Improvement in the Diagnosis of Appendicitis

Frederick Thurston Drake, MD, MPH[a,b,*] and **David Reed Flum, MD, MPH**[a,b]

[a]Department of Surgery, University of Washington, 1959 NE Pacific Street, Box 356410, Seattle, WA 98195, USA

[b]Surgical Outcomes Research Center (SORCE), Department of Surgery, University of Washington, 1107 NE 45th Street, Suite 502, Seattle, WA 98105, USA

Keywords

Appendicitis; Diagnosis; Advanced imaging; Negative appendectomy

EVIDENCE FOR IMPROVEMENT

For more than a hundred years, the diagnosis of acute appendicitis has been based on history taking, physical examination findings, and, to a lesser extent, laboratory results [1]. For experienced clinicians, an accurate diagnosis is straightforward in most patients, yet misdiagnosis stubbornly persists. There are 3 principal scenarios in which misdiagnosis occurs, and a speculative fourth:

1. appendicitis is diagnosed, the patient undergoes an operation, and non-appendiceal disease is discovered, which may or may not benefit from surgical intervention (eg, gynecologic lesions or inflammatory bowel disease of the terminal ileum); in this setting, the appendix may or may not be removed;

2. appendicitis is diagnosed, the patient undergoes an operation, and no abnormality is found; again, the appendix may or may not be removed;

3. appendicitis is not diagnosed but the patient does have an inflamed appendix (this would most often manifest as a patient who returned with persistent or recurrent appendicitis or with a perforation or other complication such as an abscess);

4. appendicitis is diagnosed that would have spontaneously resolved and is therefore clinically irrelevant (it has been suggested that increasing use of sensitive imaging such as computed tomography [CT] can detect mild inflammation of the appendix, which would otherwise resolve without antibiotics or surgical intervention) [2–4].

The first 2 forms of misdiagnosis, collectively referred to as negative appendectomy (NA) when the appendix is removed, are especially important to researchers interested in quality improvement, because they occur at the intersection of historical practice patterns that favor sensitivity over specificity and because appendicitis, straightforward to diagnose in most, is

also closely mimicked by other pelvic diseases, particularly in reproductive-aged women. The third form of misdiagnosis (missed cases of appendicitis) is undoubtedly important, but its clinical significance is complicated to define: some cases may resolve without therapy; some early cases may progress over a period of observation, allowing for detection on repeat evaluation (but without adverse impact on patients' outcomes); or, in some instances, patients with missed diagnoses may progress to perforation and re-present with complicated appendicitis. Concern for the third scenario tended to drive clinical practice toward the first 2 forms of misdiagnosis: a high number of negative explorations for suspected appendicitis were tolerated as surgeons endeavored to miss no cases, thereby averting perforation. This strategy is exemplified by Cantrell and Stafford's 1955 admonition that

> every effort should be made to establish an exact diagnosis. If, however, this is impossible and a suspicion of appendicitis exists, exploration is mandatory. It is far better to subject a moderate number of patients to a theoretically unnecessary operation than to let one patient suffer perforation [5].

Based on these beliefs, frequency of NA as high as 15% to 25% was common in surgical practice until recently. A nearly 50% NA rate was considered acceptable among pregnant women given the risks of perforation during pregnancy [6–9].

However, in the 1990s, investigators began to publish the results of studies in which CT scanning was used in the evaluation of suspected appendicitis. Rao and his colleagues in Boston focused their investigations specifically on the safe reduction of NA. Although multiple follow-up studies showed that CT for appendicitis was accurate and reliable, reducing NA was not without controversy: many surgeons still feared that attempts to reduce NA would lead to an increase in the frequency of perforation [10]. As CT became nearly ubiquitous in appendicitis, the perception persisted that it was overhyped and overused: 74% of the respondents to a large survey of the American College of Surgeons believed that the accuracy of CT scans was less than that reported in the literature and 62% believed that CT was overused in the workup for suspected appendicitis [11]. Some early reports suggested that advanced imaging was not influential in reducing NA [8,12–15]. However, evidence began to accumulate that the frequency of NA was steadily decreasing (Fig. 1) and that this relationship was associated with increased use of advanced diagnostic imaging. Moreover, these studies found no associated increase in the frequency of perforation (Fig. 2) [16–22]. There is now clear evidence that previously common NA rates of 15% to 25% should no longer be considered acceptable. In our view, this new standard has been accomplished by integrating high-quality advanced radiographic imaging into diagnostic algorithms for suspected appendicitis.

The data behind this assertion are mostly drawn from retrospective investigations with study populations of patients who had appendectomies. In terms of causal inferences, studies with this design present some challenges: there may be uncontrolled biases in terms of why certain patients were imaged, or, in longitudinal studies over time, other changes in the diagnosis of appendicitis may have accounted for, or at least contributed to, the reductions seen in NA. All of these cautions notwithstanding, a multitude of studies over the past 25 years have evaluated the relationship between imaging and NA and have shown that the use of imaging is associated with a reduction in NA [7,16–20,22–41]. This association has not

been universally shown, and there are several studies of appendectomy patients that have not found a relationship between imaging use and NA (Table 1) [8,12–15,42–46]. Several of these articles deserve mention. The 2001 article by Lee and colleagues [42], perhaps because of its unequivocal title, is frequently cited by those who argue against routine imaging in suspected appendicitis. The chief findings were that clinical examination and CT scan have equivalent sensitivity (83% vs 83.8%, respectively) and positive predictive value (PPV) (86.7% vs 83.8%) and that ultrasonography (US) performed inferiorly to both (sensitivity 35.5% and PPV 81.3%). However, these findings are limited because only 47 patients underwent CT in this group of 766 patients. Furthermore, their calculations of specificity and negative predictive value (NPV) are not reliable because such calculations cannot be based solely on patients who had an appendectomy; all patients undergoing a test are required for accurate computation of specificity and NPV.

Regarding the 2001 article by our group, the data presented were from 1987 to 1998 and showed no chronologic change in NA in Washington State, either as a percentage of all appendectomies performed (overall NA 15.5%) or as a population-based incidence. These data were based on diagnoses and procedure codes, and it was not possible to determine whether patients were being imaged more frequently (although it was clear that more imaging services were available), nor could comparisons be made between those who were and were not imaged. However, these data did clearly show, whether from lack of use or lack of effectiveness, that the purported benefits of advanced imaging were not being seen at the population level or across diverse hospital settings [8]. This population-based analysis was reinforced by a follow-up 2005 study in a large, statewide managed-health system (data from 1980–1999). The overall frequency of NA was also 15.5%, and there was no correlation between increasing use of CT and the relatively stable frequency of NA [15]. However, since these earlier studies, 2 recent analyses by our group using data abstracted directly from patient records have shown dramatic changes in the frequency of NA across Washington State: both studies demonstrated a clear reduction in NA, and both studies demonstrated a benefit for advanced imaging [22,36]. The statewide frequency of NA, estimated most recently for 2011, was 4.3%. Perhaps most striking were the reductions observed among reproductive-aged women. For these patients, the original 2001 study found that the average frequency of NA was 26.4% over the entire decade studied (1987–1998) [8]. This number decreased to 8.1% in the cohort of patients from 2006 to 2011, and large differences related to imaging use were evident: 24.7% of reproductive-aged women who did not receive advanced imaging underwent NA (essentially the same frequency of NA as in the 1990s) versus 6.9% of those women who did undergo imaging. After adjusting for age, white blood cell (WBC) count, and institution in a multivariate regression model, the odds of NA for reproductive-aged women without imaging were 3.5 times the odds of NA for those who were imaged [22].

Not all of the data on imaging and NA come from retrospective studies. We know of 6 studies published since 2000 that evaluated the relationship between imaging use and NA using prospective methodologies. In 3 of these studies, patients were randomized to a particular evaluation protocol [47–52]. All of these studies were strengthened by a design that focuses on patients suspected of having appendicitis, not retrospective evaluation of patients who underwent appendectomy.

In the most recent of these studies, 100 patients were evaluated by surgeons and assigned to a high, indeterminate, or low probability of having appendicitis. After allocation, CT scans were reviewed (if already ordered) or ordered if desired by the consulting surgeon. Seventy patients underwent appendectomy, and overall frequency of NA was 5.7%. Only 23 of the 63 highly suspicious patients had a CT scan, but in 2 of these patients, appendicitis was ruled out. Of the 27 indeterminate patients, 26 had CT scans; 11 were positive and 9 of these had appendicitis at appendectomy. Seven of the 10 patients in the low-suspicion group had CT, and 2 patients were found to have appendicitis. The investigators concluded that CT rarely changes management in patients highly suspicious for appendicitis (although it did assist surgeons in avoiding 2 unnecessary operations) but frequently changes management when the clinical diagnosis is less certain [52].

In 1 of the randomized investigations, 152 adult patients suspected of having appendicitis were randomized to a mandatory CT scan group or to a selective CT scan group, in which the emergency department (ED) physicians and consulting surgeons could order CT as dictated by their clinical certainty. The results were striking. CT scans were performed in 54 of 80 patients in the selective group and 70 of the 72 patients in the mandatory group. Frequency of NA in the mandatory group was 2.6% versus 13.9% in the selective group. The investigators did not find statistical significance in this difference (the article does not indicate which statistical tests were used to generate confidence intervals [CIs]), but they raised concern in their discussion for a type II error, noting that an 11% reduction in NA would certainly be clinically significant. The percentage of patients with appendicitis in each arm of the study was different (54.2% in the mandatory arm and 47.5% in the selective arm), and frequency of perforation differed as well (10.3% in the mandatory arm and 18.4% in the selective arm) [50]. Although these differences were not significant by the published CIs, perforation may have influenced diagnostic performance of clinical examination or CT. This finding likely represents insufficient sample size for adequate randomization, but may reflect the increased sensitivity of CT scan for mild appendicitis compared with clinical evaluation. (Some have argued, and the results of this study seem to support this argument, that routine CT scanning detects more cases of mild, possibly self-limiting cases of appendicitis, increasing the denominator in percentage calculations of perforation, which misleadingly reduces the number of perforations compared with populations in which CT scan is used less frequently and in which mild cases go undetected. For this reason, they argue that population-based incidences of perforation are a more reliable comparative metric [2–4,53]).

In another randomized study with a similar design and similar results, investigators found that NA was dramatically reduced by mandatory CT scanning (19% in the selective CT group vs 5% in the mandatory CT group) and that clinical management was altered in 26% of patients who underwent CT. In this study, CT scanning benefited both young men and reproductive-age women. Sensitivity of CT among patients in the mandatory CT arm was 94%, specificity was 100%, and accuracy was 96%. For those patients in the standard management arm (which could include US, observation, surgical exploration, or CT scan if requested by the consulting surgeon), sensitivity of standard management was 100%, but specificity was 67% and accuracy was 85% [47].

