Neutral vs positive oral contrast in diagnosing acute appendicitis with contrast-enhanced CT: sensitivity, specificity, reader confidence and interpretation time

¹D M NAEGER, MD, ²S D CHANG, MD, ¹P KOLLI, MD, ¹V SHAH, MD, ¹W HUANG, MD and ¹R F THOENI, MD

¹University of California San Francisco, Medical School, Department of Radiology and Biomedical Imaging, PO Box 1325, San Francisco, CA 94143-1325, USA, and ²Department of Radiology, University of British Columbia, Vancouver General Hospital, 899 W 12th Avenue, Vancouver BC V5Z 1M9, Canada

Objective: The study compared the sensitivity, specificity, confidence and interpretation time of readers of differing experience in diagnosing acute appendicitis with contrast-enhanced CT using neutral vs positive oral contrast agents.

Methods: Contrast-enhanced CT for right lower quadrant or right flank pain was performed in 200 patients with neutral and 200 with positive oral contrast including 199 with proven acute appendicitis and 201 with other diagnoses. Test set disease prevalence was 50%. Two experienced gastrointestinal radiologists, one fellow and two first-year residents blindly assessed all studies for appendicitis (2000 readings) and assigned confidence scores (1=poor to 4=excellent). Receiver operating characteristic (ROC) curves were generated. Total interpretation time was recorded. Each reader's interpretation with the two agents was compared using standard statistical methods.

Results: Average reader sensitivity was found to be 96% (range 91–99%) with positive and 95% (89–98%) with neutral oral contrast; specificity was 96% (92–98%) and 94% (90–97%). For each reader, no statistically significant difference was found between the two agents (sensitivities p -values >0.6; specificities p -values >0.08), in the area under the ROC curve (range 0.95–0.99) or in average interpretation times. In cases without appendicitis, positive oral contrast demonstrated improved appendix identification (average 90% vs 78%) and higher confidence scores for three readers. Average interpretation times showed no statistically significant differences between the agents.

Conclusion: Neutral vs positive oral contrast does not affect the accuracy of contrast-enhanced CT for diagnosing acute appendicitis. Although positive oral contrast might help to identify normal appendices, we continue to use neutral oral contrast given its other potential benefits.

Received 8 October 2009
Revised 2 November 2009
Accepted 9 November 2009

DOI: 10.1259/bjr/20854868

© 2011 The British Institute of Radiology

For almost two decades CT has been widely used for diagnosing appendicitis; studies have consistently demonstrated the high sensitivity and specificity of CT for this diagnosis [1–4]. Standard imaging practices include the use of intravenous iodinated contrast [5–7]. Positive (high-attenuation) oral contrast, and sometimes positive rectal contrast, have also been recommended [5, 8, 9]. The theoretical advantages of positive oral (and rectal) contrast in diagnosing appendicitis lie in their ability to highlight the lumen of the small bowel, cecum and normal appendix in cases without appendicitis and in the absence of filling of an inflamed and obstructed appendix in cases with acute appendicitis [10].

With the introduction of multidetector row CT (MDCT) and its improved visualisation of the bowel wall, combined with developments in CT enterography, the use of neutral oral contrast has become more popular. Neutral oral contrast agents such as water or VoLumen (a barium sulphate suspension, 0.1% w/v,

0.1% w/w; E-Z-EM, Inc., a Bracco-owned company, Lake Success, NY) offer the advantage of better characterisation of the bowel wall owing to differences between the low-attenuation intraluminal fluid and the contrast-enhanced wall following administration of intravenous contrast material. This difference enables improved assessment of the thickness and enhancement pattern of the bowel wall [9–13]. Sorbitol- (and other solute) containing solutions result in more bowel distention than water alone, thereby allowing for even better characterisation [11–14]. In the evaluation for appendicitis, such superior definition of the wall of the gastrointestinal tract might also improve the evaluation of the appendiceal wall. Neutral oral contrast offers additional benefits, including the lack of streak artefacts often seen on CT from high-attenuation positive oral contrast.

Despite these advantages, neutral oral contrast is not yet widely adopted for studies in which appendicitis is considered a likely diagnosis [12]. The purpose of this study was to compare the sensitivity, specificity, confidence and interpretation times of readers with differing levels of experience in diagnosing appendicitis by using neutral vs positive oral contrast with contrast-enhanced CT.

