# A highly sensitive and specific combined clinical and sonographic score to diagnose appendicitis

**Swathi B. Reddy, MD, Michael Kelleher, MD, S.A. Jamal Bokhari, MBBS, Kimberly A. Davis, MD,** *and* **Kevin M. Schuster, MD***, New Haven, Connecticut*

## AAST Continuing Medical Education Article

### Accreditation Statement

This activity has been planned and implemented in accordance with the Essential Areas and Policies of the Accreditation Council for Continuing Medical Education through the joint providership of the American College of Surgeons and the American Association for the Surgery of Trauma. The American College Surgeons is accredited by the ACCME to provide continuing medical education for physicians.

### *AMA PRA Category 1 Credits™*

The American College of Surgeons designates this journal-based CME activity for a maximum of 1 *AMA PRA Category 1 Credit*™. Physicians should claim only the credit commensurate with the extent of their participation in the activity.

Of the *AMA PRA Category 1 Credit*™ listed above, a maximum of 1 credit meets the requirements for self-assessment.

Credits can only be claimed online

AMERICAN COLLEGE OF SURGEONS
*Inspiring Quality:*
*Highest Standards, Better Outcomes*

100+*years*

### Objectives

After reading the featured articles published in the *Journal of Trauma and Acute Care Surgery*, participants should be able to demonstrate increased understanding of the material specific to the article. Objectives for each article are featured at the beginning of each article and online. Test questions are at the end of the article, with a critique and specific location in the article referencing the question topic.

### Claiming Credit

To claim credit, please visit the AAST website at http://www.aast.org/ and click on the "e-Learning/MOC" tab. You must read the article, successfully complete the post-test and evaluation. Your CME certificate will be available immediately upon receiving a passing score of 75% or higher on the post-test. Post-tests receiving a score of below 75% will require a retake of the test to receive credit.

### Disclosure Information

In accordance with the ACCME Accreditation Criteria, the American College of Surgeons, as the accredited provider of this journal activity, must ensure that anyone in a position to control the content of *J Trauma Acute Care Surg* articles selected for CME credit has disclosed all relevant financial relationships with any commercial interest. Disclosure forms are completed by the editorial staff, associate editors, reviewers, and all authors. The ACCME defines a `commercial interest' as "any entity producing, marketing, re-selling, or distributing health care goods or services consumed by, or used on, patients." "Relevant" financial relationships are those (in any amount) that may create a conflict of interest and occur within the 12'months preceding and during the time that the individual is engaged in writing the article. All reported conflicts are thoroughly managed in order to ensure any potential bias within the content is eliminated. However, if you'perceive a bias within the article, please report the circumstances on the evaluation form.

Please note we have advised the authors that it is their responsibility to disclose within the article if they are describing the use of a device, product, or drug that is not FDA approved or the off-label use of an approved device, product, or drug or unapproved usage.

### Disclosures of Significant Relationships with Relevant Commercial Companies/Organizations by the Editorial Staff

Ernest E. Moore, Editor: PI, research support and shared U.S. patents Haemonetics; PI, research support, TEM Systems, Inc. Ronald V. Maier, Associate editor: consultant, consulting fee, LFB Biotechnologies. Associate editors: David Hoyt and Steven Shackford have nothing to disclose. Editorial staff: Jennifer Crebs, Jo Fields, and Angela Sauaia have nothing to disclose.

### Author Disclosures

The authors have nothing to disclose.

### Reviewer Disclosures

The reviewers have nothing to disclose.

### Cost

For AAST members and *Journal of Trauma and Acute Care Surgery* subscribers there is no charge to participate in this activity. For those who are not a member or subscriber, the cost for each credit is $25.

### System Requirements

The system requirements are as follows: Adobe® Reader 7.0 or above installed; Internet Explorer® 7 and above; Firefox® 3.0 and above, Chrome® 8.0 and above, or Safari™ 4.0 and above.

### Questions

If you have any questions, please contact AAST at 800-789-4006. Paper test and evaluations will not be accepted.

Submitted: August 2, 2016, Revised: April 14, 2017, Accepted: April 19, 2017, Published online: April 28, 2017.
From the Department of Surgery (S.B.R., K.A.D., K.M.S.); and Department of Radiology and Biomedical Imaging (M.K., J.B.), Yale School of Medicine, New Haven, Connecticut.
This paper was presented as a podium presentation at the 75th Annual Meeting of AAST and Clinical Congress of Acute Care Surgery, September 14–17, 2016, Waikoloa, HI.
Address for reprints: Kevin M. Schuster, MD, Department of Surgery, Yale School of Medicine, 330 Cedar Street, BB310, PO Box 208062, New Haven, CT 06520-8062; email: kevin.schuster@yale.edu.

