GASTROINTESTINAL

Jamshid Dehmeshki
Steve Halligan
Stuart A. Taylor
Mary E. Roddie
Justine McQuillan
Lesley Honeyfield
Hamdan Amin

Computer assisted detection software for CT colonography: effect of sphericity filter on performance characteristics for patients with and without fecal tagging

J. Dehmeshki · J. McQuillan ·
L. Honeyfield · H. Amin
Medicsight PLC,
Berkeley Sq,
London, UK

S. Halligan (✉) · S. A. Taylor
Department of Specialist Radiology,
Level 2 Podium,
University College Hospital,
235 Euston Road,
London, NW1 2BU, UK
e-mail: s.halligan@ucl.ac.uk
Tel.: +44-20-73809010
Fax: +44-20-73809068

M. E. Roddie
Charing Cross Hospital,
Fulham Palace Road,
London, UK

Abstract The aim of this study is to investigate the effect of changing sphericity filter values on performance of a computer assisted detection (CAD) system for CT colonography for data with and without fecal tagging. Colonography data from 138 patients with 317 validated polyps were divided into those with (86) and without (52) fecal tagging. Polyp coordinates were established by three observers and datasets analysed subsequently by a proprietary CAD system used at four discrete sphericity filter settings. Prompts were compared with the known coordinates in order to determine sensitivity and specificity. Sensitivity was highest at low sphericity; of 164 polyps 6 mm or more, 144 (87.8%) were detected at sphericity 0.3, and 132 (80.1%) at sphericity 0.9. Of 42 polyps measuring 10 mm or more, 40 (95.2%) were detected at sphericity 0.3, and 36 (85.7%) at sphericity 0.9. There was no significant difference in sensitivity for tagged and un-tagged data but specificity was reduced in tagged data at low sphericity and significantly reduced in untagged data at high sphericity. CAD had a sensitivity of 95.2% for polyps measuring 1 cm or more and 87.8% for polyps 6 mm or more when used at a sphericity setting of 0.3. Higher sphericity settings increased specificity while reducing sensitivity. The bowel preparation used significantly impacts on specificity.

Keywords Colon (CT) · Colonography (CT) · Colon cancer · Computer-assisted detection

Introduction

CT colonography is an increasingly utilized technique for detection of colorectal neoplasia in both symptomatic and screening contexts [1–3]. There is considerable evidence that interpretative accuracy is correlated positively with previous experience of CT colonography [4–6] although, understandably, there are many potential users who are relatively inexperienced but still feel the technique is relevant to their clinical practice. It has been suggested that computer aided diagnosis (CAD) could potentially allow less experienced practitioners to read CT colonography because a proportion of the interpretation skills necessary to identify polyps and cancers is abrogated to the software [7]. Experienced readers might also benefit from computer-

assistance since interpretation of CT colonography is time-consuming, repetitive and fatiguing, irrespective of the abilities of the reader [7, 8].

CAD software has recently gained regulatory approval for CT colonography and some systems differ from the typical CAD paradigm in that their performance characteristics can be modified by user adjustable 'polyp enhancement filters' [9]. The fact that these filters can be adjusted by the reader to suit the individual case in question is claimed to be beneficial. For example, sensitivity for small polyps may be reduced if the clinical context requires diagnosis of large lesions only (e.g. cancer). However, while the sensitivity of such CAD systems is determined by the filter settings, it is possible that recommending a single setting to cover particular clinical situations could be

misleading because there are different categories of CT colonography data. For example, investigators may use different collimations and bowel preparation. Most notably, fecal tagging has not gained universal acceptance.

The aim of our study was to investigate the effect of changing the sphericity filter settings of a CAD system on the sensitivity and specificity of colorectal polyp detection for CT colonography, and to investigate the effect of using different categories of CT data on these performance characteristics.

Materials and methods

Data accumulation

CT colonography datasets from 138 adult patients known to have colorectal polyps were retrieved from a central research database composed of cases accumulated by four individual centers. Each of these centers had an on-going research program investigating the diagnostic performance of CT colonography via comparison with same-day optical colonoscopy and ethical permission to contribute patient data to a central database for future research was approved by the ethical review board at each individual site.

