97	< 0.0001
linear probe	3D	-1.90 (-2.29; -1.51)	< 0.0001	3.94	2.02	1.92	0.0022
linear probe	3D semi automatic	-2.31 (-2.72; -1.89)	< 0.0001	3.94	2.68	1.26	0.16

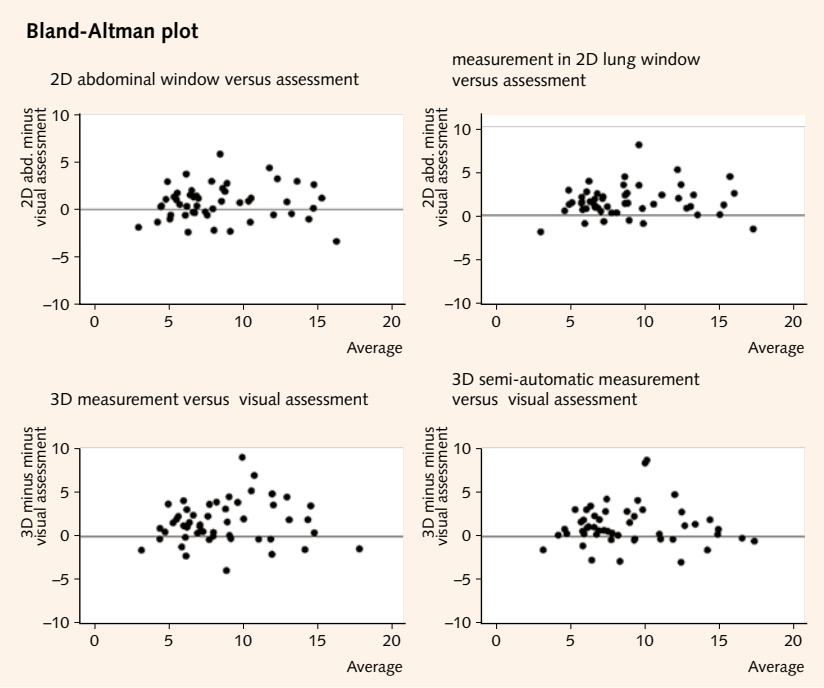


FIGURE 8.4
Most of the dots of this Bland-Altman plot are situated above the horizontal axis. This illustrates that CT measures polyp size most often larger than measurements by visual estimation.

the three colonoscopy techniques, varying from 0.65 mm (visual assessment versus 2D abdominal) to 2.32mm (linear probe versus 2D lung window). These differences were highly statistically significant. In general, the differences between CT and colonoscopy are smallest when polyps were measured in a 2D abdominal window setting and largest if the measurements were done in a 2D lung window setting or a 3D semi automatic mode.

Variability between techniques

We always applied the heterogeneous error variance structure. In Table 8.2 the variability of the various measurement techniques is displayed. Visual assessment showed more variability compared to forceps and linear probe measurements, the latter two did not differ. So, linear probe measurements are not more precise than measurements by forceps. Colonoscopy techniques in general showed more variability compared to CT colonography techniques. Among the CT colonography measurements manual 3D measurements and measurement in a 2D lung window showed the least variability. Thus, these techniques lead to less difference between observers.

Difference between observers within techniques

These differences were used to calculate the adjusted measurements i.e. adjusted for the systematic measurement differences between observers. The adjusted measurements are illustrated in Figure 8.5. The spread of dots in the Bland-Altman plots in Figure 8.5 represent the variability of the measurement techniques. The variability of the colonoscopy techniques is larger than the CT colonography techniques. The manual 3D measurement technique and the 2D measurement in lung window show the least variability.

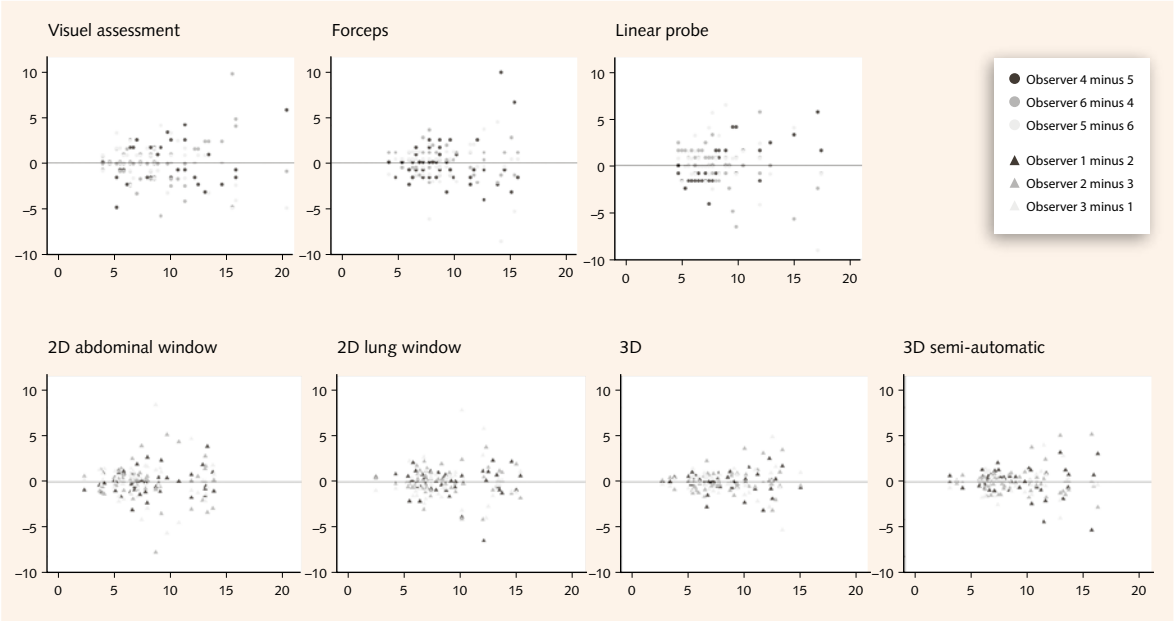


FIGURE 8.5
Bland-Altman plots of the various measurement techniques. The distribution of the dots around the horizontal axis (y=0) illustrates the variability of the measurements. The variability of the CT measurements is smaller than the colonoscopy measurements.

Discussion

This study shows that polyp size measured by linear probe does not reduce measurement variability compared to polyp size measured by forceps. The endoscopists measured polyps significantly larger when using a forceps or by visual estimation compared to using a linear probe. CT colonography observers measured polyps larger than endoscopists. Furthermore, measurement differences between observers in CT colonography were smaller, especially in manual 3D measurements and measurements in a 2D lung window, compared to measurements by optical colonoscopy.

Measurement studies are hampered by the lack of a reference standard of sufficient quality. Frequently used standards are measurement by sliding caliper after polyp removal, and endoscopic measurement by a forceps or linear probe [10-13;20-22].

Sliding caliper after polyp removal is not ideal. Vascular collapse, polyp desiccation from cautery, compression of the polyps after removal with a grasper or by suctioning through the endoscope may all contribute to a reduction in size of the reference polyp [21]. Because of the underestimation of polyp size and the fact that mainly pedunculated polyps can be removed in toto (bias), we did not consider these measurements as reference values in this study. On the other hand, endoscopy is not an ideal technique either [23;24], because the maximal diameter cannot always be displayed perpendicular to the direction of view and the endoscope images with a wide viewing angle may lead to optical distortion. Moreover, peristalsis and tortuousity of the colon will reduce the possibility to properly assess polyp size. This may lead to imperfect polyp size measurement. An in-vitro reference standard can be constructed very precisely. However, when using a phantom the influence of factors as colonic distension, difficult viewing angles or bowel movement on polyp measurement are difficult to assess. Therefore the

accuracy of different measurement techniques applied in a clinical situation may be overestimated. Because of the abovementioned drawbacks of comparing measurements with a suboptimal reference standard we have chosen for a different approach by determining variability and systematic differences of various measuring techniques. We assumed that the variability is an indication for the accuracy of the measurement technique: the smaller the difference between the observers, the better the technique [17]. However, a small variability does not rule out a large systematic error to the truth.

Conflicting results about the accuracy of CT colonography measurements in in-vivo measurement studies [10-12] have been published. In these studies in which colonoscopy was used as reference standard, underestimation [10;11] and overestimation [12] of polyp size by 2D and 3D CT colonography measurements has been reported. The reference standard in these studies was either a measurement by linear probe or forceps during colonoscopy. Our findings may explain in part the differences between previous studies that used the measurements of a gastroenterologist as reference standard. The fixed observer effects of our full model revealed that systematic differences between observers exist. The systematic differences were larger for colonoscopy measurements (up to 3mm, data not shown) than for CT colonography measurements (up to 1.1mm, data not shown). Using this model we assumed that systematic differences between observers can be avoided or repaired in practice by training. So, the operator dependency of colonoscopy measurements may be an explanation for the different conclusions drawn from the comparative measurement studies. Our study showed that polyp size is measured larger by CT colonography techniques than by colonoscopy techniques. As a consequence, in practice patients would more easily be referred for colonoscopy. We have shown that the magnitude of the difference depends on the compared techniques i.e. the mean differences were smallest in 2D abdominal window measurements and largest when using 2D lung window settings or 3D semi automatic measurements (Table 8.2 and Figure 8.3 and 8.4). However, the difference also depended on observers since structural differences between observers exist i.e. some observers systematically measure polyps larger than others. Therefore, it is not possible to determine a general adjusted CT threshold for referral.