DOI: 10.1097/TA.0000000000001551

**BACKGROUND:** Computed tomography (CT) scanning reduces the negative appendectomy rate however it exposes the patient to ionizing radiation. Ultrasound (US) does not carry this risk but may be nondiagnostic. We hypothesized that a clinical-US scoring system would improve diagnostic accuracy.

**METHODS:** We conducted a retrospective review of all patients (age, >15 years) who presented through the emergency department with suspected appendicitis and underwent initial US. A US score was developed using odds ratios for appendicitis given appendiceal diameter, compressibility, hyperemia, free fluid, and focal or diffuse tenderness. The US score was then combined with the Alvarado score. Final diagnosis of appendicitis was assigned by pathology reports.

**RESULTS:** Three hundred patients who underwent US as initial imaging were identified. Thirty-two patients with evident nonappendiceal pathology on US were excluded. In 114 (38%), the appendix was not visualized and partially visualized in 36 (12%). Fifty-seven (21.3%) had an appendectomy with 1 (1.7%) negative. Six nonvisualized appendicies underwent appendectomy, with no negative cases. Sensitivity and specificity for the sonographic score were 86% and 90%, respectively, at a score of 1.5. The combined score demonstrated 98% sensitivity and 82% specificity at 6.5, and 95% sensitivity, and 87% specificity at a score of 7.5. Sensitivity and specificity were confirmed by bootstrap resampling for validation. Area under receiver operating characteristic (ROC) curves for our new US score were similar to the ROC curve for the Alvarado score (91.9 and 91.1, $p = 0.8$). The combined US and Alvarado score yielded an area under the ROC curve of 97.1, significantly better than either score alone ($p = 0.017$ and $p < 0.001$, respectively).

**CONCLUSION:** Our scoring system based entirely on US findings was highly sensitive and specific for appendicitis, and it significantly improved when combined with the Alvarado score. After prospective evaluation, the combined US-Alvarado score might replace the need for computed tomography imaging in a majority of patients. (*J Trauma Acute Care Surg.* 2017;83: 643–649. Copyright © 2017 Wolters Kluwer Health, Inc. All rights reserved.)

**LEVEL OF EVIDENCE:** Diagnostic Test, Level III.

**KEY WORDS:** Appendicitis; ultrasound; Alvarado score.

Although acute appendicitis was at one time considered a clinical diagnosis, ultrasound (US) and computed tomography (CT) scan have become common tools in the diagnostic algorithm. Despite advances in imaging and laboratory investigation, the negative appendectomy rate is reported as high as 5% to 10%.[1] Although CT scanning reduces the negative appendectomy rate to less than 4% with the increasingly frequent use in the emergency room, it exposes the patient to higher lifetime risk of radiation-induced cancers.[2] This is a particularly salient concern for younger patients, and accordingly much of the effort to minimize CT scan has been studied in the pediatric population.[3]

Ultrasound represents a viable alternative to CT scan, but is less commonly used partly due to a high rate of inconclusive studies, up to 50% to 70% in some reports,[4–7] and the high interuser variability.[8] Accordingly, many patients with equivocal findings on US will eventually receive a CT scan, approximately 45% in one study of pediatric patients.[9] However, many of these patients with nonvisualized appendices or equivocal findings may not ultimately have appendicitis.[7,10] Multiple studies have attempted to identify patients with equivocal US findings who can safely avoid additional imaging with CT. Stewart et al.[11] found that only 5.4% of patients with a nonvisualized appendix were eventually found to have pathologically confirmed appendicitis. Similarly, Cohen et al.[10] calculated the negative predictive value (NPV) of a "nondiagnostic" US, subcategorizing nonvisualized and borderline findings and determined that a nonvisualized appendix had a NPV of 86%.

Because findings on US can be difficult to assess objectively, it would be valuable to establish a weighting system for specific findings. It may then be possible to maximize the diagnostic value of US and minimize the need for diagnostic testing using ionizing radiation. We aimed to formulate a scoring system that incorporates the well-established Alvarado score for clinical findings with a scoring system of sonographic parameters to better predict the presence of acute appendicitis. We hypothesized that our combined scoring system would outperform both the Alvarado score and US findings when examined independently.

## METHODS

After obtaining approval from the Yale Human Investigations Committee, we conducted a retrospective review of 300 medical records of patients who presented through the Yale New Haven Hospital Emergency Department with suspected appendicitis between August 2014 and June 2015. Patients were included in the study if they were older than 15 years and underwent US of the abdomen as the initial imaging study. Patients were excluded if any pathology unrelated to appendicitis but likely responsible for the patient's abdominal pain was demonstrated on US or if the patient was known to be pregnant.

