Diagnosing appendicitis: What works, what does not and where to go from here?

Simon Craig1,2,3,4 and Sarah Dalton1,5,6

1Emergency Medicine Training, Monash Medical Centre, 2School of Clinical Sciences at Monash Health, Monash University, 3Paediatric Research in Emergency Departments International Collaborative (PREDICT), 4Murdoch Children’s Research Institute, Melbourne, Victoria, 5The Children’s Hospital at Westmead and 6Clinical Excellence Commission, Sydney, New South Wales, Australia

Abstract: Although acute appendicitis is the most common reason for emergency abdominal surgery in children, diagnosis is far from straightforward. Delays in diagnosis can result in significant complications, whereas over-diagnosis can result in costly inter-hospital transfers and unnecessary surgery. This article aims to describe current evidence-based assessment of children with possible appendicitis presenting to the emergency department. We provide an overview of the literature relating to the various available diagnostic approaches, including the application of history, examination, pathology tests, imaging, and clinical decision rules.

Key words: appendicitis; diagnosis; paediatrics; ultrasonography.

Although acute appendicitis (AA) is the most common reason for emergency abdominal surgery in children,1 misdiagnosis after emergency department (ED) attendance remains a significant issue.2 Delays in diagnosis of appendicitis can result in perforation, significant morbidity3 and occasional mortality,4 whereas over-diagnosis can result in unnecessary surgery.5 A report from The Children’s Hospital at Westmead in New South Wales from 2009 describes a perforation rate of 13% and a negative appendicectomy rate of approximately 14%.6 This suggests that – using our current diagnostic approach – more than one in eight children undergoing appendicectomy have an unnecessary operation.

Recent Australian publications suggest an increasing trend for inter-hospital transfer of children with suspected appendicitis to specialised paediatric surgical services.6,7 Early involvement of paediatric surgeons has been suggested to reduce unnecessary investigations8; however, transfer of patients who do not have appendicitis results in unnecessary movement of patients and families, with attendant inconvenience, cost and duplication of assessments.

This article aims to describe current evidence-based assessment of children with possible appendicitis. We provide an overview of the literature relating to the various available diagnostic approaches, including the application of history, examination, pathology tests, imaging, and clinical decision rules (CDRs).

Clinical diagnosis of appendicitis

It is widely suggested that the diagnosis of AA should be based on clinical findings. The use of likelihood ratios (LRs) allows an estimate of the effect of how much a particular finding changes the probability of disease from a pre-test estimate. A positive likelihood ratio (LR+) of >10, or a negative likelihood ratio (LR−) of <0.1, provide a conclusive change, whereas LR+ 5–10 and LR− 0.1–0.2 provide a moderate change, and LR− 0.2 to 0.5 provide a small change respectively. Findings of LR+ 1 to 2 or LR− 0.5 to 1 rarely alter probability significantly.9

Bundy et al.10 systematically assessed the precision and accuracy of symptoms and signs for evaluation of children with possible appendicitis. In unselected children with abdominal pain, fever was demonstrated to be somewhat useful, with a positive likelihood ratio of 3.4 (95% confidence interval (CI) 2.4–4.8) and a negative likelihood ratio of 0.32 (95% CI 0.16–0.64).10

Table 1 presents a summary of the positive and negative likelihood ratios and inter-rater reliability for a number of signs and symptoms of paediatric appendicitis in those whom the disease was suspected.11,12 It can be seen that the diagnostic utility of any single clinical symptom or sign is quite low, and there are significant problems relating to poor inter-rater reliability. Even those findings with the highest positive likelihood ratios (pain...
migration and rebound tenderness) had very poor inter-rater reliability, with $\kappa < 0.4$. As a result, many investigators have examined the use of adjunctive pathology and imaging tests to assist with the diagnosis of appendicitis.