The executing endoscopists often experienced positioning of the linear probe as more difficult than the positioning of a forceps. This may be due to the eccentric mounting of the probe (Figure 8.1). As a consequence the measurement often had to be performed in the periphery of the field of view, which made measurements more difficult. Therefore using a linear probe often produces a semi-subjective estimate instead of an exact measurement.

This may explain the fact that the tool does not reduce differences in measurement between gastroenterologists. Moreover, the differences in measurement between gastroenterologists using a linear probe would be probably even larger if we had not excluded polyps that were measured larger than 20 mm (i.e. the length of the linear probe). This has created a bias against CT colonography that can more easily measure these types of large lesions. However, the referral strategy for polyps of 20 mm or polyps larger than 20mm is the same i.e. colonoscopy.

Just a few studies have evaluated variability of different techniques. Fennerty et al. showed in a phantom study that polyp size measured using a forceps significantly differed between observers [23]. To our knowledge no studies have been published that evaluated variability between equivalent observers in both CT and colonoscopy measurement techniques for the same polyps.

This study had limitations. In this study we did not have a true reference standard. We have chosen for this approach because of the abovementioned arguments. However, our data match with the data of Punwami et al. and Park et al. [8;9] in the sense that polyps are measured smaller by endoscopists than by CT colonography. The colonoscopy

observers examined a video. Therefore they could not influence the movement of the endoscopic camera, the administration of additional bowel relaxants or the insufflation of additional air. By asking the actual executing colonoscopists (not the three colonoscopy observers) to not only position the measurement tool but also try to measure polyp size, we have aimed to maximise the quality of the measurement movie.

Their commitment is reflected by the median captured video length of 33 seconds of the measurement abstract. Despite this commitment, the quality of 9% of the visual assessment movie extracts and 26% of the linear probe measurement movie extracts was rated as 'insufficient' by our observers. In our opinion this reflects the difficulty of proper placement of the endoscope and measurement tools for measurement purposes in practice.

We used a single dedicated workstation for CT colonography polyp measurement. For 3D measurement we used a surface rendered reconstruction. The images were reconstructed with a threshold of 650HU. According to Park et al. optimal surface-rendering threshold value for accurate polyp measurement is approximately -500 HU. This may have (marginally) overestimated polyp size using 3D measurement by approximately 0.3mm [25].

We used an interval of at least one day between the clustered CT colonography polyp measurements. Although there is the possibility of recall bias using this relatively short interval, the large number of randomised measurements (hundreds per observer) and the fact that the observers were not aware of the actual size measurement in mm may have ruled out recall bias to a large extent

In this study the results were based on the measurements of only three experienced CT colonography observers and only three experienced colonoscopists. Although this is more than used in most measurement studies, it may limit the generalisability. In conclusion, measurements by linear probe are not more precise when compared to forceps measurements. Moreover, CT colonography measurements are more precise and measure polyps larger than colonoscopy.

ACKNOWLEDGEMENTS – The authors thank Henk W. Venema for his critical appraisal of the manuscript.

References

1. Halligan S, Altman DG, Taylor SA, et al. CT colonography in the detection of colorectal polyps and cancer: systematic review, meta-analysis, and proposed minimum data set for study level reporting. Radiology 2005; 237:893-904
2. van Gelder RE, Birnie E, Florie J, et al. CT colonography and colonoscopy: assessment of patient preference in a 5-week followup study. Radiology 2004; 233:328-337
3. Zalis ME, Barish MA, Choi JR, et al. CT colonography reporting and data system: a consensus proposal. Radiology 2005; 236:3-9
4. McFarland EG, Levin B, Lieberman DA, et al. Revised colorectal screening guidelines: joint effort of the American Cancer Society, U.S. Multisociety Task Force on Colorectal Cancer, and American College of Radiology. Radiology 2008; 248:717-720
5. Atkin WS, Morson BC, Cuzick J. Long-term risk of colorectal cancer after excision of rectosigmoid adenomas. N Engl J Med 1992; 326:658-662
6. O'Brien MJ, Winawer SJ, Zauber AG, et al. The National Polyp Study. Patient and polyp characteristics associated with highgrade dysplasia in colorectal adenomas. Gastroenterology 1990; 98:371-379
7. Kim DH, Pickhardt PJ, Taylor AJ, et al. CT colonography versus colonoscopy for the detection of advanced neoplasia. N Engl J Med 2007; 357:1403-1412
8. Park SH, Choi EK, Lee SS, et al. Polyp measurement reliability, accuracy, and discrepancy: optical colonoscopy versus CT colonography with pig colonic specimens. Radiology 2007; 244:157-164
9. Punwani S, Halligan S, Irving P, et al. Measurement of colonic polyps by radiologists and endoscopists: who is most accurate? Eur Radiol 2008; 18:874-881
10. Burling D, Halligan S, Taylor S, et al. Polyp measurement using CT colonography: agreement with colonoscopy and effect of viewing conditions on interobserver and intraobserver agreement. AJR Am J Roentgenol 2006; 186:1597-1604
11. Jeong JY, Kim MJ, Kim SS. Manual and automated polyp measurement comparison of CT colonography with optical colonoscopy. Acad Radiol 2008; 15:231-239
12. Pickhardt PJ, Lee AD, McFarland EG, Taylor AJ. Linear polyp measurement at CT colonography: in vitro and in vivo comparison of two-dimensional and three-dimensional displays. Radiology 2005; 236:872-878
13. vanWijk C, Florie J, Nio CY, et al. Protrusion method for automated estimation of polyp size on CT colonography. AJR Am J Roentgenol 2008; 190:1279-1285
14. de Vries AH, Jensch S, Liedenbaum MH, et al. Does a computeraided detection algorithm in a second read paradigm enhance the performance of experienced computed tomography colonography readers in a population of increased risk? Eur Radiol 2009; 19:941-950
15. Marjolein H.Liedenbaum, Anne F.van Rijn, Ayso H.de Vries, et al. Using CT colonography as a triage technique after a positive Faecal Occult Blood Test in colorectal cancer screening. Gut 2009. Ref Type: In Press
16. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1:307-310
17. Quinn MF. Relation of observer agreement to accuracy according to a two-receiver signal detection model of diagnosis. Med Decis Making 1989; 9:196-206
18. Akaike H. A new look at the statistical model identification. IEEE Trans Automat Control 2008; 19:716-723 19. The Paris endoscopic classification of superficial neoplastic lesions: esophagus, stomach, and colon: November 30 to December 1, 2002. Gastrointest Endosc 2003; 58:S3-43
20. Gopalswamy N, Shenoy VN, Choudhry U, et al. Is in vivo measurement of size of polyps during colonoscopy accurate? Gastrointest Endosc 1997; 46:497-502
21. Morales TG, Sampliner RE, Garewal HS, Fennerty MB, Aickin M. The difference in colon polyp size before and after removal. Gastrointest Endosc 1996; 43:25-28
22. Schoen RE, Gerber LD, Margulies C. The pathologic measurement of polyp size is preferable to the endoscopic estimate. Gastrointest Endosc 1997; 46:492-496
23. Fennerty MB, Davidson J, Emerson SS, Sampliner RE, Hixson LJ, Garewal HS. Are endoscopic measurements of colonic polyps reliable? Am J Gastroenterol 1993; 88:496-500

24. Margulies C, Krevsky B, Catalano MF. How accurate are endoscopic estimates of size? Gastrointest Endosc 1994; 40:174- 177

25. Park SH, Choi EK, Lee SS, et al. Linear polyp measurement at CT colonography: 3D endoluminal measurement with optimized surface-rendering threshold value and automated measurement. Radiology 2008; 246:157-167

Chapter 9

Summary and General Discussion

Ayso H. de Vries

Summary

This thesis addresses different aspects of CT colonography that all clinically evaluate technical developments in CT colonography.

The *extensive* bowel preparation is often described as the most burdensome aspect by patients that have undergone colonoscopy and CT colonography [1;2]. Reducing the laxative component of the bowel preparation may increase the patient's willingness to participate in a screening or surveillance program; however image readability may diminish because more colon surface is covered by fecal material.

In **Chapter 2** 40 consecutive patients who were randomized to four mild bowel preparation schemes that had an increasing laxative component were compared.

Group 1 received 20 mg of bisacodyl, group 2 received 30 mg of bisacodyl, group 3 received 20 mg of bisacodyl (as group 1) and an additional 8.2 g of magnesium citrate, and group 4 received 30 mg of bisacodyl (as group 2) and an additional 16.4 g of magnesium citrate. All patients used a two-day low-fiber diet and received oral iodine contrast and barium for fecal tagging. The groups were compared on diagnostic quality of the images and perceived patient burden.

The results demonstrated that the diagnostic quality was not negatively affected by reducing the laxative components in the bowel preparation (i.e. bisacodyl and magnesium citrate). The diagnostic quality of the group that had undergone the least laxative bowel preparation (group 1) was still rated good or excellent. Yet, a higher dosage of laxatives was significantly associated with a higher burden of diarrhea and a perceived higher overall burden of bowel preparation. Based on these findings the conclusion was drawn that CT colonography with limited bowel preparation, using barium and ionic iodinated contrast agents for fecal tagging, requires only minimal doses of laxatives - in our study only 20 mg of bisacodyl - to obtain good diagnostic quality and to minimize patient burden. Since the image quality of the most patient-friendly bowel preparation is rated as at least "good", it is likely that the patient burden can even be further reduced.