The available history and examination findings from the medical record were reviewed to retrospectively calculate an Alvarado score.[12] In addition to recording presenting symptoms and examination findings, variables of interest included any subsequent imaging, whether a surgical consult was obtained, operative findings if undergoing appendectomy, pathology findings, total length of stay, and discharge disposition including any readmission within 7 days. Although we cannot confirm any evaluation at an outside institution, we recorded any reevaluation in an emergency department within our health system within 30 days.

Imaging reports were reviewed and images were reanalyzed by a single staff radiologist with expertise in emergency imaging. Based on review of the literature, a consensus among the investigators was developed with respect to the US parameters most likely to predict appendicitis. Odds ratios (OR) for the ability of each of the chosen parameters to predict appendicitis were calculated. Chosen parameters included appendiceal diameter, compressibility, hyperemia, and secondary signs of inflammation: free fluid and focal or diffuse tenderness. Free fluid was further categorized into no free fluid, small or trace free fluid, and moderate free fluid. Tenderness elicited by the US probe was subcategorized into no tenderness, diffuse tenderness, and focal tenderness over the appendix. This was

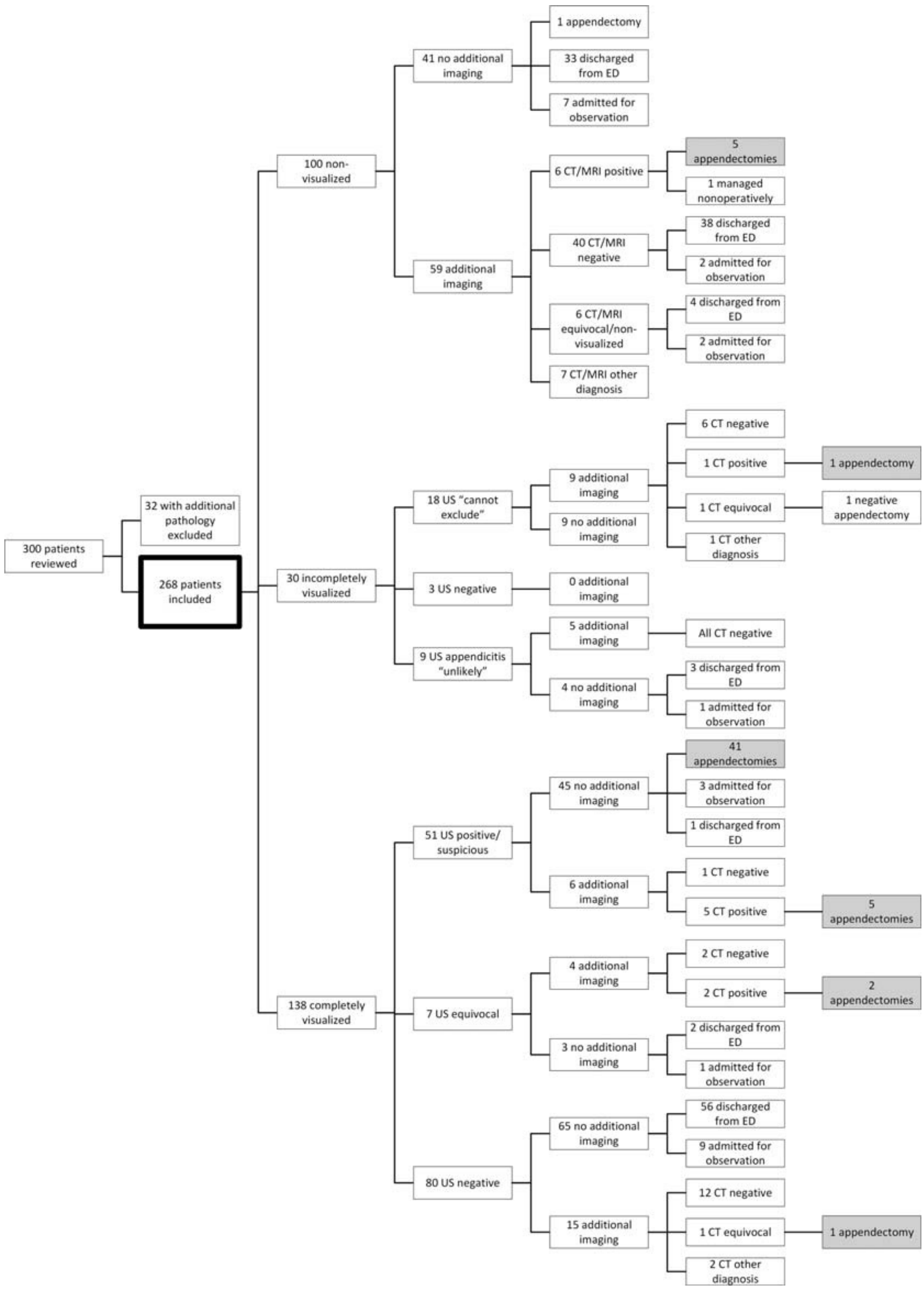


**Figure 1.** Patient flow and additional imaging. All appendectomies were positive on pathology unless otherwise indicated.

distinct from the variable derived from the physical examination finding for tenderness during the physician evaluation. If multiple physical examinations were performed, that of the surgical staff was prioritized. If no surgical staff evaluated the patient, the emergency room attending's history and examination were documented over any other emergency room staff's evaluation. If the appendix was not visualized, only secondary signs were documented. Final diagnosis of appendicitis was assigned by reviewing pathology reports or safe discharge home without representation with appendicitis during the ensuing 7 days.

**TABLE 1.** Cohort Characteristics

| Descriptives | n | %* |
|---|---|---|
| Total population | 268 | |
| Age | | |
| Median (IQR) | 23 (19–23) | |
