Effect of Cross-sectional Imaging on Negative Appendectomy and Perforation Rates in Children

PURPOSE: To compare negative appendectomy and perforation rates in children who underwent ultrasonography (US), computed tomography (CT), or no imaging before urgent appendectomy.

MATERIALS AND METHODS: All children who underwent urgent appendectomy during a 4½-year period were identified in a surgical billing database. Pathology reports were coded as negative or as showing acute inflammation or perforation. Imaging up to 14 days before appendectomy or abscess drainage was noted, and imaging-based diagnoses were compared with pathologic findings. Patient age and sex were recorded.

RESULTS: Two hundred ninety-nine children, 176 (59%) male and 123 (41%) female (mean age, 10.4 years; age range, 1–21 years), underwent urgent appendectomy. One hundred twenty-six (42%) underwent no imaging, 121 (41%) underwent US with or without CT, and 52 (17%) underwent CT only; 44 (15%) underwent both US and CT. There were significantly higher rates of appendectomy with normal pathologic findings ("negative appendectomy") in patients who underwent no imaging (14% [18 of 126]) or US (17% [20 of 121]) versus the rates in those who underwent CT only (2% [one of 52]) (P = .02 and P = .007, respectively). The negative appendectomy rate was 7% in 96 patients who underwent CT with or without prior US. The perforation rates were not significantly different.

CONCLUSION: As compared with children who underwent no preoperative imaging and those who underwent US, children who underwent CT had a significantly lower negative appendectomy rate, without a significantly higher perforation rate.

Urgent appendectomy remains the most common abdominal surgery in children, with approximately 60,000–80,000 performed in the United States each year (1). Correctly distinguishing children who have acute appendicitis from those who have nonsurgical conditions that mimic appendicitis remains a challenge (2). This is particularly true for young children, who may have nonspecific symptoms but may be unable to communicate effectively.

By using an administrative database, Wen and Naylor (3) studied rates of urgent appendectomy for presumed acute appendicitis with normal pathologic findings (hereafter, "negative appendectomy") and perforation, in-hospital deaths, and length of stay for all children who underwent appendectomy in Ontario, Canada, during an 11-year period. They reported great variation in these rates and concluded that some undefined proportion of appendectomies could be safely avoided. Traditionally, a false-negative appendectomy rate of 15%–20% has been accepted in surgical practice as an appropriate benchmark owing to the difficulty in diagnosing the condition and the serious consequences associated with missed or delayed diagnosis (1,4–8). Recently, this paradigm has been challenged because computed tomography (CT) has been shown by investigators to substantially decrease negative appendectomy rates in adults (9). Investigators (9–12) have suggestive but not statistically significant data showing lower negative appendectomy rates in children who undergo CT prior to appendectomy.
TABLE 1
Distribution of 299 Children and Adults Who Underwent Urgent Appendectomy, according to Pathology Report

<table>
<thead>
<tr>
<th>Patient Demographic</th>
<th>Negative Appendicitis (n = 39)</th>
<th>Nonperforated Appendicitis (n = 188)</th>
<th>Perforated Appendicitis* (n = 72)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male (n = 176)</td>
<td>19 (11)</td>
<td>111 (63)</td>
<td>46 (26)</td>
</tr>
<tr>
<td>Female (n = 123)</td>
<td>20 (16)</td>
<td>77 (63)</td>
<td>26 (21)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤10 y (n = 143)</td>
<td>19 (13)</td>
<td>86 (60)</td>
<td>38 (27)</td>
</tr>
<tr>
<td>&gt;10 y (n = 156)</td>
<td>20 (13)</td>
<td>102 (65)</td>
<td>34 (22)</td>
</tr>
</tbody>
</table>

Note.—Data are the number of patients. Numbers in parentheses are percentages.
* Perforation rates were based on positive pathologic findings.