In reading CT colonography examinations there are basically two approaches: it can be done two-dimensionally (2D) or three-dimensionally (3D). In practice 2D and 3D images are used complimentary.

A review method is named *primary* 2D if 2D images are used for the primary reading; 3D images are then only used for problem solving. Alternatively, the review method is named *primary* 3D if 3D rendered images are used for the primary reading of the examination with complementary 2D images for problem solving.

Using a primary 3D review method may be of restricted value in patients that have undergone a limited bowel preparation since fecal material may obscure substantial parts of the colon in both prone and supine position. Therefore, an algorithm that virtually removes fecal material (i.e. electronic cleansing) is required to improve the reading of these images.

In **Chapter 3** the effect of an electronic cleansing algorithm on lesion conspicuity, evaluation time, assessment effort and observer confidence in a *selected* patient population was reported. This feasibility study was performed using a primary 3D review method.

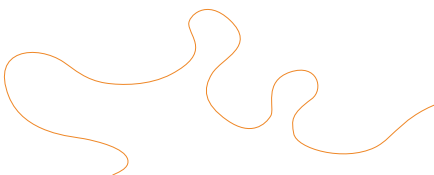
For this study two experienced CT colonography observers compared CT colonography examinations *before* and *after* electronic cleansing. Based on their findings, we discovered that lesions uncovered by electronic cleansing had a similar conspicuity compared to lesions that were already surrounded by air. Furthermore, the electronic cleansing algorithm decreased evaluation time, lowered assessment effort, and increased observer confidence compared to CT colonography *without* electronic cleansing. So, from this study we concluded that the algorithm improves primary 3D evaluation in CT colonography.

In **Chapter 4** electronic cleansing was evaluated in an *unselected* patient population that had undergone a reduced bowel preparation. This was done by comparing two different reading strategies i.e. primary uncleansed 2D and primary electronically cleansed 3D. The comparison was done in terms of accuracy, reading time and diagnostic quality. These two reading strategies are the obvious choice for examinations of *limited* preparation.
The results showed that the detection of clinically relevant polyps by six trained but relatively inexperienced observers was significantly higher when using a primary electronically cleansed 3D method. Experienced observers, however, performed overall better than inexperienced observers and performed equally well using both methods. Specificity was not affected when using primary electronically cleansed 3D. Despite its superior performance in polyp detection for the relatively inexperienced observers, primary electronically cleansed 3D was more susceptible for artifacts than the 2D alternative. This resulted in an inferior rating of the diagnostic quality of primary electronically cleansed 3D examinations. Primary 3D was less time efficient for most observers as well.

The use of oral contrast provided the possibility to reduce the laxative component of the bowel preparation in CT colonography. As mentioned above, this may reduce the patient burden of bowel preparation. Consequently, more colon surface will be covered by fecal material. And even though oral contrast is added to the patient's diet, the difference in grey values on CT of the polyp relative to its surrounding (and therefore its detectability) is often reduced.
In **Chapter 5** our objective was to determine the influence of contrast material and radiation dose on the visibility of colorectal polyps. Therefore, a study was conducted using an anthropomorphic phantom. In this study three observers determined the visibility of six millimeter polyps submerged by contrast material of various densities (grey values) scanned at different radiation doses. The study showed that to adequately visualize a polyp covered by fecal material (especially fecal material of low density) a considerably higher radiation dose is needed than to adequately visualize a polyp surrounded by air.

In **Chapter 6** the additional value of a computer aided detection (CAD) algorithm was determined. Promising results of CAD reducing the number of missed lesions for less experienced observers have been published. However, the additional value of CAD for experienced observers is still controversial. Furthermore, in previous studies performance of CAD was investigated in selected and polyp-enriched populations only. This may have had a positive effect on the observer performance.
Therefore, we determined whether CAD in a second read paradigm could improve the performance of experienced observers in a population of 170 consecutive patients at increased risk for colorectal cancer.
In this population, CAD detected one extra medium size lesion (6–9 mm) and three extra large polyps (≥ 10 mm) initially missed by the observer. Consequently, the sensitivity of CT colonography *without* CAD for patients with lesions ≥ 6 mm and ≥ 10 mm increased from respectively 80% and 64%, to 82% and 72% *with CAD*. This small increase was not statistical significant. Specificity with and without CAD remained (nearly) unchanged. Thus, although CT colonography *with* CAD detected a few more patients than CT colonography *without* CAD, it had no statistically significant positive influence on the performance of experienced observers in a population at increased risk for colorectal cancer. A drawback of this study was the relative high number of flat, difficult-to-see polyps. This drawback will be discussed in the 'General Discussion' section of this chapter.

Currently it is common sense to scan patients in both supine and prone position. This is done to improve the accuracy of polyp detection. Combining information from these scans will assist the reviewer in differentiating polyps



from untagged fecal material as most fecal material is subject to gravity; in most cases fecal material will have another position when the patient turns over.
Manual verification of findings on supine and prone positions may be a time-consuming activity since reference-points (e.g. hepatic flexure) are often not fixed. An automated matching method may facilitate this matching process.
In **Chapter 7** the feasibility of automatic prone-supine matching was evaluated using thirty-seven proven polyps as fixed points of reference. For this study we used an algorithm that was based on similarities in direction of the centre-line of the colon in both positions.
The study was done using a primary 3D display method. The results showed that in prone position all polyps could be seen within a median range of 14 millimeter (range: 0 - 59 mm) of the position calculated from the polyp's location in supine position; consequently most of the polyps (70%) were directly visible in prone view.

CT colonography is an examination that visualizes the colon but does not offer the therapeutic option of polypectomy (in contrast to optical colonoscopy). Currently, according to the American screening guidelines of CT colonography [3], all patients with a large polyp (10mm or larger) or medium sized polyp (6-9 mm) should be referred for colonoscopy. However, whether these medium sized polyps are an indication for colonoscopy is still under debate. Surveillance for growth with CT colonography has been suggested as a safe alternative [4]. Small polyps (<6mm) maybe safely left in situ because of a negligible risk of malignant transformation. Thus, size is crucial for decision making in CT colonography.
The purpose of the study in **Chapter 8** was to assess the variability of measurements within CT colonography and optical colonoscopy and secondly, to assess systematic differences in polyp size measurements between these techniques. The respective measurements were done by three experienced CT colonography observers and three experienced gastroenterologists.

Based on these measurements we concluded that:
1 Measurement differences between gastroenterologists are not smaller when using a linear probe compared to when using a forceps;
2 Gastroenterologists measured polyps significantly larger by forceps or by visual estimation than by linear probe;
3 Measurement differences between observers in CT colonography are smaller, especially using manual 3D measurements and measurements in a 2D lung window setting, compared to measurements by optical colonoscopy;
4 CT colonography observers measure polyps larger than endoscopists.

General Discussion

This thesis demonstrates that

- The image quality of CT colonography in patients that have undergone a limited bowel preparation using a two-day low-fiber diet, oral iodine contrast and barium for fecal tagging and 20mg bisacodyl is not worse compared to bowel preparation schemes with a larger laxative effect. However, limiting the laxative component does improve patient acceptance.
- For relatively inexperienced observers, electronic cleansing is a useful tool in reading CT colonography studies of limited prepped patients. For experienced observers, however, there is no additional value.
- One has to be careful with reducing radiation dose in the limited prepped patients because of a reduced contrast between the bowel wall and the lumen. The reduced contrast leads to reduced polyp conspicuity.
- Computer aided detection in a second read paradigm (i.e. as backup observer) does not significantly increase observer performance of experienced observers.
- Prone and supine matching is feasible.
- CT colonography observers tend to measure polyp size larger and with less variability than gastroenterologists.

Recently published large and well designed studies show that CT colonography has proven to be a good method for the detection of colorectal cancer and its precursors in a screening population (i.e. individuals with an average risk for colorectal cancer) [5;6]. The studies show a per-patient sensitivity of 90% for lesions of 10mm or larger [5] or a detection rate for these lesions equal to colonoscopy [6]. However, in the comparative study of CT colonography and colonoscopy in **Chapter 6** reported a per-patient sensitivity of 72% for these advanced lesions. The detection rate of the well-trained observers combined with CAD as second reader was therefore nearly 20% less than in the abovementioned studies.

An important difference between the studies was the patient population. The present study was executed in a patient population of increased risk for colorectal cancer i.e. a population with a personal or familial history of colorectal cancer. This population differed from an average risk screening population in two aspects: the relative high prevalence of lesions and the unusually high number of large flat lesions.

The latter has resulted in a relative large number of undetected lesions in the present study; mainly the flat lesions were missed by CT colonography observers and CAD. The superior performance of colonoscopy in this respect may be related to the fact that colonoscopy does not only rely on morphology (as does CT) but also other features such as changes in vascularity or mucosal structure of the colonic wall.

