Bibliographic Cite	PMID Link	Literature Type	Level of Evidence	Purpose	Population	Intervention and Outcome Measures	Results/ Recommendations	Study Limitations
Ahmed O, Rodrigues DM, Nguyen GC. Magnetic Resonance Imaging of the Small Bowel in Crohn's Disease: A Systematic Review and Meta-Analysis. Can J Gastroenterol Hepatol. 2016;2016:7857352.	27446869	Systematic Review	Moderate level of evidence	To analyze the use of MR in detecting small bowel activity as well as extramural complications in Crohn's patients.	1020 Crohn's patients were included.	There were 27 included studies, of which 19 were included in the pooled analysis. Pooled analysis of the 19 studies (1020 patients) with raw data revealed a sensitivity of 0.88 (95% CI 0.86 to 0.91) and specificity was 0.88 (95% CI 0.84 to 0.91). In regard to detecting stenosis, pooled sensitivity was 0.65 (95% CI 0.53 to 0.76) and specificity was 0.93 (95% CI 0.89 to 0.96).	MR imaging provides a reliable alternative in detecting small bowel activity in patients with Crohn's disease. Its advantages include high diagnostic accuracy and no radiation exposure while its disadvantages include high cost and limited availability.	The analysis revealed fairly high specificity in detecting stenosis, but only moderate sensitivity. Some of the limitations of our study include the varied length of time between the reference standard and MR imaging. Similarly, due to the small number of studies, we were not able to determine whether more advanced MR (such as MR with 3.0 T magnetic field strength) had any additional benefit. Finally, the large heterogeneity amongst the studies, including reference standards, radiologists experience, and results, suggests that more definitive studies might still be required.
Baolei G, Can C, Peng L, et al. Molecular imaging of abdominal aortic aneurysms with positron emission tomorgraphy: A systematic review. Eur J Vasc Endovasc Surg. 2021; 62(6):969-980.	34696984	Systematic Review	Low level of evidence	To determine the role of PET in predicting the prognosis of abdominal aortic aneurysm (AAA).	A total of 11 articles were retained and included in this review. Of the 11 studies, four were prospective, four were prospective but included retrospective growth data, two were retrospective, and one did not clarify its design. Sample sizes ranged from 14 to 151 participants, giving a total sample size of 508 individuals. The proportion of male participants was 91% and the mean or median age ranged from 68 to 78 years.	Two authors independently performed the study search, data extraction, and quality assessment following a standard method. Of the 11 studies included in this review, nine used 18F fluorodeoxyglucose (18F-FDG) PET and computed tomography (CT) imaging, whereas the remaining two used 18F-sodium fluoride (18F NaF) PET/CT and 18F-FDG PET/magnetic resonance imaging (MRI). One study was rated as low risk of bias, three studies were rated as moderate, and seven studies were rated as high.	Six studies found no significant association or correlation, and two studies found a significant negative correlation between 18F-FDG uptake and AAA expansion. Additionally, one study found that the 18F-FDG uptake was statistically positively related to the expansion rate in a specific AAA subgroup whose AAAs expanded significantly. Two studies suggested that increased 18F-FDG uptake was significantly associated with AAA repair, while the other studies either found no association between 18F-FDG uptake and AAA rupture or repair or failed to report the occurrence of clinical events. One PET/CT study that used 18F-NaF as a tracer showed that an increased tracer uptake was significantly associated with AAA growth and clinical events. Finally, the 18F FDG PET/MRI study indicated that 18F-FDG uptake was not significantly correlated with AAA expansion.	All the studies in this review were observational designs with a small sample size, characteristics which can often produce misleading results. Moreover, the majority showed a high risk of bias in study confounding, probably due to their small sample size, which make it impossible to perform multivariable analysis to adjust for the impact of confounders on the statistical analysis results.
