Effectiveness of a Staged US and Unenhanced MR Imaging Algorithm in the Diagnosis of Pediatric Appendicitis

Elizabeth H. Dibble, MD
David W. Swenson, MD
Claudia Cartagena, MS, PA-C
Grayson L. Baird, PhD
Thaddeus W. Herliczek, MD

Purpose:
To establish, in a large cohort, the diagnostic performance of a staged algorithm involving ultrasonography (US) followed by conditional unenhanced magnetic resonance (MR) imaging for the imaging work-up of pediatric appendicitis.

Materials and Methods:
A staged imaging algorithm in which US and unenhanced MR imaging were performed in pediatric patients suspected of having appendicitis was implemented at the authors’ institution on January 1, 2011, with US as the initial modality followed by unenhanced MR imaging when US findings were equivocal. A search of the radiology database revealed 2180 pediatric patients who had undergone imaging for suspected appendicitis from January 1, 2011, through December 31, 2012. Of the 2180 patients, 1982 (90.9%) were evaluated according to the algorithm. The authors reviewed the electronic medical records and imaging reports for all patients. Imaging reports were reviewed and classified as positive, negative, or equivocal for appendicitis and correlated with surgical and pathology reports.

Results:
The frequency of appendicitis was 20.5% (407 of 1982 patients). US alone was performed in 1905 of the 1982 patients (96.1%), yielding a sensitivity of 98.7% (386 of 391 patients) and specificity of 97.1% (1470 of 1514 patients) for appendicitis. Seventy-seven patients underwent unenhanced MR imaging after equivocal US findings, yielding an overall algorithm sensitivity of 98.2% (400 of 407 patients) and specificity of 97.1% (1530 of 1575 patients). Seven of the 1982 patients (0.4%) had false-negative results with the staged algorithm. The negative predictive value of the staged algorithm was 99.5% (1530 of 1537 patients).

Conclusion:
A staged algorithm of US and unenhanced MR imaging for pediatric appendicitis appears to be effective. The results of this study demonstrate that this staged algorithm is 98.2% sensitive and 97.1% specific for the diagnosis of appendicitis in pediatric patients.

© RSNA, 2017
Although appendicitis is the leading reason for emergency surgery in the pediatric population (1), its clinical diagnosis remains challenging (2,3). Diagnostic imaging therefore plays a critical role in the evaluation of pediatric patients with acute abdominal pain, but the optimal diagnostic imaging approach has been controversial (4,5).

Historically, our practice followed a staged appendicitis imaging algorithm involving initial ultrasonography (US) followed by computed tomography (CT) in patients with equivocal US findings, as described by García Peña et al (6) and recommended by the American College of Radiology (7). This algorithm reduces radiation exposure by limiting the use of CT and has been reported to be 98.6% sensitive and 90.6% specific in the diagnosis of pediatric appendicitis (8).

Almost 20 years ago, Hörrmann et al (9) suggested magnetic resonance (MR) imaging as an alternative to CT in pediatric patients with inconclusive findings at US of the appendix in an effort to minimize radiation exposure. During the past decade, advances in MR imaging have improved its utility for abdominopelvic imaging; there is a growing body of literature describing its use in the evaluation of appendicitis in pregnant women and in children (10–15). Two recent meta-analyses of MR imaging in pediatric patients with appendicitis found pooled sensitivities of 96% and 96.3% and specificities of 96% and 96.1%, respectively (16,17).

Although the results of these studies are encouraging, there is variability among described MR imaging protocols and approaches to integrating MR imaging with previous institutional practices, typically a combination of US and CT. The literature may incite enthusiasm for MR imaging as an accurate diagnostic tool for pediatric appendicitis, but there remains no established best practice. In January 2011, the members of our practice made a commitment to replace CT with unenhanced MR imaging for pediatric patients suspected of having appendicitis who have equivocal US findings. The purpose of our study was to determine, in a large cohort, the diagnostic performance of a staged algorithm that includes initial US followed by conditional unenhanced MR imaging for the imaging work-up of patients with pediatric appendicitis.

**Materials and Methods**

This Health Insurance Portability and Accountability Act–compliant retrospective study was approved by our institutional review board; the requirement to obtain informed consent was waived.