The unusually high number of flat lesions may be related to the patient's endoscopy history. Previous endoscopic screening could affect the sensitivity of CT colonography in a later stage in two ways. First, it is likely that many easy-to-see polyps and fewer hard-to-see polyps would be detected and removed at earlier colonoscopy. As a result the percentage of hard-to-see polyps increases. Second, a previous endoscopic polypectomy may have been incomplete. The remnants may be flatter than the original intact lesions resulting in a hard-to-see polyp.

Improving detection of these hard-to-see flat lesions by training, improved CT data acquisition or dedicated computer aided detection algorithms may be of limited value in improving polyp detection in this population, because the majority of these relevant lesions could not even be seen in retrospect. Therefore, the question rises whether an increased risk patient population should undergo surveillance with CT colonography instead of the current gold standard (colonoscopy). In line with this, some investigators have proposed that the use of CT colonography in a surveillance population has to be considered with caution. This proposition is endorsed by our results.

To maximize patient's willingness to participate in a surveillance or screening program for colorectal cancer, the patient burden of CT colonography needs further reduction. As shown in **Chapter 2**, this can be done by reducing the laxative component of bowel preparation that causes diarrhea and abdominal pain. The study showed that the subjective image readability was not impaired by reducing the laxative component of the preparation. Therefore it is likely that the bowel preparation in the present study (that consisted of 20 mg bisacodyl combined with a low fiber diet and barium and iodine contrast material) can be further reduced without impairment of the readability of the images.

At the time of the writing of this thesis, we no longer add any laxative agent (i.e. bisacodyl or magnesium citrate) to the patient's bowel preparation. The current bowel preparation protocol consists of four times 50 ml of oral iodine contrast (meglumine ioxithalamate, 300mgI ml-1) combined with a low fiber diet. This has not impaired the image quality of CT colonography examinations. Since iodine contrast has a mild laxative effect, the further reduction of the amount of iodine contrast is being investigated. This reduced bowel preparation scheme may further increase patient willingness to participate in a surveillance or screening program in the future.

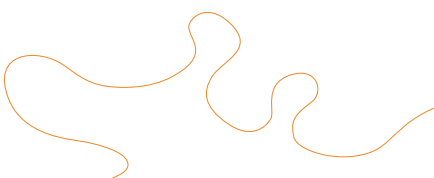
Likely, limiting the bowel preparation will result in more solid and less fluid fecal material in the colon and will also lead to a less homogeneous fecal tagging. This change in colonic content may hamper the reading of CT colonography. It may not only make new demands on the requirements of the reading skills of radiologists, but it will probably also result in a reduction of the efficacy of current electronic cleansing and CAD algorithms. Therefore future research should focus on optimizing bowel preparation schemes and the adjustment of electronic cleansing and CAD algorithms for these limited bowel preparation schemes.

Currently, most commercially available workstations are equipped with CAD software. The above-mentioned adjustments in bowel preparation will have consequences for the performance of CAD as well. The performance in terms of polyp detection and number of false-positive findings will depend on the extent of the bowel preparation *and* the presence and type of the administered oral contrast (barium or iodine). This is because polyps can be covered by fecal material (which results in fewer polyps to be detected) or fecal remains may simulate polyps (which results in more false positive findings).
Combining CAD with electronic cleansing would theoretically be able to increase CAD accuracy. However, as mentioned in **Chapter 6** electronic cleansing still generates artifacts. These artifacts interfere with an optimal use of CAD e.g. incompletely removed fecal material is detected by CAD. The number of detections done by CAD should be within reasonable limits. Therefore integrating these algorithms should be subject for future studies.

CT colonography has been recommended since 2008 in the United States as a population screening technique for colorectal cancer by the American College of Radiology, the American Cancer Society, and the U.S. Multisociety Task Force on Colorectal Cancer. The screening should be done every five years beginning at the age of fifty. Whether this will be recommended in the Netherlands remains unclear. However, if CT colonography will be part of a screening program in the Netherlands, mainly asymptomatic and healthy patients will participate. These individuals may then be scanned multiple times.
CT colonography is associated with a risk of inducing cancer and this risk is proportional to the amount of radiation. This radiation dose should therefore be kept as low as reasonably acceptable especially in screening-patients. In the past years a dose reduction in screening protocols was realized, resulting in a median effective dose of approximately 6 mSv [7]. This equals two to three times the background radiation dose. The risk of cancer induction is very small, especially when using even further optimized scan parameters. For example, for an individual aged 50-75 years, the risk of radiation-induced cancer is estimated to be in the order of 1 in 10,000 for a CT colonography procedure with 2mSv effective dose.
As shown in **Chapter 5** the pursuit for minimizing radiation dose may be impaired by the concurrent pursuit of reducing the patient burden of bowel preparation. To what extent the radiation dose can be further reduced in patients that have undergone a limited bowel preparation should be subject for further research.

In **Chapter 8** was reported that CT colonography techniques showed less variability between observers than colonoscopy techniques. Secondly, polyps measured by CT colonography techniques were measured larger than by colonoscopy techniques. As a consequence, in practice patients would more frequently be referred for colonoscopy according to the guidelines that were formulated based on colonoscopy measurements.
We have shown that the magnitude of the difference in size depends on the compared techniques (smallest in 2D abdominal window measurements and largest when using 2D lung window settings or 3D semi automatic measurements). However, the difference is also observer dependent since structural differences between observers

exist. Therefore it is not possible to determine a threshold that is valid for all observers. In order to minimize bias a personal calibration of polyp size measurements for both CT colonography readers and endoscopists can therefore be suggested.



References

1. Ristvedt SL, McFarland EG, Weinstock LB, Thyssen EP. Patient preferences for CT colonography, conventional colonoscopy, and bowel preparation. Am J Gastroenterol 2003; 98:578-585
2. van Gelder RE, Birnie E, Florie J, et al. CT colonography and colonoscopy: assessment of patient preference in a 5-week follow-up study. Radiology 2004; 233:328-337
3. McFarland EG, Levin B, Lieberman DA, et al. Revised colorectal screening guidelines: joint effort of the American Cancer Society, U.S. Multisociety Task Force on Colorectal Cancer, and American College of Radiology. Radiology 2008; 248:717-720
4. Kim DH, Pickhardt PJ, Taylor AJ, et al. CT colonography versus colonoscopy for the detection of advanced neoplasia. N Engl J Med 2007; 357:1403-1412
5. Johnson CD, Chen MH, Toledano AY, et al. Accuracy of CT colonography for detection of large adenomas and cancers. N Engl J Med 2008; 359:1207-1217
6. Kim DH, Pickhardt PJ, Taylor AJ, et al. CT colonography versus colonoscopy for the detection of advanced neoplasia. N Engl J Med 2007; 357:1403-1412
7. Liedenbaum MH, Venema HW, Stoker J. Radiation dose in CT colonography--trends in time and differences between daily practice and screening protocols. Eur Radiol 2008; 18:2222-2230

Chapter 10

Samenvatting en Algemene Discussie

Ayso H. de Vries

Samenvatting

Dit proefschrift beschrijft verschillende aspecten van CT colografie. Deze aspecten hebben gemeenschappelijk dat ze allen technische ontwikkelingen op het gebied van CT colografie evalueren.

Voordat patiënten een CT colografie onderzoek (of een coloscopie) ondergaan, vindt er een laxerende darmvoorbereiding plaats. Deze darmvoorbereiding heeft als doel de beoordeelbaarheid van het onderzoek te verhogen. Een *uitgebreide* laxerende darmvoorbereiding wordt vaak door patiënten omschreven als het meest belastende onderdeel van het CT colografie onderzoek [1, 2]. Het ligt dus voor de hand om te de patiënten belasting van de darmvoorbereiding zo veel mogelijk te beperken.
Wanneer de patiënten belasting van dit onderdeel zou worden terug gebracht, zou dit mogelijk meer personen over de streep trekken om deel te nemen aan een screenringonderzoek. Echter, de beoordeelbaarheid van het onderzoek zou kunnen afnemen omdat meer darmoppervlak is bedekt met ontlasting. Deze ontlasting zou afwijkingen aan het oog van de radioloog kunnen onttrekken.
In **hoofdstuk 2** hebben we 40 opeenvolgende patiënten gerandomiseerd voor vier verschillende *milde* darm voorbereidingen en deze vergeleken. De darmvoorbereidingen hadden een verschillend laxerend karakter.
Groep 1 kreeg 20 mg bisacodyl, groep 2 kreeg 30 mg bisacodyl, groep 3 ontving 20 mg bisacodyl (als in groep 1) èn 8,2 g magnesium citraat en groep 4 ontving 30 mg bisacodyl (als in groep 2) èn 16,4 g magnesium citraat.
De patiënten kregen daarnaast een tweedaags vezelarm dieet voorgeschreven en ontvingen oraal jodium en barium contrast om de ontlasting 'aan te kleuren' op de CT beelden. De groepen werden vergeleken op diagnostische kwaliteit van de CT beelden en patiëntenbelasting.