MATERIALS AND METHODS

Design, Setting, and Subjects

This retrospective cross-sectional study was performed at a large children’s hospital. Potential subjects were individuals aged 21 years and younger who underwent urgent appendectomy. The review period was 4½ years (January 1, 1996 to July 1, 2000). Subjects were identified by finding current procedural terminology, or CPT, codes 44950, 44960, and 56315 in the pediatric surgical billing database. Patients were excluded if chart review revealed incidental appendectomy or if the abdominal pain had been present longer than 2 weeks prior to surgery or abscess drainage followed by interval appendectomy approximately 5–10 weeks after CT-guided abscess drainage for perforated appendicitis treatment. This 2-week time limit was used to be as inclusive as possible for patients who received a delayed diagnosis of acute appendicitis.

This study received administrative approval from the hospital institutional review board for human investigation. Patient informed consent was waived.

These summer research students reviewed patient medical charts and the radiology and pathology departments’ electronic databases. The first author trained them and reviewed their coding. The pathology report was used as the reference standard for comparison with the imaging results report. The patient’s age and sex were recorded. The appendectomy pathology reports were coded as negative or as showing acute inflammation or perforation.

US and CT were performed either at presentation to the emergency department or after hospital admission for observation for possible appendicitis. Imaging studies were included in the analysis if performed within 14 days prior to either appendectomy or CT-guided abscess drainage. The US and CT reports were coded as negative or positive for appendicitis. US and CT reports describing equivocally positive findings of appendicitis were analyzed with the positive cases.

The US and CT techniques used in these patients and the imaging criteria for diagnosing appendicitis have been described previously (25). US was performed using a 5- or 7-MHz linear array transducer (XP-128, Acuson, Mountain View, Calif). Both a technologist and a CAQ, or certificate of added qualification, certified pediatric radiologist (including K.E.A. and C.J.S) used a graded-compression technique to examine the right lower quadrant of the abdomen. Positive criteria were a noncompressible appendix with a maximum diameter of greater than 6 mm and an appendicolith with or without fluid or signs of inflammation surrounding the appendix.

CT (PQ 5000; Picker, Highland Heights, Ohio) was performed using both intravenously (iothalamate meglumine, Conray 43; Mallinkrodt, St Louis, Mo) and either orally or rectally administered contrast material (3% diatrizoate meglumine solution, Gastrogrip; Bristol-Myers Squibb, Wallingford, Conn), 120 kV, 100–175 mAs, and a pitch of 1. Image collimation at CT varied depending on patient age: from 4 mm in patients younger than 18 months to 10 mm in patients older than 6 years. Collimation at the level of the appendix in these older patients was changed from 10 to 4 mm halfway through the study period. In the second half of the study period, contrast material administration for bowel imaging was changed from oral to rectal. US and CT images were interpreted by one of the six CAQ-certified pediatric radiologists by using a workstation or hard copies. When the imaging report described probable or suggestive findings of appendicitis, the report was coded as positive for appendicitis.

The negative and perforated appendicitis rates were compared between the children who underwent no cross-sectional imaging, US with or without CT, or CT only prior to urgent appendectomy. In those children with pathology and cross-sectional imaging reports, the false-positive and false-negative rates with US and CT were calculated.

To assess for trends during the study period, the data were stratified into early (in the initial 24 months) versus late (in the last 30 months) study periods and analyzed for negative appendectomy and perforation rates.

Statistical Analysis

All data were analyzed by using statistical software (SAS version 6.12; SAS, Cary, NC). The x² test for nonparametric data was performed to compare groups. When appropriate, for small samples, the Fisher exact test was performed. All tests were two-sided and indicated a significant difference if the P value was less than or equal to .05.

RESULTS

The final study cohort consisted of 299 children, 176 (59%) male patients with a mean age of 10.5 years (age range, 1–21 years) and 123 female patients with a mean age of 10.1 years (age range, 2–20 years). Table 1 shows the distribution of cases of negative, perforated, and nonperforated appendicectomy overall and by patient age and sex. The overall negative appendectomy rate was 13% (39 of 299 patients). The overall perforation rate was 24% (72 of 299 patients).

