Accuracy of Spleen Stiffness Measurement in Detection of Esophageal Varices in Patients With Chronic Liver Disease: Systematic Review and Meta-analysis

Siddharth Singh,[*] John E. Eaton,[*] Mohammad H. Murad,[‡] Hironori Tanaka,[§] Hiroko Iijima,[§] and Jayant A. Talwalkar[*]

[*]Division of Gastroenterology and Hepatology, Department of Internal Medicine, and [‡]Knowledge and Evaluation Research Unit, Mayo Clinic, Rochester, Minnesota; and [§]Division of Hepatobiliary and Pancreatic Diseases, Department of Internal Medicine, Hyogo College of Medicine, Hyogo, Japan

BACKGROUND & AIMS: **Spleen stiffness measurement (SSM) is a promising noninvasive alternative to esophagogastroduodenoscopy (EGD) that could be used in the diagnosis of esophageal varices (EV) in patients with cirrhosis. However, its overall diagnostic accuracy in various clinical settings is unknown. We conducted a systematic review and meta-analysis of studies that compared the accuracy of SSM with that of EGD in detecting EV in patients with chronic liver disease.**

METHODS: **Through a systematic search of bibliographic databases and conference proceedings, and contact with authors, we identified 12 studies that reported the accuracy of SSM, compared with EGD, in the diagnosis of any and/or clinically significant EV in adults with chronic liver disease. In a meta-analysis, we combined measures of test performance of individual studies.**

RESULTS: **Based on pooled estimates, SSM detected the presence of any EV with 78% sensitivity (95% confidence interval [CI], 75%–81%), 76% specificity (95% CI, 72%–79%), a positive likelihood ratio (LR) of 3.4 (95% CI, 2.3–4.9), a negative LR of 0.2 (95% CI, 0.1–0.4), and a diagnostic odds ratio of 19.3 (95% CI, 7.5–49.8). In a meta-analysis of 9 studies, SSM detected the presence of clinically significant EV with 81% sensitivity (95% CI, 76%–86%), 66% specificity (95% CI, 61%–69%), a positive LR of 2.5 (95% CI, 1.7–3.9), a negative LR of 0.2 (95% CI, 0.1–0.5), and a diagnostic odds ratio of 12.6 (95% CI, 5.5–28.7). There was significant heterogeneity among studies owing to differences in elastography techniques and study locations. The included studies that were at risk for spectrum bias, review bias, and disease progression bias.**

CONCLUSIONS: **Based on a meta-analysis, current techniques for measuring spleen stiffness are limited in their accuracy of EV diagnosis; these limitations preclude widespread use in clinical practice at this time.**

Keywords: Elastography; Portal Hypertension; Accuracy; Cirrhosis.

Esophageal varices (EV) are present in 50% of patients with cirrhosis, and bleeding from EV is associated with high mortality.[1,2] Endoscopic screening for EV is recommended for all patients at the time of cirrhosis diagnosis, followed by surveillance at frequent intervals depending on the size and treatment of varices.[1] However, because the point prevalence of medium/large varices, which are at highest risk of bleeding and that benefit from prophylactic therapy with β-blockers, is only 15% to 25%, the majority of cirrhotic patients who undergo screening esophagogastroduodenoscopy (EGD) either do not have varices or have small EV that do not warrant prophylactic therapy.[1] This invasive test is potentially associated with complications related to sedation and the procedure itself, and also increased costs of medical care.[3] Hence, there is great interest in developing noninvasive techniques to detect EV.[4]

Recent studies have shown that spleen stiffness correlates with hepatic fibrosis and portal hypertension in

Abbreviations used in this paper: **ARFI, acoustic radiation force impulse imaging; AUROC, area under receiver operating curve; BMI, body mass index; CI, confidence interval; EGD, esophagogastroduodenoscopy; EV, esophageal varices; LR, likelihood ratio; OR, odds ratio; ROC, receiver operating curve; RTE, real-time tissue elastography; SSM, spleen stiffness measurement; TE, transient elastography; VTTQ, virtual touch tissue quantification.**

1542-3565/$36.00
http://dx.doi.org/10.1016/j.cgh.2013.09.013

patients with chronic liver disease.[5,6] Subsequent published studies have suggested that spleen stiffness measurement (SSM) can be used to predict the presence and size of EV in patients with chronic liver disease with high diagnostic accuracy,[5,7] although these results have not been replicated universally.[8] In addition, there is considerable variability across different techniques of measuring spleen stiffness, including transient elastography (TE) and acoustic radiation force impulse imaging (ARFI), as well as the performance of SSM across different stages and etiologies of chronic liver disease.[9]

Hence, we sought to conduct a systematic review and meta-analysis to characterize the diagnostic performance of SSM as compared with EGD as the reference standard, for predicting the presence and size of EV, in patients with chronic liver diseases.