Address correspondence to: Ruedi F Thoeni, MD, University of California San Francisco, Medical School, Department of Radiology and Biomedical Imaging, PO Box 1325, San Francisco, CA 94143-1325, USA. E-mail: ruedi.thoeni@radiology.ucsf.edu

Methods and materials

Patients

This study was approved by the institutional review board and is HIPAA (Health Insurance Portability and Accountability Act) compliant. Informed consent was waived for this retrospective study. All surgical reports from April 2005 to March 2008 were reviewed to identify appendectomies performed at our institution. Cases were consecutively included into the study if pathology reports confirmed a diagnosis of appendicitis and if a contrast-enhanced CT scan with oral contrast had been performed prior to surgery. Case collection continued until 100 CT examinations with positive oral contrast (2% diatrizoate meglumine and diatrizoate sodium, Gastrografin; Bracco Diagnostics, Princeton, NJ) and 100 CT examinations with neutral oral contrast (VoLumen) were identified. Once all 200 cases had been identified, one radiologist who did not interpret the study cases re-reviewed the data before the actual readings by 5 radiologists; 1 case failed to meet pathological criteria for appendicitis on the final report and 1 case was misidentified as having neutral rather than positive oral contrast. Therefore, a total of 199 cases with appendicitis were included: 101 with positive and 98 with neutral oral contrast.

Over the same time frame, records of all abdominal and pelvic CT scans were reviewed for a clinical history of right lower quadrant and/or right flank pain, the imaging indication often provided for patients with appendicitis at our institution. Patients already identified as having pathology-proven appendicitis (see above) were not included. Cases were collected consecutively until 100 cases with positive and 100 cases with neutral oral contrast were obtained. Again, prior to interpretation, one case was found to be misclassified with regards to the type of contrast used. The case that did not meet the pathological criteria for appendicitis on the final pathology report (discussed above) was also included in this group. The final group of 201 patients consisted of 99 cases with positive and 102 cases with neutral oral contrast.

Among the 101 patients with appendicitis and positive oral contrast were 70 males (average age 32 years; range 3–68) and 31 females (average age 34 years; range 6–69). Among the 98 patients with appendicitis and neutral oral contrast were 64 males (average age 34 years; range

4–70) and 34 females (average age 38 years; range 11–72). In the 99 patients without appendicitis and positive oral contrast were 40 males (average age 37 years; range 4–66) and 59 females (average age 36 years; range 4–80). In the 102 patients without appendicitis and neutral oral contrast were 36 males (average age 36 years; range 5–69) and 66 females (average age 38 years; range 6–86).

Absence of appendicitis among the 201 cases with other diagnoses was confirmed by review of the medical record, which documented the final diagnosis at the time of discharge from the emergency department, at discharge from a hospital admission or at surgery. The various diagnoses in patients without appendicitis are listed in Table 1 and are comparable to the distribution of alternate diagnoses found in similar studies [15]. In some patients, no diagnosis could be established to explain the symptoms and “abdominal pain, not otherwise specified” was entered as the official clinical diagnosis in the medical chart (87 out of 201; 43%). The proportion of such patients was similar to that reported in the literature [16, 17]. Of these patients, 22 were lost to follow-up. For the remainder, follow-up visits excluded the diagnosis of acute appendicitis subsequently at our institution or at any affiliated clinical site (range of follow-up, 1–41 months; median, 10 months).

One of the gastrointestinal radiologists, with over 25 years experience in abdominal imaging, determined if the appendix was in a retrocecal location; a comparison was made among the four groups with regards to the proportion of appendices in this location, as it facilitates visualisation of the appendix.

Imaging parameters

Because of the retrospective nature of this study, CT imaging parameters were per standard clinical practice at the time of acquisition. Any aberrations from the routine protocol were per radiologist preference and generally minimal (additional thin sections or decubitus views). All CT studies were performed using a helical CT scanner (CT/I; GE Medical Systems, Milwaukee, WI) at 120 Kvp and 250–320 mA. Paediatric patients were imaged with a lower tube current, per our paediatric protocol. The scanning time was 0.8 s per rotation. Craniocaudal scanning was carried out with 5 mm collimation in the abdomen and 3 mm in the pelvis. All

Table 1. Diagnoses of right lower quadrant and/or right flank pain in 201 patients without appendicitis

Diagnosis	n
GI tract aetiology (gastroenteritis, infectious colitis, diverticulitis, SBO, DU, IBD, hernia constipation, GERD, neoplasm, diarrhoea, vomiting, intussusception, ischaemia) ^a	45
Ovarian aetiology (cyst, torsion, abscess, neoplasm)	19
Urinary tract aetiology (pyelonephritis, cystitis, hydronephrosis, nephrolithiasis)	17
Uterus, cervix, vagina (fibroids, cervicitis, vaginitis, endometriosis)	9
Pelvic inflammatory disease	8
Pancreatic aetiology (pancreatitis, pancreatic neoplasm)	5
Biliary aetiology (cholecystitis, cholangitis, cholelithiasis)	4
Miscellaneous (traumatic soft tissue contusion, AAA, liver metastases, radiating muscle pain)	7
Abdominal pain, otherwise not specified	87 ^b

AAA, abdominal aortic aneurysm; DU, duodenal ulcer; GERD, gastrointestinal reflux; GI, gastrointestinal; IBD, inflammatory bowel disease; SBO, small bowel obstruction. ^aIn cases where a final diagnosis could not be made, a symptom (e.g. diarrhoea) was often recorded as the final diagnosis in the medical record and is reported here. ^b65 patients with negative follow-up available in the medical record.