De resultaten lieten zien dat het beperken van de hoeveelheid laxerende middelen in de darmvoorbereiding (d.w.z. bisacodyl en magnesium citraat) geen negatief effect had op de diagnostische kwaliteit van het onderzoek. De diagnostische kwaliteit van de CT beelden van de groep die de minst belastende laxerende darmvoorbereiding had ondergaan (groep 1) was nog steeds als 'goed' of 'uitstekend' beoordeeld. In deze studie ondervonden patiënten met een hogere dosering laxeermiddelen significant meer last van de darmvoorbereiding.
Op basis van deze bevindingen hebben we geconcludeerd dat CT colografie met een beperkte darm voorbereiding en oraal contrast slechts een minimale dosis laxeermiddelen vereist. In onze studie was slechts 20 mg bisacodyl nodig voor het verkrijgen van goede diagnostische kwaliteit van de CT beelden en werd hiermee de patiënten belasting significant beperkt.
Aangezien de beeldkwaliteit van de meest patiënt-vriendelijke darm voorbereiding is beoordeeld als 'goed', is het waarschijnlijk dat de laxerende component van de voorbereiding zelfs verder kan worden terug geschroefd.

Het beoordelen van CT colografie onderzoeken kan op twee verschillende manieren plaatsvinden: het kan tweedimensionaal (2D) of driedimensionaal (3D). In de praktijk worden de technieken complementair toegepast. Bij een zogenaamde primaire 2D methode wordt bij het beoordelen van de dikke darm primair gebruik gemaakt wordt van 2D beelden. 3D beelden worden alleen gebruikt om een mogelijke afwijking te karakteriseren. Bij een primaire 3D methode is dit precies andersom.
Deze laatstgenoemde methode kan minder goed worden ingezet wanneer patiënten een beperkte darmvoorbereiding hebben ondergaan. Immers, een beperkte darmvoor-bereiding heeft een minder lege dikke darm tot gevolg, waardoor delen bedekt kunnen zijn met ontlasting. Deze ontlasting kan een een eventuele afwijking onttrekken aan het oog van de radioloog.
Een computer programma dat ontlasting 'virtueel' uit de

dikke darm verwijderd (zgn. electronic cleansing) zou dan uitkomst kunnen bieden.
In **Hoofdstuk 3** wordt het effect bepaald van 'electronic cleansing' op de zichtbaarheid van darmpoliepen, de beoordelingstijd, de inspanning van de arts en zijn/haar vertrouwen in de juistheid van de beoordeling in een *geselecteerde* patiëntenpopulatie. Deze haalbaarheidsstudie werd uitgevoerd met behulp van een primaire 3D evaluatiemethode.
Voor deze studie vergeleken twee ervaren beoordelaars CT onderzoeken vóór en na 'electronic cleansing'. Op basis van hun bevindingen konden we concluderen dat de zichtbaarheid van darmpoliepen die pas na 'electronic cleansing' zichtbaar werden (en dus onder ontlasting verborgen zaten) net zo groot was als darmpoliepen die in een leeg darmsegment zaten. Bovendien daalde met 'electronic cleansing' de evaluatie tijd, kostte het beoordelen van het onderzoek minder inspanning en had de arts meer vertrouwen in het feit dat de gehele dikke darm onderzocht was.
Op basis van deze bevindingen concludeerden wij dat 'electronic cleansing' de primaire 3D evaluatie methode verbetert.
In **Hoofdstuk 4** werd 'electronic cleansing' geëvalueerd in een *ongeselecteerde* patiëntenpopulatie. De populatie had een beperkte darm voorbereiding ondergaan en werd beoordeeld met twee verschillende beoordelingsstrategieën namelijk primair 2D zonder 'electronic cleansing' en primaire 3D met 'electronic cleansing'. Deze twee strategieën zijn de voor de hand liggende strategieën bij patiënten die een *beperkte* darmvoorbereiding hebben ondergaan. Deze twee methoden werden vergeleken ten aanzien van accuratesse, beoordelingstijd en diagnostische kwaliteit.
De resultaten toonden aan dat het opsporen van klinisch relevante poliepen door zes getrainde, maar relatief onervaren waarnemers significant hoger was wanneer gebruik gemaakt werd van 'electronic cleansing'. Ook zagen ervaren beoordelaars meer poliepen dan de onervaren beoordelaars. Er was overigens geen verschil tussen beide weergave methoden in deze waarnemers groep.
De specificiteit werd niet beïnvloed met het gebruik van de primaire 3D methode met 'electronic cleansing'.
Ondanks de betere prestaties ten aanzien van poliep detectie voor de relatief onervaren waarnemers, was de primaire 3D methode met 'electronic cleansing' gevoeliger voor artefacten dan de 2D methode. Dit resulteerde in een lagere beoordeling van de diagnostische kwaliteit van de primaire 3D methode. Ook kostte het beoordelen van de primaire 3D beelden de meeste beoordelaars meer tijd dan de 2D beelden.

Het gebruik van oraal contrast heeft het mogelijk gemaakt om de laxerende component van de darmvoorbereiding in CT colografie terug te brengen.
Echter een poliep omgeven door ontlasting met contrast materiaal is minder goed zichtbaar dan een poliep omgeven door lucht. Dit komt omdat het verschil in grijswaarden van de poliep ten opzichte van de omgeving tussen een poliep omgeven door lucht groter is dan een poliep omgeven door contrast materiaal. In **Hoofdstuk 5** werd de invloed bepaald van contrast materiaal en stralingsdosis op de zichtbaarheid van darm poliepen. Daarvoor werd een studie uitgevoerd waarbij gemaakt was van een op een dikke darm gelijkend fantoom. In deze studie bepaalden drie beoordelaars de zichtbaarheid van 6 millimeter grote poliepen. Deze poliepen waren omgeven door 'darminhoud' van verschillend röntgen dichtheden (grijswaarden). De poliepen werden gescand met verschillende stralingsdoses. De studie toonde aan dat een aanzienlijk hogere dosis straling nodig is om adequaat een poliep te visualiseren wanneer deze omgeven is door 'darminhoud' (met name wanneer deze een lage dichtheid heeft) dan wanneer een poliep omgeven is door de lucht.

Een nieuwe ontwikkeling in de radiologie is het gebruik van de computer bij het beoordelen van onderzoeken.

Deze 'computer aided detection' (CAD) algoritmes worden inmiddels bij ook in de CT colografie toegepast.
In **Hoofdstuk 6** is de toegevoegde waarde van een CAD algoritme bepaald. Veelbelovende resultaten van CAD ten aanzien van het verminderen van het aantal gemiste laesies voor *minder ervaren* waarnemers zijn gepubliceerd. Echter, de toegevoegde waarde van CAD voor *ervaren* waarnemers is nog steeds omstreden. Bovendien zijn de prestaties van CAD in eerdere studies slechts onderzocht in geselecteerde populaties met een hoge poliep prevalentie. Dit kan een positief effect hebben gehad op de prestaties van de waarnemers omdat mogelijke afwijkingen mogelijk sneller werden geduid als een poliep. Daarom is gekeken of CAD de prestaties van een ervaren waarnemer kan verbeteren. Dit is gedaan aan de hand van 170 opeenvolgend geïncludeerde patiënten met een verhoogde kans op dikke darmkanker.
In deze populatie had CAD een middelgrote poliep (6-9mm) en drie poliepen van 10mm of groter gedetecteerd die in eerste instantie waren gemist door de beoordelaar. Als gevolg van deze extra ontdekkingen steeg de sensitiviteit van CT colografie *zonder* CAD voor patiënten met poliepen groter dan 6mm van 80 naar 82% *met* CAD. Voor patiënten met poliepen van 10mm en groter steeg deze van respectievelijk 64 naar 72%. Deze kleine toename was statistisch niet significant. De specificiteit bleef door de toevoegingen van het CAD algoritme (bijna) ongewijzigd. Dus hoewel CT colografie *met* CAD enkele patiënten meer detecteerde dan CT colografie zonder CAD, had het geen significant positief resultaat op de prestaties van ervaren waarnemers in een populatie met een verhoogde kans op dikke darmkanker.
Een beperking van deze studie was het relatief hoge aantal vlakke poliepen dat moeilijk detecteerbaar blijkt. Deze beperking zal worden besproken in de sectie '*Algemene discussie*' van dit hoofdstuk.

Op dit moment is het lege artis om patiënten in zowel rug als buikligging te scannen. Dit wordt gedaan om de nauwkeurigheid van de poliep detectie te verhogen. Het combineren van beelden in rug- en buikligging kan de beoordelaar helpen poliepen te onderscheiden van ontlasting, omdat ontlasting, in tegenstelling tot poliepen, onderhevig is aan de zwaartekracht.
Het handmatig verifiëren van bevindingen in rug- en buikligging is een tijdrovende bezigheid omdat referentie punten zoals de flexura hepatica niet vastliggen maar mobile referentie punten zijn. Een geautomatiseerde matching methode kan dit proces mogelijk vergemakkelijken.
In **Hoofdstuk 7** wordt de haalbaarheid van automatische matching van rug- en buikligging geëvalueerd aan de hand van zevenendertig coloscopisch bewezen poliepen als vast referentiepunt. Voor deze studie gebruikten we een algoritme gebaseerd op de gelijkenissen ten aanzien van de richting van darmsegmenten. De met een primaire 3D-weergave techniek verrichte studie toonde aan dat de gematchte poliepen ongeveer 14 millimeter (bereik 0 – 59 mm) van de uitgerekende locatie lagen. Het grootste deel van de poliepen (70%) werd daardoor direct zichtbaar in buikligging na matching.