**Study Sample**

Our institution is a tertiary referral center that includes the state’s only pediatric hospital. We perform approximately 20,000–25,000 pediatric imaging studies annually—two-thirds of which are performed through the emergency department. Our sample included all patients aged 18 years or younger who underwent imaging for suspected appendicitis from January 1, 2011, to December 31, 2012. A total of 2180 pediatric patients underwent imaging to assess for acute appendicitis during the study period: 1023 boys (46.9%; mean age, 11.2 years; age range, 0.6–18 years) and 1157 girls (53.1%; mean age, 12.7 years; age range, 0.6–18 years) (P < .0001), with an overall mean age of 11.5 years (range, 0.6–18 years). Of the 2180 patients, 198 (9.9%) had undergone imaging outside of the algorithm by undergoing either MR imaging or CT before US, by undergoing US followed by CT, or by having equivocal US results but no subsequent MR examination. Therefore, 1982 of the 2180 patients (90.1%; 933 boys with a mean age of 10.5 years [range, 1.2–18 years] and 1049 girls with a mean age of 11.9 years [range, 6–18 years], P < .0001; overall mean age, 11.2 years [range, 1.2–18 years]) underwent imaging under the staged US and unenhanced MR imaging algorithm and thus met inclusion criteria. The frequency of appendicitis in patients who underwent imaging according to the algorithm was 20.5% (407 of 1982 patients). We postulated that the slightly greater number of girls and observed age difference between boys and girls were the result of symptom overlap between appendicitis and ovarian conditions in older girls, although we did not examine this as part of the current study. Figure 1 illustrates patient characteristics.

The electronic medical record was searched to confirm that examinations had been performed because of clinical suspicion of appendicitis. Imaging
Clinical criteria used to decide whether to perform imaging in patients suspected of having appendicitis are left to the discretion of the referring provider.

A staged imaging algorithm was implemented on January 1, 2011; in this algorithm, US and unenhanced MR imaging are performed in pediatric patients suspected of having appendicitis, with US as the initial modality followed by unenhanced MR imaging when US findings are equivocal. For this study, US findings were considered equivocal if the appendix was not identified and if either (a) there were secondary US signs of appendicitis (right lower quadrant inflammation, intraperitoneal collection, or complex free fluid) or (b) there was high clinical concern for appendicitis. Clinical criteria used to decide whether to perform unenhanced MR imaging after inconclusive US findings were left to the discretion of the referring provider. Patients with equivocal US findings who did not undergo subsequent MR imaging were considered to fall outside of the algorithm; it was assumed that clinical judgment was used in place of the imaging algorithm.

**US Technique**

US was performed with the Logic E9 or Logic E9XD system (GE Healthcare, Waukesha, Wis). A linear-array transducer (6–15 MHz and/or 9 MHz, GE Healthcare) was used to assess the appendix. US was performed by following the graded compression technique to assess the appendix for size, mural thickening, mural hyperemia, appendicolith, periappendiceal inflammation, intraperitoneal fluid, and/or collections (6,20,21). All US examinations of the appendix were performed by registered diagnostic medical sonographers experienced in pediatric US. An attending pediatric radiologist always performed real-time scanning when US findings were equivocal. An attending pediatric radiologist also frequently performed real-time scanning in equivocal cases, but, because of inconsistent documentation in the US reports, it is not possible to determine from review of the reports which patients had undergone additional real-time scanning by an attending pediatric radiologist.

**MR Imaging Technique**

MR examinations were performed with either a 1.5-T system (Magnetom Espree, Aera, or Symphony; Siemens Medical Systems, Erlangen, Germany) or a 3.0-T system (Verio; Siemens Medical Systems). Both field strengths were used throughout the study period, and patients underwent imaging with...
the most readily available MR imaging system regardless of field strength.

MR images were obtained from the inferior poles of the kidneys through the urinary bladder without the use of sedation or contrast media. The pediatric appendix MR imaging protocol varied slightly during the course of the study and was reviewed regularly by the division director to optimize visualization of the appendix and decrease total examination time; however, all examinations included at least the following pulse sequences: (a) an axial short-inversion-time inversion-recovery sequence and (b) axial and coronal T2-weighted half-Fourier acquisition single-shot turbo-spin-echo imaging without fat saturation and/or a coronal T2-weighted three-dimensional turbo-spin-echo sequence with multiplanar reconstructions (Table).

Image Interpretation

US images were evaluated by fellowship-trained pediatric radiologists (T.W.H., with 3 years of experience, and five other radiologists with 2, 12, 20, 22, and 35 years of experience). MR images were evaluated by the same fellowship-trained pediatric radiologists or by two fellowship-trained adult body radiologists with 4 and 9 years of experience, respectively.

Reference Standard

Patient electronic medical records were reviewed, including emergency room notes, admission notes, surgical reports, pathology reports, imaging examination reports, progress notes, and notes regarding follow-up visits when available.