| Gender | | |
| Female | 199 | 74.3 |
| BMI | | |
| Median (IQR) | 23.9 (21.3–28) | |
| Discharge from ED | 175 | 65.3 |
| CT or MRI | 100 | 37.3 |
| Surgery | 57 | 21.3 |
| Presenting symptoms | | |
| Nausea | 183 | 68.5 |
| Anorexia | 110 | 55.2 |
| Migration of pain to RLQ | 64 | 28.8 |
| RLQ tenderness | 225 | 90 |
| Rebound tenderness | 49 | 18.3 |
| Fever > 99.1°F | 76 | 28.4 |
| Left shift | 121 | 49.1 |

*Percent of valid total.
BMI, body mass index; RLQ, right lower quadrant.

## Calculation of Sonographic Score and Analysis

The sonographic scoring criteria were determined by the OR for a positive appendicitis diagnosis given the presence of each parameter. The points assigned were based on relative approximations of the OR rather than exact values. By assigning simple integer values, we hoped to simplify the calculation of a score in a clinical setting. The total score was calculated for each patient after assigning individual scores for each parameter. Accordingly, the maximum possible sonographic score value was 8. Negative or unrecorded findings were given a parameter score of 0. When the appendix could not be visualized the maximum score was 4 for the secondary signs (free fluid and focal tenderness elicited by the US probe).

**TABLE 2.** Sonographic Score Point Assignments

| Parameters | Points |
|---|---|
| Compressibility | |
| Compressible | 0 |
| Not compressible | 2 |
| Free Fluid | |
| None | 0 |
| Small | 0.5 |
| Moderate | 1 |
| Tenderness | |
| No tenderness | 0 |
| Diffuse | 1 |
| Focal | 2 |
| Hyperemia | |
| No hyperemia | 0 |
| Hyperemia | 1 |
| Size | |
| Not visualized or ≤5 mm | 0 |
| 6–10 mm | 1 |
| ≥11 mm | 2 |

**TABLE 3.** Visualization of the Appendix and Scores

| | | No Appendicitis | | | Appendicitis | | |
|---|---|---|---|---|---|---|---|
| Visualization | Score | n | Median | IQR | n | Median | IQR |
| Incomplete | Alvarado | 29 | 4 | 4.0–5.0 | 1 | 8.5 | 8.0–8.0 |
| | Sonographic | | 0 | 0.0–0.5 | | 2.25 | 2.5–2.5 |
| | Combined | | 4 | 4.0–6.0 | | 10.75 | 10.5–10.5 |
| No | Alvarado | 94 | 4 | 3.0–5.0 | 6 | 9 | 9.0–9.0 |
| | Sonographic | | 0 | 0.0–5.0 | | 0 | 0.0–0.0 |
| | Combined | | 4 | 3.0–5.5 | | 9 | 9.0–10.0 |
| Yes | Alvarado | 84 | 3 | 2.0–5.0 | 49 | 8 | 7.0–9.0 |
| | Sonographic | | 0 | 0.0–1.0 | | 4.5 | 3.0–5.5 |
| | Combined | | 4 | 3.0–6.0 | | 12.5 | 10.0–13.5 |