Each individual subject had undergone CT colonography followed by same-day optical colonoscopy as the reference test. Multi-detector row CT colonography was performed prior to colonoscopy in all patients according to prerequisites for good current practice established at the 5th Boston Conference on Virtual Colonoscopy [10]. In particular, full bowel purgation and prone and supine studies were acquired in all patients. There were 86 patients who had been prepared using fecal tagging via a combination of barium suspension and water-soluble contrast (Gastrografin, Schering AG, Burgess Hill, West Sussex, UK); their median age was 58 years (range 46 to 75 years). Their data was obtained at a collimation of 2.5 mm, without an intravenous muscle relaxant and using room air for colonic insufflation. These data were obtained from a single center. There were 52 patients from the remaining three centers who had been prepared without fecal tagging, median age 69 years (range 38 to 89 years). Their data was obtained at collimation ranging between 1.25 mm and 3.0 mm: 20 at 1.25 mm, 23 at 2.5 mm, 9 at 3.0 mm (reflecting differing practice at the three centers contributing this data, and the individual clinical circumstances of the patient in question). 20 mg of hyoscine butyl bromide had been administered intravenously as a muscle relaxant in this group and carbon dioxide was used for insufflation, again reflecting routine clinical practice at contributing centers.

Polyps identified at colonoscopy were measured in situ (in order to establish the reference size of each individual polyp) by comparison with adjacent open biopsy forceps. The segmental location of any polyp detected during colonoscopy was also recorded as follows: cecum, ascending colon, hepatic flexure, transverse colon, splenic flexure, descending colon, sigmoid, rectum. Polypectomy was then performed and excised lesions sent for histopathological analysis. In this way, data relating to the size, location, and nature of each polyp encountered at colonoscopy was known for each patient contributing to the accumulated database.

Establishing the reference standard

The CT datasets were transferred to a personal computer loaded with proprietary colonography software (Colon CAR 1.1, Medicsight PLC, London, UK). Three observers experienced in CT colonography interpretation (median of 300 endoscopically validated cases) read each of the 138 datasets in consensus and with full knowledge of the reference colonoscopy findings for each individual case. They attempted to identify the CT coordinates of all endoscopically proven polyps in each individual case using a combination of 2D primary read with 3D for problem solving [11]. Where the reference colonoscopy indicated that more than one polyp was present in a segment, comparison of the CT diameter of the polyp with the reference diameter was used to determine the best match. Disagreements were resolved by face-to-face discussion between the observers. In addition to establishing the location of all polyps found by colonoscopy, each reader also carefully interrogated each individual study twice in order to search for any colonoscopic false negative polyps. Such polyps were only deemed present if they clearly fulfilled well-established diagnostic criteria (i.e. well-defined, homogenous, present on both prone and supine datasets), and readers also had high diagnostic confidence. Any such polyp was again discussed in consensus. The CT image co-ordinates for each polyp identified were noted for both the prone and supine acquisitions or for a single acquisition if the polyp was not visualized on the other (for example, because it was submerged by retained fluid). Finally, one observer outlined the circumferential boundary of each polyp identified on the axial slice that best demonstrated its maximal diameter, using a mouse and freehand drawing tool embedded in the software. A binary image file of this was created and saved for each polyp identified, for both prone and supine studies where possible. This unblinded expert review panel enabled a definition of truth to be established for each individual case against which the performance of the CAD software could be judged subsequently.

Computer-assisted detection

The 138 CT datasets and the binary image files were transferred subsequently to a personal computer loaded

with CAD software (Colon CAR 1.2, Medicsight PLC, London, UK). Importantly, this software had previously encountered none of the 138 datasets. In particular, none of these datasets had been used to train the software previously during its development phase, nor was there any opportunity to train the software using these datasets during the course of the present study since the reference standard was established at a research site remote from that handling software development.