CT colografie is een onderzoeksmethode die de dikke darm visualiseert. Het biedt echter niet de mogelijkheid, in tegenstelling tot coloscopie, om een weefsel biopsie te nemen.
Volgens de Amerikaanse screeningsrichtlijnen [3] moeten alle patiënten met een grote poliep (10mm of groter) of middelgrote poliep (6-9 mm) verwezen worden voor colonoscopie. Echter, of deze middelgrote poliepen een indicatie zijn voor colonoscopie wordt nog hevig bediscussieerd. Het met behulp van CT colografie monitoren van de groei van deze poliepen is voorgesteld als een veilig alternatief [4]. Kleine (< 6mm) poliepen hoeven niet verwijderd te worden vanwege een verwaarloosbaar risico op kwaadaardige ontaarding. Uit deze richtlijnen kan men destilleren dat de grootte van de poliep van cruciaal belang voor de besluitvorming in CT colografie.

Het doel van de studie in **Hoofdstuk 8** was om de variabiliteit metingen van CT colografie en coloscopie te bepalen. Daarnaast was het doel structurele verschillen tussen de verschillende technieken te meten. De respectievelijke metingen werden verricht door drie ervaren CT colografie beoordelaars en drie ervaren maag-, darm- en leverartsen.
Op basis van de resultaten van de studie concludeerden we dat;

1 het gebruik van een liniaal tijdens de coloscopie de precisie van de meting niet verhoogd vergeleken met het gebruik van een biopsie forceps.
2 wanneer gebruik wordt gemaakt met een liniaal de meting significant kleiner uitvalt dan wanneer deze met het oog wordt uitgevoerd of met een biopsie forceps.
3 metingen tussen Radioloogen dichter bij elkaar liggen dan metingen van maag-, darm- en leverartsen.
4 Radioloogen in het algemeen poliepen groter meten dan maag-, darm- en leverartsen.

Algemene Discussie

Dit proefschrift toont aan dat:

- De diagnostische kwaliteit CT colografie bij patiënten die een voorbereidingsschema bestaande uit een vezelarm dieet, oraal jodium- en bariumcontrast en 20 mg bisacodyl niet significant slechter is dan van voorbereidingsschema's die de patiënt meer laxeren (en belasten).
- Voor relatief onervaren CT colografie beoordelaars 'electronic cleansing' (het virtueel verwijderen van ontlasting vermengd met contrast materiaal) een handig instrument is bij het beoordelen van patiënten die een beperkte darmvoorbereiding hebben ondergaan. Voor ervaren beoordelaars heeft het echter geen toegevoegd waarde.
- Men voorzichtig moet zijn met het reduceren van de stralingsdosis in een patiëntenpopulatie die een beperkte darmvoorbereiding heeft ondergaan. Het verminderde contrast tussen de darmwand en de darminhoud leidt bij lage dosis tot een verminderde zichtbaarheid van poliepen.
- Een computer aided detection (CAD) algoritme als back-up reader geen significante verbetering van de prestaties van een ervaren CT colografie beoordelaar geeft.
- Het automatisch correleren van identieke locaties in rug- en buikligging mogelijk is.
- CT colografie beoordelaars darmpoliepen groter en met een grotere precisie meten dan maag-, darm- en leverartsen.

In grote studies in screeningspopulaties is wetenschappelijk aangetoond dat CT colografie een goede methode is voor het detecteren dikke darmkanker en voorlopers hiervan [5, 6]. Een screeningspopulatie is een populatie met een normale kans op het krijgen van dikke darm kanker. Deze studies laten een sensitiviteit van 90% voor de detectie van poliepen van 10 mm of groter zien [5] of een sensitiviteit die overeenkomt met die van coloscopie [6]. Dit laatste onderzoek wordt tot op heden beschouwd als de gouden standaard.
Echter, in **Hoofdstuk 6** rapporteerden wij op basis van een vergelijkende studie van CT colografie en coloscopie een sensitiviteit van 72% voor deze grote poliepen. De detectie van deze poliepen lag dus bijna 20% lager dan in de bovengenoemde studies.
Een belangrijk verschil tussen onze studie en de bovengenoemde studies is de patiënten populatie. Onze studie werd uitgevoerd in een populatie met een verhoogde kans op dikke darmkanker. De patiënten in deze populatie hadden namelijk een persoonlijke voorgeschiedenis van dikke darmkanker of -poliepen of een familielid met een belaste voorgeschiedenis. Hierdoor verschilde onze populatie van een screeningpopulatie op twee punten namelijk de prevalentie van poliepen (die in onze studie hoger was) en het percentage vlakke poliepen (dat in onze studie ook hoger was).

Veel van deze vlakke poliepen werden niet gedetecteerd door zowel de CT colografie beoordelaars als het CAD algoritme. Een mogelijke verklaring hiervoor ligt in het feit dat detectie van poliepen bij CT colografie vooral gebaseerd is op de morfologie van een poliep. Eigenschappen als veranderingen in vaatstructuur en slijmvliesstructuur ter plaatse van de poliep worden echter, in tegenstelling tot coloscopie, niet gezien.
De relatief hoge prevalentie van vlakke poliepen kan te maken hebben met het feit dat veel patiënten uit de onderzochte populatie reeds eerder een coloscopie hadden ondergaan. Eerder uitgevoerd coloscopisch onderzoek kan er toe hebben geleid dat makkelijk-te-detecteren poliepen al verwijderd zijn. Hierdoor blijven relatief moeilijk-te-detecteren poliepen achter in de darm. Bovendien kunnen de vorige coloscopisch verrichtte poliepverwijderingen niet volledig zijn geweest. De resten van deze poliepen zouden vlakker kunnen groeien dan de poliepen die initieel niet geheel verwijderd zijn. Hierdoor zouden moeilijk detecteerbare poliepen kunnen zijn ontstaan.
Verbetering van de detectie van deze vlakke laesies door opleiding, verbeterde CT data-acquisitie en hiervoor speciaal ontwikkelde (CAD) algoritmen zal slechts een beperkte toegevoegde waarde hebben in de verbetering van poliep opsporing in deze populatie. Dit komt omdat een groot deel van de gemiste laesies zelfs in retrospectie niet kan worden teruggevonden. Sommige onderzoekers zijn van mening dat voorzichtigheid geboden is bij het controleren van deze patiënten met CT colografie. Deze stelling wordt onderschreven door onze resultaten.

Om zo veel mogelijk patiënten deel te laten nemen aan een surveillance of screeningsprogramma voor dikke darmkanker, moet de patiënten belasting worden teruggebracht. Zoals aangegeven in **Hoofdstuk 2,** kan dit worden gedaan door het verminderen van het laxerende deel van de darmvoorbereiding die diarree en buikpijn veroorzaakt.

Deze studie toonde aan dat het darmvoorbereiding schema met de minste bijwerkingen geen negatief effect had op de subjectieve beoordeelbaarheid van het onderzoek.
Op het moment dat dit proefschift is geschreven, worden de patiënten die in het Academisch Medisch Centrum een CT colografie onderzoek ondergaan niet meer voorbereid met laxerende middelen. Het huidige protocol van darmvoorbereiding bestaat uit een dosis van vier keer 50 ml oraal jodium contrast (meglumine ioxithalamate, 300mgI ml -1) in combinatie met een vezelarm dieet. Ook deze voorbereiding lijkt een goede beoordeelbaarheid van het onderzoek te geven. Omdat jodium contrast een milde laxerende werking heeft, wordt onderzocht of een verdere reductie van de dosis jodium contrast mogelijk is.
Deze ontwikkelingen kunnen een de bereidheid van de patiënt om deel te nemen aan een screenings- of surveillanceprogramma in de toekomst stimuleren.
Het verminderen van de patiënten belasting van de darmvoorbereiding zal waarschijnlijk resulteren in meer vaste en minder vloeibaar ontlasting. Daarnaast zal de ontlasting minder goed met contrast gemengd worden en dus een minder homogeen aspect hebben. Deze verandering ten aanzien van de darminhoud zou het beoordelen van de CT colografie beelden kunnen belemmeren. Het zal nieuwe eisen stellen aan de vaardigheden van CT colografie beoordelaars, bovendien zal het ook leiden tot een vermindering van de werkzaamheid van de ‘electronic cleansing’ en CAD algoritmes. Daarom moet toekomstig onderzoek zich naast het optimaliseren van de darmvoorbereiding ook richten op het aanpassen van deze algoritmen.

Momenteel zijn de meeste commercieel verkrijgbare CT colografie werkstations uitgerust met CAD software. De bovengenoemde aanpassingen ten aanzien van de darm voorbereiding zal ook gevolgen hebben voor de prestaties van CAD. De prestaties van het algoritme om enerzijds darmpoliepen te detecteren en anderzijds het aantal vals

positieven te beperken (onterecht aan de CT colografie beoordelaar getoonde 'poliepen'), is mede afhankelijk van het soort darmvoorbereiding. Dit komt enerzijds omdat met ontlasting bedekte poliepen niet te gedetecteerd kunnen worden, en anderzijds omdat ontlasting door het algoritme soms niet onderscheiden kan worden van darmpoliepen (wat resulteert in meer vals positieve bevindingen).
Het combineren van CAD met 'electronic cleansing' zou de prestaties van CAD theoretisch kunnen verhogen. Echter, zoals genoemd in **Hoofdstuk 6** genereert het algoritme artefacten die interfereren met CAD. Zo herkent het CAD soms onvolledig verwijderde ontlasting als poliep. Om het beoordelen van het CT colografie onderzoek werkbaar te houden, moet het aantal suggesties dat door CAD gedaan wordt binnen redelijke grenzen blijven. Daarom zou de integratie van deze algoritmes onderwerp moeten zijn voor toekomstige studies.

