Citation: Ogunmefun G, Hardy M and Boynes S (2016) Is magnetic resonance imaging a viable alternative to ultrasound as the primary imaging modality in the diagnosis of paediatric appendicitis? A systematic review. Radiography. Accepted for publication 2nd Jan 2016.

Copyright statement: © 2016 Elsevier. Reproduced in accordance with the publisher's self-archiving policy. This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/
INTRODUCTION

Appendicitis is defined as the inflammation of the vermiform appendix, often as a consequence of bacterial aggregation, resulting in the appendix becoming inflamed and pus filled. Appendicitis is the most common cause of acute abdominal pain requiring surgical intervention in paediatric patients. The global incidence of appendicitis in children increases from 1-2 cases per 10,000 children aged less than 4 years to 25 cases per 10,000 in children between the ages of 10 and 17 years.

In the United States of America, the annual incidence of appendicitis is 37 cases per 10,000 children aged between 0 and 14 years with approximately 70,000 cases reported among all American children annually. In the United Kingdom, acute appendicitis accounts for an estimated 34,000 hospital admissions among the general population with 20% of these cases (approximately 6,800) reported to be in children of 0-14 years of age.

The diagnosis of acute appendicitis can be very challenging. In adults, appendicitis often presents with a typical progression of symptoms: periumbilical pain progressing to nausea, right lower quadrant pain and eventually vomiting and fever. As a result, successful diagnosis can often be made on presenting clinical features and results of laboratory tests (e.g. Total Leukocyte Count (TLC), C-reactive protein; neutrophil count). In children, appendicitis may not present with such typical symptoms. Instead, while childhood appendicitis may present initially with periumbilical pain, symptom progression may lead to flatulence, bowel irregularity/diarrhoea, indigestion, and general malaise. Consequently, a substantial proportion of paediatric appendicitis cases may be misdiagnosed if clinical decision making is based on physical examination, symptoms and laboratory investigations alone. To prevent misdiagnosis and reduce negative appendectomy rates, imaging has been recommended as part of the diagnostic pathway.

Ultrasound (US) is generally the diagnostic imaging modality of choice where paediatric appendicitis is suspected as it is: readily available; has no radiation risk; is relatively fast; and, in comparison to other cross-sectional imaging modalities, is relatively inexpensive. High sensitivity (88%: 95%CI [86-90]) and specificity (94%: 95%CI [94-96]) have also been documented for US in the assessment of paediatric appendicitis. However, the focussed nature of ultrasound assessment limits its contribution in determining alternative causes of presenting symptoms and operator dependency remains an acknowledged fundamental limitation. Further, the accurate diagnosis of appendicitis using US may be restricted as a consequence of bowel gas or distension, patient obesity and a retro-caecal (deeply situated) appendix.
Computed Tomography (CT) has previously been considered the main alternative to US and has a high sensitivity and specificity in the diagnosis of appendicitis with much reduced operator dependency. A meta-analysis by Doria et al (2006) compared CT and US in the diagnosis of appendicitis and determined that in the diagnosis of paediatric appendicitis, the evidence reviewed suggested that CT had a pooled sensitivity of 94% (95% CI: 92 to 97) and a pooled specificity of 95% (95% CI: 94 to 97). However, CT also presents a far greater risk of harm to the child from exposure to ionising radiation and reaction to intravenous contrast media. As a result, the trend is not to refer paediatric patients for CT where appendicitis is suspected.

Magnetic Resonance Imaging (MRI) has been successfully adopted in the assessment of appendicitis in pregnant women offering high soft tissue contrast without the use of radiation. However, long examination times and limited clinical availability have been cited as major limitations to the wider application of MRI in the assessment of acute clinical conditions requiring timely decision making. In addition, long examination times and movement restriction requirements previously meant that sedation of children may have been considered necessary to reduce anxiety and optimise MRI (and perhaps CT) imaging outcomes. Today, the greater accessibility to MRI within the clinical radiology setting globally, and the development of new and faster imaging sequences, reduces the impact of these concerns when identifying MRI as the diagnostic imaging modality of choice. As a result, it is time to consider whether MRI should be considered a viable alternative to ultrasound as the primary imaging modality in the assessment of paediatric appendicitis. This paper reports the findings of a systematic review of the research evidence and considers whether MRI should form part of the diagnostic pathway where paediatric appendicitis is suspected and explores the optimal diagnostic scan sequences to reduce examination time. No previously published systematic reviewed has explored the value of MRI as the index test in the assessment of paediatric appendicitis and therefore this review provides a significant contribution to the evidence base.