Three readers, different from the expert readers who compiled the expert panel review, read the datasets in consensus using the CAD software, which highlighted potential polyp candidates to the reader. For each individual case a user-adjustable sphericity filter that was potentially variable between 0.0 and 1.0 via a software slider bar, was employed at four sphericity settings as follows: 0.3, 0.5, 0.7, 0.9.

In brief, the endoluminal colonic surface was extracted using a thresholding-based method. The extraction thresholds were kept constant for all datasets examined. A mathematical algorithm was then applied to this with the aim of detecting raised endoluminal objects, all of which were regarded as potential polyps. The sphericity filter aimed to facilitate discrimination between real polyps and false positive prompts, due to haustral folds for example, and did this by analyzing every voxel on the candidate surface to determine whether or not it and its neighbours formed part of the surface of a sphere. This procedure is illustrated in Fig. 1. With the sphericity enhancement filter set al 1.0, only those voxels that potentially formed part of a perfect sphere were retained as prompts, and the others were dismissed as likely false-positives. As the filter value was reduced towards zero, voxels that may form part of an increasingly less perfect sphere (a flattened oval for example) were retained as prompts. Prompts were identified by a small white square placed by the software at the region of maximal perceived sphericity and a small white cross placed at the perceived peak of the polyp (Fig. 2).

In order to establish the performance characteristics of the CAD software, in particular the number of true positive and false positive prompts at each sphericity setting, the prompts suggested by the software were compared on a case-by-case basis with the reference polyp co-ordinates and binary image file established previously by the expert readers. In this way it was possible to assign each individual software prompt as either a true positive or false positive mark. In order to count as a true-positive mark, the center of the software prompt had to overlap the circumference of polyps that had been validated by the expert panel. This was achieved by visual comparison of each potential true positive prompt with the binary image file saved during the expert panel read, which outlined the circumference of each validated polyp. Any case in which there was no overlap between a prompt and the binary image file was assigned a false-positive score. In particular,

this stipulation applied even if the software prompt was in the same colonic segment and immediately adjacent to a polyp validated by the expert panel. This procedure was performed at each of the four sphericity settings used.
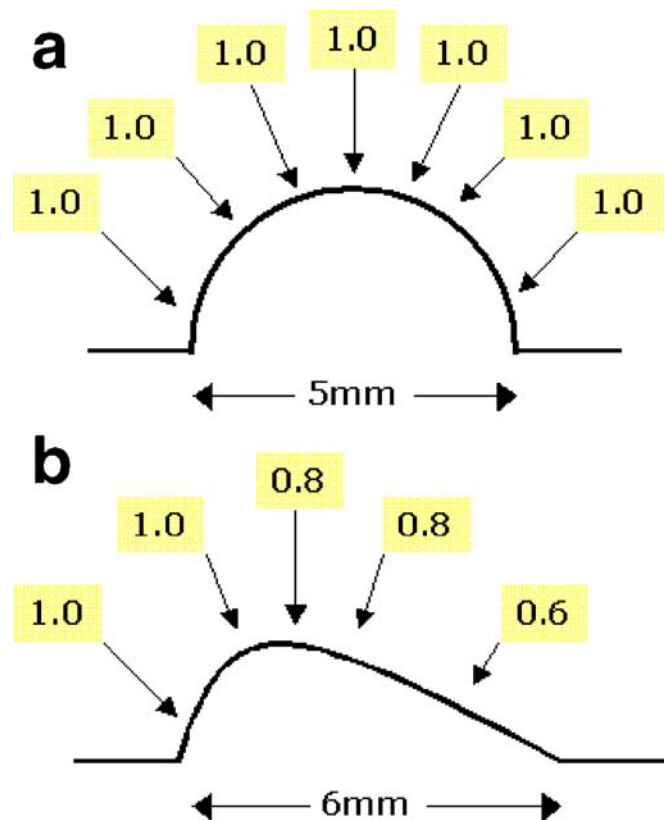