There is 1 randomized trial that did not show a benefit to the use of advanced imaging over clinical examination alone. In this 2003 study, adult patients were randomized to clinical evaluation alone versus clinical evaluation and CT scan. (Patients with an Alvarado score [see later discussion for detailed explanation of the Alvarado score] >8 were taken directly to the operating room (OR), but those with a score from 2 to 8 were randomized.) More than 20% of patients in both arms did not follow the study protocol, and there were substantially more men in the clinical evaluation arm than in the CT arm, a clear advantage for clinical examination. Frequency of NA is not reported, but the investigators found that accuracy was equivalent for the 2 evaluation protocols and concluded that routine CT was not useful in reducing misdiagnosis, even among patients with equivocal clinical findings [49]. Given the methodological weaknesses and the protocol violations of the study, this conclusion seems unwarranted.

A preponderance of observational and experimental data over the last 25 years support the assertion that NA has been reduced safely and without an associated increase in perforation; most studies find an association between reduced NA and the use of advanced imaging. Although most studies are based on observational data, the randomized trials highlighted suggest that the associations found in observational studies represent causal relationships. The experience in Washington State is instructive: in 2006, the Surgical Care and Outcomes Assessment Program (SCOAP), a statewide, voluntary collaboration of surgeons and hospitals, made a commitment to safely reduce NA among its member hospitals. The strategy for this reduction was a commitment to the routine use of high-quality, benchmarked imaging before appendectomy, especially for reproductive-aged women. (In SCOAP, "benchmark" means a performance standard that is based on the outcome and process results achieved by high-performing member institutions. When quarterly, de-identified reports are issued, member hospitals can measure and monitor their own performance relative to their peers). As the use of imaging increased over the subsequent 6 years, we observed a significant decline in the number of NA, with no change in the frequency of perforation (Fig. 3) [22]. Given earlier findings that implicated suboptimal accuracy of imaging in the persistent levels of NA [15,36], SCOAP began to monitor and benchmark diagnostic performance of imaging among member hospitals. Observational data cannot prove causality, but given the temporal relationship, the size of the association, and the numerous additional studies with similar findings (including some randomized trials), we believe that there is strong evidence to support SCOAP's emphasis on high-quality imaging and that this emphasis has driven much of the observed decline in unnecessary operations.

STRATEGIES FOR IMPROVEMENT

There are numerous tools that clinicians may use in evaluating patients with suspected appendicitis. How these tools are combined into practice depends heavily on the setting, resource availability, and clinical goals. For instance, in rural Kenya, accurate clinical assessment remains the cornerstone of diagnosis for everyone. In Washington State, diagnostic goals may vary depending on the patient: for children, minimizing exposure to ionizing radiation may mitigate the desire for superior diagnostic information; for women age 18 to 35 years, accurately ruling out nonappendiceal disease may be a priority; for

elderly patients, in whom radiation is less of a risk, the clarity and amount of information provided by CT scan may warrant its routine use. Although we advocate for the routine use of advanced imaging in our setting, because we know that this reduces misdiagnosis and unnecessary operations, in other settings and other populations, alternate strategies may be warranted. Clinicians can use the information presented below and in Tables 1–4 to formulate their own best practices. In addition to detailing some of the history of research into the diagnosis of appendicitis, the tables present diverse study designs, settings, and populations, which may be applicable to a variety of clinical scenarios.

Clinical evaluation

Surgical textbooks describe a classic presentation of appendicitis in which an initially diffuse or periumbilical abdominal pain migrates to the right iliac fossa. Symptoms frequently include nausea, vomiting, anorexia, and, less consistently, a history of fevers or chills. Examination findings include tenderness over McBurney's point (especially rebound tenderness), Rovsing sign (pain in the right lower quadrant (RLQ) when the left lower quadrant is palpated), and the psoas sign (an exacerbation of pain in the RLQ with flexion or extension of the right hip joint), all of which indicate focal peritoneal irritation. Expected laboratory findings include a leukocytosis with a left shift and an increased C-reactive protein (CRP) level [54–56].

Several clinical scoring systems have been proposed to facilitate the accurate clinical diagnosis of appendicitis. The most well-known of these is the Alvarado score (Box 1) [57]. Introduced in 1986, the score is based on 3 symptoms, 3 signs, and 2 laboratory findings. Each of these factors equals 1 point except for tenderness in the RLQ and leukocytosis, both of which earn 2 points. Alvarado recommended taking patients to surgery for a score of 7 or greater.

Box 1

A practical score for the diagnosis of appendicitis (Alvarado score)

Symptoms	Migration of abdominal pain to the RLQ	1
	Anorexia (or acetone in the urine)	1
	Nausea/vomiting	1
Signs	Tenderness in the RLQ	2
	Rebound pain	1
	Increase of temperature (37.3 °C)	1
Laboratory	Leukocytosis (>10,000)	2
	Shift to the left (in a differentiated WBC count) (eg, neutrophilia >75%)	1

Cumulative score

5–6 compatible with acute appendicitis

7–8 probable for acute appendicitis

9–10 very probable for acute appendicitis.

A score of 5 or 6 may be observed. A score of 7 or higher should proceed to appendectomy. In the original 1986 monograph, MANTRELS was the mnemonic proposed by Dr Alvarado for use of the score.

Since publication of the Alvarado score, several investigators have attempted to quantify the diagnostic accuracy of clinical evaluation as well as the individual prognostic value of various signs, symptoms, and laboratory results. One study of 766 adult and pediatric appendectomy patients found that clinical assessment was overall 75% accurate in detecting and ruling out appendicitis; sensitivity was estimated at 83% and PPV 86.7%. (As noted earlier, this study is limited in that patients who were not taken to the OR were not included in the calculations of diagnostic performance.) Individual findings identified as important included abdominal pain (99% of patients reported pain), nausea (81.7%), anorexia (72.4%), and emesis (67.7%). Findings with the greatest PPV were leukocytosis (PPV = 90.1%) and migration of abdominal pain (PPV = 91%) [42].

A more recent study prospectively examined the diagnostic performance of clinical impression, the Alvarado score, and CT scan in patients (age >15 years) who presented to the ED with suspected appendicitis. Sensitivity of clinical impression was 57.7% and was 42.2% for an Alvarado score of 8 or greater. Specificity was 71.6% and 79.1%, respectively. In this particular cohort of patients with suspected appendicitis, the 5 most common findings among those diagnosed with appendicitis were tenderness in the RLQ (92.2%), leukocytosis (75.6%), neutrophilia (73.3%), rebound tenderness (72.2%), and migratory pain (62.2%). Just more than half of the patients reported anorexia, nausea, or vomiting, and about a quarter were febrile [51].

Using 2 different methodologies (meta-analysis and a large cohort study), Andersson [1] examined the individual importance of specific clinical findings in suspected appendicitis. In the meta-analysis, receiver-operator curves (ROC) and likelihood ratios (LR) were generated from pooled data. Area under the ROC (AROC) was calculated (an AROC = 0.5 indicates that a test has no discriminatory ability). Anorexia (AROC = 0.58) and nausea (AROC = 0.56) had low discriminatory power. Migration of pain (AROC = 0.68) was the most valuable history finding. In terms of examination, tenderness alone had an AROC of 0.62, but signs of peritoneal irritation were stronger predictors of appendicitis: rebound tenderness had an AROC of 0.70 and guarding had an AROC of 0.68. Among laboratory tests, and also overall, WBC count, granulocyte count, and CRP level were the strongest discriminants of appendicitis (AROC = 0.77, 0.78, and 0.75, respectively). Combining 2 laboratory markers of inflammation had even stronger discriminatory and predictive power. The best performing combination was a WBC count greater than 10,000/L and CRP greater than 12 mg/L, with an AROC of 0.96 and LR of 23.3. These same investigators subsequently used a cohort of patients with suspected appendicitis to develop a clinical prediction tool that was validated in a separate prospective cohort (Box 2) [58]. In terms of sensitivity and specificity, this new tool performed similarly to Alvarado score; however, the Andersson score showed modest improvement in reducing the number of indeterminate results and in overall AROC. For the Andersson score, AROC was 0.93, and for the Alvarado score, AROC was 0.88 [58,59].

Box 2

The appendicitis inflammatory response score (Andersson score)

Variable	Level	Score
Pain or tenderness in RLQ		+1
Vomiting		+1
Rebound tenderness or muscular defense	Slight	+1
	Moderate	+2
	Strong	+3
WBC count	$10–14.9 \times 10^9$/L	+1
	15.0×10^9/L	+2
Proportion neutrophils	70%–84%	+1
	85%	+2
CRP concentration	10–49 mg/L	+1
	50	+2
Body temperature	38.5 °C	+1

Summation of points

0–4 = low probability for appendicitis; outpatient follow-up if unaltered general condition

5–8 = indeterminate; in-hospital active observation with rescoring, imaging, or diagnostic laparoscopy according to local practice

9–12 = high probability for appendicitis; surgical exploration recommended

Adapted from Andersson M, Andersson R. The appendicitis inflammatory response score: a tool for the diagnosis of acute appendicitis that outperforms the Alvarado score. World J Surg 2008;32(8):1843–9; with permission.

Ultrasound

US (see Table 2) has advantages over CT scanning, especially in terms of radiation exposure, but most studies have found that it performs inferiorly in terms of accuracy [25,26,28,48,60–63]. In 1 large prospective trial of US versus CT (n = 199 patients), the 2 modalities performed similarly, but completely unenhanced CT was used (ie, no oral, rectal, or intravenous [IV] contrast) and accuracy of CT in this study was inferior to most other published studies [64]. Consistently, sensitivity of US has been found inferior to its specificity [25,27,60–62,65].

To our knowledge, US has not been thoroughly evaluated in a nonacademic setting. In the 2012 SCOAP study, across a diverse group of hospitals, cumulative sensitivity of US was 47.8% [22]. Because we did not study those with suspected appendicitis (only appendectomies), our capture of true-negative results was limited, which prevented a calculation of specificity. Because we had access to information directly from patient charts, we were able to ascertain that the low sensitivity we observed resulted primarily from a large number of indeterminate US reports. (It is notable that many investigators exclude inconclusive results from their calculations of sensitivity, specificity, PPV, or NPV [48].) Two prospective trials of CT versus US in suspected appendicitis had similar findings:

indeterminate imaging results were significantly more frequent with US than with CT [48,62]. In 1 of these, nearly 75% of US studies were equivocal, and US did not rule out a single patient without appendicitis [48].

However, in clinical investigations, US often performs better than 47.8% sensitivity. Of the 12 studies presented in Table 2, all of which obtained data from patients suspected of having appendicitis (ie, not surgical patients only), sensitivity for appendicitis ranged from 76% to 98%. A rigorous 2006 meta-analysis of CT and US included 57 studies (both retrospective and prospective) and more than 13,000 patients. Studies were included only if absolute numbers of true-positive, true-negative, false-positive, and false-negative results were available, and the data were stratified by age so that adults and children could be considered separately. In children, sensitivity for US was 88% and specificity was 94%. For adults, sensitivity was 83% and specificity was 93% [61].

A well-designed Austrian study prospectively followed 350 patients evaluated by US for suspected appendicitis and determined that sensitivity was 98%, specificity was 98%, PPV was 96%, and NPV 99%. Of the 144 patients with a high clinical suspicion for appendicitis (assessment made by experienced surgeons before US), 34.8% did not have appendicitis. Among these 144 patients, US was inconclusive in only 4 patients, missed 1 case of appendicitis, and missed 2 other nonappendicitis conditions [66]. These are impressive findings, and the investigators note that their results are possible only because of long-standing close cooperation between the diagnostic imaging and surgical teams. These results exemplify what Bachur and colleagues [40] found in a recent study of US using a large, nationwide pediatric database: although there is a significant negative association, overall, between US use and NA in pediatric hospitals, among institutions that use US infrequently, there is dramatic variability in frequency of NA. The investigators attributed this finding to varying experience levels of sonographers and those interpreting the images.