Patients with acute appendicitis documented at pathologic evaluation were considered to have true-positive results. Patients who did not undergo an appendectomy at our institution within 4 weeks of initial presentation and imaging or who had a normal appendix documented on the surgical pathology report were considered to have true-negative results. The electronic medical records of patients discharged without appendectomy were reviewed for length of stay, discharge diagnosis, and subsequent appendectomy at our institution. Electronic medical record review occurred between 9 and 45 months after completion of imaging examinations.

Statistical Analysis

All analyses were conducted with software (SAS, version 9.4; SAS Institute, Cary, NC). Sensitivity, specificity, positive predictive value, and negative predictive value of the staged algorithm overall and according to its parts were examined by using generalized linear modeling assuming a binary distribution with a logit link with PROC GLIMMIX. Differences in counts were assessed by using the Fisher exact test (two tailed), given the small sample size. Differences in age were assessed by using an independent samples t test. All interval estimates were calculated for 95% confidence, and alpha was established at the .05 level.

Results

Sensitivity, Specificity, and Positive and Negative Predictive Values

US.—In total, 1905 patients had unequivocal US findings, with 386 true-positive, 1470 true-negative, 44 false-positive, and five false-negative results. The sensitivity of US with unequivocal findings was 98.7% (386 of 391 patients; 95% confidence interval [CI]: 96.96%, 99.47%), with a positive predictive value of 89.8% (386 of 430 patients; 95% CI: 86.52%, 92.30%); the specificity was 97.1% (1470 of 1514 patients; 95% CI: 96.12%, 97.83%), with a negative predictive value of 99.7% (1470 of 1475 patients; 95% CI: 99.19%, 99.86%).

MR imaging.—In total, 77 patients underwent unenhanced MR imaging when US results were equivocal, with 14 true-positive results (Fig 2), 60 true-negative results, one false-positive result, and two false-negative results. The sensitivity of MR imaging was 87.5% (14 of 16 patients; 95% CI: 60.83%, 96.93%), with a positive predictive value of 93.3% (14 of
UNEQUIVOCAL VERSUS EQUIVOCAL US FINDINGS

Of the 391 patients with appendicitis and unequivocal US findings, 291 (74%) had nonperforated appendicitis and 100 (26%) had perforated appendicitis. Of the 16 patients with appendicitis and equivocal US findings followed by MR imaging, 12 (75%) had nonperforated appendicitis and four (25%) had perforated appendicitis (P = .96).

Among patients admitted to our hospital, the median length of stay was 2 days for patients with unequivocal US findings (range, 1–30 days) and 2 days for patients with equivocal US findings (range, 1–12 days).

MR IMAGING

The median time from US to MR imaging after equivocal US findings was 227 minutes (range, 96–7,628 minutes). Sixty-eight percent (52 of 77) of studies were performed with a 1.5-T magnet, and 32% (25 of 77) of studies were performed with a 3.0-T magnet.

Among the equivocal US results in which the patient went on to MR imaging, the appendix was not seen in four of the 77 (5.2%) MR examinations, was seen in 71 (92%), and was partially seen in two (2.6%).

Seventy-five of the 77 MR images (97%) were interpreted by a fellowship-trained pediatric radiologist and two (3%) were interpreted by a fellowship-trained body radiologist.

Appendicitis was not seen at MR imaging in 62 of the 77 patients; 17 of the 62 patients (27.4%) had alternative findings identified at MR imaging: 14 patients (82.4%) had specific alternative findings, and three (17.6%) had nonspecific free fluid. Of the 14 patients with specific alternative findings at MR imaging, 12 (85.6%) had findings identified at MR imaging that became the discharge diagnosis and presumed cause of the patients' abdominal pain (ovarian or paraovarian cyst, n = 6; omental infarct, n = 2; cholelithiasis, n = 1; enteritis, n = 1; mesenteric adenitis, n = 1; musculoskeletal strain, n = 1). One MR examination depicted a possible mildly distended distal ureter, but there were no clinical findings of genitourinary tract abnormalities including urinary tract calculi at subsequent work-up.

15 patients; 95% CI: 64.04%, 99.10%); the specificity was 98.4% (60 of 61 patients; 95% CI: 88.95%, 99.78%), with a negative predictive value of 96.8% (60 of 62 patients; 95% CI: 87.73%, 99.21%).