The mean ages of the patients who underwent no imaging (10.4 years), US (10.2 years), or CT (10.4 years) were similar. However, the sex distribution differed. More males underwent no imaging prior to appendectomy, as compared
### TABLE 2
Negative Appendectomy and Perforation Rates in Children who Underwent Either Cross-sectional or No Imaging

<table>
<thead>
<tr>
<th>Imaging Examination</th>
<th>Negative Appendix (%)</th>
<th>Perforated Appendix (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>14</td>
<td>27</td>
</tr>
<tr>
<td>US with or without CT</td>
<td>17</td>
<td>32</td>
</tr>
<tr>
<td>CT with or without US</td>
<td>7*</td>
<td>29</td>
</tr>
<tr>
<td>CT only</td>
<td>21*</td>
<td>20</td>
</tr>
</tbody>
</table>

Note.—The perforation rates were not significantly different.
* CT versus no imaging, $P = .10$; CT versus US, $P = .005$.
† CT only versus no imaging, $P = .02$; CT only versus US, $P = .007$.

### TABLE 3
Cross-sectional Imaging Results with Negative Appendectomy and Perforation Rates from Pathology Reports

<table>
<thead>
<tr>
<th>Imaging Examination</th>
<th>Negative Appendix (%)</th>
<th>Perforated Appendix (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>None (n = 126)</td>
<td>18 (14)</td>
<td>29/108 (27)</td>
</tr>
<tr>
<td>US only (n = 77)</td>
<td>14 (18)</td>
<td>17/63 (27)</td>
</tr>
<tr>
<td>CT only (n = 52)</td>
<td>1 (2)</td>
<td>10/51 (20)</td>
</tr>
<tr>
<td>US with or without CT (n = 121)</td>
<td>20 (17)</td>
<td>32/101 (32)</td>
</tr>
<tr>
<td>CT with or without prior US (n = 96)</td>
<td>7 (7)</td>
<td>26/89 (29)</td>
</tr>
<tr>
<td>CT and US (n = 44)</td>
<td>6 (14)</td>
<td>16/38 (42)</td>
</tr>
</tbody>
</table>

Note.—Data are the number of patients. Numbers in parentheses are percentages.

with females (92 [52%] of 176 vs 32 [26%] of 123, respectively). Fewer children 10 years old or younger underwent cross-sectional imaging, as compared with those more than 10 years old (60 [42%] of 143 vs 103 [66%] of 156, respectively).

Table 2 summarizes the number of negative and perforated appendices according to the patients’ pathology reports. Significantly fewer negative appendectomies were noted in patients who underwent CT only (one [2%] of 52), as compared with those who underwent US (20 [17%] of 121) or no cross-sectional imaging (18 [14%] of 126; $P = .007$ and $P = .02$, respectively). There were significantly fewer negative appendectomies in the 96 patients who underwent CT with or without prior US (seven [7%] of 96), as compared with those who underwent US only (14 [18%] of 77; $P = .03$). In addition, these 96 patients had a lower negative appendectomy rate, as compared with those who underwent no imaging (18 [14%] of 126), with marginal significance ($P = .10$).

There were no significant differences in perforation rates between patients who underwent CT only (20%) prior to appendectomy and those who underwent US (32%) or no cross-sectional imaging (27%; $P = .31$ and $P = .58$, respectively) (Table 3).

Table 3 details the number of negative and perforated appendices, grouped according to patients who underwent no cross-sectional imaging (n = 126), US only (n = 77), US with or without CT (n = 121), CT only (n = 52), CT with or without prior US (n = 96), or US and CT (n = 44).

Imaging findings were compared with pathologic findings (Table 4). Ten (8%) of 121 patients had false-positive US findings, and four (4%) of 96 had false-positive CT findings, which resulted in negative appendectomies ($P = .35$). Twenty-six (21%) of 121 US findings versus two (2%) of 96 CT findings were false-negative ($P < .001$). The surgeons elected for urgent appendectomy in 10 patients with normal US findings and in three with normal CT findings (Table 4). Two of these patients with normal CT findings had prior US findings that were positive for appendicitis.

Seventeen (24%) of 72 children who had perforated appendicitis underwent CT-guided percutaneous abscess drainage and interval appendectomy; 16 of these patients underwent interval appendectomy 5–10 weeks after drainage, and one underwent appendectomy 43 weeks after diagnosis and drainage. Eleven of these 16 patients underwent laparoscopic interval appendectomy.

