CRITICAL APPRAISAL META-ANALYSIS

Reviewer: Katy Deljoui MD Eastern Virginia Medical School Journal Club
Date: 02/28/2011

Citation: Doria et al. “US or CT for Diagnosis of Appendicitis in Children and Adults? A Meta-analysis.” ***Radiology:*** Volume 241: Number 1—October 2006

Guide	Question	Comments
	Are the results valid?	
1.	Did the review explicitly address a sensible question?	-Yes: what is the diagnostic performance of CT and US in the diagnosis of acute appy in peds and adult pts?
	Was the search for relevant studies details and exhaustive? How did they search the literature?	Yes. -Used search engines like Medline, Embase, Cochrane review, etc… -Studies btw 1986 and 2004 -Looked for key words -Also searched reference lists -Also contacted experts regarding possible unpublished data.
3.	Were the primary studies of high methodological quality? Types of studies included/excluded?	No. The median score for the study quality of children was 34.4% (maximum score, 100%), and the median score for the studies of adults was 42.2% -Included prospective and retrospective studies with available primary data, kids < age 20, adults > age 13, imaging criteria, both females and males, prevalence btw 15-75% (wide prevalence range)
4.	Were the assessments of the included studies reproducible? How did they score studies? Inter-rater reliability between reviewers?	Yes. -Used a checklist to assess and score (no further info..?) -Good inter-rater reliability: 0.78 95% CI (0.64, 0.87) and between the two blinded reviewers (intraclass correlation coefficient, 0.70; 95% CI: 0.46-0.84)
II.	***What are the results?***	
1.	What are the overall results of the study? How many articles were screened? What was the prevalence of appendicitis What were the reported perforation rates? What was the pooled sensitivity, Specificity, DOR for US, CT or US and CT for peds and adult	-CT has significantly higher sensitivity than US for diagnosing appy in adults and children -Screened 229 articles, 57 met criteria (26 peds, 31 adults) -Mean prevalence of appy for peds in US and CT = 0.31. (31%) Mean prevalence appy for adults in US 0.48 (48%) and CT 0.40 (40%). -Overall min prevalence 0.15, max 0.75 Of those that reported perforation rates… -Perforated rates in kids 26.5% vs 18.5% in adults -Sensitivity kids with CT 94% vs 88% US

	populations? What did the meta regression demonstrate regarding sensitivity and specificity between CT and US? What were the rates of missed appy using US vs CT	-Sensitivity adults with CT 94% vs 83% US -Specificity kids with CT 95% vs 94% US -Specificity adults with CT 94% vs 93% US -DOR for kids with CT vs US 2.47 (p 0.02) -DOR for adults with CT vs US 3.1 (p 0.001) -Meta regression showed no evidence of heterogeneity due to confounders (study design, year, place) -Assuming prevalence 0.15 (min), US missed 10 kids and 18 adults per 1,000 -Assuming prevalence 0.75 (max), US missed 48 kids and 83 adults per 1,000 -With mean prevalence 0.31 in kids, would miss 280 appys and save 13 from CA -With mean prevalence 0.40 in adults, would miss 480 appys and save 2 from CA
2.	How precise are the results? What do the CI's look like. Do any cross 1 (line of indifference)?	-Seem fairly precise, with narrow CIs and no crossing of line of indifference
3.	Were the results similar from study to study? Was there significant heterogeneity?	-Distribution of data points shows that most studies have pretty similar results -With the exception of a few outliers, no significant heterogeneity appreciated.
III.	***Will the results help me in caring for my patients?***	
1.	How can I best interpret the results to apply them to the care of my patients? What is the rate of missed appy (using US) compared to risk of cancer (using CT) in kids being assessed for acute appy. How about adults?	-CT more sensitive than US in both peds and adults -Equal specificity… -So, if fairly certain dx and need confirmation (specific test), could use either CT or US, but US less costly and harmful -If clinical picture unclear, need screening test (sensitive), then CT would be better than US -For rates of missed appys, see section II-1 above Cancer risks clearly more acceptable in adults.
2.	Were all patient important outcomes considered? Were patient's decision/choice included in the decision regarding modalities.	No. -Not all outcomes considered. What about mortality? What about complications of perforated appy. Is risk of complications worse than that of cancer? Especially in adults who have co-morbid conditions and are more likely to get complications? -Pt's decision/choice not included in assessment.
3.	Are the benefits worth the costs and potential risks?	Yes. -As stated above. Risk of CA in kids seems worse than risk of perforated appy. Specificity appears to be similar though no adjustment for higher BMI's which could sig impact specificity

		difference between the two technologies as obesity is not a common confounder to CT. Sensitivity of CT is higher in all tested.

Additional issues

Most surgeons were aware of imaging results, which could bias management decisions.

No mention of BMI or obesity rates were considered which is known to impact sensitivity of US..

Clinical Bottom Line: US may be an appropriate initial testing modality for consideration in patients suspected of having acute appendicitis. Difference is sensitivity however favors CT particularly in higher prevalence populations and a negative US may not be sufficiently sensitive to avoid missing diagnoses

APPENDIX: QUADAS CRITERIA FOR EVALUATING DIAGNOSTIC ACCURACY STUDIES

Table 2The Quadas tool

Item		Yes	No	Unclear
1.	Was the spectrum of patients representative of the patients who will receive the test in practice?	( )	( )	( )
2.	Were selection criteria clearly described?	( )	( )	( )
3.	Is the reference standard likely to correctly classify the target condition?	( )	( )	( )
4.	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	( )	( )	( )
5.	Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?	( )	( )	( )
6.	Did patients receive the same reference standard regardless of the index test result?	( )	( )	( )
7.	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	( )	( )	( )
8.	Was the execution of the index test described in sufficient detail to permit replication of the test?	( )	( )	( )
9.	Was the execution of the reference standard described in sufficient detail to permit its replication?	( )	( )	( )
10.	Were the index test results interpreted without knowledge of the results of the reference standard?	( )	( )	( )
11.	Were the reference standard results interpreted without knowledge of the results of the index test?	( )	( )	( )
12.	Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	( )	( )	( )
13.	Were uninterpretable/intermediate test results reported?	( )	( )	( )
14.	Were withdrawals from the study explained?	( )	( )	( )

Source: Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J, Whiting P, et al. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Medical Research Methodology 2003;3:25. Used with permission.