Acute appendicitis: meta-analysis of diagnostic performance of CT and graded compression US related to prevalence of disease
van Randen A, Bipat S, Zwinderman AH, Ubbink DT, Stoker J, Boermeester MA

CRD summary
This well-conducted review directly compared computed tomography (CT) and compression graded ultrasonography for the diagnosis of acute appendicitis. The authors’ conclusion that CT performed better than ultrasonography, is a valid interpretation of the available data, but these data were sparse.

Authors’ objectives
To compare the accuracy of graded compression ultrasonography and computed tomography (CT) for the diagnosis of acute appendicitis, and to examine how diagnostic performance (in terms of post-test probability) varies with disease prevalence.

Searching
MEDLINE, EMBASE, CINAHL and the Cochrane Library were searched from inception to February 2006. Search strategies were reported in an on-line appendix (but the URL link provided did not work, December 2009). The bibliographies of included studies were screened for further relevant articles.

Study selection
Prospective cohort studies that compared ultrasonography and CT in the same population (of more than ten patients) with suspected appendicitis, or right lower quadrant pain, were eligible for inclusion. Included studies were required to use histopathology results at surgery and/or clinical follow-up as the reference standard, and to report sufficient data for the calculation of 2x2 contingency tables (i.e. numbers of true positive (TP), false negative (FN), false positive (FP) and true negative (TN) results). Studies evaluating only children were excluded.

Included studies used a range of CT and ultrasonography techniques (details reported). The reference standards included surgery and clinical follow-up of between one and 13 months, with the exception of one study where the reference standard was not reported. The age of participants ranged from three to 89 years, the proportion of female participants ranged from 53 to 84%, and the prevalence of appendicitis ranged from 13 to 77%.

Studies were assessed for relevance by two reviewers.

Assessment of study quality
Two reviewers independently assessed the methodological quality of included studies using the Quality Assessment of Diagnostic Accuracy Studies Assessment (QUADAS) tool; disagreements were resolved by consensus.

Data extraction
Technical details of the index tests (CT and ultrasonography), including observer experience, and 2x2 contingency tables for ultrasonography and CT compared with the reference standard, were extracted for each study. These data were used to calculate the main outcome measures (sensitivity, specificity, positive likelihood ratios and negative likelihood ratios and their 95% confidence intervals) presented in the report.

Data were extracted independently by two reviewers and disagreements were resolved by consensus.

Methods of synthesis
Sensitivity and specificity values for ultrasonography and CT from each included study were plotted in receiver operating characteristic (ROC) space. A bivariate, random-effects, non-linear regression model was used to generate summary estimates of sensitivity and specificity with 95% confidence intervals (CIs). These summary estimates were used to calculate positive and negative likelihood rations with 95% confidence intervals. Summary accuracy estimates,
for ultrasonography and CT, were compared using a z-test for paired data, with a p value of <0.05 taken to indicate significant difference. The small number of studies precluded assessment of the impact of study quality on measures of test performance.

Pre-test probability or prevalence of acute appendicitis was estimated for the following patient groups presenting at the emergency department: acute abdominal pain; suspected appendicitis or right lower quadrant pain; and those more than 65 years of age with acute abdominal pain. Overall mean prevalence for each of these groups was estimated using data from studies identified by the original searches that were conducted between 1996 and 2006, included more than 90 patients with acute abdominal pain or right lower quadrant pain and reported data per diagnosis. The post-test probabilities, for positive and negative test results in each of these groups, were calculated from the prevalence estimates and the summary likelihoods derived from the meta-analysis.

Results of the review
Six studies, with a total of 671 participants, were included in the review. All ultrasonography and computed tomography (CT) results were assessed independently from one another and from the reference standard. However, the reference standard was not assessed independently of the index test results. The time between ultrasonography and CT was not reported in two studies and was three hours or less in all other studies.

Diagnostic value of computed tomography (CT): The summary estimates for the sensitivity CT in diagnosing appendicitis was 0.91 (95% CI 0.84 to 0.95) and for specificity was 0.90 (95% CI 0.85 to 0.94). The summary estimate of the positive likelihood for CT was 9.29 (95% CI 6.86 to 12.6) and for the negative likelihood ratio was 0.10 (95% CI 0.06 to 0.17).

Diagnostic value of graded compression ultrasonography: The summary estimate for the sensitivity of ultrasonography was 0.78 (95% CI 0.67 to 0.86) and for specificity was 0.83 (95% CI 0.76 to 0.88). The summary estimate of the positive likelihood for ultrasonography was 4.5 (95% CI 3.03 to 6.68) and for the negative likelihood ratio was 0.27 (95% CI 0.17 to 0.43).

Summary sensitivity (p<0.017), specificity (p<0.037), positive likelihood (p=0.011) and negative likelihood (p=0.013) values differed significantly between CT and ultrasonography, in favour of CT.

Overall post-test probabilities, following a positive test, were 90% for CT and 82% for ultrasonography. Overall post-test probabilities, following a negative test, were 9% for CT and 21% for ultrasonography. CT showed significantly better diagnostic performance at every clinically relevant prevalence. Post-test probabilities, following a positive test, decreased with lower prevalence.

Authors' conclusions
CT performed better than graded compression ultrasonography in diagnosing acute appendicitis. There was an important relationship between pre-test probability (prevalence) and diagnostic performance.

CRD commentary
The review addressed a clearly defined research question, which was fully defined by appropriate inclusion criteria. A number of relevant sources were searched to identify studies for inclusion and measures were taken to minimise error and bias at all stages of the review process. The methodological quality of included studies was assessed using an appropriate, validated tool and a summary of this assessment was provided. The small number of studies precluded investigation of the impact of components of study quality, or other relevant variables (e.g. index test techniques) on measures of test performance. The methods used to generate summary estimates of test performance represented the best available option. The authors' conclusions are valid, based on the available data. However, the small number and size of studies may limit the robustness of the conclusions.

Implications of the review for practice and research
Practice: The authors stated that CT is recommended in patients suspected of acute appendicitis. Given the radiation load associated with CT, graded compression ultrasonography is recommended as the primary investigation for young
female patients and patients of slight build.

**Research:** The authors stated that future studies should take into account the differences in pre-test probability associated with different populations and physical and laboratory findings, and consider the added value of each imaging modality.

**Funding**
Not stated.

**Bibliographic details**

**PubMedID**
18682583

**DOI**
10.1148/radiol.2483071652

**Original Paper URL**
http://radiology.rsna.org/content/249/1/97.abstract

**Indexing Status**
Subject indexing assigned by NLM

**MeSH**
Acute Disease; Adolescent; Adult; Appendicitis /diagnosis /epidemiology; Female; Humans; Male; Prevalence; Probability; Sensitivity and Specificity; Tomography, X-Ray Computed; Ultrasonography /methods

**AccessionNumber**
12008106619

**Date bibliographic record published**
03/02/2009

**Date abstract record published**
23/12/2009

**Record Status**
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.