The time trend analysis of earlier versus later periods showed that the negative appendectomy rate decreased slightly in children who underwent no imaging (7% [nine of 55] vs 13% [nine of 72]) but decreased more significantly from 27% (11 of 41) to 9% (three of 36) in those who underwent US only ($P = .08$). Children who underwent CT with or without prior US had a stable negative appendectomy rate of 8% (two of 26) in the early period and 7% (five of 69) in the later period. The perforation rates increased from 22% (12 of 55) to 24% (17 of 72) in patients who underwent no imaging, decreased from 27% (11 of 41) to 17% (six of 36) in the US-only group, and decreased from 31% (eight of 26) to 20% (14 of 69) in patients who underwent CT with or without prior US. CT in patients clinically suspected of having appendicitis was first performed at our institution at the start of the study period, so that there were fewer CT scans and more US images in the earlier half of the study versus in the later half.

### DISCUSSION

The overall 13% negative appendectomy rate in our study is similar to those in other reports in the literature (1–7). The overall perforation rate of 27% is lower than the average value of 35% reported in the literature (1–7).

The results of the present study dem-
onstrate a significantly lower negative appendectomy rate in children who underwent CT prior to appendectomy when compared with those who underwent US or no cross-sectional imaging. We believe that this has the potential to change the treatment of patients suspected of having appendicitis. At our institution, pediatric surgeons have changed from ordering predominantly US examinations to ordering predominantly CT examinations. The surgeon must weigh the options of observing a child with acute abdominal pain, thus risking perforation of the inflamed appendix, versus unnecessarily removing a normal appendix. There is clearly a tradeoff between these two undesirable outcomes: To lower the perforation rate, surgeons must perform more surgeries, which raises the negative appendectomy rate. Conversely, to lower the negative appendectomy rate, surgeons choose observation rather than surgery, which raises the perforation rate.

CT has the potential to shift this long-accepted paradigm in the evaluation of children suspected of having appendicitis. CT has high sensitivity, which may allow it to be used as a triage tool to rule out appendicitis in these patients as it has in adults. Our results show that patients who underwent CT prior to urgent appendectomy underwent significantly fewer negative appendectomies (2%), as compared with those who underwent US (17%) or no cross-sectional imaging (14%). There were also fewer negative appendectomies in the patients who underwent CT with or without prior US (7%).

The higher negative appendectomy rate (14%) when both US and CT were performed may be explained in part by discordant results between the two imaging examinations (25). Further, we hypothesize that the patients who underwent both examinations may have had atypical and severe clinical presentations that suggested appendicitis despite the imaging findings.

For the past decade, US has been performed to improve diagnostic accuracy in children and adults suspected of having appendicitis (13-19). However, to our knowledge there is little evidence showing improved outcomes with US in children or adults. Several investigators (20-24) have shown no improvement in length of stay, time to surgery, cost of care, or negative appendectomy and perforation rates with the use of US in individuals suspected of having appendicitis. Our study also showed no improvement in negative appendectomy and perforation rates in children who underwent US, as compared with those who underwent no imaging.

Lower negative appendectomy and perforation rates have been reported in recent adult and pediatric studies of CT (9,10,12). A study of adults by Rao et al (9) showed statistically significant reductions in these rates. Although the results were not statistically significant, a recent pediatric study by Garcia-Pena et al (10) showed a decrease from 12% to 6% in the negative appendectomy rate in children who underwent CT, as compared with children who underwent no imaging.

The studies by Rao et al (9) and Garcia-Pena et al (10) shared two important techniques: They used multidisciplinary teams of emergency department physicians, surgeons, and radiologists who developed an “up-front” (performed in the emergency department, not later during observation) limited CT protocol in the emergency department, with specific entry criteria; and in neither study were appendectomy outcomes compared with US and CT outcomes. In contrast, in the current study, the results of cross-sectional imaging were compared in children in the emergency department and in those admitted for variable periods of observation. Karakas et al (12) reported a 13% negative appendectomy rate in children who underwent no imaging; 5%, with CT; and 8%, with US. The differences were not statistically significant because of the small sample size.

