Elastography for the diagnosis of severity of fibrosis in chronic liver disease: a meta-analysis of diagnostic accuracy

Tsochatzis EA, Gurusamy KS, Ntaoula S, Cholongitas E, Davidson BR, Burroughs AK

CRD summary
This well conducted review found that elastography had good sensitivity and specificity for cirrhosis, but should be applied cautiously to everyday clinical practice because there was no validation of the stiffness cut-offs for the various stages. These conclusions are likely to be reliable, although the possibility of publication bias should be considered.

Authors’ objectives
To evaluate the performance of elastography for the diagnosis of fibrosis in chronic liver disease.

Searching
MEDLINE, EMBASE, Science Citation Index expanded, Cochrane Hepato-Biliary Group Controlled Trials Register and Cochrane Central Register of Controlled Trials (CENTRAL) were searched to May 2009. Two relevant conference proceedings were searched from 2007 to 2009. Reference lists of retrieved studies and reviews were handsearched. The full search strategies were available as a web appendix and included a diagnostic filter. No language restrictions were applied. The review was restricted to published studies available as abstracts or full-text articles.

Study selection
Studies that assessed elastography (index test) for the diagnosis or monitoring of the severity of liver fibrosis against liver biopsy (reference standard) in at least 10 patients were eligible for inclusion. Studies had to report sufficient data to construct a 2x2 table of test performance based on a defined cut-off point for liver stiffness. Where necessary, authors were contacted for additional information. Studies in which the time interval between elastography and liver biopsy was more than three months were excluded.

The aetiology of liver disease in the included studies was chronic hepatitis B or C, alcoholic liver disease, non-alcoholic fatty liver disease and chronic hepatitis C HIV coinfection. Mean age, where reported, ranged from 29 to 68 years. Mean body mass index (BMI) ranged from 22 to 28. Most studies used the METAVIR histological system; other systems used were Scheuer, Kleiner, Brunt, Ishak and Desmet.

Two reviewers independently assessed studies for inclusion. Disagreements were resolved through referral to two further reviewers.

Assessment of study quality
Two reviewers independently assessed study quality using the 14-item QUADAS tool and whether the index test was performed according to the manufacturers instructions. Liver biopsy was rated as an acceptable reference standard if the specimen was at least 15mm long and included at least six portal tracts. Disagreements were resolved through referral to two further reviewers.

Data extraction
Two reviewers independently extracted data to construct 2x2 tables of test performance and calculated sensitivity, specificity and diagnostic odds ratios (DOR). Disagreement were resolved through referral to two further reviewers.

Methods of synthesis
Summary sensitivity and specificity, with 95% confidence intervals, and summary receiver operating characteristic (SROC) plots were estimated using the bivariate/hierarchical SROC (HSROC) models. Estimates of the post-test probability of disease were calculated from the summary sensitivity and specificity for three different pre-test probabilities of disease: 25%, 50% and 75%. Data were pooled separately according to fibrosis stage.

Subgroup analyses were conducted on high versus low study quality, fibrosis stage, aetiological diagnoses, different stiffness tresholds thresholds a specific stage of fibrosis, initial diagnosis versus monitoring of fibrosis, treatments, body
mass index, range of transaminase and study country. Subgroup analyses were conducted by investigating whether DORs differed across subgroups and using meta-regression analysis with the logarithmically transformed DOR as the dependent variable. Heterogeneity was assessed using the $I^2$ and $X^2$ statistics. Sensitivity analysis was conducted by excluding studies of low methodological quality and those published only as abstracts.

Results of the review
Forty studies (7,723 participants) were included in the review. No studies were found to be free of any risk of bias. Nine studies had both acceptable index test and reference standard quality.

Detection of fibrosis stage 4 (cirrhosis) (30 studies): Summary sensitivity was 83% (95% CI 79% to 86%) and specificity was 89% (95% CI 87% to 91%). Thresholds ranged from 9.0 to 26.5 kPa. There was statistically significant heterogeneity in the DOR ($p=0.002, I^2=49%$).

Detection of fibrosis stage 3 (24 studies): Summary sensitivity was 82% (95% CI 78% to 86%) and specificity was 86% (95% CI 82% to 89%). Thresholds ranged from 7.3 to 15.4 kPa. There was no evidence of heterogeneity in the DOR.

Detection of fibrosis stage 2 (31 studies): Summary sensitivity was 79% (95% CI 74% to 82%) and specificity was 78% (95% CI 72% to 83%). Thresholds ranged from 4.0 to 10.1 kPa. There was statistically significant heterogeneity in the DOR ($p<0.001, I^2=67%$).

Detection of fibrosis stage 1 (10 studies): Summary sensitivity was 78% (95% CI 73% to 83%) and specificity was 83% (95% CI 72% to 90%). Thresholds ranged from 4.9 to 8.8 kPa. There was no evidence of heterogeneity in the DOR.

Results of subgroup analyses were reported.

Authors’ conclusions
Elastography theoretically had good sensitivity and specificity for cirrhosis (and less for lesser degrees of fibrosis). However, it should be applied cautiously to everyday clinical practice as there was no validation of the stiffness cut-offs for the various stages. Such validation was required before elastography was considered sufficiently accurate for non-invasive staging of fibrosis.

CRD commentary
The review addressed a focused question supported by clearly defined inclusion criteria. The literature search included relevant databases and included attempts to reduce language bias. Use of a diagnostic filter and restriction to published studies raised the possibility of publication bias. This was assessed in the review, but the methods used were not appropriate for diagnostic accuracy data. Appropriate steps were taken to minimise bias and errors at all stages of the review.

Study quality was assessed using appropriate criteria and was considered in the analysis; the assessment results were not reported in detail, although the authors stated they were available as a web appendix this could not be located. A statistically robust analysis was conducted using appropriate methods that included assessment and investigation of heterogeneity and results were clearly presented.

This was generally a well conducted review and the results are likely to be reliable, although the possibility of publication bias should be considered.

Implications of the review for practice and research
Practice: The authors stated the elastography should be applied cautiously to everyday clinical practice because there was no validation of the stiffness cut-offs for the various stages.

Research: The authors stated that validation of the stiffness cut-offs was required before elastography was considered sufficiently accurate for non-invasive staging of fibrosis. They also stated a need for further high-quality studies that compared elastography measurements with quantitative histological assessments of fibrosis, such as collagen proportionate area or evaluate these with respect to clinical outcomes.

Funding
No direct funding.

**Bibliographic details**

**PubMedID**
21146892

**DOI**
10.1016/j.jhep.2010.07.033

**Original Paper URL**
http://www.journal-of-hepatology.eu/article/S0168-8278%2810%2900825-1/abstract

**Indexing Status**
Subject indexing assigned by NLM

**MeSH**
Biopsy; Chronic Disease; Elasticity Imaging Techniques; Liver /pathology; Liver Cirrhosis /diagnosis /pathology; Liver Diseases /diagnosis /pathology; Severity of Illness Index

**AccessionNumber**
12011002339

**Date bibliographic record published**
29/06/2011

**Date abstract record published**
29/01/2012

**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.