METHOD

A search of Medline, Cinahl and PubMed central databases and Google Scholar was undertaken supplemented by hand searching of key imaging journals (e.g. British Journal of Radiology; Radiography), review of reference lists, author searching and review of the NICE (National Institute for Health & Care Excellence) evidence base for existing guidelines. Citations were identified using the following key search terms (and their alternatives): Magnetic Resonance Imaging (MRI; MR; Nuclear Magnetic
Resonance (NMR)); Appendicitis (vermiform appendix, epiphylitis). The search was limited to primary research studies published from January 2005 to April 2015 to take account of the recent advances in MRI pulse sequences and the broader clinical application of the technology.

Following the identification of all potentially relevant research studies, the titles and abstracts were screened to determine whether they met the inclusion/exclusion criteria. These criteria were derived from the primary research question “is MRI a viable alternative to Ultrasound as the primary imaging modality in the diagnosis of paediatric appendicitis” and are listed in Table 1. Decision making around inclusion was based on the “rule out” principle with papers only being rejected where the reviewer was certain of their lack of relevance. At each stage, if uncertainty existed over whether a paper should be included in the review, the paper was retained. The full text of all retained articles was examined to make the final decision on inclusion/exclusion.

**Table 1: Inclusion/Exclusion criteria**

All retained full text papers were independently assessed for quality using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2)\(^2\) by at least 2 authors and data were extracted directly into a Microsoft Excel\(^2\) spread sheet using a purposefully pre-designed extraction framework to promote consistency. Any inconsistencies in opinion re inclusion or paper quality were resolved through discussion and consensus agreement. Paper quality was documented using an adaptation of the QUADAS-2 assessment checklist summary criteria\(^2\) (Table 2). A summary value was awarded to each paper in Table 3 with ‘High’ representing a study with low risk of bias and low concerns regarding applicability of study findings, ‘Average’ representing a study with an unclear risk of bias and unclear concerns regarding applicability, and ‘Low’ representing a study with a high risk of bias and high concerns regarding applicability. Only papers considered to be of high or average quality were retained in the final evaluation (Table 3).

Data analysis was by descriptive synthesis and comparison of extracted data. Meta-analysis or pooling of extracted data was not appropriate due to the diversity of study designs, variations in clinical characteristics and technical parameters, and differences in how diagnostic accuracy was determined.

**Table 2: QUADAS-2 appraisal of included studies**
Results

A flow chart detailing the review process is provided in Figure 1. Details of the included studies are summarized in Table 3 with quality assessment grading identified in the final column. Details of the index test characteristics and extracted outcome measures are detailed in Tables 4 and 5 respectively.

Figure 1: Flow diagram of included studies

Table 3: Characteristics and quality assessment (QA) of included studies

Table 4: Index test characteristics extracted from included studies

Table 5: Extracted outcome measure data

Sensitivity estimates

All the included studies used pathology following surgery, or clinical follow-up of those patients receiving conservative care, as the reference standard. Sensitivity across the included studies ranged from 92%-100%. Only one study (Rosines et al) employed Gadolinium contrast agent (Gadobenate Dimeglumine (Bracco diagnostics) administered intravenously with a weight based dosage of 0.2ml per kg) and they reported an examination sensitivity of 94%, which is lower than the sensitivities reported by Herliczek et al, Johnson et al, Moore et al, and Thieme et al. The study by Bayraktutan et al explicitly set out to evaluate the accuracy of Diffusion Weighted Imaging (DWI) and reported a sensitivity of 78%, 81%, and 92% for DWI, conventional MRI sequences, and combined DWI and conventional sequences respectively.