When the appendix can be visualized, US performs well [66,67]. Furthermore, in specialized centers in which US for suspected appendicitis is emphasized, high-quality results are possible. Observational data suggest that performance of US in some community settings lags behind the results achieved in clinical studies.

Computed Tomography

CT is more widely used in the workup of suspected appendicitis and there is extensive literature on the subject (see Table 3). In addition to their assessment of US, Doria and colleagues [61] also performed a meta-analysis for CT scan. The results showed a notable improvement in sensitivity for CT over US. For adults, pooled sensitivity and specificity of CT were both 94%. For children, sensitivity was 94% and specificity 95%. In a systematic review of CT scan in adults, Neumayer and Kennedy [9] found that sensitivity ranged from 77% to 100% and that specificity ranged from 83% to 100%. The results reported in this systematic review were derived from studies with diverse patient populations and in which different CT protocols were used (for instance, without enteral contrast, completely unenhanced, or with both enteral and IV contrast). Some studies considered only patients in whom clinical findings were not convincing enough to warrant immediate appendectomy (so-called atypical patients). It is likely that this variability in study population and imaging

protocols (which is also reflected in Table 3) at least partially explains the wide differences in results. CT technology has also become more advanced over the last 20 years (for example 4-detector or single-detector scanners common in the early days of CT for appendicitis have given way to 64-multidetector helical scanners), leading some to suggest an association between improved technology and improved diagnostic performance [68]. Indeed, several studies of CT published since the Neumayer review and Doria meta-analysis have reported superior diagnostic performance [51,65,69,70], including 1 study that used low-dose radiation protocols [71]. One of these studies evaluated a large number of patients with suspected appendicitis (2871 consecutive adults) and followed those who did not undergo surgery with a rigorous follow-up protocol. CT scan sensitivity was 98.5%, specificity was 98%, NPV was 99.5%, and PPV was 93.9% [69]. As CT technology has improved, and as radiologists and surgeons have grown more comfortable and skillful in interpreting scans for appendicitis, many of the sensitivities and specificities reported from earlier in the era of CT for suspected appendicitis should be considered obsolete.

CT scan is also adept at revealing alternative diagnoses when appendicitis is not the cause of patients' RLQ pain. Of the 23 studies cited in the Neumayer review article, 13 articles reported the percentage of patients for whom an alternative diagnosis was uncovered by CT scan; this ranged from 6% to 36%, with a median of 21% [9]. In the large recent trial by Pickhardt and colleagues [69], alternative diagnoses were identified or suggested by CT in 42.1% of the patients who did not undergo appendectomy and who did not have CT findings suggestive of appendicitis. These CT findings represented a wide variety of clinical conditions, including other gastrointestinal conditions, gynecologic diseases, and urinary tract diseases.

Magnetic resonance imaging

Innovation in the use of magnetic resonance imaging (MRI) in the diagnosis of appendicitis has been primarily focused on its use in pregnant patients and in children (2 populations at particular risk from exposure to ionizing radiation) [72]. However, there are also studies on the use of MRI in general patients presenting with acute RLQ pain or suspected appendicitis (see Table 4). A recent meta-analysis combined the results of 8 previous studies (for a total of 363 patients). Studies were included only if the total number of true-positive, true-negative, false-positive, and false-negative results were available or derivable. Of the 363 patients included, the appendix could not be identified in only 8 patients. The investigators reported a sensitivity of 97% (95% CI 92%–99%) and specificity of 95% (95% CI 94%–99%). This meta-analysis included both enhanced and unenhanced scans [73].

In a prospective study by Cobben and colleagues [74], US was followed by MRI in all enrolled patients (n = 138). In 36 of these patients, US showed no abnormalities and did not visualize the appendix. Of these 36 patients, MRI showed that 8 had acute appendicitis and uncovered an alternative diagnosis in 4 patients. Even among patients who were clinically believed likely to have appendicitis (patients were stratified before imaging), MRI ruled out appendicitis in 22 patients, preventing 22 unnecessary operations. In 12 patients, the MRI was equivocal (ie, the appendix was not visualized), but the appendix had been visualized and noted to be normal on US in 10 of these patients. The investigators speculated that the

MRI could not visualize the appendix in these 12 patients because they were thin and had little intra-abdominal fat. Although MRI is a relatively expensive imaging modality, the investigators estimated a substantial cost savings from reductions in unnecessary operations and inpatient observation. According to the investigators, 1 of the benefits of the protocol used in this study (T1, T2, and T2 fat-suppression breath-hold sequences) is that the sequences are "easy to use and learn and are quick," which has the advantage of making "MRI more user-friendly [and may] lead to its more frequent use out of hours."

Diagnostic laparoscopy

In the early 1990s, diagnostic laparoscopy arose as a tool in the diagnosis of appendicitis. This was a time in which advanced imaging was not yet routinely used (investigation into imaging was happening roughly concomitantly with the trials involving diagnostic laparoscopy). A Cochrane review updated in 2010 evaluated both the therapeutic and diagnostic benefits of laparoscopy in appendicitis. In this analysis of diagnostic laparoscopy, studies were included only if the laparoscopic procedure was explicitly designed to address diagnosis and that (at least in cases where another diagnosis was evident on laparoscopy) the macroscopically normal appendix was left in situ. In this meta-analysis, which included 14 studies, the weighted relative risk (RR) of having a normal appendix removed for patients who received diagnostic laparoscopy versus patients who proceeded directly to therapeutic appendectomy was 0.37 (however, the 95% CI was 0.13–1.01). Among reproductive-aged women, the reduction in NA was more impressive, with an RR of 0.20 (95% CI 0.11–0.34). The RR of not having a diagnosis made was also significantly reduced by diagnostic laparoscopy (0.27, 95% CI 0.17–0.44) [75].

Special populations

Pregnant patients—Appendectomy is the most common nonobstetric surgical procedure performed in pregnant women [76]. Diagnosis of appendicitis in pregnant women is challenging, because many of the typical findings of appendicitis (eg, nausea, vomiting, abdominal discomfort, increased WBC count) are also present in a normal pregnancy; furthermore, the appendix may be displaced by the developing fetus and enlarging uterus [77]. Accurate and timely diagnosis is crucial, because perforation is associated with increased fetal and maternal complications, including fetal demise [78]. This risk is why the surgical literature once indicated that NA rates of 50% were acceptable in pregnant women [9]. However, the risks to a fetus of an operation and exposure to general anesthesia are not trivial. Avoiding NA when possible is important [79], and efforts have been focused on improving the accuracy with which appendicitis is diagnosed in pregnant patients.

CT is used when appendicitis is suspected in pregnant women, but, because of the risks of ionizing radiation, many clinicians prefer to use CT only when US or MRI is inconclusive [80,81]. Estimates of fetal radiation exposure from CT scans for appendicitis range from 19.9 to 43.6 mGy depending on the trimester. A radiation dose less than 50 mGy is considered safe from deterministic effects (threshold-dependent and dose-dependent effects), and it is the point at which there is a less than 1% risk of stochastic effects (dose-independent effects without a threshold). However, malignancies are generally considered related to the stochastic effects of radiation, and Long and colleagues [81] state "there is an

estimated doubling in the overall risk for childhood cancer in a fetus exposed to 30.0 mGy of radiation during an MDCT appendix protocol examination," citing RR data from the International Commission on Radiologic Protection. Because of the potential for teratogenicity and carcinogenicity, the 2007 American College of Radiology Appropriateness Criteria guidelines recommend MRI and US over CT in the evaluation of RLQ pain in pregnant patients.

US, using the graded compression technique, is the most frequently used modality for evaluating suspected appendicitis in pregnant women. Many studies report poor performance for US, related in most cases to nonvisualization of the appendix [78]. In a recent study of 99 women in the second or third trimester, US failed to locate the appendix in 97% of patients. Overall, 7 of the 99 patients had appendicitis, and only 2 of these patients had a positive US [78]. Other recent studies have reported equally dismal results for US in pregnant patients: nonvisualization ranged from 88% to 96% [82,83].

Because of this problem, MRI has emerged as a compelling option to evaluate suspected appendicitis in pregnant patients [84]. In a systematic review, Basaran and Basaran [81] compared CT scan with MRI and showed equivalent performance for both modalities in the second and third trimester of pregnancy (for both modalities, studies included mostly patients with a previous normal or inconclusive US because, in many instances, patients with conclusive US are taken directly to the OR without proceeding to CT or MRI). For CT, pooled sensitivity was 85.7% and specificity was 97.4%. For MRI, pooled sensitivity was 80% and specificity was 99% [85]. In a 2011 meta-analysis of MRI that includes 2 additional large studies not included in the Basaran review, pooled sensitivity for MRI was 91%, specificity 98%, PPV 86%, and NPV 99%. In this meta-analysis, as in most of the articles that provided data, scans in which the appendix was not seen but there was no evidence of RLQ inflammation were included as negative results for the purposes of calculating performance characteristics.

Pediatric patients—The use of imaging in pediatric patients to reduce NA has shown some variability (see Table 1). Some studies have shown improvements with increased imaging [7,26,27,33,40,41] and others have shown no reduction in NA [13,28,43]. A recent nationwide study of 52,426 pediatric patients (all after appendectomy) found that there was no association between CT scan use and NA but that the increasing use of US or US and CT (at the institutional level) was inversely associated with the frequency of NA [40]. Other studies have shown improved performance for CT over US in pediatric patients [26,33,61].

As with pregnant patients, ionizing radiation is an important concern in children. Although US may not perform as well as CT, studies based in centers that emphasize sonographic evaluation of suspected appendicitis have generated impressive results [66]. A 2001 study of US in children reported 93% sensitivity, 94.5% specificity, 86% PPV, and 97% NPV (of 2056 US evaluations for suspected appendicitis). These investigators concluded that US could improve accuracy of diagnosing appendicitis in children, especially when repeat US is used for initially indeterminate studies [27].

Recognizing the benefits of US in terms of radiation exposure and the benefits of CT in diagnostic accuracy, several high-quality studies have prospectively assessed an imaging protocol in which US is the first modality used followed by CT scan only for equivocal US. All showed acceptable sensitivity and specificity [63,65,67,86–88]. Three of these studies were performed specifically in pediatric populations: in the first study, from 1999, the algorithm achieved 94% sensitivity, 94% specificity, a PPV of 90%, and an NPV of 97% [86]. The 2 more recent studies are both from the same center [67,87]. The 2011 article reported sensitivity of 98.6%, specificity of 90.6%, a PPV of 84.6%, and an NPV of 93.3%. Frequency of NA was 8.1%. Of the 631 children who followed the imaging protocol, 52.6% were spared a CT scan and exposure to radiation. There was 1 (0.16%) missed case of appendicitis overall, (although there were 3 false-negative results from the overall protocol; 2 of these patients were taken to the OR based on clinical suspicion). Notably, nearly half of the patients who presented with possible appendicitis did not follow the imaging protocol. However, a strength of the study is the detailed description of outcomes for these noncompliant patients. In that group, there were no missed cases of appendicitis and 7.6% NA [87].