### Pathology testing

The most commonly ordered investigations to assist with the diagnosis of appendicitis are a full blood examination – to demonstrate total white cell count or presence of a neutrophilia – and a C-reactive protein (CRP). More recent studies have also evaluated procalcitonin. These tests have been investigated extensively, and their diagnostic utility is well described (Table 2). However, studies have found variable performance using different cut-offs for optimal sensitivity and specificity. Yu et al. conducted a recent systematic review, which concluded poor sensitivity for white cell count, CRP and procalcitonin.\(^{19}\) In addition, recent work has demonstrated that the results of laboratory investigations fluctuate over the course of the illness, with different test characteristics and utility depending on the duration of abdominal pain.\(^{14,20–22}\)

Although of limited utility in the ED setting, children without appendicitis often demonstrate normalisation of inflammatory markers on repeated testing. Those with normal values of both CRP and white cell count are less likely to have appendicitis.\(^{23}\) In addition, patients with complicated or perforated appendicitis tend to have higher levels of inflammatory markers than those with uncomplicated disease.\(^{13,20}\)

Many other urinary and serum tests for AA have been investigated, including various inflammatory cytokines (interleukin-6, interleukin-8, tumour necrosis factor – alpha), serum lactoferrin and calprotectin, and urinary leucine-rich $\alpha$-2-glycoproteins. However, although promising, clinical utility of these tests is yet to be demonstrated.\(^{11}\)

### Imaging

Ultrasound and computed tomography (CT) of the abdomen are the main imaging modalities used to assist in the diagnosis of AA. In the last 15 years, there has been a dramatic increase in the use of CT scans for the workup of paediatric abdominal pain in North America\(^{24}\) with no significant changes in the number of hospital admissions or patients diagnosed with appendicitis.\(^{25}\) A recent US study found that earlier involvement of paediatric surgeons reduced CT utilisation from around 39% to 18%.\(^{26}\) Conversely, overall CT rates in Australian children are trending downwards,\(^{27}\) with a recent survey finding that not one of 145 emergency physicians practicing in Australasian EDs affiliated with a paediatric emergency research network would order a CT scan when presented with clinical scenarios involving a child with intermediate risk for appendicitis.\(^{28}\)

Ultrasound is suggested to be useful for confirmation of suspected appendicitis (reported specificity ranges from 88% to 98%); however, it is unable to definitively exclude disease (reported sensitivity ranges from 78% to 100%).\(^{29}\) In addition, it requires skilled operators and may not be available after-hours. Failure to visualise the appendix on ultrasound can result in a non-diagnostic scan, which is more common in obese children and those with low clinical probability of appendicitis.\(^{30}\) In a large study of over 1800 children with suspected appendicitis, Bachur et al. demonstrated that the sensitivity of ultrasound increases with increasing duration of abdominal pain.\(^{31}\)

CT has become the imaging modality of choice in many US centres due to reported sensitivity and specificity greater than 95% and greater availability than ultrasound.\(^{32}\) It also has the advantage of similar sensitivity early and late in the disease process.\(^{33}\) CT utilisation appears to be higher after-hours\(^{34}\) and in non-paediatric hospitals. However, there are significant concerns regarding the risk of radiation exposure in children and the subsequent development of malignancy.\(^{35}\) Recent publications have recognised this concern and have investigated various clinical protocols aimed at reducing the exposure of children with suspected appendicitis to CT radiation.\(^{26,32,35,36}\)

Two recent studies suggest a ‘radiation-free’ approach to imaging in suspected appendicitis, with the use of magnetic resonance imaging (MRI) after non-diagnostic ultrasound. Herlitzcek et al. reviewed 60 children who had an inconclusive ultrasound for suspected appendicitis followed by an MRI. They demonstrated a sensitivity for MRI of 100% and specificity of 96%.\(^{37}\) Aspelund et al. introduced a new imaging protocol using ultrasound followed by MRI and compared this with their previous CT-based approach. They found equivalent clinical outcomes – no difference in perforation rate, time to operation, negative appendicectomy rate or hospital length of stay.\(^{38}\)

