Shear wave elastography imaging for detecting malignant lesions of the liver: a systematic review and pooled meta-analysis

Yang Jiao\textsuperscript{1*}, Fajin Dong\textsuperscript{1*}, Hui Wang\textsuperscript{2}, Lei Zhang\textsuperscript{1}, Jinfeng Xu\textsuperscript{1}, Jing Zheng\textsuperscript{1}, Haibo Fan\textsuperscript{2}, Hanjing Gan\textsuperscript{1}, Lixin Chen\textsuperscript{1}, Min Li\textsuperscript{3}

\* the authors shared the first authorship
\textsuperscript{1}Department of Ultrasonography, Second Clinical College of Jinan University, Shenzhen People’s Hospital, Shenzhen, \textsuperscript{2}Department of Infectious Disease, People’s Hospital of Rizhao, Rizhao, \textsuperscript{3}Department of Personnel, Second Clinical College of Jinan University, Shenzhen People’s Hospital, Shenzhen, China

Abstract

Aim: To investigate the clinical utility of shear wave elastography (SWE) imaging in the identification of malignant and benign lesions of the liver lesions by conducting a meta-analysis. Material and methods: The Cochrane library, Embase and Pubmed were searched for relevant studies with publication data through February 2016. Studies evaluating the diagnostic accuracy of SWE in the identification of malignant and benign lesions of the liver using SWE technology were selected. The cytology, histology or clinical imaging was used as the reference standard. The pooled sensitivity, specificity, diagnostic odds ratio, likelihood ratio, and the area under hierarchical summary receiver operating characteristic curve (HSROC) were used to examine the diagnostic accuracy. Results: A total of 9 cohort studies involving 1046 liver lesions (malignant 679) from 968 patients were identified. All of the 9 studies were prospective studies. The pooled sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, and diagnostic odds ratio of SWE in differentiating malignant and benign liver lesions were 82.2\% (95\% CI: 73.4–88.5), 80.2\% (95\% CI: 73.3–85.7), 4.159 (95\% CI: 2.899–5.966), 0.222 (95\% CI: 0.140–0.352), and 18.749 (95\% CI: 8.746–40.195), respectively. The area under the HSROC curve was 87\% (95\% CI: 84–90). Conclusions: This meta-analysis indicates that SWE is useful in evaluating the stiffness of liver lesions and in differentiating between malignant and benign lesions. Due to the high sensitivity, specificity, and diagnostic odds ratio, SWE can be considered as a useful complement to conventional ultrasonography.

Keywords: Elastography, diagnostic accuracy, shear wave elastography, liver cancer, meta-analysis

Introduction

Liver cancer is the 6\textsuperscript{th} most frequently diagnosed cancer and the 3\textsuperscript{rd} most common cause of cancer-related death worldwide. The number of new cases is estimated to be 564,000 per year (398,000 in men), and the increase will likely continue for some decades [1]. This trend is a result of many cancer risks including infection with the hepatitis virus (B and C), exposure to aflatoxin, alcohol, and cigarette consumption [2]. A precise evaluation of malignant liver lesions at an early stage is very important for improving patient outcomes.

Focal liver lesions (FLLs) are common findings during abdominal examinations. Ultrasonography (US) is generally used as a first-line imaging tool to distinguish FLLs. Currently, non-invasive methods such as US, especially contrast-enhanced ultrasound (CEUS), magnetic resonance imaging (MRI), and computed tomography (CT) have been used to detect FLLs and have a high-level of diagnostic accuracy in assessing the morphology of these lesions [3-6]. Nonetheless, each of these technologies has its limitations.

Liver biopsy has been considered the gold standard for differentiating malignant from benign lesions. How-
ever, it is an invasive procedure that can lead to several complications. In fact, serious complications occur in 1 to 5% of patients [7]. In addition, needle biopsy samples contain only about 1/50,000 of the entire liver, and their diagnostic value is limited by sampling variability [8].

Parenchyma elasticity is related to tissue composition, which is often altered by malignant and inflammatory processes, as well as a number of other diseases. Compared with normal surrounding tissue, the malignant tissue is generally harder. Thus, measuring tissue elasticity can potentially aid in the diagnosis of different pathologic changes. Tissue stiffness can be measured by elastography US, which has been widely used to detect lesions in superficial organs, such as the thyroid gland, breast, or prostate gland [9,10].