SUMMARY

Our attempts to systematically improve accuracy in the evaluation of patients with suspected appendicitis are, in some ways, hindered by the fact that the condition is so frequently straightforward to diagnose. Careful history-taking and physical examination are reliable in most patients. However, establishing the diagnosis with these skills alone remains vulnerable to conditions that masquerade as acute appendicitis. A substantial body of clinical research over the last quarter-century has shown that improved accuracy is possible. Strategies for improvement include the use of diagnostic scoring systems, laboratory makers such as CRP, diagnostic laparoscopy, and advanced imaging modalities such as CT, MRI, and US. How clinicians use these strategies depends on many factors related to practice setting, the population served, and clinical goals. In children, for instance, the desire to limit exposure to ionizing radiation competes with the greater anatomic detail that a CT scan can provide; at the same time, many hospitals that treat children do not have the resources to maintain the sort of full-time, highly sophisticated abdominal US programs that achieve the highest rates of diagnostic accuracy in clinical studies.

Trade-offs have to be made, but improvement is possible in almost all groups of patients: the clinical community should no longer settle for a 15% NA rate when 5% is clearly possible without adverse consequences. Many clinicians will be faced with the task of evaluating patients suspected of having acute appendicitis. A deliberate, proactive, and, ideally, benchmarked strategy for improving diagnosis should be the standard to which we hold ourselves and the promise we deliver to our patients.

References

1. Andersson RE. Meta-analysis of the clinical and laboratory diagnosis of appendicitis. Br J Surg. 2004; 91(1):28–37. [PubMed: 14716790]
2. Andersson R, Hugander A, Thulin A, et al. Indications for operation in suspected appendicitis and incidence of perforation. BMJ. 1994; 308(6921):107–10. [PubMed: 8298378]

3. Andersson RE. The natural history and traditional management of appendicitis revisited: spontaneous resolution and predominance of prehospital perforations imply that a correct diagnosis is more important than an early diagnosis. World J Surg. 2007; 31(1):86–92. [PubMed: 17180556]

4. Livingston EH, Woodward WA, Sarosi GA, et al. Disconnect between incidence of nonperforated and perforated appendicitis: implications for pathophysiology and management. Ann Surg. 2007; 245(6):886–92. [PubMed: 17522514]

5. Cantrell JR, Stafford ES. The diminishing mortality from appendicitis. Ann Surg. 1955; 141(6):749–58. [PubMed: 14377318]

6. Rao PM, Feltmate CM, Rhea JT, et al. Helical computed tomography in differentiating appendicitis and acute gynecologic conditions. Obstet Gynecol. 1999; 93(3):417–21. [PubMed: 10074991]

7. Rao PM, Rhea JT, Rattner DW, et al. Introduction of appendiceal CT: impact on negative appendectomy and appendiceal perforation rates. Ann Surg. 1999; 229(3):344–9. [PubMed: 10077046]

8. Flum DR, Morris A, Koepsell T, et al. Has misdiagnosis of appendicitis decreased over time? A population-based analysis. JAMA. 2001; 286(14):1748–53. [PubMed: 11594900]

9. Neumayer L, Kennedy A. Imaging in appendicitis: a review with special emphasis on the treatment of women. Obstet Gynecol. 2003; 102(6):1404–9. [PubMed: 14662234]

10. Aanning HL. Negative appendectomy and perforation rates in the SCOAP trial. Ann Surg. 2009; 249(4):699. [author reply: 699–700]. [PubMed: 19348033]

11. Sarkaria IS, Eachempati SR, Weyant MJ, et al. Current surgical opinion of computed tomography for acute appendicitis. Surg Infect (Larchmt). 2004; 5(3):243–52. [PubMed: 15684795]

12. McDonald GP, Pendarvis DP, Wilmoth R, et al. Influence of preoperative computed tomography on patients undergoing appendectomy. Am Surg. 2001; 67(11):1017–21. [PubMed: 11730216]

13. Partrick DA, Janik JE, Janik JS, et al. Increased CT scan utilization does not improve the diagnostic accuracy of appendicitis in children. J Pediatr Surg. 2003; 38(5):659–62. [PubMed: 12720164]

14. Perez J, Barone JE, Wilbanks TO, et al. Liberal use of computed tomography scanning does not improve diagnostic accuracy in appendicitis. Am J Surg. 2003; 185(3):194–7. [PubMed: 12620554]

15. Flum DR, McClure TD, Morris A, et al. Misdiagnosis of appendicitis and the use of diagnostic imaging. J Am Coll Surg. 2005; 201(6):933–9. [PubMed: 16310698]

16. Jones K, Pena AA, Dunn EL, et al. Are negative appendectomies still acceptable? Am J Surg. 2004; 188(6):748–54. [PubMed: 15619494]

17. Wagner PL, Eachempati SR, Soe K, et al. Defining the current negative appendectomy rate: for whom is preoperative computed tomography making an impact? Surgery. 2008; 144(2):276–82. [PubMed: 18656636]

18. Coursey CA, Nelson RC, Patel MB, et al. Making the diagnosis of acute appendicitis: do more preoperative CT scans mean fewer negative appendectomies? A 10-year study. Radiology. 2010; 254(2):460–8. [PubMed: 20093517]

19. Raja AS, Wright C, Sodickson AD, et al. Negative appendectomy rate in the era of CT: an 18-year perspective. Radiology. 2010; 256(2):460–5. [PubMed: 20529988]

20. Guller U, Rosella L, McCall J, et al. Negative appendicectomy and perforation rates in patients undergoing laparoscopic surgery for suspected appendicitis. Br J Surg. 2011; 98(4):589–95. [PubMed: 21259233]

21. Seetahal SA, Bolorunduro OB, Sookdeo TC, et al. Negative appendectomy: a 10-year review of a nationally representative sample. Am J Surg. 2011; 201(4):433–7. [PubMed: 21421095]

22. Drake FT, Florence MG, Johnson MG, et al. Progress in the diagnosis of appendicitis: a report from Washington State's. Ann Surg. 2012; 256(4):586–94. [PubMed: 22964731]

23. Colson M, Skinner KA, Dunnington G. High negative appendectomy rates are no longer acceptable. Am J Surg. 1997; 174(6):723–6. [discussion: 726–7]. [PubMed: 9409605]

24. Stroman DL, Bayouth CV, Kuhn JA, et al. The role of computed tomography in the diagnosis of acute appendicitis. Am J Surg. 1999; 178(6):485–9. [PubMed: 10670858]

25. Horton MD, Counter SF, Florence MG, et al. A prospective trial of computed tomography and ultrasonography for diagnosing appendicitis in the atypical patient. Am J Surg. 2000; 179(5):379–81. [PubMed: 10930484]

26. Applegate KE, Sivit CJ, Salvator AE, et al. Effect of cross-sectional imaging on negative appendectomy and perforation rates in children. Radiology. 2001; 220(1):103–7. [PubMed: 11425980]

27. Dilley A, Wesson D, Munden M, et al. The impact of ultrasound examinations on the management of children with suspected appendicitis: a 3-year analysis. J Pediatr Surg. 2001; 36(2):303–8. [PubMed: 11172421]

28. Bendeck SE, Nino-Murcia M, Berry GJ, et al. Imaging for suspected appendicitis: negative appendectomy and perforation rates. Radiology. 2002; 225(1):131–6. [PubMed: 12354996]

29. Naoum JJ, Mileski WJ, Daller JA, et al. The use of abdominal computed tomography scan decreases the frequency of misdiagnosis in cases of suspected appendicitis. Am J Surg. 2002; 184(6):587–9. [discussion: 589–90]. [PubMed: 12488178]

30. Yetkin G, Basak M, Isgor A, et al. Can negative appendectomy rate be decreased by using spiral computed tomography without contrast material? Acta Chir Belg. 2002; 102(5):334–7. [PubMed: 12471766]

31. Brandt MM, Wahl WL. Liberal use of CTscanning helps to diagnose appendicitis in adults. Am Surg. 2003; 69(9):727–31. [discussion: 731–2]. [PubMed: 14509318]

32. DeArmond GM, Dent DL, Myers JG, et al. Appendicitis: selective use of abdominal CT reduces negative appendectomy rate. Surg Infect (Larchmt). 2003; 4(2):213–8. [PubMed: 12906722]

33. Kaiser S, Mesas-Burgos C, Soderman E, et al. Appendicitis in children–impact of US and CT on the negative appendectomy rate. Eur J Pediatr Surg. 2004; 14(4):260–4. [PubMed: 15343467]

34. McGory ML, Zingmond DS, Nanayakkara D, et al. Negative appendectomy rate: influence of CT scans. Am Surg. 2005; 71(10):803–8. [PubMed: 16468523]

35. Morse BC, Roettger RH, Kalbaugh CA, et al. Abdominal CT scanning in reproductive-age women with right lower quadrant abdominal pain: does its use reduce negative appendectomy rates and healthcare costs? Am Surg. 2007; 73(6):580–4. [discussion: 584]. [PubMed: 17658095]

36. Cuschieri J, Florence M, Flum DR, et al. Negative appendectomy and imaging accuracy in the Washington State Surgical Care and Outcomes Assessment Program. Ann Surg. 2008; 248(4):557–63. [PubMed: 18936568]

37. Guss DA, Behling CA, Munassi D. Impact of abdominal helical computed tomography on the rate of negative appendicitis. J Emerg Med. 2008; 34(1):7–11. [PubMed: 18155380]

38. Piper HG, Rusnak C, Orrom W, et al. Current management of appendicitis at a community center–how can we improve? Am J Surg. 2008; 195(5):585–8. [discussion: 588–9]. [PubMed: 18367143]

39. Webb EM, Nguyen A, Wang ZJ, et al. The negative appendectomy rate: who benefits from preoperative CT? AJR Am J Roentgenol. 2011; 197(4):861–6. [PubMed: 21940573]

40. Bachur RG, Hennelly K, Callahan MJ, et al. Advanced radiologic imaging for pediatric appendicitis, 2005–2009: trends and outcomes. J Pediatr. 2012; 160(6):1034–8. [PubMed: 22192815]

41. Bachur RG, Hennelly K, Callahan MJ, et al. Diagnostic imaging and negative appendectomy rates in children: effects of age and gender. Pediatrics. 2012; 129(5):877–84. [PubMed: 22508920]

42. Lee SL, Walsh AJ, Ho HS. Computed tomography and ultrasonography do not improve and may delay the diagnosis and treatment of acute appendicitis. Arch Surg. 2001; 136(5):556–62. [PubMed: 11343547]

43. Weyant MJ, Eachempati SR, Maluccio MA, et al. The use of computed tomography for the diagnosis of acute appendicitis in children does not influence the overall rate of negative appendectomy or perforation. Surg Infect (Larchmt). 2001; 2(1):19–23. [PubMed: 12594877]

44. Vadeboncoeur TF, Heister RR, Behling CA, et al. Impact of helical computed tomography on the rate of negative appendicitis. Am J Emerg Med. 2006; 24(1):43–7. [PubMed: 16338508]