| Table 1 Summary likelihood ratios (LRs) and inter-rater reliability ($\kappa$) for history and physical examination findings in children with suspected appendicitis |
|---|---|---|---|
| Sign or symptom | LR + (95% CI)\(^{11}\) | LR– (95% CI)\(^{11}\) | $\kappa$ (95% CI)\(^{12}\) |
| History | | | |
| Fever | 1.2 (1.1–1.4) | 0.53 (0.29–0.97) | |
| Anorexia | 1.4 (1.2–1.6) | 0.57 (0.44–0.73) | 0.32 (0.24 to –0.40) |
| Vomiting | 1.4 (1.3–1.6) | 0.57 (0.47–0.69) | 0.84 (0.80–0.89) |
| Right lower quadrant pain | 1.2 (1.0–1.5) | 0.56 (0.43–0.73) | 0.48 (0.39–0.58) |
| Pain migration | 2.06 (1.63–2.60) | 0.52 (0.40–0.69) | 0.37 (0.29–0.45) |
| Examination findings | | | |
| Rebound tenderness | 3.0 (2.3–3.9) | 0.28 (0.14–0.55) | 0.32 (0.24–0.40) |
| Right iliac fossa tenderness | 1.3 (1.1–1.4) | 0.45 (0.35–0.59) | 0.45 (0.36–0.54) |
| Rectal tenderness | 2.3 (1.3–4.1) | 0.70 (0.56–0.87) | |

© 2015 The Authors
Journal of Paediatrics and Child Health © 2015 Paediatrics and Child Health Division (Royal Australasian College of Physicians)
The above review of variables from history, physical examination, pathology and imaging demonstrates that no one individual component is reliably predictable for AA, leading to the development of CDRs. CDRs are decision-making tools for clinicians that contain at least three predictive variables from history, physical examination and simple diagnostic tests that aim to provide the probability of an outcome or diagnosis.

The two best known and most studied CDRs for AA are the Alvarado score and the Paediatric Appendicitis Score (PAS). In their original derivation studies, these scores were used as rule-in tests to predict the likelihood of appendicitis and were designed to reduce the number of negative appendicectomies. Although both CDRs performed well in their derivation studies, subsequent validations studies have not supported this, leading to the conclusion that neither is clinically sufficient to diagnose AA in children.

These findings are supported by a recent study that demonstrated clinical judgement by a senior surgeon to be superior to the performance of a CDR. The ongoing development of CDRs for possible appendicitis has aimed to stratify children with acute abdominal pain into clinical risk groups for AA. Individual approaches have been variable, either using different cut-off numbers for PAS or Alvarado, or through the development of alternative CDRs.

Kulik concluded that a high-quality, well-validated and consistently high-performing CDR for AA cannot be identified. The most promising CDR identified in the review by Kulik was that of Kharbanda (see Table 3). This rule was specifically designed to identify a group of children with acute abdominal pain at sufficiently low risk that further investigation for possible AA may not be required. The original CDR performed well but has since been refined with the modified version yet to be externally validated.

Despite the theoretical benefit of collating multiple indicators into a single prognostic score, existing iterations of the published CDRs for appendicitis have not been demonstrated to be superior to overall clinical judgement.

There are also significant limitations when applying published CDRs for AA to the Australasian setting. All rules have limited intra- and inter-observer reliability, none are locally derived or validated and all depend on the results of blood tests necessitating venesection. Recent work has aimed to use CDRs to stratify risk prior to imaging protocols, but none have been widely adapted to date.

**Putting it all together**

There is no debate that patients with very low risk of appendicitis based on history and physical examination may be discharged from the ED without further diagnostic testing.
whereas those with obvious appendicitis require early surgical referral.

However, much uncertainty remains with respect to the best approach to children with intermediate risk of disease. Figure 1 provides our view of the major decision points in the diagnosis of appendicitis, from initial assessment and risk stratification to surgical referral or further testing, depending on clinical suspicion. Local application of this approach is complicated by the difference between individual practice settings, as many EDs do not have ready access to immediate paediatric surgical review.

Options for managing children with intermediate risk of appendicitis include observation without the input of paediatric surgical services, imaging, transfer for paediatric surgical review, or discharge home and repeat evaluation at a paediatric surgical centre for persistent symptoms. Currently available evidence suggests there is not one single approach that guarantees an accurate diagnosis every time. Local approaches are likely to depend upon the availability of paediatric surgical consultation.