Cumulative sensitivity and specificity were studied by calculating receiver operated characteristic (ROC) curves for the three scoring systems. The three scoring systems were the Alvarado score, the sonographic score, and the Combined Alvarado-Ultrasound Score. Differences in area under the ROC curves (AUCs) were calculated using the nonparametric methods of DeLong et al.[13] Bootstrap resampling was used to internally validate our data set. Sampling was unrestricted and random with replicates. One thousand replicate data sets were created with a sample size equal to our original data set (n = 268). Sensitivity and specificity for the combined score were calculated in each simulated data set (by group analyses) as well as mean sensitivities and specificities with 95% confidence intervals (CIs), which were compared with our raw data. All statistical analyses were conducted on deidentified data using SAS 9.4 (SAS institute, Cary, NC).

# RESULTS

We identified 300 patients being evaluated for suspected appendicitis, who underwent US as initial imaging. Thirty-two patients with evident nonappendiceal pathology on US were excluded from analysis. Of the remaining 268 patients, 74% were females and the median age was 23 years. Fifty-seven (21.3%) patients ultimately had an appendectomy during the same admission with one (1.7%) negative operation (Fig. 1).

Ninety-nine (36.9%) of our patients received additional imaging after their US, either CT or magnetic resonance imaging (MRI) (Table 1). Thirteen (13%) of these patients had a final diagnosis of positive appendicitis. Sixty-three percent of patients with a nonvisualized appendix underwent additional imaging,

**TABLE 4.** Sensitivities and Specificities for the Combined Score

| Combined Score | Sensitivity (%) | Specificity (%) | PPV (%) | NPV (%) |
|---|---|---|---|---|
| 6.5 | 98.2 | 82.1 | 59.1 | 99.4 |
| 7 | 96.4 | 84.4 | 62.1 | 98.9 |
| 7.5 | 94.6 | 86.8 | 65.4 | 98.4 |
| 8 | 92.9 | 91.0 | 73.2 | 98.0 |
| 8.5 | 91.1 | 94.8 | 82.2 | 97.6 |

**TABLE 5.** ROC Statistics

**ROC Association Statistics**

| | Mann-Whitney | | | |
|---|---|---|---|---|
| | AUC | Standard Error | 95% Wald Confidence Limits | |
| Sonographic score | 0.912 | 0.029 | 0.856 | 0.967 |
| Alvarado score | 0.919 | 0.019 | 0.882 | 0.956 |
| Combined score | 0.971 | 0.014 | 0.944 | 0.999 |

and 15% of patients with an incompletely visualized appendix received a CT scan (Fig. 1). Twenty-six patients underwent additional imaging even though the appendix was completely visualized on US. From the total population, 175 (65.3%) were discharged from the hospital without any intervention. Of the 175 patients who were discharged from the ED without intervention, 34 patients were reevaluated in the emergency room during the following 30 days. Twenty-six of these patients were evaluated for abdominal pain related to their initial presentation. Fifteen of these patients initially had a nonvisualized or an incompletely visualized appendix. None of the patients who returned for reevaluation were diagnosed with appendicitis.

The appendix was not visualized in 100 (37.3%) patients and partially visualized in 30 (11.1%) patients. Fourteen (47%) of the patients with an incompletely visualized appendix underwent additional imaging, as did 59 (59%) of the patients with a nonvisualized appendix. Of the patients with an incompletely or nonvisualized appendix, two and six respectively, ultimately underwent appendectomy (Fig. 1). All of these patients had acute appendicitis on pathology. Of the 138 patients with a completely visualized appendix on US, 41 (29.7%) underwent an appendectomy without any additional imaging being performed. Of these same 138 patients, 25 (18.1%) received additional imaging, eight of whom ultimately underwent appendectomy. The single patient in our population with a negative appendectomy had an incompletely visualized 7 mm appendix with hyperemia, as well as a CT scan with equivocal findings. In all, there were six cases of appendicitis that were missed (appendix not visualized or incompletely visualized) by US. One of these missed cases was a retrocecal appendix, found to be gangrenous upon surgical exploration. One other case was found to be perforated appendicitis. In both of these cases, no free fluid was identified on US, and presence or absence of tenderness elicited by the probe was not documented in the radiology report.

A maximum diameter of 11 mm or greater had 100% specificity for acute appendicitis (OR for less than 5 mm or nonvisualized relative to 11 mm or greater of 0.001 by firth logistic regression). The OR for focal tenderness elicited by the US probe and for noncompressibility of the appendix were 176 and 153, respectively. The sum of the points assigned for each parameter was compiled into a sonographic score (Table 2). The sonographic score was then added to the Alvarado score to compute a combined score.