Fig. 1 Diagrammatic representation of the effect of sphericity filtration on the detection of polyps by CAD. Following thresholding and segmentation, all raised endoluminal objects are detected by a mathematical algorithm. All of these are potential polyps until differentiated from normal colonic structures by assessment of their sphericity. Each individual voxel on the surface of each polyp candidate is analysed to determine whether it potentially lies on the surface of a sphere (rather than on a haustral ridge for example). A voxel lying on the surface of a perfect sphere is assigned the highest sphericity value of 1.0 whereas a value of zero indicates that no element of the voxel could potentially form part of a sphere. **a** Diagrammatic representation of a perfectly spherical polyp, with voxels sitting on the polyp surface assigned a sphericity value of 1.0. A sphericity filter value of 1.0 would allow this polyp to be highlighted to the observer as long as its diameter is perceived by the software to be 4 mm or larger. **b** Diagrammatic representation of a flattened polyp. While some voxels on the polyp surface reach a sphericity of 1.0, others do not, indicating a flattened structure. With the sphericity filter set to 1.0 or 0.8 the voxels potentially forming part of a perfect sphere do not reach a minimum diameter of 4 mm and so the polyp is not highlighted. However, when the sphericity filter is reduced to 0.6 the perceived diameter crosses the 4 mm threshold and the polyp is highlighted to the reader
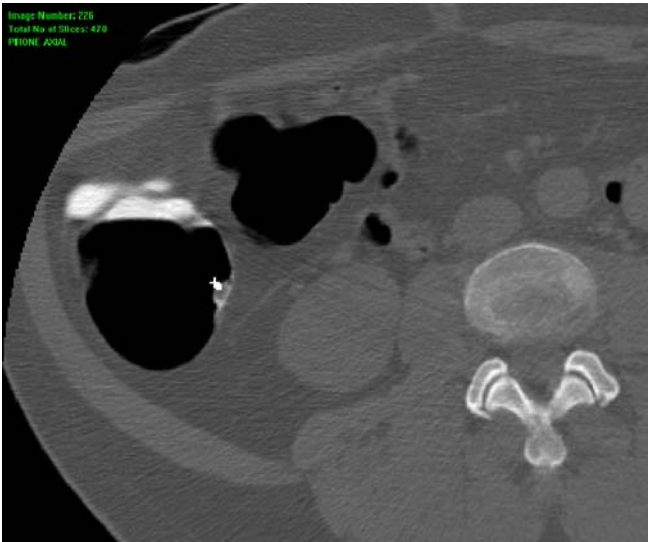
Fig. 2 CT colonography. Prone axial view from a patient with a known 9 mm ascending colon adenoma. The computer-assisted detection system has successfully identified the polyp, placing a small white square at the region of maximal perceived sphericity and a small white cross at the perceived peak of the polyp

Analysis

The datasets were divided into two groups; 86 cases in which positive fecal tagging had been used for patient preparation and 52 in which it had not. A per-polyp analysis was performed. The numbers of true-positive and false-positive CAD prompts were determined for tagged and un-tagged data, both for all polyps and for polyps with a reference diameter of 6 mm, 7 mm, 8 mm, 9 mm, and 10 mm or more. Significant differences in the proportion of polyps detected at equivalent sphericity values were determined by Fisher's Exact Test and the Mann-Whitney U test statistic was used to determine significant differences in the number of false-positive prompts per patient, split by bowel preparation group. Probability was assigned at the 5% level and analysis was performed using Arcus

Quickstat Biomedical (Version 1.2, Research Solutions, Cambridge, UK).

Results

There were 317 polyps validated by the expert panel in the 138 patients. Forty-two of these measured 10 mm or larger and 122 measured 6 mm to 9 mm. Patients with fecal tagging had 172 polyps whose median size was 5 mm overall (range 2 to 35 mm). Of these, 11 measured 10 mm or more and 66 measured 6 mm to 9 mm. Patients without fecal tagging had 145 polyps whose median size was 6 mm overall (range 1 to 41 mm). Of these, 31 measured 10 mm or more and 56 measured 6 mm to 9 mm.