D'Souza N, Hicks G, Beable R, et al. Magnetic resonance imaging (MRI) for diagnosis of acute appendicitis. Cochrane Database Syst Rev. 2021; 12(12):CD012028.	34905621	Systematic Review and Meta-analysis	High level of evidence	To determine the dagnostic accuracy of MRI for detecting appendicitis in all patients.	Included were studies that compared the outcome of an MRI scan for suspected appendicitis with a reference standard of histology, intraoperative findings, or clinical follow-up. A total of 58 studies with sufficient data for meta-analysis were included, including a total of 7462 participants (1980 with and 5482 without acute appendicitis).	Three study team members independently filtered search results for eligible studies. Authors independently extracted study data and assessed study quality using the Quality Assessment of Studies of Diagnostic Accuracy - Revised (QUADAS-2) tool. Authors also used the bivariate model to calculate pooled estimates of sensitivity and specificity.	Estimates of sensitivity ranged from 0.18 to 1.0; estimates of specificity ranged from 0.4 to 1.0. Summary sensitivity was 0.95 (95% confidence interval (CI) 0.94 to 0.97); summary specificity was 0.96 (95% CI 0.95 to 0.97). Sensitivity and specificity remained high on subgroup analysis for pregnant women (sensitivity 0.96 (95% CI 0.88 to 0.99); specificity 0.97 (95% CI 0.95 to 0.98); 21 studies, 2282 women); children (sensitivity 0.96 (95% CI 0.95 to 0.97); specificity 0.96 (95% CI 0.92 to 0.98); 17 studies, 2794 children); and adults (sensitivity 0.96 (95% CI 0.93 to 0.97); specificity 0.93 (95% CI 0.80 to 0.98); 9 studies, 1088 participants), as well as dilerent scanning techniques. In a hypothetical cohort of 1000 patients, there would be 12 false-positive results and 30 false-negative results. Methodological quality of the included studies was poor, and the risk of bias was high or unclear in 53% to 83% of the QUADAS-2 domains.	The significant limitation to the review was the overall methodological weakness of the included studies and low standards of reporting. Although concern for applicability was low, risk of bias was high in the assessment. Essentially, this means that whilst these studies were conducted within a relevant clinical setting using typical patients with suspected appendicitis, the summary estimates may not be representative of the accuracy of MRI for diagnosing appendicitis in clinical practice.
Fernando SM, Tran A, Cheng W, et al. Accuracy of presenting symptoms, physical examination, and imaging for diagnosis of ruptured abdominal aortic aneurysm: Systematic review and meta-analysis. Acad Emerg Med. 2022; 29(4):486-496.	35220634	Systematic Review and Meta-analysis	Moderate level of evidence	To evaluate the accuracy of presenting symptoms, physical examination signs, computed tomography with angiography (CTA), and point-of-care ultrasound (PoCUS) for diagnosis of ruptured abdominal aortic aneurysm (rAAA).	Included were 20 studies (2,077 patients), with 11 of these evaluating signs and symptoms, seven evaluating CTA, and five evaluating PoCUS.	The primary reference standard used in all studies was intraoperative diagnosis or death from rAAA. Because PoCUS cannot detect rupture, authors secondarily assessed its accuracy for the diagnosis of AAA, using the reference standard of intraoperative or CTA diagnosis. The authors used GRADE to assess certainty in estimates.	Pooled sensitivities of abdominal pain, back pain, and syncope for rAAA were 61.7%, 53.6%, and 27.8%, respectively (low certainty). Pooled sensitivity of hypotension and pulsatile abdominal mass were 30.9% and 47.1%, respectively (low certainty). CTA had a sensitivity of 91.4% and specificity of 93.6% for diagnosis of rAAA (moderate certainty). In the secondary analysis, PoCUS had a sensitivity of 97.8% and specificity of 97.0% for diagnosing AAA in patients\ suspected of having rAAA (moderate certainty). The authors conclude that classic clinical symptoms associated with rAAA have poor sensitivity, and their absence does not rule out the condition. CTA has reasonable accuracy, but misses some cases of rAAA. PoCUS is a valuable tool that can help guide the need for urgent transfer to a vascular center in patients suspected of having rAAA.	This review also has limitations. Most importantly, because the majority of studies in our review only recruited patients with confirmed rAAA (i.e., only positive cases), we were unable to accurately generate a pooled incidence of rAAA among ED patients presenting with abdominal pain, nor were we able to perform a test–treatment threshold analysis. Second, we evaluated these various tools independently. In practice, however, providers typically use combinations of signs and tests to arrive at a diagnosis.Third, our methodologic approach was limited by a lack of available data, and while we preferred to use the bivariate HSROC model, this was not possible in all instances, and univariate analyses had to be performed instead in some cases. Fourth, there was evidence of statistical heterogeneity, as evidenced by a higher I^2 value for some meta-analyses.