45. Huynh V, Lalezarzadeh F, Lawandy S, et al. Abdominal computed tomography in the evaluation of acute and perforated appendicitis in the community setting. Am Surg. 2007; 73(10):1002–5. [PubMed: 17983068]

46. Frei SP, Bond WF, Bazuro RK, et al. Appendicitis outcomes with increasing computed tomographic scanning. Am J Emerg Med. 2008; 26(1):39–44. [PubMed: 18082779]

47. Walker S, Haun W, Clark J, et al. The value of limited computed tomography with rectal contrast in the diagnosis of acute appendicitis. Am J Surg. 2000; 180(6):450–4. [discussion: 454–5]. [PubMed: 11182396]

48. Wilson EB, Cole JC, Nipper ML, et al. Computed tomography and ultrasonography in the diagnosis of appendicitis: when are they indicated? Arch Surg. 2001; 136(6):670–5. [PubMed: 11387006]

49. Hong JJ, Cohn SM, Ekeh AP, et al. A prospective randomized study of clinical assessment versus computed tomography. Surg Infect (Larchmt). 2003; 4(3):231–9. [PubMed: 14588157]

50. Lee CC, Golub R, Singer AJ, et al. Routine versus selective abdominal computed tomography scan in the evaluation of right lower quadrant pain: a randomized controlled trial. Acad Emerg Med. 2007; 14(2):117–22. [PubMed: 17192450]

51. Kim K, Rhee JE, Lee CC, et al. Impact of helical computed tomography in clinically evident appendicitis. Emerg Med J. 2008; 25(8):477–81. [PubMed: 18660392]

52. Santos DA, Manunga J Jr, Hohman D, et al. How often does computed tomography change the management of acute appendicitis? Am Surg. 2009; 75(10):918–21. [PubMed: 19886134]

53. Andersson RE. Proportion of cases of perforated appendicitis: a bad measure of quality of care. Arch Surg. 2011; 146(7):885. [author reply: 885–6]. [PubMed: 21768441]

54. Schwartz, SI.; Brunicardi, FC. Schwartz's principles of surgery. New York: McGraw-Hill; 2010. p. 1online resource (p. xxi, 1866)

55. Sack U, Biereder B, Elouahidi T, et al. Diagnostic value of blood inflammatory markers for detection of acute appendicitis in children. BMC Surg. 2006; 6:15. [PubMed: 17132173]

56. Noh H, Chang SJ, Han A. The diagnostic values of preoperative laboratory markers in children with complicated appendicitis. J Korean Surg Soc. 2012; 83(4):237–41. [PubMed: 23091796]

57. Alvarado A. A practical score for the early diagnosis of acute appendicitis. Ann Emerg Med. 1986; 15(5):557–64. [PubMed: 3963537]

58. Andersson M, Andersson RE. The appendicitis inflammatory response score: a tool for the diagnosis of acute appendicitis that outperforms the Alvarado score. World J Surg. 2008; 32(8): 1843–9. [PubMed: 18553045]

59. de Castro SM, Ünlü C, Steller EP, et al. Evaluation of the appendicitis inflammatory response score for patients with acute appendicitis. World J Surg. 2012; 36(7):1540–5. [PubMed: 22447205]

60. Balthazar EJ, Birnbaum BA, Yee J, et al. Acute appendicitis: CT and US correlation in 100 patients. Radiology. 1994; 190(1):31–5. [PubMed: 8259423]

61. Doria AS, Moineddin R, Kellenberger CJ, et al. US or CT for diagnosis of appendicitis in children and adults? a meta-analysis. Radiology. 2006; 241(1):83–94. [PubMed: 16928974]

62. Keyzer C, Zalcman M, De Maertelaer V, et al. Comparison of US and unenhanced multi-detector row CT in patients suspected of having acute appendicitis. Radiology. 2005; 236(2):527–34. [PubMed: 16040910]

63. Poortman P, Oostvogel HJ, Bosma E, et al. Improving diagnosis of acute appendicitis: results of a diagnostic pathway with standard use of ultrasonography followed by selective use of CT. J Am Coll Surg. 2009; 208(3):434–41. [PubMed: 19318006]

64. Poortman P, Lohle PN, Schoemaker CM, et al. Comparison of CT and sonography in the diagnosis of acute appendicitis: a blinded prospective study. AJR Am J Roentgenol. 2003; 181(5):1355–9. [PubMed: 14573433]

65. Gaitini D, Beck-Razi N, Mor-Yosef D, 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(5):1300–6. [PubMed: 18430847]

66. Rettenbacher T, Hollerweger A, Gritzmann N, et al. Appendicitis: should diagnostic imaging be performed if the clinical presentation is highly suggestive of the disease? Gastroenterology. 2002; 123(4):992–8. [PubMed: 12360459]

67. Ramarajan N, Krishnamoorthi R, Barth R, et al. An interdisciplinary initiative to reduce radiation exposure: evaluation of appendicitis in a pediatric emergency department with clinical assessment

supported by a staged ultrasound and computed tomography pathway. Acad Emerg Med. 2009; 16(11):1258–65. [PubMed: 20053244]

68. Kessler N, Cyteval C, Gallix B, et al. Appendicitis: evaluation of sensitivity, specificity, and predictive values of US, Doppler US, and laboratory findings. Radiology. 2004; 230(2):472–8. [PubMed: 14688403]

69. Pickhardt PJ, Lawrence EM, Pooler BD, et al. Diagnostic performance of multidetector computed tomography for suspected acute appendicitis. Ann Intern Med. 2011; 154(12):789–96. [PubMed: 21690593]

70. van Randen A, Lameris W, van Es HW, et al. Profiles of US and CT imaging features with a high probability of appendicitis. Eur Radiol. 2010; 20(7):1657–66. [PubMed: 20119730]

71. Platon A, Jlassi H, Rutschmann OT, et al. Evaluation of a low-dose CT protocol with oral contrast for assessment of acute appendicitis. Eur Radiol. 2009; 19(2):446–54. [PubMed: 18797875]

72. Inci E, Hocaoglu E, Aydin S, et al. Efficiency of unenhanced MRI in the diagnosis of acute appendicitis: comparison with Alvarado scoring system and histopathological results. Eur J Radiol. 2011; 80(2):253–8. [PubMed: 20655156]

73. Barger RL, Nandalur KR. Diagnostic performance of magnetic resonance imaging in the detection of appendicitis in adults: a meta-analysis. Acad Radiol. 2010; 17(10):1211–6. [PubMed: 20634107]

74. Cobben L, Groot I, Kingma L, et al. A simple MRI protocol in patients with clinically suspected appendicitis: results in 138 patients and effect on outcome of appendectomy. Eur Radiol. 2009; 19(5):1175–83. [PubMed: 19137303]

75. Sauerland S, Jaschinski T, Neugebauer EA. Laparoscopic versus open surgery for suspected appendicitis. Cochrane Database Syst Rev. 2010; (10):CD001546. [PubMed: 20927725]

76. Wei PL, Keller JJ, Liang HH, et al. Acute appendicitis and adverse pregnancy outcomes: a nationwide population-based study. J Gastrointest Surg. 2012; 16(6):1204–11. [PubMed: 22402956]

77. Wang PI, Chong ST, Kielar AZ, et al. Imaging of pregnant and lactating patients: part 2, evidence-based review and recommendations. AJR Am J Roentgenol. 2012; 198(4):785–92. [PubMed: 22451542]

78. Lehnert BE, Gross JA, Linnau KF, et al. Utility of ultrasound for evaluating the appendix during the second and third trimester of pregnancy. Emerg Radiol. 2012; 19(4):293–9. [PubMed: 22370694]

79. Ito K, Ito H, Whang EE, et al. Appendectomy in pregnancy: evaluation of the risks of a negative appendectomy. Am J Surg. 2012; 203(2):145–50. [PubMed: 21784406]

80. Patel SJ, Reede DL, Katz DS, et al. Imaging the pregnant patient for nonobstetric conditions: algorithms and radiation dose considerations. Radiographics. 2007; 27(6):1705–22. [PubMed: 18025513]

81. Long SS, Long C, Lai H, et al. Imaging strategies for right lower quadrant pain in pregnancy. AJR Am J Roentgenol. 2011; 196(1):4–12. [PubMed: 21178041]

82. Israel GM, Malguria N, McCarthy S, et al. MRI vs. ultrasound for suspected appendicitis during pregnancy. J Magn Reson Imaging. 2008; 28(2):428–33. [PubMed: 18666160]

83. Cobben LP, Groot I, Haans L, et al. MRI for clinically suspected appendicitis during pregnancy. AJR Am J Roentgenol. 2004; 183(3):671–5. [PubMed: 15333354]

84. Pedrosa I, Zeikus EA, Levine D, et al. MR imaging of acute right lower quadrant pain in pregnant and nonpregnant patients. Radiographics. 2007; 27(3):721–43. [discussion: 743–53]. [PubMed: 17495289]

85. Basaran A, Basaran M. Diagnosis of acute appendicitis during pregnancy: a systematic review. Obstet Gynecol Surv. 2009; 64(7):481–8. [quiz: 499]. [PubMed: 19545456]

86. Garcia Pena BM, Mandl KD, Kraus SJ, et al. Ultrasonography and limited computed tomography in the diagnosis and management of appendicitis in children. JAMA. 1999; 282(11):1041–6. [PubMed: 10493202]

87. Krishnamoorthi R, Ramarajan N, Wang NE, et al. Effectiveness of a staged US and CT protocol for the diagnosis of pediatric appendicitis: reducing radiation exposure in the age of ALARA. Radiology. 2011; 259(1):231–9. [PubMed: 21324843]

88. Poletti PA, Platon A, De Perrot T, et al. Acute appendicitis: prospective evaluation of a diagnostic algorithm integrating ultrasound and low-dose CT to reduce the need of standard CT. Eur Radiol. 2011; 21(12):2558–66. [PubMed: 21805194]

89. Lowe LH, Penney MW, Stein SM, et al. Unenhanced limited CTof the abdomen in the diagnosis of appendicitis in children: comparison with sonography. AJR Am J Roentgenol. 2001; 176(1):31–5. [PubMed: 11133533]

90. Balthazar EJ, Megibow AJ, Siegel SE, et al. Appendicitis: prospective evaluation with high-resolution CT. Radiology. 1991; 180(1):21–4. [PubMed: 2052696]

91. Malone AJ Jr, Wolf CR, Malmed AS, et al. Diagnosis of acute appendicitis: value of unenhanced CT. AJR Am J Roentgenol. 1993; 160(4):763–6. [PubMed: 8456661]

92. Lane MJ, Katz DS, Ross BA, et al. Unenhanced helical CT for suspected acute appendicitis. AJR Am J Roentgenol. 1997; 168(2):405–9. [PubMed: 9016216]

93. Rao PM, Rhea JT, Novelline RA, et al. Helical CT combined with contrast material administered only through the colon. AJR Am J Roentgenol. 1997; 169(5):1275–80. [PubMed: 9353441]

94. Lane MJ, Liu DM, Huynh MD, et al. Suspected acute appendicitis: nonenhanced helical CT in 300 consecutive patients. Radiology. 1999; 213(2):341–6. [PubMed: 10551210]

