Research Article

EASL EUROPEAN ASSOCIATION FOR THE STUDY OF THE LIVER | JOURNAL OF HEPATOLOGY

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

E.A. Tsochatzis[1], K.S. Gurusamy[2], S. Ntaoula[1], E. Cholongitas[1], B.R. Davidson[2], A.K. Burroughs[1,*]

[1]*The Royal Free Sheila Sherlock Liver Centre and Division of Surgery, Royal Free Hospital, London NW3 2QG, UK;* [2]*University Department of Surgery, Royal Free Hospital, London NW3 2QG, UK*

Background & Aims: Transient elastography is a non-invasive method, for the assessment of hepatic fibrosis, developed as an alternative to liver biopsy. We studied the performance of elastography for diagnosis of fibrosis using meta-analysis.

Methods: MEDLINE, EMBASE, SCI, Cochrane Library, conference abstracts books, and article references were searched. We included studies using biopsy as a reference standard, with the data necessary to calculate the true and false positive, true and false negative diagnostic results of elastography for a fibrosis stage, and with a 3-month maximum interval between tests. The quality of the studies was rated with the QUADAS tool.

Results: We identified 40 eligible studies. Summary sensitivity and specificity was 0.79 (95% CI 0.74–0.82) and 0.78 (95% CI 0.72–0.83) for F2 stage and 0.83 (95% CI 0.79–0.86) and 0.89 (95% CI 0.87–0.91) for cirrhosis. After an elastography result at/over the threshold value for F2 or cirrhosis ("positive" result), the corresponding post-test probability for their presence (if pre-test probability was 50%) was 78%, and 88% respectively, while, if values were below these thresholds ("negative" result), the post-test probability was 21% and 16%, respectively. No optimal stiffness cut-offs for individual fibrosis stages were validated in independent cohorts and cut-offs had a wide range and overlap within and between stages.

Conclusions: Elastography theoretically has good sensitivity and specificity for cirrhosis (and less for lesser degrees of fibrosis); however, it should be cautiously applied to everyday clinical practice because there is no validation of the stiffness cut-offs for the various stages. Such validation is required before elastography is considered sufficiently accurate for non-invasive staging of fibrosis.

Introduction

Liver fibrosis represents the final common outcome of chronic liver injury and is often progressive, eventually evolving into cirrhosis [1]. Cirrhosis is the severest form of fibrosis with the worst clinical outcomes.

Currently, histological examination of a liver biopsy specimen is the reference standard for the diagnosis, staging, and monitoring of liver fibrosis [2]. However, it is invasive and can lead to fatal bleeding [2].

Transient elastography is a non-invasive method of quantifying fibrosis developed as an alternative to liver biopsy. Ultrasound elastography analyses ultrasound frequency waves which are related to the elasticity (deforming capacity) of the liver. It is simple, highly reproducible, and can be completed in 10 min in an outpatient setting with no significant expertise [3]. Magnetic resonance elastography involves measuring the elasticity of liver tissues using complex algorithms [4]. An increasing number of studies have evaluated the accuracy of elastography in the staging of fibrosis and compared it to liver biopsy.

In the present study, we used meta-analysis to assess the performance of elastography in the diagnosis of severity of liver fibrosis using liver biopsy as the reference standard.

Methods

Criteria for the selection of studies

We included full papers and abstracts, without language restrictions that (1) evaluated elastography in the diagnosis of severity of liver fibrosis or monitoring thereof, using liver biopsy as the reference standard, and (2) reported on data necessary to calculate the true positive false positive, true negative and false negative diagnostic results of elastography for the diagnosis of a fibrosis stage based on a defined cut-off point for liver stiffness. If such data were unavailable, the corresponding author was contacted via e-mail to provide them; if he/she failed to reply, the study was excluded. We also excluded studies reporting on <10 patients, or in which the maximum time interval between performing elastography and liver biopsy was >3 months.

Literature search

MEDLINE (Pubmed), Embase, Science Citation Index expanded, and The Cochrane Hepato-Biliary Group Controlled Trials Register, the Cochrane Central Register of Controlled Trials in The Cochrane Library were searched until May 2009. Conference abstracts from the American Association for the Study of Liver Diseases (AASLD) and the European Association for the Study of the Liver (EASL) were manually searched from 2007 to 2009. Reference lists of identified studies and reviews were also hand-searched. Search terms for each database are shown in the Appendix.

Received 11 March 2010; received in revised form 16 July 2010; accepted 20 July 2010; available online 24 September 2010

* Corresponding author. Address: The Royal Free Sheila Sherlock Liver Centre and Division of Surgery, Royal Free Hospital, Pond Street, Hampstead, London NW3 2QG, UK. Tel.: +44 2074726229; fax: +44 2074726226.
E-mail address: Andrew.Burroughs@royalfree.nhs.uk (A.K. Burroughs).

Study selection and data extraction

The identified studies were screened independently by E.T. and S.N., and then verified reciprocally. The data were extracted independently by E.T. and S.N. using a predefined form. Any differences in study selection or data extraction were resolved by K.S.G. and A.K.B. The following data were extracted: country of study's origin, year of publication, patient number, patients' epidemiological and laboratory characteristics, aetiology of liver disease, technical failures in undertaking liver biopsy or elastography, liver stiffness cut-offs for stages of fibrosis, histological score used, true positive, false positive, true negative, and false negative elastography results and methodological quality.

Methodological quality

The quality of the studies was assessed independently by E.T. and S.N. using the Quality Assessment of Studies of Diagnostic Accuracy included in Systematic Review (QUADAS) assessment tool which contains 14 questions [5]. We also assessed if the index test (elastography) was performed according to the manufacturer's instructions i.e. at least ten successful measurements with a success rate of at least 60% and was thus likely to correctly classify the target condition. Liver biopsy was rated as an acceptable reference standard if the specimen was $\geqslant$15 mm long and included $\geqslant$6 portal tracts [2].