Kabir SA, Kabir SI, Sun R, et al. How to diagnose an acutely inflamed appendix; a systematic review of the latest evidence. Int J Surg. 2017;40:155-62.	28279749	Systematic Review	Moderate level of evidence	To systematically report and analyse the latest evidence on the different approaches used in diagnosing appendicitis.	The study included ultimate diagnoses of appendicitis. After applying inclusion and exclusion criteria, a total of 58 studies were selected for final review.	Two independent researchers screened title and abstracts, 3222 articles were considered irrelevant. A third independent reviewer reviewed equivocal cases. Selections were based on the PRISMA Flow methodology. Included studies comprised of randomized controlled trials, meta-analyses, systematic reviews, retrospective studies, case series and case reports.	In summary, in adults, raised Alvarado scores and laboratory markers (WCC, CRP) all contribute to the suspicion of appendicitis. When alone, none of them are able to predict the diagnosis in a valid or reliable way. Subsequent surgical intervention should therefore not be based on either of them alone. However, when used in combination they show greater promise. A precise algorithm for the diagnosis of appendicitis based on a combination of these variables will prove to be useful. We believe also that many novel markers will be adopted and utilised successfully in the future. Further research is warranted to determine the effectiveness of these markers, and to continue searching for undiscovered potential markers. CT remains the best radiological modality for diagnosing appendicitis but radiation exposure and long-term cancer risks are a major concern. The use of USS-CT pathways or even USS-MRI pathways increases diagnostic certainty without always having to expose unclear cases to radiation. The alternative use of repeat USS may reach a sensitivity of 100%. The precise sequence and threshold for imaging pathways remains are yet to be determined.	N/A
Rud B, Vejborg TS, Rappeport ED, et al. Computed tomography for diagnosis of acute appendicitis in adults. Cochrane Database Syst Rev. 2019; Nov 19; 2019(11):CD009977.	31743429	Systematic Review	High level of evidence	To evaluate the accuracy of CT for diagnosing appendicitis in adults with suspected appendicitis. Secondary objectives were to compare the accuracy of contrast-enhanced versus non-contrast-enhanced CT, to compare the accuracy of low-dose versus standard-dose CT, and to explore the influence of CT-scanner generation, radiologist experience, degree of clinical suspicion of appendicitis, and aspects of methodological quality on diagnostic accuracy.	Authors included prospective studies that compared results of CT versus outcomes of a reference standard in adults (> 14 years of age) with suspected appendicitis. We excluded studies recruiting only pregnant women; studies in persons with abdominal pain at any location and with no particular suspicion of appendicitis; studies in which all participants had undergone ultrasonography (US) before CT and the decision to perform CT depended on the US outcome; studies using a case-control design; studies with fewer than 10 participants; and studies that did not report the numbers of true-positives, false-positives, false-negatives, and true-negatives. Authors identified 64 studies including 71 separate study populations with a total of 10,280 participants (4583 with and 5697 without acute appendicitis).	Two review authors independently screened and selected studies for inclusion. Two review authors then independently collected the data from each study and evaluated methodological quality according to the Quality Assessment of Studies of Diagnostic Accuracy - Revised (QUADAS-2) tool. A bivariate random-effects model was used to obtain summary estimates of sensitivity and specificity.	