Detection rates for the 317 validated polyps at the four different sphericity settings applied are shown in Table 1. There was a clear trend for increased sensitivity at lower sphericity settings. For example, CAD detected 213 (67.2%) polyps overall at sphericity 0.3 compared to 184 (58%) at sphericity 0.9 (Table 1). Of the 164 polyps measuring 6 mm or more, 144 (87.8%) were detected at a sphericity of 0.3 and 132 (80.1%) at sphericity 0.9. Of the 42 polyps measuring 10 mm or more, 40 (95.2%) were detected at sphericity 0.3 and 36 (85.7%) at sphericity 0.9. For both tagged and un-tagged data the median diameter of polyps detected was greater than that for polyps that were missed (Table 1).

However, sensitivity was apparently reduced generally in patients with fecal tagging when compared to sensitivities at equivalent sphericity settings in patients with un-tagged bowel preparation (Table 1). For example, the lowest sensitivity was achieved in both datasets when the highest sphericity setting (0.9) was used but the equivalent values were 0.47 versus 0.72 for tagged and un-tagged data respectively. Similarly, the highest sensitivity was achieved in both sets of data when the lowest sphericity setting (0.3) was used, but again the greatest individual value was achieved in un-tagged data, with a value of 0.80 versus 0.56 (Table 1).

Table 1 CAD detection rates and false positive prompts for tagged (n=172 polyps) and un-tagged (n=145 polyps) data obtained at four different sphericity filter settings

Sphericity filter	Data	Polyps detected (%)	Median size (range) mm	Polyps missed (%)	Median size (range) mm	FP median (range)	p*
0.3	Tagged	97 (56.4)	6 (2–35)	75 (43.6)	5 (2–8)	65 (15–219)	0.009
	Untagged	116 (80)	7 (1–41)	29 (20)	4 (2–12)	46 (8–141)	
0.5	Tagged	96 (55.8)	6 (2–35)	76 (44.2)	5 (2–8)	58.5 (13–201)	0.014
	Untagged	116 (80)	7 (1–41)	29 (20)	4 (2–12)	44 (8–131)	
0.7	Tagged	90 (52.3)	6 (2–35)	82 (47.7)	5 (2–8)	39 (0–145)	0.140
	Untagged	111 (76.6)	7 (1–41)	34 (23.4)	4 (2–12)	33 (8–107)	
0.9	Tagged	80 (46.5)	6.5 (3–35)	92 (53.5)	5 (2–33)	12.5 (1–50)	0.033
	Untagged	104 (71.7)	7 (1–41)	41 (28.3)	5 (2–35)	16 (4–60)	

*p=Mann-Whitney U test statistic

In order to determine if the apparent reduction in sensitivity for tagged data (at equivalent sphericity filter settings) was due to an interaction with the size of the polyps present in each dataset, a further analysis was performed that was restricted to polyps of equivalent sizes: 6 mm, 7 mm, 8 mm, 9 mm, and 10 mm or more (Table 2). This analysis revealed that there was no significant difference between the proportion of polyps detected at equivalent sphericity settings when tagged and un-tagged data were compared (Table 2). There was some suggestion that detection of 8 mm polyps was enhanced in un-tagged data but this value was on the 5% probability level and most likely contingent upon multiple testing (Table 2).

As expected, specificity decreased in line with increased sensitivity, with both tagged and un-tagged data showing best specificity at sphericity 0.9 and worst at 0.3 (Table 1). However, there were significantly more false-positive detections in tagged data at lower sphericity values (i.e. 0.3 and 0.5) and significantly more false positive detections in untagged data at the highest sphericity level of 0.9 (Table 1).

Discussion

In the classic CAD paradigm, studies are read first without computer-assistance and then again with, i.e. the CAD system acts as a surrogate for a second reader and systems have gained regulatory approval in this context. However, there are potential disadvantages to this approach notwithstanding the temptation to activate the system before a careful and thorough primary read has been performed [12]. This is because the classic CAD implementation can be viewed as a 'black-box', i.e. there is little opportunity for the observer to influence the detection characteristics of the system. For example, a typical CAD implementation cannot be calibrated on-the-fly to take account of the characteristics of the individual clinical case in question and the potential for interaction with the reader is limited as a consequence.