Specificity estimates

Specificity across the included studies ranged from 89%-100%. The study by Rosines et al, employing Gadolinium contrast agent, reported a specificity of 100% as did the study by Bayraktutan et al. The study by Thieme et al reported the lowest specificity at 89% with all other studies reporting sensitivities of 96% or higher. The study by Bayraktutan et al evaluating the accuracy of DWI reported a specificity of 67%, 100% and 100% for DWI, conventional MRI sequences, and combined DWI and conventional sequences respectively.
Variations in diagnostic sequence pathway and scan time

Scan time is dependent on several variables including magnetic field strength, RF coil type and pulse sequences used. All studies, with the exception of Herliczek et al, used a multichannel phased array coil. Herliczek et al did not indicate the type of coil used. The studies by Johnson et al and Herliczek et al utilised a 3T magnet although Herliczek et al combined this with a 1.5 T magnet. All other studies used 1.5T magnets which remains the most common magnet field strength in the UK. With the exception of Herliczek et al, all studies standardised their within study imaging sequences but there was significant variation between the studies with respect to the sequences employed. Four studies, Bayraktutan et al, Herliczek et al, Orth et al, and Rosines et al included T1 sequences together with T2 weighted images. However, contrast enhanced T1 sequences were only employed by Rosines et al. All studies utilized at least two or more T2 weighted sequences. With the exception of Bayraktutan et al and Orth et al, single shot sequences featured heavily in the studies. Balanced gradient echo sequences were used by Rosines et al and Thieme et al. Diffusion Weighted Imaging (DWI) was employed in three studies (Bayraktutan et al, Orth et al and Thieme et al). Fat saturation was used by all studies except Rosines et al and Thieme et al and a volume 3D T2-weighted FSE SPACE, a volumetric isotropic acquisition, was utilized by Herliczek et al.

Given the variation in field strength and sequences used it is not surprising that scan time varied among the included studies. Four studies (Johnson et al, Moore et al, Orth et al and Thieme et al) recorded scan times of less than 20 minutes per patient. Johnson et al reported a median scan time of 5 minutes 40 seconds with all examinations completed within 8 minutes and 45 seconds. This study was performed using a 3T magnet and four fast T2 weighted sequences. The longest time was reported by Herliczek et al with a mean image acquisition time of 30.5 minutes but with a reported time range of 10-66 minutes, this may be influenced by a small number of lengthy examinations. No median time was reported to confirm this. Bayraktutan et al and Rosines et al did not provide exact data on scan time.

Discussion

The patient pathway, modality and techniques assessment of acute appendicitis in children has been the subject of debate internationally. Contemporary pre-operative imaging modalities in the evaluation of paediatric appendicitis are currently ultrasound (US) as the primary tool and CT as the secondary
complimentary imaging modality as recommended by the Royal College of Radiologists in the UK. An ionising radiation-free cross sectional imaging pathway of US, selectively followed by MRI where US findings are inconclusive, is a new and novel approach to the diagnosis of paediatric appendicitis that is being considered by the imaging community. However, with the acknowledged limitations of ultrasound and the greater accessibility to faster MRI technologies, it is likely that questions will soon be asked around preferred primary imaging approach. This systematic review is the first to consider whether MRI is a viable alternative to US as the primary imaging modality in the diagnosis of paediatric appendicitis.

Seven primary research studies met the inclusion criteria for this review. Reported sensitivities for MRI were high across all included studies ranging from 92%-100% with three studies reporting 100% sensitivity. Reported specificities were also high across included studies ranging from 89%-100% with two studies reporting 100% specificity. Reported positive (PPV) and negative (NPV) predictive values were also relatively high across the studies, despite the varying prevalence of appendicitis within the cohorts examined, suggesting that the diagnostic outcome of MRI examination is accurate and reliable.

A diagnosis of appendicitis was based on defined criteria and all studies considered an enlarged outer diameter of the appendix to be an important criterion. However, the numerical measure used to determine an enlarged diameter varied slightly across included studies with Bayraktutan et al and Johnson et al fixing their diagnostic boundary at ≥ 6mm while Rosines et al and Herliczek et al fixed the diameter at ≥ 7mm. The remaining three studies did not define the numerical value for the diameter but instead mentioned a thickened appendiceal wall, presumably assessed subjectively on image review. With the diagnostic boundary set at ≥ 6mm, Bayraktutan et al and Johnson et al reported sensitivities of 92% and 100% and specificities of 100% and 99% respectively. With the diagnostic boundary set at ≥ 7mm, Rosines et al and Herliczek et al reported sensitivities of 94%, 100% and 96%, 100% respectively. It can be therefore be inferred that the 1mm difference in appendiceal outer diameter measurement made very little difference to reported sensitivity and specificity values and as such, the smaller value (≥ 6mm) is recommended to reduce risk of false negative findings.