95. Wijetunga R, Tan BS, Rouse JC, et al. Diagnostic accuracy of focused appendiceal CT in clinically equivocal cases of acute appendicitis. Radiology. 2001; 221(3):747–53. [PubMed: 11719671]

96. Christopher FL, Lane MJ, Ward JA, et al. Unenhanced helical CT scanning of the abdomen and pelvis changes disposition of patients presenting to the emergency department with possible acute appendicitis. J Emerg Med. 2002; 23(1):1–7. [PubMed: 12217464]

97. Ege G, Akman H, Sahin A, et al. Diagnostic value of unenhanced helical CT in adult patients with suspected acute appendicitis. Br J Radiol. 2002; 75(897):721–5. [PubMed: 12200239]

98. Hershko DD, Sroka G, Bahouth H, et al. The role of selective computed tomography in the diagnosis and management of suspected acute appendicitis. Am Surg. 2002; 68(11):1003–7. [PubMed: 12455796]

99. in't Hof KH, van Lankeren W, Krestin GP, et al. Surgical validation of unenhanced helical computed tomography in acute appendicitis. Br J Surg. 2004; 91(12):1641–5. [PubMed: 15386320]

100. Mun S, Ernst RD, Chen K, et al. Rapid CT diagnosis of acute appendicitis with IV contrast material. Emerg Radiol. 2006; 12(3):99–102. [PubMed: 16362812]

101. Dearing DD, Recabaren JA, Alexander M. Can computed tomography scan be performed effectively in the diagnosis of acute appendicitis without the added morbidity of rectal contrast? Am Surg. 2008; 74(10):917–20. [PubMed: 18942613]

102. Chabanova E, Balslev I, Achiam M, et al. Unenhanced MR imaging in adults with clinically suspected acute appendicitis. Eur J Radiol. 2011; 79(2):206–10. [PubMed: 20347539]

103. Johnson AK, Filippi CG, Andrews T, et al. Ultrafast 3-T MRI in the evaluation of children with acute lower abdominal pain for the detection of appendicitis. AJR Am J Roentgenol. 2012; 198(6):1424–30. [PubMed: 22623558]

104. Moore MM, Gustas CN, Choudhary AK, et al. MRI for clinically suspected pediatric appendicitis: an implemented program. Pediatr Radiol. 2012; 42(9):1056–63. [PubMed: 22677910]

Key points

- Although acute appendicitis is often straightforward to diagnose, clinical evaluation is susceptible to conditions that mimic the signs and symptoms of appendicitis.

- Diagnosis of appendicitis has improved over the last 25 years through the use of such strategies as clinical scoring systems, laboratory studies, diagnostic laparoscopy, and advanced imaging studies (US, CT, and MRI).

- Which strategy, or combination of strategies, clinicians choose to employ depends on practice setting, resource availability, patient population, and clinical goals.

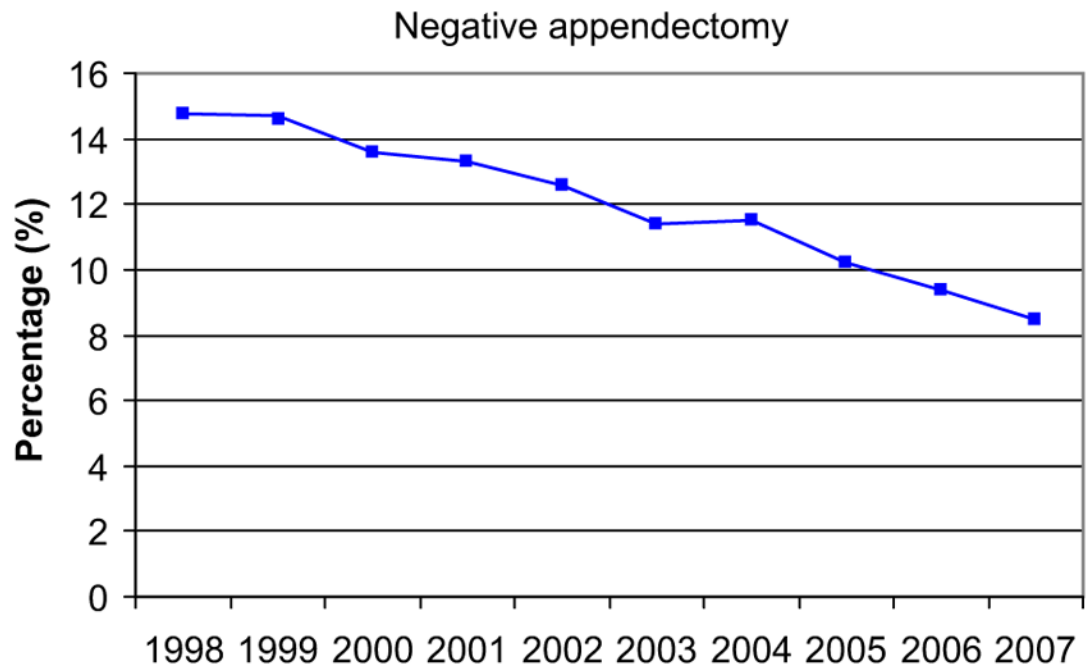## Negative appendectomy

Fig. 1.
Decrease in negative appendectomies (1998–2007). (*Reprinted from* Seetahal SA,
Bolorunduro OB, Sookdeo TC, et al. Negative appendectomy: a 10-year review of a
nationally representative sample. Am J Surg 2011;201(4):435; with permission.)

Perforated or Gangrenous Appendicitis
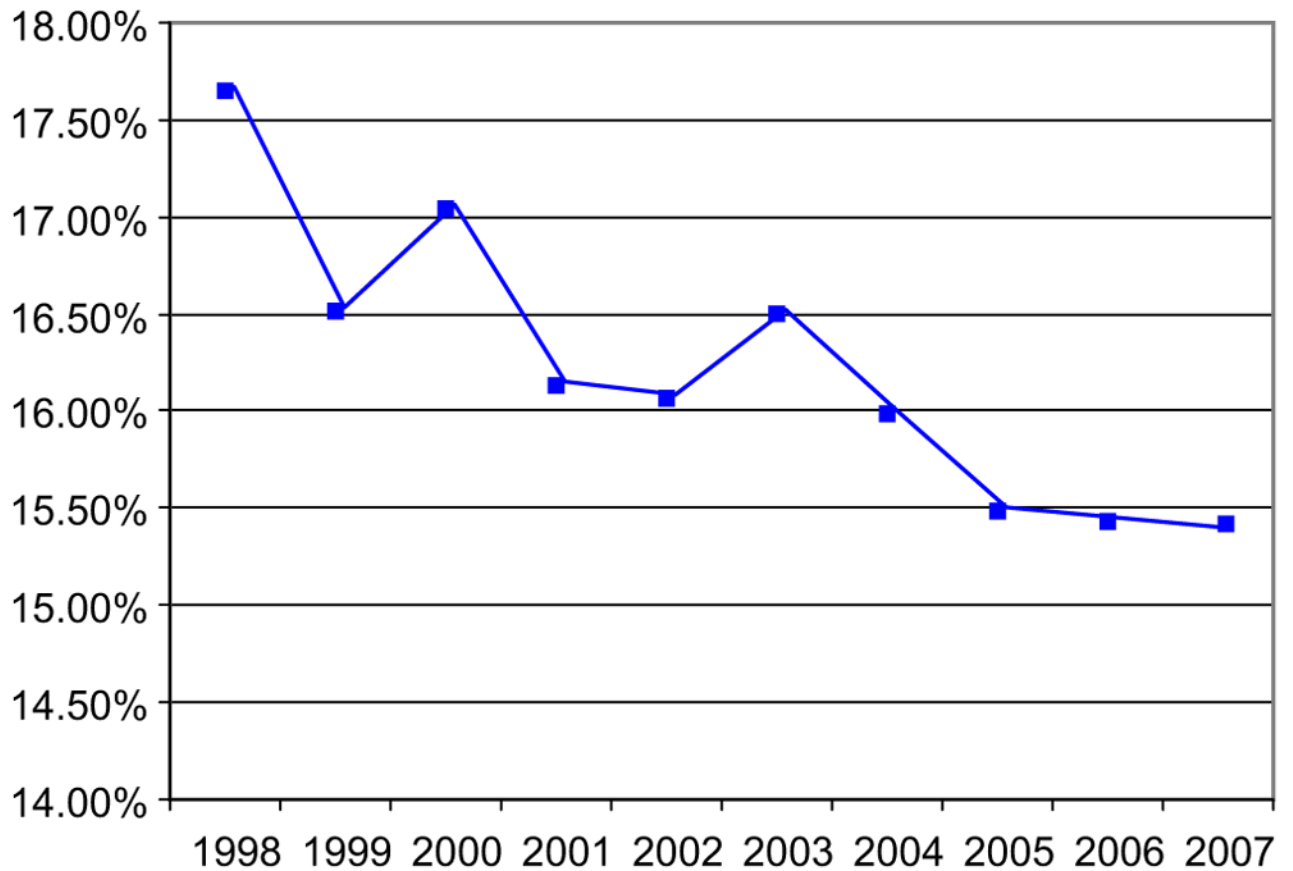
Fig. 2.
Incidence of perforated or gangrenous appendicitis (1998–2007). (*Reprinted from* Seetahal
SA, Bolorunduro OB, Sookdeo TC, et al. Negative appendectomy: a 10-year review of a
nationally representative sample. Am J Surg 2011;201(4):436; with permission.)

Fig. 3.
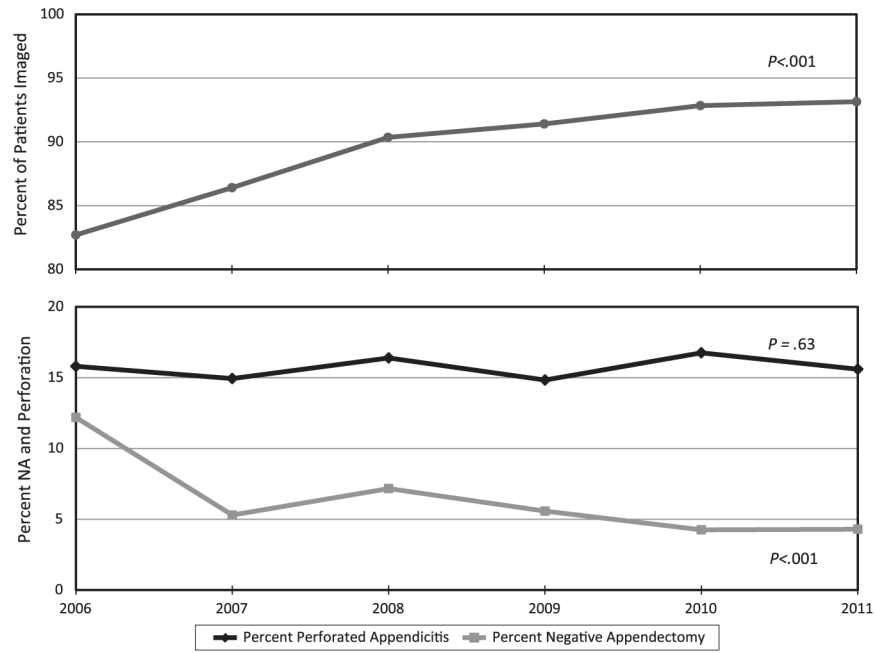Frequency of imaging use, NA, and perforation in SCOAP hospitals, 2006–2011. Use of imaging, NA, and perforation are expressed as annual percentages. The Cochran-Armitage test was used to evaluate for trends over time.