scanning parameters remained unchanged for the duration of the study with the pitch varying from 1–1.5 with minimal change to the designated collimation. Iohexol (150 ml Omnipaque; Amersham Health, GE Healthcare, Princeton, NJ) was administered intravenously at a rate of 4 ml s⁻¹ with a 70 s scan delay. Scanning was started after ingestion of 900–1200 ml of positive oral contrast (Gastrografin) administered in 450 ml doses at 30 min and 20 min and in a dose of 300 ml at 10 min prior to the study. A 2% solution of Gastrografin was used by mixing 12 ml of the iodinated contrast material with 450 ml of water. Neutral oral contrast agent (900–1350 ml of VoLumen) was given in 450 ml doses at 30 min, 20 min and 10 min prior to the CT study. Cases with positive oral contrast administration were mostly from the first 1.5 years of the 3 year study period owing to the standard protocols used at the time. Approximately 1.5 years into the retrospectively reviewed period, VoLumen was introduced and quickly became the standard oral agent. Although radiologists maintained the ability to give positive or neutral oral contrast as they saw fit throughout the entire 3 year period, most followed the standard protocols at the time for imaging of suspected appendicitis.

Image interpretation

Five readers blinded to the patients' name, age, detailed clinical history, prior diagnoses and results of the initial read of the CT independently interpreted each study. All cases were randomised and reviewed on a Picture Archiving and Communication System (PACS) station with all annotations removed. The only information provided to the readers was a history of right lower quadrant pain and/or right flank pain. The readers consisted of 2 subspecialised abdominal radiologists, 1 with more than 25 years experience (Reader 1) and 1 with 9 years experience (Reader 2), a board-certified fellow in abdominal imaging (Reader 3) and 2 first-year radiology residents with 6 months of radiology training (Readers 4 and 5). The abdominal radiologist with 9 years experience (Reader 2) had never used VoLumen before this study and the abdominal imaging fellow (Reader 3) had used VoLumen only for the 6 months of his fellowship training. The two first-year residents (Readers 4 and 5) started their training when VoLumen was the routinely used oral contrast agent, although they were also familiar with positive oral contrast.

The 5 readers reviewed all 400 cases in a randomised, blinded fashion. For each case, each reviewer recorded the ability to identify the appendix and the presence or absence of seven specific findings supportive of the diagnosis of acute appendicitis: appendiceal dilation (>6 mm), appendiceal wall thickening (>2 mm), appendiceal wall hyperenhancement, peri-appendiceal fat stranding, peri-appendiceal abscess, evidence of perforation (missing wall sign) and presence of an appendicolith. Each reader was asked to categorise the study as "positive" or "negative" for appendicitis and rank the level of confidence in the diagnosis (1, poor; 2, moderate; 3, good; 4, excellent). Evidence of pathology other than the presence or absence of appendicitis was also recorded, thereby requiring evaluation of the remainder of the scan series. Each reader confirmed that the study was of

diagnostic quality and recorded the total interpretation time by listing the starting and ending times. To evaluate the sensitivity for appendicitis in this study *vs* that in our clinical practice, the original reports were reviewed by the radiologist who did not interpret the study cases and the diagnoses recorded.

Statistical analysis

All statistics were calculated using Stata, version 8.0 (College Station, TX). A *p*-value threshold for statistical significance of 0.05 was used for all analyses. Between the positive and neutral oral contrast groups, the proportion of male patients and retrocecal appendices was compared with the χ^2 statistic; patients' age was compared with a one-way ANOVA test and the proportions of patients with identifiable appendices were compared with the Fisher's exact statistic. This latter test was chosen because of the low number of appendices that were not identified. Each reader's sensitivity for appendicitis was calculated for positive and neutral oral contrast cases separately, defined as the number of correctly identified cases of appendicitis among all surgically and histologically proven appendicitis cases. The specificity for each reader was calculated for cases with positive and neutral oral contrast, defined as the number of correctly identified cases without appendicitis divided by the total number of cases without appendicitis. The 95% confidence interval was calculated using a binomial exact distribution. This test was selected because the calculated sensitivities and specificities were close to 100%. The sensitivity and specificity of each reader was compared between cases with positive and neutral oral contrast using the Fisher's exact statistic.

The confidence score data (values ranging from 1 to 4) were used to generate receiver operator characteristic (ROC) curves. The confidence scores were combined with the categorisation of "acute appendicitis" or "no appendicitis" to create eight levels of diagnostic certainty and standard ROC curves were generated for each reader, for the 200 cases in which positive oral contrast was administered and for the 200 cases in which neutral oral contrast was administered. Areas under the curve were calculated using non-parametric methods, which require fewer assumptions about the nature of the true underlying curves. The 95% confidence intervals were calculated for each reader and for each type of contrast and comparisons were made using standard methods [18].

For each reader, the confidence scores for positive and neutral oral contrast groups were compared using the Wilcoxon rank sum test, a non-parametric statistic which maintains the ordinal nature of the data. The Wilcoxon rank sum test was also used to compare confidence score data between correctly and incorrectly classified cases.