The desire for a more 'radiologist-driven' approach has led to the development of CAD systems that allow their performance characteristics to be manipulated via adjustable filters. Such filters are an inherent part of any CAD algorithm, and are used to reduce false-positive prompts, but allowing the user to titrate the sensitivity/specificity to suit the clinical case in question is a relatively new concept. For example, sensitivity for small polyps may be reduced if the clinical context requires diagnosis of large lesions only (e.g. symptomatic cancer), or where the referring clinician ascribes little importance to small polyps. We investigated the effect of different values applied to a filter based on the sphericity of the target lesion and were able to show that increasing the numerical value of the filter decreased the number of small polyps detected, with corresponding reductions in the false-positive rate. For example, if the user is only interested in detecting polyps of 1 cm diameter or more, we found that maximal performance was achieved at a sphericity setting of 0.7. Increasing sphericity beyond this resulted in missed polyps compared to lower values but values less than 0.7 resulted in increased false-positive prompts for no detection gains at the target diameter. As a consequence, users of this type of software need to be aware of the benefits and trade-offs at each potential filter setting in order to maximize the advantages of the software.

It should be noted that the experimental design we used meant that the software had not been exposed previously to the polyps on which it was tested. Furthermore, there was no opportunity for the software to 'learn' from these during the study, not least because the reference standard was established at a site remote from that handling software development. Some internal validation schemes require

Table 2 True-positive and false-negative rates for CAD at four different sphericity values in both tagged and un-tagged datasets, for polyps with a maximal diameter of 6 mm, 7 mm, 8 mm, 9 mm and 10 mm or more

Maximal polyp diameter (mm)		Polyps detected number (%)							
		Sphericity filter value							
		0.3	p*	0.5	p*	0.7	p*	0.9	p*
6	Tagged data (*n*=28)	19 (68)	0.16	19 (68)	0.16	18 (64)	0.2	18 (64)	0.2
	Un-tagged data (*n*=18)	16 (89)		16 (89)		15 (83)		15 (83)	
7	Tagged data (*n*=18)	14 (78)	0.18	14 (78)	0.18	14 (78)	0.18	14 (78)	0.4
	Un-tagged data (*n*=19)	18 (95)		18 (95)		18 (95)		17 (89)	
8	Tagged data (*n*=9)	8 (89)	0.39	8 (89)	0.39	6 (67)	0.05	6 (67)	0.26
	Un-tagged data (*n*=14)	14 (100)		14 (100)		14 (100)		13 (93)	
9	Tagged data (*n*=11)	11 (100)	0.31	11 (100)	0.31	11 (100)	0.31	10 (91)	0.21
	Un-tagged data (*n*=5)	4 (80)		4 (80)		4 (80)		3 (60)	
10 or more	Tagged data (*n*=11)	11 (100)	1	11 (100)	1	11 (100)	1	10 (91)	1
	Un-tagged data (*n*=31)	29 (94)		29 (94)		29 (94)		26 (84)	

*Fisher's exact test

that the software be trained using an accumulated dataset from which a single polyp has been excluded. It is then determined whether the software can detect the excluded polyp, and the procedure is repeated excluding each polyp in turn [13–15]. While this paradigm maximizes scarce resources since identical data is used for both training and testing the software, it may result in an over-optimistic assessment of performance [16]. Bogoni and co-workers [17] used a methodology similar to the present study, whereby accumulating data is randomized into 'training' and 'testing' partitions, a well-recognised procedure in the development and testing of CAD systems [12, 18]. However, they tested a CAD system whose performance was non-interactive [17]. Detection rates for 11 medium polyps (defined as 6–9 mm), and 10 large polyps (defined as 10 mm or larger) were 81.8% and 100%, respectively [17]. Using a much larger dataset (164 polyps measuring 6 mm or more versus 21) we found corresponding sensitivities of 85.2% and 95.2% at our most sensitive sphericity setting of 0.3.