Table 1

Selected studies of NA from 1999 to 2012, including retrospective reviews, large database analyses, prospective observational studies, and randomized trials

Reference	Study design, population, and size	Findings
Rao et al [7], 1999	Retrospective review of pediatric and adult appendectomy patients. Patients from pre-CT era (n = 493) compared with those of CT era (n = 209). Only 59% of patients in CT era underwent CT scanning.	All patients: for pre-CT era, NA = 20%, and for CT era, NA = 7%. For female patients, 31% vs 11%, $P<.001$, for male patients, 10% vs 5%, $P = .057$, and for pediatric patients, 13% vs 7%, $P = .22$.
Walker et al [47], 2000	RCT of mandatory CT scan vs standard management in adults with suspected appendicitis (n = 108).	Mandatory CT scan: NA = 5%. Standard management: NA = 19%.
Applegate et al [26], 2001	Retrospective review of surgical database for appendectomy patients (age <21 y) at large children's hospital (n = 299).	Overall NA = 13%. For CT alone, NA = 2%, for US, NA = 17%, for CT with or without US, NA = 7% and for no imaging, NA = 14%.
Dilley et al [27], 2001	Retrospective review of pediatric appendectomy patients comparing those with and without US at a large children's hospital (n = 1007).	With US, NA = 15.6%; without US, NA = 18.3% ($P<.05$).
Flum et al [8], 2001	Retrospective cohort of appendectomy patients from a statewide discharge database (n = 85,790). No direct data on imaging use. NA determined by ICD-9 coding. 1987 to 1998.	Overall NA was 15.5% with no change seen over time. Among RA women, NA was 26.4%.
McDonald et al [12], 2001	Retrospective, before-and-after study of appendectomy patients (age 12 y) in selective CT era vs routine CT era (n = 226).	For selective CT era, NA = 15%, and for routine CT era, NA = 13% ($P = .7$). With CT, NA = 6.7% and without CT, NA = 16.9% ($P = .05$).
Wilson et al [48], 2001	Prospective, single-arm study. Adults and children with suspected appendicitis. Clinical decision made, then all patients had CT and US. Based on imaging results, original plan reevaluated (n = 99).	Overall: NA = 10%. For females, imaging reduced NA from potential 50%–17%. For males, NA = 8% and imaging did not change management.
Naoum et al [29], 2002	Retrospective chart review of adult appendectomy patients (n = 194). Comparison made before and after a guideline encouraging CT use.	CT scan use increased from 32% to 84%. NA before guideline = 25% and NA after guideline was 6%.
DeArmond et al [32], 2003	Retrospective review of children and adult appendectomy patients. 1995 compared with 1998 (n = 778) Selective CT strategy in place over both periods.	1995: for CT, NA = 19% and for no CT, NA = 13% ($P=$ NS). 1998: for CT, NA = 6% and for no CT, NA = 7% ($P=$ NS).
Partrick et al [13], 2003	Retrospective review of pediatric appendectomy patients at a single large children's hospital (n = 616). Comparison made for CT, US, or no imaging. NA based on pathology report. 1997 to 2001.	CT used increased substantially and US use decreased substantially over the study period, but no trend in NA. For no imaging, NA = 8%, for CT, NA = 7% and for US, NA = 11%.
Perez et al [14], 2003	Retrospective review of children and adult appendectomy patients. 1994 (n = 100) compared with 1999–2000 (n = 118).	1994: NA = 12%; 1999–2000: NA = 17.8% ($P = .32$). They report a 4-fold increase in the use of CT scan.
Jones et al [16], 2004	Retrospective review of appendectomy patients, age 14–84 y (n = 389); time-based analysis of NA as CT use increased.	Progressive decrease in NA from 17% (2000) to 9% (2001) to 2% (2002). NA was 12% without CT and 7% with CT scan.
Kaiser et al [33], 2004	Retrospective review of consecutive pediatric appendectomy patients for 1991, 1994, 1997 and 2000. Large children's hospital (n = 600).	CT and US use increased steadily, and NA decreased: 23% (1991), 8.7% (1994), 8.0% (1997), and 4.0% (2000).
Lee et al [50], 2007	RCT of adult patients presenting with suspected appendicitis, randomized to mandatory or selective CT scan (n = 152).	Mandatory CT group, NA = 2.6% vs 13.9% in selective CT group. ($P=$ NS).
Huynh et al [45], 2007	Retrospective chart review of appendicitis patients at a community hospital; adults and children, age 2 to 89 (n = 550).	For CT, NA = 11%. For no CT, NA = 13% ($P=$ NS).

Reference	Study design, population, and size	Findings
Morse et al [35], 2007	Retrospective chart review of appendectomy patients age >16 y (n = 439). Single institution.	For CT, NA = 9%; for no CT, NA = 24%. Among RA women: for CT, NA = 17% and for no CT, NA = 42%. Among men: for CT, NA = 5% and for no CT, NA = 15%.
Cuschieri et al [36], 2008	Review of multi-institution, prospectively collected data from consecutive appendectomy patients (children and adults) from 2006 to 2007 (n = 3540). NA defined by pathology report findings.	Overall NA = 5.6%. With no imaging, NA = 10%, with US, NA = 8%, and with CT, NA = 4.5%. Accuracy of imaging (based on pathology-radiology report concordance) correlated with frequency of NA.
Guss et al [37], 2008	Retrospective chart review of adult appendectomy patients. Single institution. Period A, 1990–1994 (n = 316) and Period B, 2001–2004 (n = 477).	Period A: 12% had nonhelical CT, NA = 15.5%. Period B: 82% had helical CT, NA = 7.9%.
Piper et al [38], 2008	Retrospective review of appendectomy patients (age 5–86 y) at 2 community hospitals (n = 134).	Overall NA = 9%. With imaging, NA = 3%, without imaging, NA = 10% (no P value reported).
Wagner et al [17], 2008	Retrospective review of appendectomy patients (pediatric and adult) at a university hospital (n = 1425); time-based methodology as well as comparison of those with and without CT.	In 2000, NA = 10.6% and in 2006, NA = 5.0%. With CT, NA = 7.2%. Without CT, NA = 14.0%. Major reduction in NA for women; less impact among men, pediatric patients.
Coursey et al [18], 2010	Retrospective chart review of adult appendectomy patients from 1998 to 2007 (n = 925). University hospital setting.	NA declined, 16.7%–8.7%. Large decrease for RA women (43% vs 7%) but not for men. CT use increased from 18.5% to 94.2%.
Raja et al [19], 2010	Retrospective chart review of adult appendectomy patients from 1990 to 1994 (n = 971) and 2003 to 2008 (n = 637). University hospital setting.	In 1990, NA = 23.0% and in 2007, NA = 1.7%. Large decrease in NA for both male and female patients. CT scan use increased from 1% in 1990 to 97.5% in 2007.
Seetahal et al [21], 2011	Retrospective review of adult appendectomy patients in the NIS from 1998 to 2007 (n = 475,651). Determination of NA based on ICD-9 coding. No data on imaging use available.	NA = 14.7% in 1998 and decreased steadily to 8.5%. Women accounted for most NA over entire period (71.3%).
Bachur et al [40], 2012	Retrospective evaluation of nationwide pediatric hospital database (PHIS). For calculation of NA, records obtained for appendectomy patients. Data from 2005 to 2009. (n = 52,426).	From 2005 to 2009, frequency of NA decreased significantly from 4.4% to 3.1%. No relationship between CT use and NA, but inverse relationship between US or US + CT use and frequency of NA.
Drake et al [22], 2012	Review of prospectively collected data on consecutive appendectomy patients (age 15 y) from 2006–2011 (n = 19,327). Multi-institution.	Overall: if imaged, NA = 4.5%, if not imaged, NA = 15.4%. RA women: if imaged, NA = 6.9%, if not imaged, NA = 24.7%. Males: if imaged, NA = 3%, if not imaged, NA = 10%.

Statistical significance may be assumed unless otherwise noted in the table.

Abbreviations: ICD-9, International Classification of Diseases, Ninth Revision; NIS, National Inpatient Sample; PHIS, Pediatric Health Information System; RA, reproductive age; RCT, randomized controlled trial.

Table 2

Selected studies evaluating diagnostic performance of US, from 1994 to 2009

Reference	Study design	N	Sensitivity (%)	Specificity (%)
Balthazar et al [60], 1994	Prospective. Consecutive patients, age 15–82 y, with "suggestive but not typical clinical and laboratory findings of appendicitis" evaluated by both CT and US. Clinical follow-up for nonsurgical patients.	100	76	91
Horton et al [25], 2000	Prospective. Patients, age 18–65 y, with possible appendicitis but missing 1 or more classic clinical findings. Subjects randomized to CT or US. No description of follow-up protocol for nonsurgical patients.	40 (US)	76	90
Dilley et al [27], 2001	Retrospective review of all rule-out appendicitis US cross-referenced to surgical and pathology reports. Pediatric patients. Study includes repeat US. No description of follow-up protocol for nonsurgical patients.	2056	89[a] 93[b]	95[a] 95[b]
Lowe et al [89], 2001	Retrospective. Consecutive children with suspected appendicitis underwent CT and compared with a historical cohort of consecutive children who underwent US. 12 children with negative imaging studies lost to follow-up and excluded.	78 (US)	100	88
Rettenbacher et al [66], 2002	Prospective. Patients, age 5–92 y, with suspected appendicitis. All underwent US. Clinical follow-up for nonsurgical patients.	350	98[c] 97[d]	98[c] 100[d]
Poortman et al [64], 2003	Prospective. Adult and pediatric patients with suspected appendicitis. Subjects underwent both US and CT. For nonsurgical patients median follow-up was 13 mo.	226	79	78
Kessler et al [68], 2004	Prospective. Patients, age 15–83 y, with suspected appendicitis. Follow-up in nonsurgical patients was chart review and telephone call at least 6 mo from discharge. Specific US findings were evaluated for diagnostic performance. Appendiceal diameter ≥6 mm was most sensitive and specific finding for appendicitis; sensitivity and specificity for this finding reported here.	125	98	98
Keyzer et al [62], 2005	Prospective. Patients, age 16–81 y, referred for CT for suspected appendicitis; all patients underwent both CT and US. Follow-up obtained from medical record and telephone call at 1 mo.	94	77	87
Doria et al [61], 2006	Meta-analysis (adults): study design described in text.	N/A	83[e]	93[e]
Doria et al [61], 2006	Meta-analysis (children): study design described in text.	N/A	88[f]	94[f]
Gaitini et al [65], 2008	Retrospective. Adult patients with suspected appendicitis. 132 patients also underwent CT. Clinical follow-up (inpatient only) for nonsurgical patients; indeterminate results excluded from analysis (n = 17).	420	74	97
Poortman et al [63], 2009	Prospective. Adult patients with symptoms of appendicitis underwent US followed by CT for negative or indeterminate US. Clinical follow-up and inpatient observation for nonsurgical patients.	151 (US)	77	86

In the Dilley article, sensitivity and specificity were calculated for all (initial and repeat) US examinations performed during the study[a] and recalculated[b] based only on final US findings (ie, the initial US if the findings were conclusive, or the definitive US in those patients who had multiple studies). In the study by Rettenbacher and colleagues, sensitivity and specificity were calculated for appendicitis[c] and for appendicitis + alternative diagnosis.[d] Doria and colleagues calculated 95% CIs for their performance metrics: [e]sensitivity 95% CI 78%–87%, specificity 95% CI 90%–96%, [f]sensitivity 95% CI 86%–90%, specificity 95% CI 92%–95%.