Field strengths of 1.5T and 3.0T were employed by included studies. Johnson et al utilised a field strength of 3.0T and reported a sensitivity of 100% and specificity of 99%. Herliczek et al used a
combination of 1.5T and 3.0T field strengths and reported a sensitivity of 100% and 96% but distinction was not made as to any variation in sensitivity or specificity between the 2 field strengths. At 3.0T, there is an increase in Signal to Noise Ratio (SNR) and Contrast to Noise Ratio (CNR) which may lead to improved image resolution and shortened scan times. These gains in SNR are more pronounced in T2-weighted images than T1 sequences. However, at 3.0T there is potentially greater magnetic susceptibility and chemical shift artefact with radio frequency pulse inhomogeneity also being a potential source of artefact in large or obese children and young people. Due to the limited data available within the included studies in relation to field strength variations, it is not possible to determine the added value of 3.0T MRI imaging over and above 1.5T MRI with respect to diagnostic accuracy in paediatric appendicitis and further work is required to explore this.

The use of Gadolinium in the examination of appendicitis has been advocated within Baert as enhanced T1 weighted sequences are useful for subtle mucosal enhancement and establishing complications with appendicitis. However, only Rosines et al employed Gadolinium contrast agent within this review and as their reported sensitivity and specificity values were comparable to those reported by other authors, it appears unlikely that Gadolinium is a requirement for optimising the diagnosis of paediatric appendicitis. This suggestion is supported by the finding that non-contrast enhanced MRI is equivalent to contrast enhanced CT in the assessment of the appendix in pregnant women. However, given the wide variation in technical parameters between studies included in this review, further work is required to confirm the contribution of Gadolinium contrast media within the context of paediatric appendicitis before recommendations for its adoption or exclusion can be made with confidence.

A diverse range of pulse sequences were used within the included studies. Bayraktutan et al evaluated the improvement in diagnostic accuracy as a result of including DWI as an adjunct to conventional sequences. DWI demonstrates an inflamed appendix and surrounding fat as high signal, secondary to a restriction in water diffusion. DWI has previously been shown to have high sensitivity and specificity values and therefore useful in discriminating between perforated and non-perforated appendicitis.

T2-weighted sequences were used by all studies and, when combined with ultra-fast sequences, are noted to greatly improve assessment of lesions in the small bowel. T2 weighted sequences are very sensitive to fluid filled pathology such as oedema which appears as a hyper-intense signal. The T2 FSE
3D (SPACE) sequence utilized by Herliczek et al is an approach to volumetric acquisition that allows retrospective reformatting in multiple orientations of the appendix\textsuperscript{44}. Fat saturated sequences demonstrate oedema or inflammatory processes as a hyper-intense signal within the wall or adjacent to fat\textsuperscript{43,45}.

T1 weighted sequences utilised by four studies in this review can be useful in demonstrating the normal appendix\textsuperscript{36,45}. T1 weighted GRE sequences were employed by Orth et al and Rosines et al and may be useful in the diagnostic work up for acute appendicitis as gas or negative oral contrast within the appendix (suggesting luminal patency) will demonstrate susceptibility artefact on these sequences\textsuperscript{46}. The unenhanced T1 weighted sequence has been shown to be of little use in demonstrating an inflamed appendix. Evidence from the diagnostic outcome measures of the primary articles within this review supports this view as the diagnostic outcomes of the three studies which did not employ any T1 weighted sequence (Johnson et al, Moore et al, Thiemme et al) were comparable to those studies that did. Enhanced T1 weighted sequences add to the imaging time, thereby raising the risk of perforation, and also increase the risk of Nephrogenic Systemic Fibrosis (NSF) in patients with impaired renal function\textsuperscript{47}. Johnson et al, Thieme et al and Bayraktutan et al also made use of parallel imaging to reduce scan time. Parallel imaging enables faster image acquisition resulting in fewer motion artefacts and improving resolution\textsuperscript{48}. The remaining studies did not make reference to the use of parallel imaging and therefore its contribution to diagnosis and examination speed are unclear.