The median Alvarado score was 4 (interquartile range [IQR], 3.0–4.0), from a maximum possible of 10; however, for patients who underwent surgery, it was 8. The median sonographic score overall was 0 (IQR, 0–1.5), due to the many studies with a nonvisualized appendix. However, in patients who had surgery, the median sonographic score was 4 from a maximum possible of 8. The median combined score (Alvarado score + sonographic score) was 5 (IQR 3.0–8.0), from a maximum total of 18. In contrast, among patients who had surgery, the median combined score was 12. Table 3 reviews the median scores by degree of visualization of the appendix on US.

The sensitivity and specificity of the sonographic and combined scores were calculated at each point interval. The optimal sensitivity and specificity for the sonographic score alone were 86% and 90%, respectively, at a score of 1.5. These improved to 98% sensitivity and 82% specificity at a combined score of 6.5 and 95% sensitivity and 87% specificity at a score of 7.5 (Table 4). The positive predictive value (PPV) and NPV at 7.5 were 65% and 98%, respectively. In our bootstrap simulated validation data set, the mean sensitivity and mean specificity at a combined score of 7.5 were 94.7% (95% CI, 94.5–94.8) and 86.8% (95% CI, 86.6–86.9), respectively.

Area under receiver operating characteristic (ROC) curves were calculated for each curve (Table 5). The AUC for our sonographic score was similar to that of the ROC curve for the Alvarado scores (91.9 and 91.1, respectively, $p$ = 0.8). The combined score yielded an AUC of 97.1. This was significantly better than either the sonographic score or Alvarado score alone ($p$ = 0.017 and $p$ < 0.001, respectively).

## DISCUSSION

Although appendicitis is a diagnosis made primarily based on clinical findings, CT scanning has become ubiquitous, especially in cases where the presentation or examination findings are atypical. Although US does not carry the risk of radiation exposure associated with CT scan, the quality of an US is known to be dependent on the sonographer's skill and experience and carries a lower sensitivity and specificity.[8] Thus, a method of maximizing the utility of the information gained from an US study should greatly facilitate clinical decision-making.

To this end, a few earlier studies have aimed to integrate US findings with the Alvarado score to determine a threshold at which point further imaging or surgery is required. Toprak and colleagues retrospectively divided pediatric patients into groups based on Alvarado score and also based on US findings, including secondary signs of inflammation.[5] They found that there were only four of 122 patients who had a nonvisualized appendix and also a high Alvarado score, ultimately requiring surgical intervention, concluding that an Alvarado score greater than 6 even with a nonvisualized appendix required further investigation. Leeuwenburg et al.[14] formulated a clinical decision rule to assist in determining which patients with inconclusive US findings could be safely discharged home without subsequent CT or MRI. In the pediatric population, Fallon et al.[15] developed a radiologic reporting score based on sonographic parameters alone to predict appendicitis with a high PPV.

We successfully integrated objectively measured sonographic findings with an established clinical score to predict appendicitis with both a high specificity and sensitivity. Although US alone is known to be highly specific for appendicitis, 95.9% in one recent meta-analysis, it has an overall lower sensitivity, 83.7% in the same study, which has limited its usefulness for the diagnosis of appendicitis. Using our combined clinical and sonographic score, we were able to predict appendicitis with 98% sensitivity, markedly improving on the literature accepted sensitivity. More importantly, our combined score was a significantly better predictor of appendicitis than either the Alvarado score or the sonographic score alone. As such, it both validates the utility of US as an adjunct to examination findings and represents an improvement upon the current standard clinical decision making rule. Additionally, the sonographic component of our score showed an equally high sensitivity and specificity as that of the Alvarado score alone, as evidenced by the nearly identical AUC. This implies that even in a situation where history and physical examination are not useful, US could be valuable as the sole diagnostic method.

One limitation of US as a primary mode of imaging for appendicitis is its high rate of nonvisualization. The high sensitivity and specificity of our combined score, however, was demonstrated even with the inclusion of patients with nonvisualized and partially visualized appendices likely indicating complimentary roles of the clinical and US components of the score. There is evidence that incomplete visualization in the absence of other signs of appendicitis is not an indication for additional imaging. One recent study reported that in pediatric patients without appendicitis, there was a nonvisualization rate of 97%.[16] However, this figure varies, and in pediatric patients, as many as 15% of patients with an incompletely visualized appendix have been reported to have appendicitis.[7] In our population, only 6.9% of the patients in our study with a nonvisualized appendix or an incompletely visualized appendix required surgery. Provided there is otherwise good performance of our combined scoring system there should be few false negatives.