Estimates of sensitivity ranged from 0.72 to 1.0 and estimates of specificity ranged from 0.5 to 1.0 across the 71 study populations. Summary sensitivity was 0.95 (95% confidence interval (CI) 0.93 to 0.96), and summary specificity was 0.94 (95% CI 0.92 to 0.95). At the median prevalence of appendicitis (0.43), the probability of having appendicitis following a positive CT result was 0.92 (95% CI 0.90 to 0.94), and the probability of having appendicitis following a negative CT result was 0.04 (95% CI 0.03 to 0.05). In subgroup analyses according to contrast enhancement, summary sensitivity was higher for CT with intravenous contrast (0.96, 95% CI 0.92 to 0.98), CT with rectal contrast (0.97, 95% CI 0.93 to 0.99), and CT with intravenous and oral contrast enhancement (0.96, 95% CI 0.93 to 0.98) than for unenhanced CT (0.91, 95% CI 0.87 to 0.93). Summary sensitivity of CT with oral contrast enhancement (0.89, 95% CI 0.81 to 0.94) and unenhanced CT was similar. Results show practically no differences in summary specificity, which varied from 0.93 (95% CI 0.90 to 0.95) to 0.95 (95% CI 0.90 to 0.98) between subgroups. Summary sensitivity for low-dose CT (0.94, 95% 0.90 to 0.97) was similar to summary sensitivity for standard-dose or unspecified-dose CT (0.95, 95% 0.93 to 0.96); summary specificity did not differ between lowdose and standard-dose or unspecified-dose CT. No studies had high methodological quality as evaluated by the QUADAS-2 tool. Major methodological problems were poor reference standards and partial verification primarily due to inadequate and incomplete follow-up in persons who did not have surgery. The authors conclude that the sensitivity and specificity of CT for diagnosing appendicitis in adults are high. Unenhanced standard-dose CT appears to have lower sensitivity than standard-dose CT with intravenous, rectal, or oral and intravenous contrast enhancement.	In some study reports, the reporting quality made it difficult to assess whether data collection was conducted prospectively or retrospectively. In most of these situations, authors contacted the corresponding author and excluded the study if they received no reply. However, for some studies, judgments may have been too liberal. In general, they accepted studies as having prospective data collection if study authors used the term 'prospective' or 'consecutive' to characterise the data collection, and if they found no clear-cut evidence to suggest the contrary. Another limitation was that authors did not distinguish between uncomplicated and complicated acute appendicitis as separate target conditions.

Taylor MR, Lalani N. Adult small bowel obstruction. Acad Emerg Med. 2013;20(6):528-44.	23758299	Meta-Analysis; Review	Moderate level of evidence	The primary objective was to perform a systematic review and meta-analysis of the history, physical examination, and imaging modalities associated with the diagnosis of (small bowel obstruction)SBO. The secondary objectives were to identify the prevalence of SBO in prospective ED-based studies of adult abdominal pain and to apply Pauker and Kassirer's threshold approach to clinical decision-making to the diagnosis and management of SBO	To be included in this review, prospective studies were required to have 1) bedside US performed by EPs, 2) enrollment of adult patients with symptoms/signs suggestive of AAAs, and 3) comparison/confirmation of results. We searched MEDLINE and EMBASE with the PubMed interface for articles from 1965 through November 2011 (see Appendix A for complete MEDLINE and EMBASE search strategies). We also searched the Cochrane Central Register of Controlled Trials and the Cochrane Review addressing the topic of emergency bedside US in the diagnosis of AAA. The searches were conducted with the assistance of a medical librarian. Review of the titles and abstracts of the search results were conducted independently by two authors (ER and NM) and disagreements were adjudicated by a third author (RS). Bibliographies of the included articles were also reviewed.	