SUPPLEMENTARY MATERIALS

Literature Search and Selection

We conducted an electronic literature search to identify suitable studies from the Ovid-MEDLINE (U.S. National Library of Medicine) and EMBASE (Elsevier) databases until June 30, 2018. The following search terms were used: (liver OR chronic liver disease OR chronic hepatitis OR fibrosis OR cirrhosis) AND (shear wave elastography OR shear wave speed OR ARFI OR acoustic radiation force impulse) AND (failure OR unreliable OR reproducibility OR repeatability OR reliability OR intraclass correlation OR variation OR variability OR agreement OR interobserver OR intraobserver OR interreader OR intrareader). The search was limited to English articles. The bibliographies of the identified studies were manually screened for additional related studies.

We identified 805 studies through the database search (n = 800) and bibliography review (n = 5), and excluded 192 duplicates. During the screening based on titles and abstracts, 482 studies were excluded, including 184 that were not in the field of interest, 235 letters/editorials/conference abstracts, 12 articles in press, 32 reviews, and 19 articles with ineligible populations. The full text of the remaining 131 articles was reviewed, and 97 were further excluded: 71 studies not in the field of interest, 1 case report, 1 editorial, 1 article in Chinese, 2 studies including pediatric populations, 3 with potentially overlapping cohorts, and 18 with insufficient outcome data. Finally, 34 articles were included in our systematic review and meta-analysis.

Quality Assessment

The methodological quality of the selected studies was assessed by one reviewers using questionnaires and criteria provided by the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2). Since there was no standard tool of quality assessment for a meta-analysis of proportions, we tailored QUADAS-2 to characteristics of our study by omitting the following domains: reference standard and flow & timing. As a result, we assessed the quality of eligible studies regarding the risk of bias and applicability of the patient selection and index test. All studies satisfied more than half the tailored questionnaires (Supplementary Table 1). In terms of the patient selection domain, there was high risk of bias in 1 study owing to potential inappropriate exclusion. Other 9 studies showed risk of bias because of their unclear patient samplings. In terms of the index test domain, there was unclear risk of bias in 5 studies since the test threshold was not clearly pre-specified.