Systematic review and meta-analysis of emergency ultrasonography for blunt abdominal trauma

Stengel D, Bauwens K, Sehouli J, Porzsolt F, Rademacher G, Mutze S, Ekkernkamp A

Authors’ objectives
To evaluate how precise and reliable ultrasonography is as a primary tool for the assessment of injury in blunt abdominal trauma.

Searching
MEDLINE, EMBASE and the Cochrane Library were searched using the search terms listed in the review. The German and American National Medicine Library websites were also searched. The references of retrieved articles were examined and an extensive manual search of relevant publications and books was performed. Articles written in any language were considered. Authors and other experts were contacted for information on original and unpublished data.

Study selection
Study designs of evaluations included in the review
Diagnostic accuracy studies that compared ultrasound with a reference standard, whether blinded or not, were eligible for inclusion in the review. The included studies used consecutive enrolment (18 studies), non-consecutive enrolment (11 studies), and one was a case-control study. In only four of the studies was the evaluator blinded to the diagnostic test.

Specific interventions included in the review
Studies of any ultrasonography techniques were eligible for inclusion in the review. The authors did not specify any inclusion criteria for the review in relation to the ultrasound technique. The included studies were all of emergency ultrasound.

Reference standard test against which the new test was compared
The authors did not specify any inclusion criteria for the review in relation to the ‘gold’ standard. The ‘gold’ standard employed in most of the included studies was laparotomy with diagnostic peritoneal lavage (DPL), computed tomography (CT) and observation. Some studies did not include observation, some employed just DPL and CT, others just CT, and one study used 99m-technetium scintigraphy.

Participants included in the review
Studies of patients with suspected blunt abdominal trauma were considered for inclusion in the review.

Outcomes assessed in the review
The authors did not specify any inclusion criteria for the review in relation to the outcomes. The included studies reported the detection of organ injury and the detection of free intraperitoneal fluid as outcomes. Prevalence, sensitivity, specificity, accuracy and likelihood ratios were calculated for the included studies and reported in the review.

How were decisions on the relevance of primary studies made?
The authors do not state how the papers were selected for the review, or how many of the reviewers performed the selection.

Assessment of study quality
The trials were rated on the basis of two main potential sources of bias: spectrum composition bias and work-up bias; specifically, ‘Were patients enrolled consecutively’ and ‘Was a reference standard applied independently and regardless of the results of the ultrasound test’. The trials were graded according to levels of evidence. Two of the authors applied the quality rating to the included studies.

Data extraction
The authors do not state how the data were extracted for the review, or how many of the reviewers performed the data extraction.

The categories of data extracted were: bibliographic details; 'gold' standard; type of enrolment; whether blind or not; the number of patients; and the results.

Methods of synthesis

How were the studies combined?

Weighted negative predictive values (NPVs) were regressed to the prevalence of either organ injury or free fluid and confidence intervals (CIs) were calculated. All 2x2 tables were adjusted for continuity. Receiver operating characteristic (ROC) curves were estimated for each trial and the common effect size estimated using summary ROC curves (see Other Publications of Related Interest nos.1-2). Positive and negative likelihood ratios (PLR and NLR, respectively) were calculated. The studies were pooled using a fixed-effect model ((see Other Publications of Related Interest nos.3-4) except where there was heterogeneity, when a random-effects model was used (see Other Publications of Related Interest no.5).

A multivariate linear regression analysis was performed to assess a possible stage migration effect, i.e. the possibility of the ultrasound technique improving over time and affecting the findings of the systematic review. The interaction of the pooled ROC curve with the line where sensitivity equals specificity was computed (Q). End points had to have been stated in more than one trial.

How were differences between studies investigated?

Heterogeneity was calculated using the chi-squared statistic for k minus one degrees of freedom.

Results of the review

Thirty studies (n=9,047) were included. Of these, 11 studies (n=2,819; 6 level IIb evidence, 5 level IIIb) investigated the detection of organ injury and 19 (n=6,492; 14 level IIb evidence, 5 level IIIb) investigated the detection of free intraperitoneal fluid.

Organ injury.

Q was estimated as 0.91 (95% CI: 0.76, 1.07) using inverse variance weights and 0.90 (95% CI:0.75, 1.05) using robust regression. The NPV ranged from 0.72 to 0.99.

