

Appendiceal ultrasound: the importance of determinacy

Larry A. Binkovitz¹ · Kyle M. L. Unsdorfer² · Prabin Thapa³ · Amy B. Kolbe¹ ·
Nathan C. Hull¹ · Shannon N. Zingula¹ · Kristen B. Thomas¹ · James L. Homme⁴

Received: 29 June 2015 / Accepted: 2 July 2015 / Published online: 18 August 2015
© Springer-Verlag Berlin Heidelberg 2015

“The first job of the radiologist is to minimize doubt.”
— William E. Shiels, 1954–2015

We appreciate the thoughtful comments of Drs. Trout and Larson [1] concerning our article in this issue of *Pediatric Radiology* and the opportunity to further this discussion with a few additional thoughts. There are striking similarities between our two studies with respect to design and results but there are substantial differences in the conclusions drawn [2, 3].

Patients referred for US evaluation for acute appendicitis have already been stratified into an indeterminate risk group based on their clinical data. Those thought to have a very high probability of acute appendicitis may be sent directly to surgery and those thought to have a very low probability of appendicitis often are not imaged at all. It is for those patients whose probability of acute appendicitis is indeterminate that clinicians seek our help with US imaging. When clinicians order an imaging test to assist in establishing or excluding the diagnosis of acute appendicitis, they should know the likelihood that the test will provide a determinate result and how accurate that result will be. We presented our appendiceal US data in two ways so as to clearly provide this information to our clinicians. The intention-to-diagnose analysis categorizes

indeterminate results as missed cases based on the final outcome. Contrary to Trout and Larson’s statement, this analysis is specifically endorsed by Fedko et al. [4] because it allows for “transparent reporting of all results and determination of diagnostic yield” and likelihood ratios and informs physicians what proportion of appendiceal US examinations will not yield determinate results. The intention-to-diagnose method does underestimate measures of the diagnostic performance of US such as accuracy and sensitivity for the determinate results. To account for this, we performed a second analysis using the standard binary approach that excludes indeterminate studies because this analysis relates to clinicians the accuracy of appendiceal US when a determinate result is given.

In our opinion, indeterminate results do not yield useful information. Trout and Larson think they do. They assert that indeterminate results reflect a range of probabilities that a given patient has appendicitis *and* that this information is meaningful to clinicians. While we did find a narrow range of prevalences of appendicitis in our three indeterminate groups, we disagree that this information is clinically useful. Our findings do not support their assertion on two grounds.

First, the prevalence of appendicitis in our study group, children with abdominal pain referred for US, was 18.5%. This was the pretest probability for acute appendicitis. The overall prevalence of appendicitis in patients with positive and negative US results was 87% and 1%, respectively, and these are the post-test probabilities. They indicate that US is a very good test at discriminating the presence or absence of appendicitis when a definitive result is given. For indeterminate results, the prevalence of appendicitis was 14.2% overall. This is the post-test probability and it is not statistically different from the pre-test probability of 18.5% for the overall group or for each subset of indeterminate US results, $P > 0.05$. Thus in our practice indeterminate US results do not change the likelihood that a patient does or does not have appendicitis

✉ Larry A. Binkovitz
binkovitz.larry@mayo.edu

¹ Department of Diagnostic Radiology, Mayo Clinic,
200 First St. SW, Rochester, MN 55905, USA

² NEOMED, Rootstown, OH, USA

³ Department of Biostatistics, Mayo Clinic,
Rochester, MN, USA

⁴ Department of Emergency Medicine, Mayo Clinic,
Rochester, MN, USA

and provide clinicians with no predictive information. From the clinicians' perspective, *an indeterminate US result is the same as if the US examination had not been performed at all.*

Second, there is a cost to the patient when radiologists equivocate in appendiceal US reporting and for too long this cost has not been addressed. When the US report is indeterminate, clinicians will find determinacy in other ways and in many practices this means a CT study or an exploratory surgery. Our experience, as well as that of other investigators including Larson et al. [3] bears this out. When our US reports were indeterminate, our patients had four times the rate of follow-up CTs and twice the rate of negative laparotomies. Similar numbers have been reported by Nielsen et al. [5]. From their reported data, Larson et al. [3] found that indeterminate US reports were 2.7 times more likely to have follow-up CT studies (55/123) than determinate US results (208/1,234), $P < 0.0001$. Additionally, our negative laparotomy rate was significantly higher when an indeterminate US result was given than when a determinate result was given, 6.9% and 3.5%, respectively. Larson et al. [3] reported a 17.9% negative laparotomy rate but did not indicate the rate for determinate and indeterminate US results separately.