The average interpretation times per reader for each oral contrast group were calculated and compared using the Wilcoxon rank sum test.

Results

Patients

The difference in the proportion of male *vs* female patients was not statistically significant between the positive *vs* neutral oral contrast groups with appendicitis

Table 2. Identification of the appendix

	Proven appendicitis		<i>p</i> -Value ^a	No appendicitis		<i>p</i> -Value ^a
	Positive oral contrast (Gastrografin)	Neutral oral contrast (VoLumen)		Positive oral contrast (Gastrografin)	Neutral oral contrast (VoLumen)	
Reader 1	100/101 (99%)	98/98 (100%)	1.00	95/99 (96%)	92/102 (90%)	0.17
Reader 2	97/101 (96%)	92/98 (94%)	0.53	89/99 (90%)	79/102 (77%)	0.02
Reader 3	98/101 (97%)	96/98 (98%)	1.00	88/99 (89%)	80/102 (78%)	0.06
Reader 4	99/101 (98%)	96/98 (98%)	1.00	88/99 (89%)	81/102 (79%)	0.08
Reader 5	94/101 (93%)	92/98 (94%)	1.00	84/99 (85%)	68/102 (67%)	0.003
Average of readers	488/505 (96.6%)	474/490 (96.7%)	1.00	444/495 (89.7%)	400/510 (78.4%)	<0.001

^aFisher's exact test.

($p=0.55$) and without appendicitis ($p=0.46$), although the appendicitis groups had a higher number of males (70 and 64) compared with the groups without appendicitis (40 and 36); $p<0.001$ (χ^2 statistic). The age between all four groups was comparable with no statistically significant difference ($p=0.7$). When the appendix was identified and based on the designations from Reader 1, a retrocecal location of the appendix was noted in 44 of 100 (44%) patients with appendicitis and Gastrografin, in 48 of 98 (49%) patients with appendicitis and VoLumen, in 49 of 95 (52%) patients without appendicitis and Gastrografin and in 40 of 92 (43%) patients without appendicitis and VoLumen. The differences between the groups were not statistically significant ($p=0.62$).

Ability to identify the appendix

The proportion of cases where the appendix was identified is reported for each reader in Table 2. In patients with appendicitis, the overall rate of identifying the appendix was high for all readers regardless of the type of oral contrast used (average 96.6% for Gastrografin *vs* 96.7% for VoLumen; $p=1.00$) and higher than in cases without appendicitis ($p<0.001$). For cases without appendicitis, all readers identified the appendix more frequently when Gastrografin was used (average 89.7% *vs* 78.4%, $p<0.001$; two out of the five readers demonstrated statistically significant differences when analysed individually).

Sensitivity and specificity for diagnosing acute appendicitis

The sensitivity for diagnosing acute appendicitis in 101 surgically and histologically proven acute appendicitis

cases examined with positive oral contrast and in 98 cases with neutral oral contrast is reported for each reader in Table 3 (Figures 1 and 2). The sensitivity for appendicitis in the original interpretations of our study cases was 99% for positive oral contrast cases and 97% for neutral ($p=0.4$), statistically not different from the results obtained in our current study: 95% and 96% (average sensitivities of all readers), respectively. The specificity in this study for correctly diagnosing the absence of appendicitis in the 201 cases of right lower quadrant and/or right flank pain are reported in Table 4 for each reader and both types of oral contrast agent.

Receiver operator characteristic curves and areas under the curve

The recorded confidence score data was used to generate standard ROC curves for each reader. Two curves are compared for each reader, one generated from the 200 cases in which positive oral contrast was administered and one from the 200 cases in which neutral oral contrast was administered (Figure 3). No comparison between the areas under the curve for Gastrografin and VoLumen yielded statistically significant differences ($p=0.92, 0.96, 0.75, 0.18$ and 0.33 for Readers 1–5, respectively; $p=0.94$ for the averaged area under the curve comparison).

Confidence scores

Confidence scores were analysed between the positive and neutral oral contrast groups for cases with and without acute appendicitis. In cases with appendicitis, no reader demonstrated statistically significant differences in the confidence scores (p -values range from 0.3 to 1.0).

Table 3. Sensitivity for diagnosing acute appendicitis per reader for positive vs neutral oral contrast

	Sensitivity (%) ^a	Sensitivity (%) ^a	<i>p</i> -Value ^b
	Positive oral contrast (Gastrografin) ($n=101$ proven appendicitis)	Neutral oral contrast (VoLumen) ($n=98$ proven appendicitis)	
Reader 1	98 (93–99.7)	97 (91–99)	0.7
Reader 2	94 (88–98)	93 (86–97)	0.8
Reader 3	96 (90–99)	97 (91–99)	1.00
Reader 4	99 (95–99.9)	98 (93–99.8)	0.6
Reader 5	91 (84–96)	89 (81–94)	0.6
Average of readers	($n=505$) 96 (93–97)	($n=490$) 95 (92–97)	0.5 (χ^2)

^aBinomial exact 95% confidence intervals used for confidence interval reporting. ^bFisher's exact test.