If a child is discharged after an evaluation for possible appendicitis, careful explanation of the possibility of an evolving condition, the potential for misdiagnosis of early disease and the need for repeated evaluation should be emphasised. Written and verbal instructions should be provided, with definite plans for follow-up, and specific instructions on what to do if symptoms persist or worsen.

Areas for future research include the further investigation of applying CDRs for risk stratification, the incorporation of imaging into risk stratification pathways and the development of CDRs that do not include blood testing. Additional study should also compare diagnostic approaches and outcomes in centres with high CT use (e.g. North America) versus low CT use (e.g. Australasia).

**Conclusion**

The diagnosis of appendicitis in children remains challenging, with no individual approach currently providing adequate sensitivity or specificity to confidently predict the need for appendicectomy. To date, the best performing variables are rebound tenderness and a high WCC and CRP in children with short duration of symptoms; however, the test characteristics of these variables appear insufficient to reliably rule in or rule out the disease. CDRs for possible appendicitis show some promise, particularly when used to stratify risk groups and indicate need for further imaging. However, none are currently proven to perform sufficiently well to replace other diagnostic approaches.

**Acknowledgement**

This paper was supported in part by a National Health and Medical Research Council Centre of Research Excellence Grant.

### Table 3 Predictors and performance of clinical decision rules for possible appendicitis

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Alvarado</th>
<th>PAS</th>
<th>Modified Lindberg</th>
<th>Lintula</th>
<th>Kharbanda</th>
<th>Van den Broek</th>
</tr>
</thead>
<tbody>
<tr>
<td>History – predictors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right iliac fossa (RIF) pain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Migratory RIF pain</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Intensity of pain</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration of pain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Progression of pain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nausea/vomiting</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Anorexia</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Examination – predictors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right lower quadrant tenderness</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Rebound tenderness</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Guarding</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdominal rigidity</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cough/percussion/hopping tenderness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decreased bowel sounds</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temperature increased</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unable to walk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investigation – predictors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White blood cell (WBC) &gt; 10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neutrophilia†</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rule performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.72–0.93</td>
<td>0.82–1.00</td>
<td>0.86–0.90</td>
<td>0.85</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.49–0.98</td>
<td>0.65–0.96</td>
<td>0.62–0.87</td>
<td>0.95</td>
<td>0.34</td>
<td>0.17</td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>0.45–0.97</td>
<td>0.54–1.00</td>
<td>0.88–0.96</td>
<td>0.85</td>
<td>0.47</td>
<td>0.9</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>0.81–0.92</td>
<td>0.88–1.00</td>
<td>0.65–0.66</td>
<td>0.95</td>
<td>1</td>
<td>0.96</td>
</tr>
<tr>
<td>LR−</td>
<td>0.09–0.34</td>
<td>0.00–0.27</td>
<td>0.16</td>
<td>0.16</td>
<td>0</td>
<td>0.01</td>
</tr>
</tbody>
</table>

†Definition of neutrophilia varies between studies: PAS > 7.5/mm³, Alvarado > 75% of WBC, Kharbanda > 6.75 × 10⁹/μL.
General Considerations
Should a clinical decision rule be used?
Which CDR? Are blood tests mandatory?
When should it be used?
Who should be using it?
Is there a role for serial re-assessment using a CDR?

High-risk patients
Do I have to wait for blood test results before referral?
Do they need surgical review in the ED prior to surgical admission?

Intermediate-risk patients
Should surgical review occur prior to decision on imaging?
Which imaging modality should be used?
Where should observation occur?

Low-risk patients
Who should perform follow-up?
Should blood tests be repeated?
Patient education is crucial.
Missdiagnosis is common, and illness can progress.

Fig. 1  Conceptual framework of the diagnostic approach to acute appendicitis in children, demonstrating areas of uncertainty.

for Pediatric Emergency Medicine (GNT1058560), Canberra, ACT, Australia.

References
5 Debnath J. Acute appendicitis: is removal of a normal appendix still existing and can we reduce its rate? Saudi J. Gastroenterol. 2010; 16: 122.
9 Jaeschke R, Guyatt GH, Sackett DL. Users’ guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. JAMA 1994; 271: 703–7.