This study builds on prior efforts by incorporating a clinical evaluation with a standardized report of sonographic findings. In addition to efforts to stratify US findings based on clinical scores[17,18] there have been several attempts to standardize the interpretation of specific sonographic findings. Trout et al.[19] measured appendiceal diameters in children to form a logistic predictive model identifying size cutoffs for positive, negative, and equivocal categories for appendicitis. In another investigation conducted in pediatric patients, sonographic findings were assembled to categorize a US study as diagnostic or nondiagnostic.[15] Their scoring criteria were based entirely on sonographic findings, including secondary signs of inflammation, such as presence of free fluid and periappendiceal fat edema. The score developed by these authors, however, was an effort to standardize radiologic reporting rather than a clinical decision-making tool.[15]

In the current study, certain sonographic parameters were more predictive of appendicitis than others. For instance, a maximal appendiceal diameter above 10 mm had a predictive value of 100%. Likewise, a noncompressible appendix was associated with a 4.5 times greater odds of appendicitis. Trace-free fluid, although a positive predictor, had a lower OR for appendicitis (as compared with no free fluid), possibly indicative of another intra-abdominal process unrelated to the appendix. There are other secondary signs of inflammation, such as identification of lymph nodes and periappendiceal fat, which could be included in this scoring system in future studies but were too inconsistently reported to include in this study.

Essentially all patients in the modern emergency room will receive some type of confirmatory imaging before pursuing surgical intervention. When US results are equivocal, additional imaging is usually pursued. In our population, 46.7% of the patients with an incompletely visualized appendix ultimately received a CT scan or MRI, as did 59% of patients with a nonvisualized appendix. Although this combination of imaging modalities may allow for more confident diagnosis, it continues to expose patients to unnecessary radiation and warrants continued efforts to validate the diagnostic value of US. Even when information obtained from additional imaging is ignored, our scoring system based entirely on clinical examination and US findings continued to be highly sensitive and specific for an appendicitis and offers an opportunity to forego unnecessary imaging even in the case of an unclear clinical presentation. Ideally, this score would be used by an emergency room provider before ordering a CT and before obtaining a surgical consult.

Our score is an attempt to improve the accuracy of US when used in conjunction with clinical findings. Because the excellent performance of CT scan is well known, we believe that it is clinically significant that we can use US in a clinician selected patient group to achieve equivalent performance characteristics. After we complete our ongoing prospective validation study, we hope to significantly reduce the rate of CT scans performed for suspected appendicitis in our institution. We conditionally recommend limiting imaging to US except in cases with a borderline score and an unclear clinical presentation.

### Limitations

There are several limitations to our study. Due to the retrospective nature of our image review, it was not possible to actively assess for compressibility of the appendix or tenderness. As such, we relied on the sonographer's record-keeping and documentation in the radiology reports at the time of evaluation, and in many cases, these parameters were not recorded. Therefore, there may have been an underlying selection bias in reporting these findings. As we go forward with prospective validation of the score, we are using a checklist to make certain that many parameters as possible are evaluated by the sonographer and radiologist diad for more complete data.

A CT scan is often immediately performed in patients with a high body mass index because of the concern that US will not identify the appendix. Our study does not include patients who proceeded to CT scan without an US, and as such, there is a possibility that our sensitivity and specificity could have been affected by this selection bias. Additionally, we did not have any cases of patients discharged from the emergency room after a negative US who returned to our institution with appendicitis. Because we did not contact patients, we cannot be certain that they did not require additional imaging or intervention at another institution.

## CONCLUSION

We have demonstrated in this initial evaluation of a combined clinical and sonographic scoring system that very high sensitivity and specificity can be achieved for diagnosis of appendicitis without the use of ionizing radiation. Although our ongoing prospective validation will be important, this scoring system and some if its key components including appendiceal size and compressibility can be used to make decisions regarding treatment.

## AUTHORSHIP

S.B.R., S.A.J.B., K.A.D., K.M.S. participated in the study design. S.B.R., M.K., K.M.S. participated in acquisition of data. S.B.R., M.K., and K.M.S. participated in data analysis. S.B.R., M.K., S.A.J.B., K.A.D., and K.M.S. participated in the drafting and critical revision.

## DISCLOSURE

The authors declare no conflicts of interest.