Figure 1. A 27-year-old male patient with positive oral contrast and proven appendicitis correctly diagnosed by all readers. This axial CT scan shows a dilated appendix with wall thickening and increased enhancement of the wall (arrows) associated with mild peri-appendiceal stranding indicative of appendicitis. The lumen of the appendix does not contain positive oral contrast probably owing to inflammation.

Figure 2. A 32-year-old female patient with neutral oral contrast and proven appendicitis correctly diagnosed by all readers. This axial CT scan demonstrates a dilated appendix containing fluid in its lumen and increased enhancement of the appendiceal wall indicating appendicitis (arrows). Stranding of the fat surrounding the inflamed appendix is also seen.

In cases without appendicitis, three readers registered statistically significant higher confidence scores with positive oral contrast (Reader 2, mean 3.5 *vs* 3.3, $p=0.05$; Reader 3, 3.6 *vs* 3.3, $p<0.001$; Reader 5, 2.7 *vs* 2.4, $p=0.01$). The two remaining readers demonstrated higher confidence scores with neutral oral contrast, although the differences were not statistically significant (Reader 1, mean 3.7 *vs* 3.5, $p=0.3$; Reader 4, 3.9 *vs* 3.8, $p=0.12$).

False-positive and false-negative diagnoses

The number of false-positive and false-negative interpretations is listed per reader in Table 5. All readers demonstrated lower confidence in their incorrect diagnoses compared with their correct diagnoses ($p<0.001$ for all comparisons). Of the 101 appendicitis cases with positive oral contrast, 5 cases were misinterpreted by multiple readers: 1 case by all 5 readers, 1 case by 3 readers and 3 cases by 2 readers. Of the 98 appendicitis cases with neutral oral contrast, 5 cases were misinterpreted by multiple readers: 1 case by 4 readers, 2 cases by 3 readers and 2 cases by 2 readers. In the majority of these cases, the pathology report indicated appendicitis with a mild degree of inflammation. The final diagnoses in false-positive cases in which more than one

reader misclassified the case are listed in Table 6 (Figures 4 and 5).

Time of interpretation

The average interpretation time per case is shown in Table 7, presented per reader and per oral contrast medium. One reader (Reader 4) registered a statistically significant shorter average interpretation time with neutral oral contrast in cases with appendicitis, although the absolute difference was relatively small at 19 s. All other average interpretation time differences were not statistically significant and all were less than 22 s between the two oral contrast media. Although not statistically significant, the two readers with little experience with VoLumen (Readers 2 and 3) registered shorter times with Gastrografin and the two readers who began their training using primarily VoLumen registered shorter times with VoLumen (Readers 4 and 5). The average difference in interpretation times for all readers was 4 s in cases with appendicitis (neutral oral contrast being faster; $p=0.4$) and 5 s in cases without appendicitis (positive oral contrast being faster; $p=0.2$). Average interpretation times were shorter in cases with appendicitis *vs* those without appendicitis, irrespective of contrast media ($p<0.001$).

Table 4. Specificity for the diagnosis of acute appendicitis (*i.e.* correctly diagnosing absence of appendicitis) per reader for positive *vs* neutral oral contrast

	Specificity (%) ^a Positive oral contrast (Gastrografin) (<i>n</i> =99 cases without appendicitis)	Specificity (%) ^a Neutral oral contrast (VoLumen) (<i>n</i> =102 cases without appendicitis)	<i>p</i> -Value ^b
Reader 1	97 (91–99)	95 (89–98)	0.7
Reader 2	98 (93–99)	93 (86–97)	0.17
Reader 3	97 (91–99)	90 (83–95)	0.08
Reader 4	92 (85–96)	95 (89–98)	0.4
Reader 5	95 (89–98)	97 (92–99)	0.5
Average of readers	(<i>n</i> =495) 96 (94–97)	(<i>n</i> =510) 94 (92–96)	0.24 (χ^2)

^aBinomial exact 95% confidence intervals used for confidence interval reporting. ^bFisher's exact test.

Figure 3. Receiver operating characteristic curves for all readers and for the average of the readers. 2 curves were generated for each reader, 1 using data from the 200 patients who received positive oral contrast and 1 from the 200 patients who received neutral oral contrast. Areas under the curve, 95% confidence intervals and the p -value comparing the areas under the curves are reported in each legend.