Statistical analysis and data synthesis

Data were combined using the hierarchical summary receiver operator characteristics (HSROC) method and the bivariate normal random-effects analysis of sensitivity and specificity within the METANDI module [6,7] in the STATA 10 statistical software (Statacorp LP, Texas, USA). METANDI performs meta-analysis of diagnostic test accuracy studies in which both the index test under study and the reference test (gold standard) are dichotomous. It fits a two-level mixed logistic regression model, with independent binomial distributions for the true positives and true negatives within each study, and a bivariate normal model for the logit transforms of sensitivity and specificity between studies [6]. Currently, these methods are considered more reliable than the Littenberg and Moses meta-analytical method of diagnostic accuracies [8]. The Metandiplot, which uses a HSROC plot, was used for graphical representation.

We evaluated pre-test probabilities of 25%, 50%, and 75% versus corresponding post-test probabilities following a "positive" or "negative" elastography result based on the summary sensitivity and specificity. "Positive" elastography results were defined as all results above the optimal liver stiffness threshold for a given fibrosis stage, given in each individual study, while "negative" test results were all results below the same threshold.

Exploration of heterogeneity, subgroup and sensitivity analysis

We planned to perform the following subgroup analyses: high versus low methodological quality, different stages of fibrosis (scoring systems were converted to comparable stages in METAVIR [9]), aetiological diagnoses, different stiffness cut-offs for a specific stage of fibrosis, initial diagnosis versus monitoring of fibrosis, different specific treatments such as interferon therapy which may alter the scoring of liver biopsy, different ranges of body mass index (BMI) (<18.5, 18.5–24.9, 25–29.9, and $\geqslant$30), different ranges of transaminases (normal range, between normal, and $\geqslant$3 or >3 times the upper limit of normal), and country of the study's origin.

We computed the diagnostic odds ratio (DOR) as a single indicator of elastography performance in each study and disease stage [10]. We used the test of interaction to assess if the diagnostic odds ratio (DOR) was statistically different between subgroups by computing the DOR ratio and the 95% confidence intervals of the DOR ratio between the subgroups [11].

We explored heterogeneity in each fibrosis stage by computing Higgin's I^2 and χ^2 tests for heterogeneity using the generic inverse variance method of meta-analysis of DOR [12]. We considered an I^2 value of more than 30% and a χ^2 p value of 0.10 to indicate statistically significant heterogeneity. We further explored heterogeneity by using meta-regression of logarithmically transformed DOR (ln DOR) for continuous variables that could potentially cause heterogeneity and used test of interaction for binary variables in each fibrosis stage. The number of studies in each fibrosis stage was not sufficient to perform a multiple meta-regression analysis.

We performed sensitivity analyses by excluding studies of low methodological quality and studies solely published as abstracts from the main analysis.

Publication bias

We performed a funnel plot; visual assessment was used to assess the bias, as well as the methods by Deeks et al. [13].

Results

Description of studies

We identified 1289 references. The reference flow is shown in Fig. 1. The inclusion criteria were fulfilled in 43 studies (35 full papers, eight abstracts) [3,4,14–54]. From these, three studies were excluded, as they evaluated magnetic resonance elastography [4], real time elastography [27] or used a probe especially built for children [23]. Important characteristics of the included studies are shown in Table 1. The risk of bias of these studies is detailed in Supplementary Fig. 1.

Transient elastography for the detection of fibrosis METAVIR stages 1–4

The diagnostic accuracy of transient elastography for cirrhosis (METAVIR $F = 4$) was evaluated in 30 studies: the summary sensitivity was 0.83 (95% CI 0.79–0.86) and specificity was 0.89 (95% CI 0.87–0.91). The mean optimal cut-off point of liver stiffness was 15 ± 4.1 kPa (median 14.5), however, values ranged from 9.0–26.5 kPa.

The diagnostic accuracy for $F = 3$ was evaluated in 24 studies, with summary sensitivity 0.82 (95% CI 0.78–0.86) and summary specificity 0.86 (0.82–0.89). The mean optimal cut-off was 10.2 ± 1.9 kPa (median 9.6), with a range of 7.3–15.4 kPa.

The detection of $F = 2$ was evaluated in 31 studies, with summary sensitivity 0.79 (95% CI 0.74–0.82) and specificity 0.78 (95% CI 0.72–0.83). The mean optimal cut-off was 7.3 ± 1.4 kPa (median 7.2) with a range of 4.0–10.1.

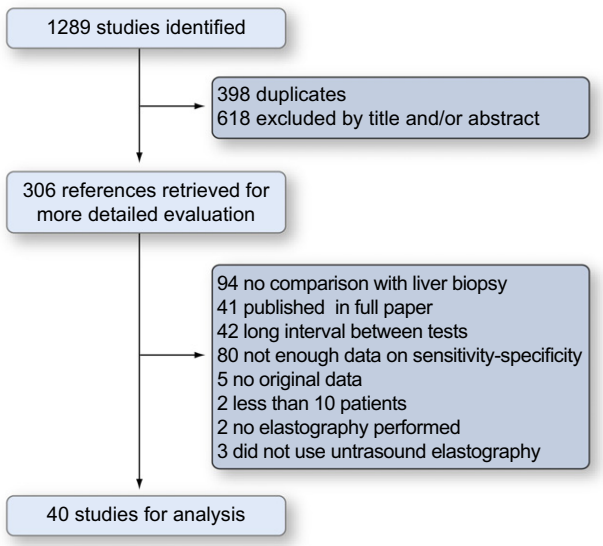

Fig. 1. Flow diagram of search results and study selection.