METHODS: MEDLINE, EMBASE, major emergency medicine (EM) textbooks, and the bibliographies of selected articles were scanned for studies that assessed one or more components of the history, physical examination, or diagnostic imaging modalities used for the diagnosis of SBO. The selected articles underwent a quality assessment by two of the authors using the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool. Data used to compile sensitivities and specificities were obtained from these studies and a meta-analysis was performed on those that examined the same historical component, physical examination technique, or diagnostic test. Separate information on the prevalence and management of SBO was used in conjunction with the meta-analysis findings of computed tomography (CT) to determine the test and treatment threshold.	The prevalence of SBO in the ED was determined to be approximately 2% of all patients who present with abdominal pain. Having a previous history of abdominal surgery, constipation, abnormal bowel sounds, and/or abdominal distention on examination were the best history and physical examination predictors of SBO. X-ray was determined to be the least useful imaging modality for the diagnosis of SBO, with a pooled positive likelihood ratio (+LR) of 1.64 (95% confidence interval [CI] = 1.07 to 2.52). On the other hand, CT and magnetic resonance imaging (MRI) were both quite accurate in diagnosing SBO with +LRs of 3.6 (5- to 10-mm slices, 95% CI = 2.3 to 5.4) and 6.77 (95% CI = 2.13 to 21.55), respectively. Although limited to only a select number of studies, the use of ultrasound (US) was determined to be superior to all other imaging modalities, with a +LR of 14.1 (95% CI = 3.57 to 55.66) and a negative likelihood ratio (-LR) of 0.13 (95% CI = 0.08 to 0.20) for formal scans and a +LR of 9.55 (95% CI = 2.16 to 42.21) and a -LR of 0.04 (95% CI = 0.01 to 0.13) for beside scans. Using the CT results of the meta-analysis for the 5- to 10-mm slice subgroup as well as information on intravenous (IV) contrast reactions and nasogastric (NG) intubation management, the pretest probability threshold for further testing was determined to be 1.5%, and the pretest probability threshold for beginning treatment was determined to be 20.7%. The authors conclude that potentially useful aspects of the history and physical examination were limited to a history of abdominal surgery, constipation, and the clinical examination findings of abnormal bowel sounds and abdominal distention. CT, MRI, and US are all adequate imaging modalities to make the diagnosis of SBO. Bedside US, which can be performed by EPs, had very good diagnostic accuracy and has the potential to play a larger role in the ED diagnosis of SBO. More ED-focused research into this area will be necessary to bring about this change	There were several limitations of this meta-analysis. First, it is possible that some studies relating to SBO diagnostics were missed given the strategy of our search. Second, we limited our searches to generalized SBO in adults and therefore our meta-analysis. The quality of the studies in this meta-analysis was highly variable and was subject to several biases. Eventual clinical outcome is fraught with bias, however, as many variables could play into what ultimately happens to a patient in the hospital. One of the limitations to the pooled meta-analysis groups is the large heterogeneity seen in the studies. Some of this was controlled for by removing certain outliers, but was not always completely eliminated. The nature of diagnostic imaging studies, especially CT scans, lends itself to heterogeneity given the very wide range of machines and scanning techniques used, as well as the tools used for interpretation. Furthermore, the benefits and risks of NG placement did not take into account patient preference or pain relief and were derived from lower-quality, potentially biased primary studies. Risk of bias - one or more key results (state which ones in the comments section) were based on studies with a majority having a high risk of bias No test for heterogeneity was performed on all of the studies so it is impossible to know if it exists in this analysis. Since the studies included in this meta-analysis were cross-sectional or retrospective in nature the risk of bias is possible. The study population may be different in the detection and analysis of this analysis.