## REFERENCES

1. Man E, Simonka Z, Varga A, Rarosi F, Lazar G. Impact of the Alvarado score on the diagnosis of acute appendicitis: comparing clinical judgment, Alvarado score, and a new modified score in suspected appendicitis: a prospective, randomized clinical trial. *Surg Endosc*. 2014;28(8):2398–2405.
2. Brenner DJ, Hall EJ. Computed tomography—an increasing source of radiation exposure. *N Engl J Med*. 2007;357(22):2277–2284.
3. Kotagal M, Richards MK, Chapman T, Finch L, McCann B, Ormazabal A, Rush RJ, Goldin AB, Safe and Sound Campaign. Improving ultrasound quality to reduce computed tomography use in pediatric appendicitis: the Safe and Sound campaign. *Am J Surg*. 2015;209(5):896–900; discussion.
4. Mostbeck G, Adam EJ, Nielsen MB, Claudon M, Clevert D, Nicolau C, Nyhsen C, Owens CM. How to diagnose acute appendicitis: ultrasound first. *Insights Imaging*. 2016;7(2):255–263.
5. Krishnamoorthi R, Ramarajan N, Wang NE, Newman B, Rubesova E, Mueller CM, Barth RA. 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):231–239.
6. Blitman NM, Anwar M, Brady KB, Taragin BH, Freeman K. Value of focused appendicitis ultrasound and Alvarado score in predicting appendicitis in children: can we reduce the use of CT? *AJR Am J Roentgenol*. 2015; 204(6):W707–W712.
7. Ross MJ, Liu H, Netherton SJ, Eccles R, Chen PW, Boag G, Morrison E, Thompson GC. Outcomes of children with suspected appendicitis and incompletely visualized appendix on ultrasound. *Acad Emerg Med*. 2014; 21(5):538–542.
8. Pinto F, Pinto A, Russo A, Coppolino F, Bracale R, Fonio P, Macarini L, Giganti M. Accuracy of ultrasonography in the diagnosis of acute appendicitis in adult patients: review of the literature. *Crit Ultrasound J*. 2013;5(Suppl 1):S2.
9. Ramarajan N, Krishnamoorthi R, Gharahbaghian L, Pirrotta E, Barth RA, Wang NE. Clinical correlation needed: what do emergency physicians do after an equivocal ultrasound for pediatric acute appendicitis? *J Clin Ultrasound*. 2014;42(7):385–394.
10. Cohen B, Bowling J, Midulla P, Shlasko E, Lester N, Rosenberg H, Lipskar A. The non-diagnostic ultrasound in appendicitis: is a non-visualized appendix the same as a negative study? *J Pediatr Surg*. 2015;50(6):923–927.
11. Stewart JK, Olcott EW, Jeffrey RB. Sonography for appendicitis: nonvisualization of the appendix is an indication for active clinical observation rather than direct referral for computed tomography. *J Clin Ultrasound*. 2012;40(8):455–461.
12. Alvarado A. A practical score for the early diagnosis of acute appendicitis. *Ann Emerg Med*. 1986;15(5):557–564.
13. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837–845.
14. Leeuwenburgh MM, Stockmann HB, Bouma WH, Houdijk AP, Verhagen MF, Vrouenraets B, Cobben LP, Bossuyt PM, Stoker J, Boermeester MA, OPTIMAP Study Group. A simple clinical decision rule to rule out appendicitis in patients with nondiagnostic ultrasound results. *Acad Emerg Med*. 2014;21(5):488–496.
15. Fallon SC, Orth RC, Guillerman RP, Munden MM, Zhang W, Elder SC, Cruz AT, Brandt ML, Lopez ME, Bisset GS. Development and validation of an ultrasound scoring system for children with suspected acute appendicitis. *Pediatr Radiol*. 2015;45(13):1945–1952.
16. Garcia Pena BM, Mandl KD, Kraus SJ, Fischer AC, Fleisher GR, Lund DP, Taylor GA. Ultrasonography and limited computed tomography in the diagnosis and management of appendicitis in children. *JAMA*. 1999;282(11): 1041–1046.
17. Jones RP, Jeffrey RB, Shah BR, Desser TS, Rosenberg J, Olcott EW. Journal Club: the Alvarado score as a method for reducing the number of CT studies when appendiceal ultrasound fails to visualize the appendix in adults. *AJR Am J Roentgenol*. 2015;204(3):519–526.
18. Toprak H, Kilincaslan H, Ahmad IC, Yildiz S, Bilgin M, Sharifov R, Acar M. Integration of ultrasound findings with Alvarado score in children with suspected appendicitis. *Pediatr Int*. 2014;56(1):95–99.
19. Trout AT, Towbin AJ, Fierke SR, Zhang B, Larson DB. Appendiceal diameter as a predictor of appendicitis in children: improved diagnosis with three diagnostic categories derived from a logistic predictive model. *Eur Radiol*. 2015;25(8):2231–2238.

 