We hypothesized that different categories of CT data would affect the performance characteristics of the CAD software, and that the filter settings might have to be adjusted by the user to take account of this. Specifically, we investigated whether patient preparation using positive fecal tagging required different filter values than non-tagged data in order to achieve equivalent sensitivity. This information is important because it allows users to make informed judgments concerning filter settings. Under ideal circumstances, batch file analysis would be performed on local data obtained from each center using the CAD software so that detailed recommendations for filter settings could be made. However, this is clearly not possible in practice, not least because most potential purchasers of the software will not have accumulated validated data upon which batch analysis can be performed. Because of this, any recommendations made must be based on centralized studies such as this.

Our initial results suggested that CAD sensitivity was less when investigating tagged data at equivalent sphericity levels but we could find no significant difference when the distribution of polyp sizes between the two datasets was accounted for. However, there was a significant tendency for decreased specificity with tagged data at the most sensitive sphericity settings. The opposite applied at low sensitivity settings, with significantly more false-positive detections in un-tagged data (although absolute specificity was much improved when compared to lower sphericity settings). The reasons underlying this observation are unclear but might relate to the fact that the CAD system tested in the present study was not developed to account for tagged data specifically, although a proportion of tagged data was used to provide polyps during the development phase. The most likely explanation is that there were differences in the amount of particulate residue attributable to the preparation: small particles of retained residue will

be perceived as larger than reality when a tagged preparation is used because of partial volume effect. In contrast, standard preparations may leave larger particles [19], whose detection may become significant when detection of smaller particles has largely been eliminated by using a less sensitive sphericity setting. In any event, there is increasing awareness that CAD systems need to account for the patient preparation employed if performance characteristics are to be optimized [20]; the results of our study support this observation. Also, at the time of testing there was no facility to limit directly the number of false-positive prompts contingent upon the perceived diameter of the lesion identified by the CAD system, other than the fact that a minimum perceived diameter of 4 mm had to be reached (contingent on the setting of the sphericity filter). Work is currently underway to develop a further user-adjustable filter that allows the observer to specify the minimum size of polyp to be reported, for example, 1 cm—prompts with a diameter less than this will be discarded by the software.

Our study does have limitations. We assimilated data from four centers who obtained their data in different ways. For example, one center used positive oral contrast while others did not, and the effect of this on CAD was the focus of our study. However, the data was heterogenous in many other ways. For example a range of collimations, insufflation methods and gas, and spasmolytic were used. It would be tempting to perform subset analyses to investigate the effect of each of these on CAD detection rates but we did not have adequate statistical power to do this and so chose not to; only nine patients had data with a collimation of 2.5 mm without tagging for example. Any hypothesis testing on such small subgroups would likely provide meaningless results, especially given the confounder of polyp size on detection. It is unfortunate that we were unable to make these comparisons and further work in this area is required. However, it is important to be aware that detection rates will be over-optimistic if the data on which CAD is tested is too homogenous [12, 16]. By way of example, different radiologists at different hospitals will inevitably use different preparations, scanners, and technical parameters to acquire their data, even when working within accepted guidelines for best practice [10]. Patient subgroups will also differ, and the prevalence and morphology of polyps may vary depending on local referral practice. It is therefore actually desirable that CAD software is tested on data that is heterogenous since the more heterogenous the data, the more generalizable will be the results to centers and purchasers that have not participated in the development of the software [12, 16]. It should also be noted that we did not attempt to determine the detection characteristics of the software for cancer. While it could easily be argued that detection of cancer is more important than detection of polyps, the intended market for CAD products is usually for screening, where the intention is to detect significant adenomas with the aim

of cancer prevention. Also, the very low prevalence of established, invasive cancer in these populations makes powering studies of their detection by CAD problematic since the event rate is so low, e.g. there were only two cancers in one study of 1,233 screened [1].

In summary, CAD had a sensitivity of 95.2% for polyps measuring 1 cm or more and 87.8% for polyps 6 mm or more when used at a sphericity setting of 0.3. Higher sphericity settings reduced the number of false-positive prompts in exchange for reduced sensitivity. While we could find no difference in the sensitivity of CAD when used in tagged and un-tagged data, the preparation used significantly impacts on specificity.