Wu LM, Xu JR, Gu HY, et al. Is magnetic resonance imaging a reliable diagnostic tool in the evaluation of active Crohn's disease in the small bowel? J Clin Gastroenterol. 2013;47(4):328-38.	23340059	Meta-Analysis; Review	Moderate level of evidence	To evaluate the overall diagnostic accuracy of magnetic resonance imaging (MRI) in assessing the activity of Crohn's disease (CD) in the small bowel.	An electronic search yielded 630 primary studies, of which 601 were excluded after reviewing the title and abstract. Twelve articles were excluded after reviewing the full article. Therefore, a total of 17 studies (19 populations) with 725 patients, who fulfilled all of the inclusion criteria, were considered for the analysis.	Two reviewers searched MEDLINE, EMBASE, and other electronic databases to identify studies in which MRI imaging was evaluated for assessing the activity of CD in the small bowel from January 2001 to September 2011. Bivariate random effects metaanalytic methods were used to estimate summary, sensitivity, specificity, and receiver operating characteristic curves.	MRI had a pooled sensitivity of 0.87 [95% confidence interval (CI): 0.77, 0.93] and a pooled specificity of 0.91 (95% CI: 0.81, 0.96). Overall, likelihood ratio (LR)+ was 9.5 (95% CI: 4.4, 20.6) and LR was 0.14 (95% CI: 0.08, 0.26). In patients with high pretest probabilities, MRI enabled confirmation of active CD; in patients with low pretest probabilities, MRI enabled exclusion of active CD. Worst-case-scenario (pretest probability, 50%) posttest probabilities were 90% and 13% for positive and negative MRI results, respectively. The authors conclude that a limited number of small studies suggest that MRI has high sensitivity and specificity for diagnosis of active CD in the small bowel; MRI will likely also prove to be suitable as the primary modality for active CD imaging surveillance.	The authors report several possible limitations. Authors attempted to examine publication bias using the Deeks funnel plot, and no publication bias was found. However, potential publication bias may still exist, because small studies with optimistic results may be published more easily than small studies with unfavorable results. Moreover, only included studies published in English, which might invoke the so-called "Tower of Babel" bias, which refers to the fact that investigators working in a language other than English could be sending only studies with positive results to international journals. Furthermore, the interpretation of MRI scans was performed qualitatively in the majority of the studies, and blinding in 4 studies was either unclear or absent. So there is a risk of subjective interpretation, but it is more likely to be in favor of MRI, and its diagnostic accuracy might be even lower.
Yoon HM, Suh CH, Cho YA, et al. The diagnostic performance of reduced-dose CT for suspected appendicitis in paediatric and adult patients: A systematic review and diagnostic meta-analysis. Eur Radiol. 2018; 28(6):2537-2548.	29327290	Systematic Review and Meta-analysis	Moderate level of evidence	To evaluate the diagnostic performance of reduced-dose CT for suspected appendicitis.	Fourteen original articles with a total of 3,262 patients were included. Studies or subsets of studies that investigated the diagnostic performance of reduced-dose CT for suspected appendicitis in paediatric and adult patients were eligible for inclusion in the analysis. Studies were excluded if any of following criteria were met: (1) case reports or case series that involved <10 patients; (2) conference abstracts, letters, editorials, reviews, meta-analyses, consensus statements and guidelines; (3) studies that focused on topics other than using reduced-dose CT for evaluating suspected appendicitis; (4) studies with insufficient data for evaluating the diagnostic performance of reduced-dose CT for suspected appendicitis; and (5) existence of studies with partially overlapping patient populations.	A systematic search of the MEDLINE and EMBASE databases was carried out through to 10 January 2017. Studies evaluating the diagnostic performance of reduced-dose CT for suspected appendicitis in paediatric and adult patients were selected. Pooled summary estimates of sensitivity and specificity were calculated using hierarchical logistic regression modelling. Meta-regression was performed.	For all studies using reduced-dose CT, the summary sensitivity was 96 % (95 % CI 93–98) with a summary specificity of 94 % (95 % CI 92–95). For the 11 studies providing a head-to-head comparison between reduced-dose CT and standard-dose CT, reduced-dose CT demonstrated a comparable summary sensitivity of 96 % (95 % CI 91–98) and specificity of 94 % (95 % CI 93–96) without any significant differences (p=.41). In meta-regression, there were no significant factors affecting the heterogeneity. The median effective radiation dose of the reduced-dose CT was 1.8 mSv (1.46–4.16 mSv), which was a 78 % reduction in effective radiation dose compared to the standard-dose CT. The authors conclude that reduced-dose CT shows excellent diagnostic performance for suspected appendicitis.	First, nine of 14 included studies were retrospective, resulting in a high risk of bias in patient selection. Second, the decision threshold of indeterminate cases was considered as positive in eight studies and negative in one study, and not reported in four studies.