Staged algorithm.—In total, the staged algorithm was followed in 1982 of the 2180 of patients (90.9%), with 400 true-positive, 1530 true-negative, 45 false-positive, and seven false-negative results. The algorithm false-positive rate was 2.3% (45 of 1982 patients). Of the 45 false-positive algorithm results, four patients went on to appendectomy, presumably because clinical judgment altered management in 41 cases (eg, a pediatric surgeon deemed appendicitis unlikely). The negative appendectomy rate was 0.2% (four of 1982 patients). Overall, the staged algorithm had a sensitivity of 98.3% (400 of 407 patients; 95% CI: 96.44%, 99.18%), with a positive predictive value of 89.9% (400 of 445 patients; 95% CI: 86.72%, 92.37%); the specificity was 97.1% (1530 of 1575 patients; 95% CI: 96.19%, 97.86%), with a negative predictive value of 99.5% (1530 of 1537 patients; 95% CI: 99.05%, 99.78%).
Another MR examination depicted an omental infarct, but the discharge diagnosis was gastroenteritis.

Discussion

Our staged algorithm of US followed by unenhanced MR imaging in patients with equivocal US findings was 98.2% sensitive and 97.1% specific for acute appendicitis. This algorithm has similar effectiveness to the US and contrast agent–enhanced CT algorithm described in 2011 by Krishnamoorthi et al (8), which was 98.6% sensitive and 90.6% specific, but our algorithm does not expose patients to ionizing radiation, thereby following the “as low as reasonably achievable,” or ALARA, principle (22,23). The sensitivity and specificity with unequivocal US as reported by Krishnamoorthi et al were 98.7% and 86.6%, respectively, versus 98.2% and 97.1% in our study, and they had a 47.2% utilization rate of CT after equivocal US findings versus our 13% MR imaging utilization rate. They reported a sensitivity and specificity of CT after equivocal US findings of 98.4% and 93.6%, respectively, versus 87.5% and 98.4% with MR imaging in our study.

Our algorithm also does not expose patients to intravenous contrast media, an advantage over imaging algorithms involving both staged US with CT and staged US with contrast-enhanced MR imaging. Aspelund et al (24) demonstrated the effectiveness of an algorithm involving US and contrast-enhanced MR imaging in 397 patients, reporting 100% sensitivity and 99% specificity; however, 16.6% of patients underwent initial CT (5.0%), initial MR imaging (3.0%), or CT after equivocal US findings (8.6%) (24). In addition, 56.4% of US studies had equivocal results and 15% of patients with equivocal US findings did not undergo additional imaging. The equivocal US rate likely resulted in increased MR imaging utilization (43% of patients in the US and MR imaging group compared with 13% in our study, when US of the appendix was performed only in the presence of an attending pediatric radiologist during restricted hours). The average hospital stay for our patients was 2 days, similar to 43.4 hours in their study. Our algorithm has similar sensitivity, specificity, and length of stay, plus lower MR imaging utilization and lack of exposure to intravenous contrast media.

Like Aspelund et al, Trout et al (25) reported a low diagnostic US yield of 24.4% for pediatric appendicitis. The authors concluded that “US does not perform as purported in the literature. . . . To improve the diagnostic performance of this modality, involvement by experienced personnel and/or additional training is needed.” Our study shows that high diagnostic accuracy is possible when US is performed in a subspecialized pediatric imaging center.

The utilization rate of MR imaging after US varies in the literature. Thieme et al (26) reported a utilization rate of 53% after inconclusive or negative US findings versus our rate of 13% after equivocal US findings. In their study, US was performed 24 hours per day by trainees in radiology, likely resulting in a higher rate of equivocal results (42% vs 13% in our study). In addition, MR imaging was performed in all patients with negative US findings. Dillman et al (27) assessed the effectiveness of unenhanced MR imaging after equivocal US findings and reported an MR imaging utilization rate of 15%, which is similar to our rate of 13%. Their US examinations were also performed under the direct supervision of pediatric radiologists. The diagnostic utility in their study (85%) was similar to that in our study (87%). Although Dillman et al (27) compared the performance of CT and MR imaging after equivocal US findings rather than the effectiveness of a staged algorithm, the sensitivity and specificity of MR imaging were similar to ours—at 94.4% and 100%, respectively.