Discussion

The patients in the two oral contrast groups, both with and without appendicitis, were similar in sex and age. The proportion of cases identified with retrocecal appendices was also similar for the various groups, which excludes a possible bias owing to an overabundance of retrocecal

appendices in one of the groups. Our consecutively acquired cases without appendicitis represent a case distribution typically seen in clinical practice, with most aetiologies related to gastrointestinal, gynaecological or urinary problems. The proportion of patients (43%) with the diagnosis "abdominal pain, otherwise not specified" was similar to that reported in the literature, which ranges

Table 5. False-negative and false-positive CT results using positive and neutral oral contrast for diagnosing appendicitis

	False-negative interpretations		False-positive interpretations	
	Positive oral contrast (Gastrografin)	Neutral oral contrast (VoLumen)	Positive oral contrast (Gastrografin)	Neutral oral contrast (VoLumen)
Reader 1	2	3	3	5
Reader 2	6	7	2	7
Reader 3	4	3	3	10
Reader 4	1	2	8	5
Reader 5	9	11	5	3
Total	22 (out of 505 interpretations)	26 (out of 490 interpretations)	21 (out of 495 interpretations)	30 (out of 510 interpretations)
Proportion of appendices visualised	13/22 (59%)	16/26 (62%)	21/21 (100%)	27/30 (90%)

between 25% and 40% [16, 17] and reflects the reality of a busy county hospital emergency room.

Our study demonstrates that the ability to correctly diagnose acute appendicitis does not depend on the type of oral contrast material, regardless of the level of experience of the radiologist. No statistically significant differences were noted in sensitivity, specificity or area under the ROC curve between positive and neutral oral contrast. This lack of difference held true for all five independent readers representing various levels of experience and for all five readers averaged together. Our average sensitivity and specificity with positive oral contrast were similar to numbers reported in the literature, which reached 88–99% for sensitivity and 86–97% for specificity [1, 5, 19, 20]. No data are available today that address sensitivity and specificity in diagnosing appendicitis with neutral oral contrast. It is noteworthy that in patients with appendicitis, our average sensitivities of 96% with positive oral contrast and 95% with neutral oral contrast were similar to the sensitivities found in our institution's original CT reports for the same cases (99% for positive and 97% for neutral oral contrast, respectively). This indicates that the readers in this study did not achieve artificially high sensitivities owing to the study design, which included almost 50% of cases positive for appendicitis. In fact, the original interpretations had slightly higher but statistically not significantly different sensitivities, possibly because more clinical information was available. Considering the similarity in sensitivities between inexperienced and experienced readers, this study confirms the overall strength of CT in evaluating for appendicitis.

In our study, a normal appendix was more frequently identified with Gastrografin than with VoLumen. CT identification of a normal appendix is clinically important, as it unequivocally confirms the absence of appendicitis. It appears, therefore, that for identifying a normal appendix Gastrografin has an advantage over VoLumen. Moreover, in cases without appendicitis three out of five readers registered statistically significant higher confidence scores with Gastrografin, again suggesting an advantage of Gastrografin. However, the higher confidence with Gastrografin in non-appendicitis cases was related to unfamiliarity with neutral oral contrast (VoLumen) in two readers. By contrast, for patients with appendicitis no statistically significant difference was found for identifying the appendix or for the confidence score in diagnosing appendicitis. The proportion of normal appendices identified using positive oral contrast in this study is within the range of 36–100% reported in the literature [3, 5, 21].

A small number of cases of appendicitis were missed, particularly those with subtle findings. These false-negative diagnoses were encountered in almost equal numbers for both types of oral contrast materials. Multiple readers called false-positives in cases where inflammatory processes involved the right flank and right lower quadrant and mimicked changes that occur with appendicitis. Other than the inflammatory changes, no clear trend for false-positive diagnoses could be established for either type of contrast agent. For both groups of diagnoses and both types of oral contrast, the confidence scores were lower in cases where an erroneous diagnosis was made

Table 6. False-positive cases: diagnoses in cases in which more than one reader misclassified the case

Group	Diagnosis	Number of readers that misclassified case
Positive oral contrast	Pelvic inflammatory disease	3
	Perforated cecal cancer	3
	Diverticulitis	2
	Perforated duodenal ulcer	2
	RLQ pain, not otherwise specified (1 year follow-up)	2
Neutral oral contrast	Gastroenteritis	4
	RLQ pain, not otherwise specified (no follow-up)	3
	RLQ pain, not otherwise specified (no follow-up)	3
	Gastroenteritis	2
	RLQ pain, not otherwise specified (18 month follow-up)	2
	RLQ pain, not otherwise specified (7 month follow-up)	2

RLQ, right lower quadrant.

Figure 4. A 19-year-old woman with pelvic inflammatory disease and false-positive diagnosis of appendicitis by three of the five readers (Reader 1, an abdominal radiologist; Reader 3, a fellow; Reader 4, one of the first-year residents). This axial CT scan was performed with positive oral contrast. The appendix (arrowheads) is air-filled, but there is stranding in the mesentery (long arrow) and peritoneal enhancement (short arrow).

than in cases that were correctly diagnosed (p -values of <0.001 in all comparisons).

Overall, the interpretation times were similar between positive and neutral oral contrast groups. A statistically insignificant trend for slightly faster interpretation times was observed when readers were interpreting cases obtained with oral contrast with which they were more familiar.