It is fair to ask whether there might be a cost to patients when radiologists use reporting schemes that do not include indeterminate categories. This is especially a concern for us at the Mayo Clinic because during our 5-year study period, no child with suspected appendicitis who underwent US evaluation was discharged with undiagnosed acute appendicitis. This was true regardless of the US result and highlights that the diagnosis of appendicitis is made with the synthesis of data from multiple sources, clinical, laboratory and imaging, which may at times be contradictory. Nielsen et al. [5] have found that the introduction of a US reporting scheme that did not include indeterminate categories *decreased* both the rate of follow-up CT studies and negative laparotomies while *increasing* US sensitivity and accuracy [5]. The results from our study identified ways to increase determinacy and accuracy [2] and we have begun to implement these changes while monitoring for negative outcomes.

Two additional points are worth emphasizing. Larson et al. [3] found that inclusion of indeterminate interpretive categories resulted in an increase in US accuracy, from 94.1% to 96.8%. It should be noted that their increase in accuracy resulted from exclusion of the indeterminate US results from the binary statistical analysis and this increase represents a limitation of the binary analysis rather than a true improvement in US performance. We believe it is a mistake to focus on ways to marginally improve appendiceal US accuracy because it

misses the larger point. *US accuracy in isolation is not a meaningful outcome.* Whether the inclusion of indeterminate interpretive categories provides any real benefit to patients will not be expressed in terms of slightly improved accuracy but rather in terms of clinical outcomes such as fewer follow-up CT studies or lower negative laparotomy rates.

Last, we believe that the rate of indeterminate results can be radiologist-specific and the tendency to equivocate or to be determinative is made particularly evident with appendiceal US. It is this tendency toward equivocation, and the negative outcomes associated with it, that motivated our study. That is why we investigated which patient or system factors might allow an increase in determinate US reporting. We were surprised to find no increase in either accuracy or determinacy despite increasing appendiceal visualization rates over the course of our study. These findings are similar to those of Chang et al. [6]. Although efforts toward increased appendiceal visualization and accuracy are important, the actual improvement in patient outcomes that they may achieve is not entirely straightforward. We believe that patients will have a greater benefit from increased determinacy in US reporting and that decreasing indeterminate results as much as possible is an important goal for pediatric radiologists.

Conflicts of interest None

References

1. Trout AT, Larson DB (2015) Appendiceal ultrasound: the importance of conveying probability of disease. *Pediatr Radiol* INSERT DOI MS 15-00288. doi:10.1007/s00247-015-3428-3
2. Binkovitz LA, Unsdorfer KM, Thapa P et al (2015) Pediatric appendiceal ultrasound: accuracy, determinacy and clinical outcomes. *Pediatr Radiol* INSERT DOI MS 15-00069R1. doi:10.1007/s00247-015-3432-7
3. Larson DB, Trout AT, Fierke SR et al (2015) Improvement in diagnostic accuracy of ultrasound of the pediatric appendix through the use of equivocal interpretive categories. *AJR Am J Roentgenol* 204: 849–856
4. Fedko M, Bellamkonda VR, Belliolio MF et al (2014) Ultrasound evaluation of appendicitis: importance of the 3×2 table for outcome reporting. *Am J Emerg Med* 32:346–348
5. Nielsen JW, Boomer L, Kurtovic K et al (2015) Reducing computed tomography scans for appendicitis by introduction of a standardized and validated ultrasonography report template. *J Pediatr Surg* 50: 144–148
6. Chang ST, Jeffrey RB, Olcott EW (2014) Three-step sequential positioning algorithm during sonographic evaluation for appendicitis increases appendiceal visualization rates and reduces CT use. *AJR Am J Roentgenol* 203:1006–1012