Abbreviation: N/A, not applicable.

Table 3

Selected studies evaluating diagnostic performance of CT, from 1991 to 2011

Reference	Study design	N	Sensitivity (%)	Specificity (%)
Balthazar et al [90], 1991	Prospective. Patients with suspected appendicitis. CT scan with oral and IV contrast.	100	98	83
Malone et al [91], 1993	Prospective. Patients, age 4–91 y, with RLQ pain. Unenhanced CT scan from L3 to symphysis pubis.	211	87	97
Balthazar et al [60], 1994	Prospective. Consecutive patients, age 15–82 y, with "suggestive but not typical clinical and laboratory findings of appendicitis" evaluated by both CT and US. Clinical follow-up for nonsurgical patients.	100	96	89
Lane et al [92], 1997	Prospective. Adult patients with suspected appendicitis. Unenhanced helical CT from T12 to symphysis pubis. Clinical follow-up for nonsurgical patients.	109	90	97
Rao et al [93], 1997	Prospective. Patients, age 6–75 y, with suspected appendicitis. Focused, helical CT (appendiceal CT) with rectal contrast only. Follow-up for nonsurgical patients was review of outpatient records and telephone calls at 1 wk and 2 mo.	100	98	98
Lane et al [94], 1999	Prospective. Patients, age 8–86 y, with suspected appendicitis. Unenhanced, thin-slice helical CT scan. Clinical follow-up for nonsurgical patients.	300	96	99
Stroman et al [24], 1999	Prospective. Patients, age 13–89 y, with equivocal clinical findings for appendicitis. IV and oral contrast CT scan. Nonsurgical patients followed up by chart review.	107	92	85
Horton et al [25], 2000	Prospective. Patients, age 18–65 y, with possible appendicitis but missing 1 or more classic clinical finding; subjects randomized to CT or US. Focused, noncontrast CT from L3 to cecum. No description of follow-up protocol for nonsurgical patients.	49 (CT)	97	100
Walker et al [47], 2000	Randomized controlled trial of CT vs standard management. Adult patients with suspected appendicitis. Focused, helical CT scan with rectal contrast only. Telephone follow-up for nonsurgical patients.	128	94	100
Lowe et al [89], 2001	Retrospective. Consecutive children with suspected appendicitis underwent CT and compared with a historical cohort of consecutive children who underwent US. Unenhanced CT scan from inferior liver edge to symphysis pubis. 12 children with negative imaging studies lost to follow-up and excluded.	72	97	100
Wijetunga et al [95], 2001	Prospective. Patients, age 14–81 y, with suspected appendicitis but equivocal clinical findings. Focused, helical thin-slice CT scan of RLQ with oral contrast only. Follow-up for nonsurgical patients was telephone call and chart review over a period of 1–8 mo.	100	93	97
Wilson et al [48], 2001	Prospective. Patients, age 4–81 y, with suspected appendicitis. All subjects underwent limited pelvic CT with rectal contrast only and RLQ US. Nonsurgical patients followed up at 1 and 7 d.	99	94[a] 66[b]	92[a] 67[b]
Christopher et al [96], 2002	Prospective. Patients, age 5–77 y, with suspected appendicitis (those with obvious need for surgery excluded). Unenhanced helical CT. Telephone follow-up at 6–8 wk for nonsurgical patients.	101	77	94
Ege et al [97], 2002	Prospective. Patients, age 16–69 y, with RLQ pain but an equivocal clinical impression for appendicitis. Unenhanced thin-slice CT scan from L3 to symphysis pubis; clinical follow-up for nonsurgical patients.	296	96	98
Hershko et al [98], 2002	Prospective. Patients, age 15–83 y, with equivocal clinical impression for appendicitis. Helical CT scan with IV and oral contrast.	198	91	92
Poortman et al [64], 2003	Prospective. Patients (children and adults) with suspected appendicitis underwent both US and single–detector, helical, unenhanced CT scan. For nonsurgical patients, median follow-up was 13 mo.	226	76	83
in 't Hof et al [99], 2004	Prospective. Adults with suspected appendicitis. All subjects underwent unenhanced helical CT from T11 to symphysis pubis followed by a diagnostic/therapeutic laparoscopy. Patients followed to 6 wk.	113	95	100

Reference	Study design	N	Sensitivity (%)	Specificity (%)
Keyzer et al [62], 2005	Prospective. Patients, age 16–81 y, referred for CT for suspected appendicitis; all patients underwent unenhanced 4-MDCT and US. Follow-up obtained from medical record and telephone call at 1 mo.	94	87	92
Doria et al [61], 2006	Meta-analysis (adults): study design described in text.	N/A	94[c]	94[c]
Doria et al [61], 2006	Meta-analysis (children): study design described in text.	N/A	94%[d]	95%[d]
Mun et al [100], 2006	Retrospective. Patients, age >17 y, referred for CT for acute abdominal pain or suspected appendicitis. Helical CT scan with IV contrast only. Patients with equivocal CT studies excluded from analysis (n = 4). No in-person follow-up; patients excluded if no documented clinical evaluation within 1 wk (35 of original 271 patients thus lost to follow-up).	173	100	97
Dearing et al [101], 2008	Retrospective. Patients, age 2–78 y, with suspected appendicitis but equivocal clinical findings. 16-MDCT scan with various contrast regimens (no rectal contrast). No description of follow-up protocol.	254	90	98
Gaitini et al [65], 2008	Retrospective. Consecutive adult patients referred for US for suspected appendicitis with 132 patients also undergoing follow-up 16-MDCT (oral and IV contrast). Clinical follow-up (inpatient only) for nonsurgical patients; indeterminate CT results not included in analysis (n = 1).	132 (CT)	100	99
Kim et al [51], 2008	Prospective. Patients, age >15 y, with suspected appendicitis. 16-MDCT with IV contrast only; there were no indeterminate CT scans. Telephone follow-up at 3 mo for nonsurgical patients.	157	98	96
Platon et al [71], 2009	Prospective. Adult patients with suspected appendicitis. Low-dose CT with oral contrast only followed by standard-dose CT and IV contrast (4-MDCT). Discharge diagnosis used for nonsurgical patients.	86	100[e] 95[f]	96[e] 96[f]
Poortman et al [63], 2009	Prospective. Adult patients with suspected appendicitis. US was followed by IV-enhanced 8-MDCT if US negative or indeterminate. Clinical follow-up and inpatient observation for nonsurgical patients.	60 (CT)	100	100
Van Randen et al [70], 2010	Prospective. Adult patients with suspected appendicitis. Subjects underwent both US and IV-enhanced abdominal CT. Diagnostic performance reflects presence of 2 or more essential CT findings.	942	96	95
Pickhardt et al [69], 2011	Retrospective. Consecutive adult patients referred for CT for suspected appendicitis or for signs/symptoms of appendicitis (ie, RLQ pain). Most patients received IV and oral contrast. Initially, 4-MDCT and 8-MDCT used; by end of study 8-MDCT, 16-MDCT, and 64-MDCT used. For nonsurgical patients, 12-mo clinical follow-up.	2871	99	98

Unless otherwise noted, CT scans were obtained of the entire abdomen and pelvis (ie, nonfocused). In the study by Wilson and colleagues, sensitivity and specificity were calculated with equivalent scans excluded[a] and with equivalent scans included.[b] Doria and colleagues calculated 95% CIs for their performance metrics: for adults,[c] sensitivity 95% CI 92%–95% and specificity 95% CI 94%–96%; for children,[d] sensitivity 95% CI 92%–97% and specificity 95% CI 94%–97%. Platon and colleagues calculated performance metrics for standard radiation dose CT[e] and for low radiation dose CT.[f]

Abbreviations: IV, intravenous; MDCT, multidetector CT; N/A, not applicable; RLQ, right lower quadrant.

Table 4

Selected studies evaluating diagnostic performance of magnetic resonance imaging, from 2009 to 2012

Reference	Study design	N	Sensitivity	Specificity
Cobben et al [74], 2009	Prospective. Patients, age 6–80 y, with clinically suspected appendicitis. Breath hold, unenhanced MRI with T1, T2, and T2 fat-suppressed sequences. Nonsurgical patients followed up in clinic at 1 wk and followed as outpatient for at least 2 y.	138	100	98
Barger and Nandalur [73], 2010	Meta-analysis. 8 prospective and retrospective articles included; studies included if the absolute numbers of true-positive, true-negative, false-positive, and false-negative results were reported.	363	97[a]	95[a]
Chabanova et al [102], 2011	Prospective. Adult patients with clinically diagnosed appendicitis. Unenhanced MRI with T1, T2, and fat-suppressed sequences. All patients underwent surgery.	48	87[b]	61
Inci et al [72], 2011	Prospective. Patients, age 14–72 y, with clinically suspected appendicitis. Unenhanced MRI with T1, T2, and fat-suppressed sequences. Clinical follow-up for 1–3 mo for nonsurgical patients.	85	96.5	66.7[c]
Johnson et al [103], 2012	Prospective. Pediatric patients, age 4–17 y, with acute abdominal pain suspicious for appendicitis. All subjects underwent MRI as well as CT or US. (MR sequences = single-shot turbo spin-echo T2 in 3 planes and T2 fat-suppressed axial sequence.) No specific mention of follow-up protocol for nonsurgical patients.	42	100	99
Moore et al [104], 2012	Retrospective. Pediatric patients, age 3–17 y, with suspected appendicitis. Unenhanced MRI with single-shot turbo spin-echo coronal T2, axial T2, and axial fat saturation T2, as well as a coronal SPAIR. Follow-up for nonsurgical patients based on chart review and discharge diagnosis.	208	97.6	97.0

Barger and colleagues calculated 95% CIs for their meta-analysis; for sensitivity,[a] 95% CI is 92%–99% and for specificity,[a] it is 94%–99%. In the study by Chabanova and colleagues, although MRI was 87% sensitive for appendicitis,[b] it was 100% sensitive in detecting pelvic abnormalities (ie, the radiologist called appendicitis but other disease found in the operating room). Of the 4 patients with no pathologic abnormalities, 3 were correctly identified by MRI, yielding a specificity of 75%. In the study by Inci and colleagues, by our calculation, specificity[c] is 89.3% (true negatives/[true negatives + false positives]); the investigators report 25 true-negative, 2 false-negative, 55 true-positive, and 3 false-positive cases ("probable appendicitis" was treated as a positive result in this study). Sensitivity of 96.5% is consistent with the raw numbers reported.

Abbreviation: SPAIR, spectral adiabatic inversion recovery.