This systematic review has focused on the ability of MRI to enable accurate diagnosis of acute paediatric appendicitis. Acute Appendicitis is a progressive clinical condition occurring over a period of 24-36 hours\textsuperscript{49}. The number of sequences employed greatly influences the scan time per patient. By limiting the number of sequences acquired to four, as well as employing T2 ultra-fast sequences, parallel imaging and 3T field strength, Johnson et al were able to report the fastest scan times without evidence of compromising diagnostic accuracy compared to other studies in this review. While an abdominal US examination typically takes 20-30 minutes in children\textsuperscript{50}, US examination for the evaluation of paediatric appendicitis may take even longer as a child may become uncooperative and experience pain with the rebound technique at McBurney’s point\textsuperscript{50}. The findings of this review suggest that if the number of MRI sequences can be restricted and tailored to T2 fluid filled and fat suppressing sequences with a phase array coil, then it is possible that MRI examination time for assessment of the paediatric appendix may be shorter than an US examination without compromising accuracy. This focussed approach to MRI
assessment of the appendix may reduce its contribution to determining differential diagnoses (e.g. Crohn’s disease, endometriosis, Meckel’s diverticulum) 51 but this argument could also be made for other focussed examinations across imaging modalities specifically directed to answering the clinical question posed.

A previous restriction on the use of MRI for acute paediatric conditions has been the increased use in sedation to reduce anxiety and improve compliance with movement restrictions. In this review, only Bayraktutan et al sedated patients with 58% of participants receiving sedation. This may be as a consequence of the wide age range of participants (0-14 years) as the remaining included studies restricted the lower age range to 3 years minimum and permitted parents/guardians to accompany the child into the magnet room. Given the comparability in diagnostic outcomes across studies reviewed, it would appear that sedation is not required in the assessment of acute appendicitis in children aged 3 years and over and should no longer be a perceived barrier to access.

Limitations/Strength

This systematic review was undertaken as part of a Master of Science award at the University of Bradford. The initial review of literature and evaluation was completed independently by GO under the supervision of MH. For purpose of publication, the review process was repeated with independent evaluation by SB and mediation by MH to ensure rigour of systematic review process. Inclusion was limited to papers published between 2005 and April 2015 to accommodate innovations in MRI technology and wider application within clinical practise. Heterogeneity was noted across the included studies which prevented meta-analysis of the data. Despite this limitation, analysis of the data has suggested that MRI may well be a viable alternative to ultrasound in the diagnosis of acute appendicitis in children.

Practice Recommendation

In order to facilitate the application of evidence into practice, the evidence from this review has been summarised into the series of recommendations below. Further evaluation of the application of these recommendations in practice is required for validation.

- Where MRI is available 24 hours per day, the evidence from this review suggests that MRI may be designated as the primary imaging tool to investigate suspected paediatric acute appendicitis
as the diagnostic accuracy (sensitivity and specificity) is comparable to those reported for CT and greater than those reported for Ultrasound.

- The use of contrast agents is not necessary for the diagnosis of appendicitis in children.
- The use of sedation is not required routinely in children aged 3 years or over and, where possible, parental attendance in the magnet room is to be encouraged to reduce child anxiety and promote compliance.
- Where the outer diameter of the appendix is used as a diagnostic decision making boundary, a measure of ≥ 6mm should be applied.
- A focussed sequence pathway using a phased array coil and parallel imaging is recommended to optimise diagnosis and minimise scan time. From the evidence reviewed, the optimum sequence selection is:
  - FSE/SSFSE T2 axial
  - FSE/SSFSE T2 coronal
  - FSE/SSFSE T2 sagittal
  - FSE/SSFSE T2 axial fat saturation
  - FSE/SSFSE DWI (axial and coronal b500)

**Conclusion**
The evidence from this review suggests that MRI has high sensitivity and specificity in relation to the accurate diagnosis of acute appendicitis in children, comparable to those reported for contrast enhanced CT and greater than Ultrasound. Consequently, we conclude that is a viable alternative to ultrasound and may be adopted as the primary imaging modality of choice, where accessibility is not restricted, or used to complement ultrasound where findings are indeterminate or inconclusive. Practice recommendations have been provided based upon the evidence reviewed and these require validation within the practice setting.

**Conflict of interest**
The authors have no conflict of interest to disclose
References


33. Saito JM, Yan Y, Evashwick TW, Warner BW, Tarr PI. Use and Accuracy of Diagnostic Imaging by Hospital Type in Paediatric Appendicitis. Paediatrics 2012; 131 (1)

34. Royal College of Radiologists. iRefer. Making the best use of clinical radiology. 7th Ed. UK: Royal college of Radiologists, 2011