We did not assess the value of Gastrografin *vs* VoLumen for diagnoses other than appendicitis. The value of neutral contrast, and in particular VoLumen, for assessing the bowel wall in neoplastic [22, 23], inflammatory and infectious disease of the gastrointestinal tract [12, 13, 24], and possibly even for small bowel obstruction [25], is documented in the literature. Comparing the slight advantage of positive oral contrast in confidently excluding appendicitis *vs* the possible benefits of neutral oral contrast in evaluating alternative diagnoses is beyond the scope of this study. At our institution, neutral oral contrast (VoLumen) is now routinely used because of the potential diagnostic benefits in a large number of conditions; however, positive oral contrast is still administered for specific indications (*e.g.* suspected

Figure 5. A 45-year-old male patient with positive oral contrast and a false-positive diagnosis of appendicitis by three of the five readers (Reader 3, a fellow; Readers 4 and 5, two first-year residents). The appendix appears enlarged (thin arrows) and the cecal wall is nodular and thickened (thick arrows). At surgery a cecal carcinoma with micro-perforation was found, which might have caused the apparent appendiceal inflammation. The positive oral contrast did not reach the cecum by the time of the scan.

post-operative leak or perforation) and at the radiologist's discretion. We now also routinely use neutral oral contrast agent for cases of suspected appendicitis, because our study showed no significant differences in the diagnostic accuracy for appendicitis and because of its potential benefits in evaluating for alternative diagnoses.

This study has several limitations. The retrospective nature of this study resulted in slight heterogeneity of the scanning parameters, but the variations were minor and similar between the two groups. The test set characteristics in this study did not exactly reflect clinical reality, as our test set was derived from a disease prevalence of 50%, which is higher than radiologists will encounter in their daily practice. Therefore, the results in this study might not be transferable to the general

Table 7. Time of interpretation

Time of interpretation, mean (IQR) (s)				Time of interpretation, mean (IQR) (s)		
Proven appendicitis				No appendicitis		
	Positive oral contrast (Gastrografin) $n=101$	Neutral oral contrast (VoLumen) $n=98$	p -Value ^a		Positive oral contrast (Gastrografin) $n=99$	Neutral oral contrast (VoLumen) $n=102$
Reader 1	135 (120–150)	131 (90–150)	0.7	Reader 1	143 (120–150)	148 (120–180)
Reader 2	154 (120–180)	174 (120–210)	0.10	Reader 2	158 (90–210)	174 (120–210)
Reader 3	77 (60–90)	80 (60–90)	0.7	Reader 3	91 (60–120)	113 (60–150)
Reader 4	132 (60–120)	113 (60–120)	0.04	Reader 4	161 (60–240)	157 (120–180)
Reader 5	152 (60–180)	131 (60–180)	0.16	Reader 5	173 (120–240)	159 (120–180)
Average of readers	130 (60–150)	126 (60–150)	0.4	Average of readers	145 (90–180)	150 (90–180)

IQR, interquartile range; range of times in parentheses.

^aWilcoxon rank sum.

population; however, the uniformity and homogeneity of these results support the interpretation that for the diagnosis of appendicitis it is acceptable to use neutral rather than positive oral contrast agents. A prospective study or, ideally, a study in which the same patients were scanned twice would lessen any possible bias. However, such a study would be impractical to perform in acutely ill patients and difficult to justify when applying the usual precepts of medical research ethics, as scanning the same patient twice would expose patients to additional radiation. Had this study included a large number of readers from each career stage (resident, fellow, attending) conclusions could have been drawn about these groups as a whole; however, this was not possible with our limited number of readers.

Conclusion

This study demonstrates that the use of positive (Gastrografin) *vs* neutral (VoLumen) oral contrast does not affect the diagnostic utility of contrast-enhanced CT in identifying acute appendicitis irrespective of the experience or the familiarity of the reader with the type of oral contrast used. Although positive oral contrast might help in the identification of a normal appendix, we consider this slight advantage to be insufficient to overcome the potential benefits of using neutral oral contrast, which we currently use in our daily practice.