References

1. Pickhardt PJ, Choi JR, Hwang I, Butler JA, Puckett ML, Hildebrandt HA et al (2003) Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. N Engl J Med 349:2191–2200
2. Yee J, Akerkar GA, Hung RK, Steinauer-Gebauer AM, Wall SD, McQuaid KR (2001) Colorectal neoplasia: performance characteristics of CT colonography for detection in 300 patients. Radiology 219:685–692
3. Fenlon HM, Nunes DP, Schroy PC 3rd, Barish MA, Clarke PD, Ferrucci JT (1999) A comparison of virtual and conventional colonoscopy for the detection of colorectal polyps. N Engl J Med 341:1496–1503
4. Johnson CD, Toledano AY, Herman BA et al (2003) Computerized tomographic colonography: performance evaluation in a retrospective multicenter setting. Gastroenterol 125:688–695
5. Burling D, Halligan S, Altman DG et al (2006) Polyp measurement and size categorisation by CT colonography: effect of observer experience in a multicentre setting. Eur Radiol 16:1737–1744
6. 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
7. Mani A, Napel S, Paik DS, Jeffrey RB, Yee J, Olcott EW, Prokesch R, Davila M, Schraedley-Desmond P, Beaulieu CF (2004) Computed tomography colonography: Feasibility of computer-aided polyp detection in a "first-reader" paradigm. J Comput Assist Tomogr 28:318–326
8. Johnson CD, Harmsen WS, Wilson LA, Maccarty RL, Welch TJ, Ilstrup DM et al (2003) Prospective blinded evaluation of computed tomographic colonography for screen detection of colorectal polyps. Gastroenterology 125:311–319
9. Taylor SA, Halligan S, Burling D et al (2006) Performance of computer-assisted reader software for polyp detection in comparison to expert readers. AJR 186:696–702
10. Barish MA, Soto JA, Ferucci JA (2005) Consensus on current clinical practice of virtual colonoscopy. AJR 184:786–792
11. Macari M, Milano A, Lavelle M, Berman P, Megibow AJ (2000) Comparison of time-efficient CT colonography with two- and three-dimensional colonic evaluation for detecting colorectal polyps. Am J Roentgenol 174:1543–1549
12. Wagner RF, Beiden SV, Campbell G, Metz CE, Sacks WM (2002) Assessment of medical imaging and computer-assist systems: lessons from recent experience. Acad Radiol 9:1264–1277
13. Yoshida H, Nappi J, MacEnwaney P, Rubin DT, Dachman AH (2002) Computer-aided diagnosis scheme for detection of polyps at CT colonography. Radiographics 22:963–979
14. Jerebko AK, Malley JD, Franaszek M, Summers RM (2003) Multiple neural network classification for detection of colonic polyps in CT colonography datasets. Acad Radiol 10:154–160
15. Nappi JJ, Frimmel H, Dachman AH, Yoshida H (2004) Computerized detection of colorectal masses in CT colonography based on fuzzy merging and wall-thickness analysis. Med Phys 31:860–872
16. Altman DG, Royston P (2000) What do we mean by validating a prognostic model? Statist Med 19:453–473
17. Bogoni L, Cathier P, Dundar M, Jerebko A, Lakare S, Liang J, Periaswamy S, Baker ME, Macari M (2005) Computer-aided detection (CAD) for CT colonography: a tool to address a growing need. Br J Radiol 78: S57–S62
18. Dodd LE, Wagner RF, Armato SG, McNitt-Grey, Bieden S, Chan HP, Gur D, McLennan G, Metz CE, Petrik N, Sahiner B, Sayre J (2004) Assessment methodologies and statistical issues for computer-aided diagnosis of lung noduled in computed tomography: contemporary research topics relevant to the lung image database consortium. Acad Radiol 11:462–475
19. Taylor SA, Halligan S, Goh V, Morley S, Atkin W, Bartram CI (2003) Optimizing bowel preparation for multi-detector row CT colonography: effect of Citramag and Picolax. Clin Radiol 58:723–732
20. Summers RM, Franaszek M, Miller MT, Pickhardt PJ, Choi JR, Schindler WR (2005) Computer-aided detection of polyps on oral contrast-enhanced CT colonography. AJR 184:105–108