There was significant heterogeneity for the LRs so a random-effects model was used for pooling. The PLR and NLR were 21.99 (95% CI: 15.08, 32.06) and 0.23 (95% CI: 0.18, 0.28), respectively. The related post-test probabilities were also presented in the review.

Free intraperitoneal fluid.

Q was estimated as 0.89 (95% CI: 0.73, 1.05) using inverse variance weights and 0.89 (95% CI:0.76, 1.03) using robust regression. The NPV ranged from 0.72 to 0.99.

There was significant heterogeneity for the LRs so a random-effects model was used for pooling. The PLR and NLR were 36.28 (95% CI: 28.97, 45.44) and 0.24 (95% CI: 0.20, 0.28), respectively. The related post-test probabilities were also presented in the review.

Subgroup analyses.

For studies involving paediatric patients (4 studies, n=501), there was no distinction between recognition of free fluid and organ injury. Q was estimated as 0.88 (95% CI: 0.33, 1.42) using inverse variance weights and 0.87 (95% CI:0.76, 0.99) using robust regression. For ultrasonography in the presence of haemodynamic instability (3 studies, n=155), Q was estimated as 0.89 (95% CI: 0.13, 1.64) using robust regression. For ultrasonography versus different 'gold' standards, higher Q values were computed comparing ultrasonography with diagnostic peritoneal lavage than with CT. These analyses were not reported well but a significant finding appears to be that, compared with CT, the ability of
ultrasonography to exclude lacerations remained below clinically acceptable threshold limits (NLR 0.21, 95% CI: 0.16, 0.29).

Linear regression analysis failed to demonstrate any statistically- significant change in the diagnostic accuracy of emergency ultrasound over the preceding 20 years.

Authors' conclusions
In their abstract, the authors concluded 'Despite its high specificity, ultrasonography has an unexpectedly low sensitivity for the detection of both free peritoneal fluid and organ lesions. In clinically suspected abdominal trauma, another assessment (e.g. helical computed tomography) must be performed regardless of the initial ultrasound findings'.

In the text, the overall summary was 'Relatively low sensitivity and NPVs, as well as negative likelihood ratios and associated post-test probabilities out of clinically acceptable range, prevent belief in negative sonographic findings, whereas high specificity and positive LRs above 10 mean that positive results prove the presence of intra-abdominal injury'.

CRD commentary
This review addressed an appropriate question although the inclusion and exclusion criteria were poorly defined in the published paper. The conduct of the review was described adequately, although further details on how many of the reviewers selected the papers and whether or not the paper selection, quality testing and data extraction processes were subject to independent duplication, would have been useful. The literature search appeared reasonably comprehensive with three main databases searched, supplemented with appropriate handsearches. Only limited details of the primary studies were presented in the review; probably due to the fact that it contained 30 studies.

The data were pooled using meta-analyses. Whilst this was in principle correct, the pooling of all studies irrespective of the 'gold' standard used might have been a source of the significant heterogeneity seen. Furthermore, the effect of the different 'gold' standards was not well discussed in the text, with different terminology used to that in the tables and a lack of clarity regarding which studies were included. Also, there was no sensitivity analysis or investigation of any differences between the few blinded studies and the remaining studies. Some of the subgroup analyses that were performed suffered from a small number of studies.

The authors' conclusions appear to go beyond the findings of their systematic review. However, the summary of the review's findings in their discussion does follow from the data as presented.

Implications of the review for practice and research
Practice: The authors state that 'In clinically suspected abdominal trauma, another assessment (e.g. helical computed tomography) must be performed regardless of the initial ultrasound findings'.

Research: The authors did not state any implications for further research.

Bibliographic details

PubMedID
11442520

DOI
10.1046/j.0007-1323.2001.01777.x

Other publications of related interest


**Indexing Status**
Subject indexing assigned by NLM

**MeSH**
Abdominal Injuries /ultrasonography; Ascitic Fluid; Clinical Trials as Topic; Emergencies; Humans; Prospective Studies; Sensitivity and Specificity; Wounds, Nonpenetrating /ultrasonography

**AccessionNumber**
12001001751

**Date bibliographic record published**
31/03/2003

**Date abstract record published**
31/03/2003

**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.