Sinds 2008 bevelen de American College of Radiology, de American Cancer Society, en de US Multisociety Task Force on Colorectal Cancer CT colografie in de Verenigde Staten aan als screeningsmethode voor dikke darmkanker. Deze instituten adviseren om vanaf een leeftijd van vijftig jaar elke vijf jaar te screenen. Of dit ook in Nederland gaat gebeuren is vooralsnog onduidelijk. Echter, als CT colografie onderdeel gaat uitmaken van een screeningsprogramma zullen asymptomatische en met name gezonde personen deelnemen aan dit programma. Deze personen zullen dan waarschijnlijk meermalen worden gescand.
CT colografie gaat mogelijk gepaard met een risico van kanker inductie en aangenomen wordt dat dit risico evenredig is met de hoeveelheid straling waaraan de persoon wordt blootgesteld. De stralingsdosis moet daarom in een screeningsprogramma zo laag mogelijk worden gehouden, zeker bij gezonde personen. In de afgelopen jaren was deze effectieve dosis in screeningsprotocollen ongeveer 6 mSv [7]. Dit is gelijk aan twee tot drie maal de achtergrond straling. Het risico van kanker inductie is dus zeer klein, vooral bij het gebruik nog verder geoptimaliseerd scan parameters. Het risico met een dergelijke stralingsdosis voor een persoon in de leeftijd van 50-75 jaar geschat in de orde van 3 op 10.000 voor een CT colografie.
In **Hoofdstuk 5** hebben we laten zien dat het nastreven van een zo laag mogelijke stralingsdosis van het onderzoek het beperken van de darmvoorbereiding in de weg kan staan. In hoeverre de stralingsdosis verder gereduceerd kan worden in patiënten met een beperkte darmvoorbereiding is stof voor verder onderzoek.

De studie in **Hoofdstuk 8** toont aan dat metingen uitgevoerd met CT colografie minder variabiliteit tussen waarnemers laten zien dan metingen met colonoscopie. Ook worden poliepen met CT colografie groter gemeten dan met colonoscopie. Als gevolg daarvan, zouden patiënten in de praktijk vaker worden verwezen voor colonoscopie omdat voor een eventuele verwijzing de grootte van de poliep het uitgangpunt is. Deze drempelwaarden werden overigens op basis van coloscopische metingen bepaald.
Echter, het verschil is ook persoonsafhankelijk waarbij er structurele verschillen bestaan tussen waarnemers. Daarom is het helaas niet mogelijk een verwijsdrempel te definiëren die voor alle radiologen van toepassing is. Om de bias te minimaliseren zou een persoonlijke kalibratie voor poliep grootte voor zowel radiologen als maag-, darm- en leverartsen een oplossing zijn.

Referenties

1. Ristvedt SL, McFarland EG, Weinstock LB, Thyssen EP. Patient preferences for CT colografie, conventional colonoscopy, and bowel preparation. Am J Gastroenterol 2003; 98:578-585
2. van Gelder RE, Birnie E, Florie J, et al. CT colografie and colonoscopy: assessment of patient preference in a 5-week follow-up study. Radiology 2004; 233:328-337
3. McFarland EG, Levin B, Lieberman DA, et al. Revised colorectal screening guidelines: joint effort of the American Cancer Society, U.S. Multisociety Task Force on Colorectal Cancer, and American College of Radiology. Radiology 2008; 248:717-720
4. Kim DH, Pickhardt PJ, Taylor AJ, et al. CT colografie versus colonoscopy for the detection of advanced neoplasia. N Engl J Med 2007; 357:1403-1412
5. Johnson CD, Chen MH, Toledano AY, et al. Accuracy of CT colografie for detection of large adenomas and cancers. N Engl J Med 2008; 359:1207-1217
6. Kim DH, Pickhardt PJ, Taylor AJ, et al. CT colografie versus colonoscopy for the detection of advanced neoplasia. N Engl J Med 2007; 357:1403-1412
7. Liedenbaum MH, Venema HW, Stoker J. Radiation dose in CT colografie--trends in time and differences between daily practice and screening protocols. Eur Radiol 2008; 18:2222-2230

List of Publications

Liedenbaum MH, **de Vries AH**, Gouw CI, van Rijn AF, Bipat S, Dekker E, Stoker J. CT colonography with minimal bowel preparation: evaluation of tagging quality, patient acceptance and diagnostic accuracy in two iodine-based preparation schemes. Eur Radiol 2009 August 26.

Jensch S, Bipat S, Peringa J, **de Vries AH**, Heutinck A, Dekker E, Baak LC, Montauban van Swijndregt AD, Stoker J. CT colonography with limited bowel preparation: prospective assessment of patient experience and preference in comparison to optical colonoscopy with cathartic bowel preparation. Eur Radiol 2009 July 23.

Liedenbaum MH, van Rijn AF, **de Vries AH**, Dekker HM, Thomeer M, van Marrewijk CJ, Hol L, Dijkgraaf MG, Fockens P, Bossuyt PM, Dekker E, Stoker J. Using CT colonography as a triage technique after a positive faecal occult blood test in colorectal cancer screening. Gut 2009 September;58(9):1242-9.

de Vries AH, Liedenbaum MH, Bipat S, Truyen R, Serlie IW, Cohen RH, van Elderen SG, Heutinck A, Kesselring O, de MW, te SL, Wiersma T, Stoker J. Primary uncleansed 2D versus primary electronically cleansed 3D in limited bowel preparation CT-colonography. Is there a difference for novices and experienced readers? Eur Radiol 2009 August;19(8):1939-50.

Liedenbaum MH, **de Vries AH**, Halligan S, Bossuyt PM, Dachman AH, Dekker E, Florie J, Gryspeerdt SS, Jensch S, Johnson CD, Laghi A, Taylor SA, Stoker J. CT colonography polyp matching: differences between experienced readers. Eur Radiol 2009 July;19(7):1723-30.

de Vries AH, Jensch S, Liedenbaum MH, Florie J, Nio CY, Truyen R, Bipat S, Dekker E, Fockens P, Baak LC, Stoker J. Does a computer-aided detection algorithm in a second read paradigm enhance the performance of experienced computed tomography colonography readers in a population of increased risk? Eur Radiol 2009 April;19(4):941-50.

Serlie IW, **de Vries AH**, van Vliet LJ, Nio CY, Truyen R, Stoker J, Vos FM. Lesion conspicuity and efficiency of CT colonography with electronic cleansing based on a three-material transition model. AJR Am J Roentgenol 2008 November;191(5):1493-502.

de Vries AH, Venema HW, Florie J, Nio CY, Stoker J. Influence of tagged fecal material on detectability of colorectal polyps at CT: phantom study. AJR Am J Roentgenol 2008 October;191(4):1101.

Jensch S, **de Vries AH**, Pot D, Peringa J, Bipat S, Florie J, van Gelder RE, Stoker J. Image quality and patient acceptance of four regimens with different amounts of mild laxatives for CT colonography. AJR Am J Roentgenol 2008 July;191(1):158-67.

van Wijk C, Florie J, Nio CY, Dekker E, **de Vries AH**, Venema HW, van Vliet LJ, Stoker J, Vos FM. Protrusion method for automated estimation of polyp size on CT colonography. AJR Am J Roentgenol 2008 May;190(5):1279-85.

Jensch S, **de Vries AH**, Peringa J, Bipat S, Dekker E, Baak LC, Bartelsman JF, Heutinck A, Montauban van Swijndregt AD, Stoker J. CT colonography with limited bowel preparation: performance characteristics in an increased-risk population. Radiology 2008 April;247(1):122-32.

van Wijk C, van R, V, Vos FM, Truyen R, **de Vries AH**, Stoker J, van Vliet LJ. Detection of protrusions in curved folded surfaces applied to automated polyp detection in CT colonography. Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv 2006;9(Pt 2):471-8.

Nio CY, **de Vries AH**, Stoker J. Perceptive errors in CT colonography. Abdom Imaging 2007 September;32(5):556-70.

de Vries AH, Truyen R, van der PJ, Florie J, van Gelder RE, Gerritsen F, Stoker J. Feasibility of automated matching of supine and prone CT-colonography examinations. Br J Radiol 2006 September;79(945):740-4.

