Certainty in Medicine: Dream On, Especially for Appendicitis Diagnosis
by Dr Bud Wiedermann MD, MA, Evidence eMended Editor, Grand Rounds

Uncertainty is inherent in medical practice, but that doesn't mean we shouldn't strive for accuracy in our medical decision-making. This study on sonographic diagnosis of appendicitis points to the difficulties involved in decision making based on relatively subjective criteria.


This study was a retrospective chart review of sonograms obtained for suspected appendicitis at 1 large children's hospital over a 1-year period. The investigators particularly focused on studies originally interpreted as equivocal, as defined by a scoring system developed at the same institution. Utilizing information from 157 equivocal sonograms of children in whom adequate follow up was available, the authors found that 3 sonographic factors were associated with an operative (as opposed to histologic) diagnosis of appendicitis: loss of mural stratification, peri-appendiceal fat inflammation, and presence of appendicolith. More interesting, I thought, was the "fine print" in the study; i.e. you need to read the entire study to appreciate the nuances.

First, without reading this particular study, we all know that sonography is one of the most subjective of imaging studies in medicine. The results are highly dependent on the skill of the individual performing the study, as well as on the skill of the radiologist interpreting it. Because it is so dependent on human factors, the error rate is less predictable than, for example, a plain chest radiograph, where the technical aspects of obtaining the image are less variable.

In addition to the problems inherent in any retrospective study, I took notice that the authors chose to use operative diagnosis of appendicitis (along with clinical follow up) as their gold standard for diagnosing appendicitis in these equivocal cases. However, I would think there must be some subconscious tendency of surgeons to err on the side of diagnosing appendicitis in the OR; no one likes to think they performed an operation that didn't help the patient. In this study, of 49 patients with equivocal sonograms who underwent appendectomy, 38 had a pathologic diagnosis of appendicitis, while 48 had an operative diagnosis (i.e. determined by the surgeon in the OR) of appendicitis. I wish they had evaluated the results based on pathologic diagnosis as well.

Again, only by reading the entire study does one get a sense of the uncertainty inherent in sonography. The authors took care to have 2 radiologists, unaware of the original readings and patient outcomes, re-interpret the equivocal sonogram readings. (A third radiologist broke any ties resulting from the 2 having disparate interpretations.) They measured degree of agreement between the 2 radiologists using a well-accepted statistical test, Cohen's kappa statistic.

Without boring you with the arithmetic, Cohen's kappa measures the degree of agreement among different
raters, taking into account the likelihood of agreement occurring by chance. Kappa ranges from 0 (no agreement) to 1 (perfect agreement), but there is some controversy about what a "good" kappa statistic really should be. There are standardized ratings for kappa, such as that by Landis and Koch which groups ratings into slight (0-0.2), fair (0.21-0.4), moderate (0.41-0.6), substantial (0.61-0.8), and almost perfect (0.81-1) agreement. However, these categories are based on nothing more than personal opinion, and my own approach is to ignore such cookie-cutter categories and look at the big picture. For sonography as a screening test for appendicitis, I'd want really good inter-rater reliability; otherwise, your child might be sent to the OR or home based solely on which radiologist or ultrasound technician was on duty the night they had their sonogram. In the current study, the kappa statistic for various elements of the sonogram ranged from 0.3 to 0.49 for most sonographic findings (the 1 decent kappa stat was 0.73 for appendix size in mm). In my view, that's not great.

Mostly, the limitations of sonography impact decision-making of surgeons and emergency department clinicians, rather than primary care providers. I certainly hope they are aware of the uncertainty inherent in sonography, and I certainly hope to see further studies from the ambitious group of authors of the current study. I'm also intrigued that the new era of artificial intelligence in medicine may be the next big breakthrough for vagaries in radiology and similar fields.

Further Reading

- Biomarkers in Children With Appendicitis
- A Simple Blood Test for Appendicitis?
- Grand Rounds on Facebook
- AAP Journals On Twitter