References

- Anderson BA, Salem L, Flum DR. A systematic review of whether oral contrast is necessary for the computed tomography diagnosis of appendicitis in adults. *Am J Surg* 2005;190:474–8.
- Mun S, Ernst RD, Chen K, Oto A, Shah S, Mileski WJ. Rapid CT diagnosis of acute appendicitis with IV contrast material. *Emerg Radiol* 2006;12:99–102.
- Rao PM, Rhea JT, Novelline RA, McCabe CJ, Lawrason JN, Berger DL, et al. Helical CT technique for the diagnosis of appendicitis: prospective evaluation of a focused appendix CT examination. *Radiology* 1997;202:139–44.
- Gaitini D, Beck-Razi N, Mor-Yosef D, Fischer D, Ben Itzhak O, Krausz MM, et al. Diagnosing acute appendicitis in adults: accuracy of color Doppler sonography and MDCT compared with surgery and clinical follow-up. *AJR Am J Roentgenol* 2008;190:1300–6.
- Jacobs JE, Birnbaum BA, Macari M, Megibow AJ, Israel G, Maki DD, et al. Acute appendicitis: comparison of helical CT diagnosis focused technique with oral contrast material versus nonfocused technique with oral and intravenous contrast material. *Radiology* 2001;220:683–90.
- Kaiser S, Finnbogason T, Jorulf HK, Soderman E, Frenckner B. Suspected appendicitis in children: diagnosis with contrast-enhanced versus nonenhanced helical CT. *Radiology* 2004;231:427–33.
- Iwahashi N, Kitagawa Y, Mayumi T, Kohno H. Intravenous contrast-enhanced computed tomography in the diagnosis of acute appendicitis. *World J Surg* 2005;29:83–7.
- Hershko DD, Awad N, Fischer D, Mahajna A, Guralnik L, Israelit SH, et al. Focused helical CT using rectal contrast material only as the preferred technique for the diagnosis of suspected acute appendicitis: a prospective, randomised, controlled study comparing three different techniques. *Dis Colon Rectum* 2007;50:1223–9.
- Stroman DL, Bayouth CV, Kuhn JA, Westmoreland M, Jones RC, Fisher TL, et al. The role of computed tomography in the diagnosis of acute appendicitis. *Am J Surg* 1999;178:485–9.
- Kharbanda AB, Taylor GA, Bachur RG. Suspected appendicitis in children: rectal and intravenous contrast-enhanced versus intravenous contrast-enhanced CT. *Radiology* 2007;243:520–6.
- Arslan H, Etlik O, Kayan M, Harman M, Tuncer Y, Temizoz O. Peroral CT enterography with lactulose solution: preliminary observations. *AJR Am J Roentgenol* 2005;185:1173–9.
- Megibow AJ, Babb JS, Hecht EM, Cho JJ, Houston C, Boruch MM, et al. Evaluation of bowel distention and bowel wall appearance by using neutral oral contrast agent for multi-detector row CT. *Radiology* 2006;238:87–95.
- Paulsen SR, Huprich JE, Fletcher JG, Booya F, Young BM, Fidler JL, et al. CT enterography as a diagnostic tool in evaluating small bowel disorders: review of clinical experience with over 700 cases. *Radiographics* 2006;26:641–57; discussion 657–62.
- Macari M, Megibow AJ, Balthazar EJ. A pattern approach to the abnormal small bowel: observations at MDCT and CT enterography. *AJR Am J Roentgenol* 2007;188:1344–55.
- Raman SS, Lu DS, Kadell BM, Vodopich DJ, Sayre J, Cryer H. Accuracy of nonfocused helical CT for the diagnosis of acute appendicitis: a 5-year review. *AJR Am J Roentgenol* 2002;178:1319–25.
- Graff LG 4th, Robinson D. Abdominal pain and emergency department evaluation. *Emerg Med Clin North Am* 2001;19:123–36.
- Powers RD, Guertler AT. Abdominal pain in the ED: stability and change over 20 years. *Am J Emerg Med* 1995;13:301–3.
- DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837–45.
- Mittal VK, Goliath J, Sabir M, Patel R, Richards BF, Alkalay I, et al. Advantages of focused helical computed tomographic scanning with rectal contrast only vs triple contrast in the diagnosis of clinically uncertain acute appendicitis: a prospective randomized study. *Arch Surg* 2004;139:495–9; discussion 499–500.
- Rhea JT, Halpern EF, Ptak T, Lawrason JN, Sacknoff R, Novelline RA. The status of appendiceal CT in an urban medical center 5 years after its introduction: experience with 753 patients. *AJR Am J Roentgenol* 2005;184:1802–8.
- Rao PM, Rhea JT, Novelline RA, Mostafavi AA, Lawrason JN, McCabe CJ. Helical CT combined with contrast material administered only through the colon for imaging of suspected appendicitis. *AJR Am J Roentgenol* 1997;169:1275–80.
- Antoch G, Kuehl H, Kanja J, Lauenstein TC, Schneemann H, Hauth E, et al. Dual-modality PET/CT scanning with negative oral contrast agent to avoid artifacts: introduction and evaluation. *Radiology* 2004;230:879–85.
- Prabhakar HB, Sahani DV, Fischman AJ, Mueller PR, Blake MA. Bowel hot spots at PET-CT. *Radiographics* 2007;27:145–59.
- Hara AK, Leighton JA, Heigh RI, Sharma VK, Silva AC, De Petris G, et al. Crohn disease of the small bowel: preliminary comparison among CT enterography, capsule endoscopy, small-bowel follow-through, and ileoscopy. *Radiology* 2006;238:128–34.
- Zhang LH, Zhang SZ, Hu HJ, Gao M, Zhang M, Cao Q, et al. Multi-detector CT enterography with iso-osmotic mannitol as oral contrast for detecting small bowel disease. *World J Gastroenterol* 2005;11:2324–9.