Decreasing the negative appendectomy rate by using CT did not produce a significantly higher perforation rate. However, the slightly different perforation rates in patients who underwent US only (27%), CT only (20%), or CT and US (42%) may be explained in two ways: First, the surgeon may have referred the sicker patients for imaging, with the sicker undergoing both US and CT. Second, the patients who underwent imaging may have been more delayed in undergoing surgery, as compared with those who did not undergo cross-sectional imaging. We do not know the severity of illness in the individual patients or the reason for the surgeon’s referral for cross-sectional imaging; these were limitations of our study. However, the perforation rate serves as a crude measure of the severity of illness in the groups compared.

Finally, our study results show the discrepancy between the imaging results and the surgeon’s decision to perform surgery in the patients (Table 4). As compared with the US report, the CT report was more consistent with the pathology report and the surgeon’s decision to perform surgery. There were four (4%) of 96 false-positive CT studies, as compared with 10 (8%) of 121 false-positive US studies. Two (2%) of 96 CT studies were false-negative, as compared with 26 (21%) of 121 US studies (P < .001). Ten of 121 patients with negative US findings underwent negative appendectomy, which suggests that the surgeons lacked confidence in these US results. In contrast, the surgeons may have had more confidence in the CT result because they acted more in parallel with it. Only three patients with negative CT results underwent negative appendectomy, and two of these already had a positive US result.

Although our study sample size was small, it provides insight into the effect of imaging on clinician diagnostic and therapeutic thinking and patient outcomes in children who have appendicitis. These issues are described in Fryback and Thornbury’s (26) six-tier hierarchic model of diagnostic imaging efficacy. Level 1 describes the image-quality measurement of an imaging examination; level 2, the sensitivity, specificity, and accuracy of an imaging examination; level 3, how the imaging examination might influence the clinician’s diagnostic confidence; level 4, how an imaging examination might affect patient treatment; level 5, patient outcomes (such as negative appendectomy and perforation rates); and level 6, societal effects of an imaging examination. Effectiveness at lower levels is necessary but not sufficient to ensure efficacy at higher levels. Therefore, high sensitivity and specificity of an examination (such as those reported for US for appendicitis) do not ensure that patient outcomes at level 5 (such as the negative appendectomy and perforation rates in our study) will be improved.

There were several limitations to our study. The most important was the selection bias of this observational study. We did not ascertain the severity of illness in the patients. The reason for the surgeons to refer a child for CT or US versus for no imaging may have differed because of how sick the patient was. This is suggested by the higher perforation rates, although not statistically significant, in the patients who underwent both US and CT. However, even in the presence of possible selection bias, the study provides information for radiologists about how the surgeons used US and CT results differently.

Another limitation of our study was the order in which imaging examina-
tions were performed. When both CT and US were performed, US was always performed prior to CT. The radiologist was not blinded to the US result and often interpreted both studies. Although this practice could have biased the results in favor of CT in the 44 patients who underwent both examinations, we believe that it did not. The 14% negative appendectomy rate in this group was nearly as high as the 18% in the US-only group. The 42% perforation rate in the patients who underwent both US and CT was the highest of all groups. Bias from prior knowledge of the US result does not explain the favorable results obtained with CT alone, nor does it influence the lack of improved outcome in patients who underwent US, as compared with those who underwent no imaging prior to surgery.

The CT technique varied during the 4½-year study period. Variability in technique (such as wider collimation, used early in the study) and experience (less experience in interpreting CT findings early in the study) did not affect the results favoring CT, as compared with those favoring US or no imaging. In addition, patients who underwent CT had a stable negative appendectomy rate of 8% in the early period and 7% in the later period. This suggests that our results are robust and generalizable; they should be reproducible in other radiology practices.

There is a need for prospective studies of cross-sectional imaging, preferably multinational ones, to assess the possible delay of surgery and the costs of care, adjusting for the severity of illness of the patients suspected of having appendicitis. Such studies will provide both a larger sample size and generalizability.

In conclusion, as compared with children who underwent no preoperative imaging and with those who underwent US, those who underwent CT for clinical suspicion of appendicitis had a significantly lower rate of negative appendectomy without a higher rate of perforated appendicitis. There were similar higher rates of negative appendectomy in the patients who underwent no imaging or US versus those who underwent CT. The perforation rates were not significantly different in the three groups.

Acknowledgment: We acknowledge the excellent work of Katie O'Brien, who performed the medical record reviews for this research.

References