Two recent meta-analyses found pooled MR imaging sensitivities of 96% and 96.3% and specificities of 96% and 96.1% for pediatric appendicitis (16,17). Contrast-enhanced MR imaging was 94%-99.2% sensitive and 95.7%-100% specific (17,24,28,29). Unenhanced MR imaging was 92.1%-99.2% sensitive and 88.3%-99.4% specific (13-15,17,26,27,30-32). Of 77 patients undergoing unenhanced MR imaging after equivocal US findings in our study, 16 had appendicitis. Unenhanced MR imaging after equivocal US findings was 87.5% sensitive and 98.4% specific in our study. Our false-negative and false-positive MR imaging results may have been due to reader inexperience because this algorithm was newly introduced at our institution in 2011. One false-negative MR imaging result was interpreted as a normal appendix, but the patient underwent exploratory laparotomy for worsening abdominal pain and peritoneal signs; the postprocedure note described an incidental appendectomy, and the final surgical report described an indurated but not inflamed appendix. The final pathology report described focal minimal inflammation suggestive of early appendicitis. The other false-negative MR imaging result was interpreted as showing no evidence of acute appendicitis with a prominent appendix measuring 8–9 mm. The patient underwent laparotomy because of diffuse worsening abdominal pain and clinical suspicion for perforated appendicitis. Surgical findings included appendicitis with purulent ascites. Pathologic review revealed acute appendicitis and periappendicitis. Our single false-positive MR imaging report described a minimally thickened appendix with edema in the wall, suggestive of early appendicitis. The patient’s condition improved with conservative treatment; the discharge diagnosis was abdominal pain. The similar diagnostic performances of enhanced MR imaging (94%-99.2% sensitive and 95.7%-100% specific [17,24,28,29]) and unenhanced MR imaging (92.1%-99.2% sensitive and 88.3%-99.4% specific [13-15,17,26,27,30-32]) suggest that use of contrast material may not be necessary; cessation of its use would eliminate the cost plus the risks of gadolinium deposition in the brain (33) and contrast-related nephrogenic systemic fibrosis (34).

Moore et al (14) postulated that an algorithm of US followed by MR imaging might be the best imaging algorithm for pediatric right lower quadrant pain. We, too, believe that a staged imaging
algorithm of unenhanced MR imaging following equivocal US findings is best practice for imaging in suspected pediatric acute appendicitis. Although some authors have suggested that MR imaging may supplant CT in imaging of pediatric appendicitis (15) and may be appropriate as a first-line test (16), our study demonstrated excellent results with US alone but an even higher diagnostic yield when the US and unenhanced MR imaging staged algorithm was used, with an MR imaging utilization rate of only 13%. We believe that an algorithm of US followed by unenhanced MR imaging optimizes patient care, patient safety, efficiency, and resource utilization. MR imaging as the first-line modality for suspected pediatric appendicitis may have sufficient throughput at institutions with that capability, and a recent study by Petkovska et al (35) found unenhanced MR imaging to be 97.0% sensitive and 99.3% specific for acute appendicitis in patients aged 3–50 years. A negative MR imaging result might allow for earlier discharge, and MR imaging in equivocal cases can add value by providing alternative diagnoses. A staged US and MR imaging algorithm would not be an efficient use of resources in this pediatric plus adult population, given the inherent challenges of performing US in larger patients. Because of high MR imaging cost and high US sensitivity and specificity in pediatric patients, expectant management in equivocal US cases is another potentially cost-effective approach. Further research is needed to demonstrate the efficiency and cost-effectiveness of our algorithm versus other algorithms for pediatric right lower quadrant pain in different clinical settings.

Our study has several limitations, including the retrospective design, limited clinical follow-up for true-negative findings, range in time from US to MR imaging, and variation in MR imaging readers, field strength, and protocols as our practice evolved. Although possible, it is unlikely that our patients with true-negative results underwent appendectomy elsewhere after discharge because our institution is the only pediatric hospital in the state. The median time from equivocal US findings to MR imaging was less than 4 hours, but we did not exclude patients with longer lag times if MR imaging was performed for suspected appendicitis. We included all pediatric MR examinations performed after US findings that were equivocal for appendicitis, whether interpreted by a pediatric or body radiologist. Nevertheless, our results are representative of real-world outcomes and therefore may be more reflective of variations in practice patterns. Although a comparison of the usefulness of 1.5- and 3.0-T systems would be of interest because of the shorter acquisition time and greater signal-to-noise ratio of 3.0-T systems could benefit smaller, younger patients, we did not attempt to compare different field strengths. We performed imaging in patients with inconclusive US findings on the next available MR imaging system at our institution regardless of field strength. Finally, our protocol varied during the study period due to intermittent review by the division director in an attempt to optimize visualization of the appendix and decrease total examination time. Establishing the optimum protocol was not the goal of this study; because all examinations included at least single-plane short-inversion-time inversion-recovery and two planes of T2-weighted single-shot turbo-spin-echo and/or T2-weighted three-dimensional turbo-spin-echo sequences, we do not believe the variation in protocol altered the effectiveness of the algorithm during the study period.

The results of our study demonstrate that a staged algorithm of US followed by unenhanced MR imaging can be highly sensitive and specific for pediatric appendicitis. When MR imaging is available, the role of CT in the assessment of suspected pediatric appendicitis should be minimized.


References