Methods

This systematic review was conducted following guidance provided by the Cochrane handbook for systematic reviews of diagnostic test accuracy,[10] and is reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.[11] The process followed an a priori–established protocol.

Search Strategy

With the assistance of an expert librarian, we first performed a systematic search of PubMed, EMBASE, Web of Science, and Cochrane Library from database inception through March 31, 2013, for all relevant articles on the assessment of spleen stiffness for diagnosis of EV. Medical subject heading terms used in the search included "stiff*," "elast*," AND "spleen." The title and abstract of studies identified in the search were reviewed by 2 authors independently (S.S. and J.E.E.) to exclude studies that did not answer the research question of interest (details of selection criteria are described later). The full text of the remaining articles was examined to determine whether it contained relevant information. Next, the bibliographies of the selected articles, as well as review articles on the topics, were searched manually for additional articles. Third, a manual search of abstracts from major gastroenterology and hepatology conferences between 2008 and 2012 (American Association for the Study of the Liver annual meeting, European Association for the Study of the Liver annual meeting, and Digestive Diseases Week) was performed for additional studies on the topic. In case of missing or incomplete data, the primary authors' of the studies were contacted for additional information.

Selection Criteria

Studies included in this meta-analysis were observational studies that met the following inclusion criteria: (1) performed in patients with intrinsic chronic liver diseases, due to any etiology with or without evidence of portal hypertension or cirrhosis, (2) provided adequate description of SSM using either ultrasound-based or magnetic resonance–based elastography, as well as (3) assessment of EV based on upper endoscopy (EGD) as the gold standard, and (4) provided sufficient data (either in the primary article or after contact with study authors) to allow estimation of test performance (sensitivity, specificity, prevalence of EV in the study population). Inclusion was not otherwise restricted by study size, language, or publication type. When there were multiple publications from the same cohort, data from the most recent comprehensive report were included. Discrepancies in article selection were resolved by joint re-evaluation of the article and through consensus with a senior reviewer (J.A.T.).

Data Abstraction and Quality Assessment

The following data from each study were abstracted: (1) study characteristics: primary author; time period of study/year of publication; and country of study; (2) patient characteristics: age, sex, body mass index (BMI), underlying etiology of the chronic liver disease (viral vs nonviral), stage of liver disease (noncirrhotic, compensated cirrhosis, decompensated cirrhosis), and Child–Pugh class; (3) spleen stiffness assessment: technique (TE, ARFI, magnetic resonance elastography, real-time tissue elastography [RTE], virtual touch tissue quantification [VTTQ]), diagnostic threshold (or cut-off) corresponding to maximum sensitivity and specificity values from the receiver operator curve (ROC); (4) outcomes reported: presence or absence of EV, assessment and definition of clinically significant EV; (5) test performance of SSM: sensitivity, specificity, prevalence of outcome of interest in study (to impute numbers of true-positive, true-negative, false-positive, and false-negative results), and area under ROC (AUROC).

The quality assessment of included studies was performed by 2 investigators independently (S.S. and J.E.E.) using the quality assessment of diagnostic accuracy studies (QUADAS) questionnaire, which was designed to assess the internal and external validity of diagnostic accuracy studies included in systematic reviews.[12] This tool is a 14-item instrument that allows for the identification of important design elements in diagnostic accuracy studies such as patient spectrum, the presence or absence of observer blinding and verification bias, handling of indeterminate results, and reporting of patient loss to follow-up evaluation. Each item was scored as "yes" if reported (1 point) or as "no" if not reported, or as "unclear" if there is no adequate information in the article to make an accurate judgment (0 points). A score of 10 or higher was considered suggestive of a high-quality study.