Other publications

de Vries AH, van Gelder RE, Stoker J. 3D Imaging: Invaluable for the Correct Diagnosis? In: Lefere P., Grijspeerdt S. editor. Virtual Colonoscopy, a Practical Guide 1st and 2nd ed. Springer Berlin Heidelberg: 2005 and 2010 ISBN 978-3-540-22865-3/ 978-3-540-79879-8

de Vries AH, Vos FM, Serlie IWO, Stoker J. Unfolded Cube Projection of the Colon. In: Neri E, Caramella D, Bartolozzi C, editor. Image Processing in Radiology. 1st ed. Springer Berlin Heidelberg: 2008 ISBN 978-3-540-25915-2

Dankwoord

Voor de totstandkoming van dit proefschrift ben ik veel mensen dank verschuldigd. Zonder hun medewerking was dit proefschrift niet tot stand gekomen. In het bijzonder wil ik de patiënten bedanken die aan de verschillende studies hebben meegewerkt. Zij zijn het fundament waarop dit proefschrift is gebouwd. Een aantal personen zou ik graag in het bijzonder willen bedanken:

Allereerst mijn promotor **prof. dr. J.Stoker:**
Beste Jaap, ik kan me nog goed herinneren hoe ik als geneeskundestudent op de afdeling Radiologie onder jouw vleugels begon aan mijn wetenschappelijke stage. Zonder jouw gedrevenheid, enthousiasme en snelle feedback in die tijd, had ik de stap naar het artsonderzoekerschap niet gemaakt. Ik ben nu, zeker achteraf, blij dat ik de stap toch gemaakt hebt. En zie hier! Bedankt voor je fijne begeleiding.

Mijn co-promotoren **dr. H.W. Venema en dr.E. Dekker:**
Beste Henk, met enige weemoed denk ik terug aan de tijd van scannen in Utecht. De vaardigheden waarmee jij de belletjes uit ons colon fantoom hebt gekregen, is op zich al een artikel waard. Dankzij jouw bijna compromisloze vasthoudendheid hebben we mijns inziens een mooi artikel neergezet. Maar wat ik misschien nog wel meer waardeer is je avontuurlijkheid. Ik ken niet veel 60-plussers die in hun eentje naar Madrid fietsen. Het is echt een voorrecht om jouw promovendus te mogen zijn.
Lieve Evelien, ook jouw steun in de laatste fase van deze marathon heb ik erg gewaardeerd. Je snelle feedback heeft ervoor gezorgd dat de periode waarin ik de opleiding en de afrondingsfase van het proefschrift tot een minimum kon beperken. Daarnaast was jij voor mij het aanspreekpunt op de afdeling maag-, darm- en leverziekten. Nooit had ik het gevoel dat ik stoorde als ik je belde en je had altijd tijd. Hartelijk dank daarvoor.

Mijn paranimfen **Sebastiaan Jensch** en **Martijn Kruijt Spanjer:**
Beste Sebastiaan, Onze tijd samen als student en later als onderzoeker in het AMC was mij een waar genoegen. Niet alleen de samenwerking op de afdeling, maar zeker ook de uitstapjes naar o.a. Boston en Chicago waren een feestje. Dat we nu samen gaan promoveren is voor mij een fantastische afsluiting van deze periode. Ik ben blij dat ik dat met jou mag doen. Samen straks weledelzeergeleerd!
Beste Martijn, amice! We hebben de afgelopen 15 jaar samen een aantal fantastische avonturen beleefd. Of het nou wild kamperen in Toscane was, een clash met de autoriteiten in Rusland, een defecte 'bomba de agua' in Spanje, een bijna gestolen fiets in Nijmegen, een paspoortcontrole in Marokko, een weekend op Terschelling, een potje mokken onder de Eiffeltoren of gewoon een biertje in Groningen, ik had ze niet willen missen. Het is daarom is het volstrekt logisch dat jij nu ook mijn paranimf bent om ook dit avontuur succesvol af te ronden. Duff, in 2013 the Long Way Down, ik kan niet wachten!

Mijn commissieleden:
Graag wil ik de commissie leden Prof. dr. P. Fockens, Prof. dr. W.M. Prokop, Prof. dr. J.S. Laméris, Prof. dr. L.J. van Vliet en Dr. F.M. Vos bedanken voor het beoordelen van dit proefschrift en het plaatsnemen in de promotie commissie. Daarnaast wil ik dr. K. Tytgat bedanken voor het opponeren.

'Mijn' research nurse **Anneke Heutinck:**
Lieve Anneke, jouw gestructureerde manier van werken, plichtsbesef en trouw hebben het werven van patiënten, de dataverwerking en de analyses bijna tot een feestje gemaakt. Ik ken weinig mensen die zo veel in hun mars hebben en toch zo bescheiden zijn.

Mijn **medeonderzoekers** met wie ik onderzoekslief en -leed heb gedeeld:
Beste Rogier, van jou heb ik het stokje overgenomen. Hoewel onze overlap als onderzoekers erg kort was, bood jouw authentieke en duidelijke mening in veel zaken een prima houvast om te beginnen. Dank hiervoor.
Beste Jasper, ik ken weinig mensen die hun privé leven zo goed kunnen combineren met hun werkende bestaan. Met plezier heb ik met je samen gewerkt en ik hoop dat we dat in de toekomst blijven doen.
Lieve Karin, veel gezucht en gesteun kwam er vaak uit de kamer tegenover me, maar ondertussen had jij je zaakjes prima voor elkaar. Geen pretenties, goedlachs, hard werken maar wel weten wat ècht belangrijk is. Je bent een fantastische collega.
Lieve Adrienne, dank voor het uitblazen op je kamer als het even tegen zat. Jij moet nog even, maar ik heb er alle vertrouwen in dat je je onderzoekscarrière succesvol afsluit. Zet hem op!
Lieve Shandra, bedankt voor je begeleiding in de statistiek en de hulp bij het afronden van dit boekje.
Lieve Marjolein, de samenwerking met jou heb ik als heel prettig ervaren. Jouw precisie van werken, gedrevenheid, slimheid en charme maakten de samenwerking buitengewoon prettig.

Natuurlijk wil ik de ook de *Powerpuff Girls* (Nicole, Maaike en Maartje), Annette, Roos, Sanna, Sandra, Manon, Wouter, Jochem, Joppe, Christiaan en Frank bedanken voor de prettige sfeer, de tripjes naar de koffie automaat en de gezellige lunches.

Philips Healthcare:
Roel Truyen, Iwo Serlie en Frans Gerritsen, bedankt voor de constructieve bijeenkomsten, prettige samenwerking, snelle service bij problemen en de etentjes op rekening van 'Uncle Phil'.

Laboranten:
Ik wil graag alle laboranten bedanken die mijn patiënten hebben gescand. Martin Poulus wil ik graag in het bijzonder bedanken voor het bijbrengen van de fijne kneepje van het scannen en de prettige samenwerking.

Radiologen (i.o):
Graag wil ik Rutger Cohen, Saskia van Elderen, Oskar Kesselring, Wouter de Monyé, Lambertus te Strake, Tjeerd Wiersma, Yung Nio en Anje Spijkerboer bedanken voor het beoordelen van de vele CT onderzoeken.

De afdeling IT:
Natuurlijk kan de IT in dit lijstje niet ontbreken. Jean-Paul, Martin, Jan en Onno. Dank voor de snelle service als er weer eens een software probleempje was.

Mijn familie:
Dat de bevalling van een dergelijk boekwerk ook niet aan mijn directe omgeving ongemerkt voorbij is gegaan moge duidelijk zijn, daarom wil ik mijn familie, schoonfamilie en goede vrienden bedanken. Zonder jullie interesse, steun en vertrouwen was dit grote project niet mogelijk geweest.

Last but not least, lieve Aukje, je bent de liefde van mijn leven en mijn allerliefste, bedankt voor de steun die jij al vele jaren voor mij bent. Zeker in de laatste fase van mijn promotie is jouw bijdrage van onschatbare waarde geweest. Met jou is geen berg te hoog en zee te diep. Ik hou van jou!

Curriculum Vitae

Ayso Harmen de Vries werd op 10 maart 1976 geboren in de gemeente Noordoostpolder. Zijn ouders verhuisden al snel naar Nijmegen waar hij op de Nutsschool en later op het Canisius College Mater Dei zijn eerste carrièrestappen zette.

Na zijn eindexamen vertrok hij in 1994 naar het bruisende Groningen om daar de studie Geneeskunde te beginnen. In 1997 heeft hij deze studie een jaar onderbroken om eerst een paar maanden als taxichauffeur geld te verdienen en daarna te genieten van een zeven maanden lange reis door Australië en Nieuw Zeeland. Na dit intermezzo werd de studie werd weer voortgezet. Na de milleniumwisseling deed hij zijn coschappen in Enschede, waarna een keuze coschap urologie aan de voet van de Kilimanjaro volgde. Tenslotte deed hij zijn wetenschappelijke stage op de afdeling Radiologie in het Academisch Medisch Centrum te Amsterdam waarbij de haalbaarheid van dosis reductie bij CT urinewegstenen werd bepaald. Dit was onder leiding van dr. J. Stoker. Na het afronden van deze stage ontving hij in september 2002 zijn artsenbul.

Na ruim een jaar als AGNIO urologie gewerkt te hebben begon hij in maart 2004 als arts-onderzoeker op de plek waar hij eerder als student gezeten had. In de jaren die daarop volgden heeft hij onder begeleiding van prof. dr. J.Stoker klinisch onderzoek gedaan naar technische ontwikkelingen op het gebied van CT colografie. Van een deel het werk waaraan hij toen begon heeft u nu het resultaat in handen. In december 2008 is hij in AMC begonnen aan de opleiding radiologie onder leiding van dr. O.M. van Delden en prof. dr. J.S. Lamerís.

Ayso woont samen met Aukje van Wissen.