Shear wave elastography (SWE) is a new technique that is based on shear waves and has been implemented in diagnostic ultrasound systems [11]. SWE estimates the speed of shear waves to provide a quantification estimate of tissue stiffness. The technique has the advantage of being able to image liver tissue stiffness in real time because the shear waves are generated by US impulses. Moreover, SWE imaging is guided using B-mode images with a higher frame rate. This method can provide a more accurate assessment of liver tissue stiffness due to the advantages of SWE and B-mode image guidance [12].

A series of studies have assessed the accuracy of elastography in the differential diagnosis of benign and malignant FLLs. In this study, we performed a systematic review and meta-analysis to evaluate the diagnostic accuracy of SWE in differentiating malignant from benign FLLs.

**Material and methods**

**Search strategy**

We searched the Cochrane library, Embase, and PubMed for relevant studies up to February 2016. The literature search was performed by MeSH and free words, without language restrictions. Moreover, the reference lists of the retrieved systematic and narrative reviews were also manually searched to identify additional relevant studies. The search strategies in Cochrane library, Pubmed, and Embase are as follows: (Liver Neoplasms or Carcinoma, Hepatocellular) and ((Elasticity Imaging Techniques) or (acoustic radiation force impulse imaging or Virtual touch quantification or shearwave velocity or SWV or VTQ) or (shearwave elastography or SWE) or (“supersonic shear imaging” or SSI or SuperSonic Imagine or SuperSonic) or (Aixplorer ultrasound or Aixplorer or S2000)), ((((((Elasticity Imaging Techniques) OR acoustic radiation force impulse imaging) OR ARFI) OR Virtual touch quantification) OR VTQ OR Shearwave elastography) OR SWE) OR Supersonic shear imaging) OR SuperSonic Imagine) OR SuperSonic OR Aixplorer ultrasound) OR Aixplorer) OR S2000)) AND Liver Neoplasms, ‘elastography’/exp OR elastography OR (acoustic AND radiation AND force AND impulse AND imaging) OR (virtual AND touch AND quantification) OR swv OR (shearwave AND elastography) OR swe OR (supersonic AND shear AND imaging) OR ssi OR (supersonic AND imagine) OR supersonic OR (aixplorer AND ultrasound) OR s2000 AND (liver AND cancer OR ‘liver cell carcinoma’), respectively.

**Study selection**

The inclusion criteria were: prospective or retrospective cohort design; study population of at least 20 patients; evaluation of the malignant and benign FLLs by SWE; cytology, histology or clinical imaging used as the reference standard; providing the true-positive (TP), true-negative (TN), false-positive (FP), and false-negative (FN) at the per nodule level. If the studies did not provide data to directly construct 2×2 contingency tables, we calculated sensitivity, specificity, negative and positive predictive values from the reported. Studies such as reviews, meta-analyses, letters, abstracts, case reports, or editorials were excluded. Two authors (L Zhang and H Wang) with a similar level of experience and expertise independently selected the eligible studies. Discrepancies between two authors were resolved by a third author (JF Xu), who rechecked the search results and the assessment process.

**Data extraction and quality assessment**

The following information was independently extracted by two authors (L Zhang and H Wang) using a standardized form: the first author’s surname, year of publication, origin of the study, study design, number of patients, number of nodules available for analysis, mean diameter of the nodules, inclusion criteria for nodule size, diagnostic criteria, reference standard, TP, FN, FP, TN, sensitivity and specificity.

The methodological quality of eligible studies was assessed by the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool [13]. The QUADAS-2 tool consists of 4 key domains and 3 applicability concerns. The key domain concerns were the patient selection, index test, reference standard, flow and timing. Applicability concerns were structured in a way similar to the risk bias sections but did not include flow and timing. For each item of the QUADAS-2 tool, if a study is judged as “low” on all domains relating to bias or applicability, it is appropriate to have an overall judgment of “low risk of bias” or “low concern regarding applicability” for that study. If a study is judged “high” or “unclear” in one or
more domains, then it may be judged “at risk of bias” or as having “concerns regarding applicability.”

**Statistical analysis**

All the analyses were performed using the midas module of Stata statistical software, version 12.0 (Stata Corp, College Station, TX, USA) and RevMan software, version 5.3.5 (The Nordic Cochrane Center, The Cochrane Collaboration, Copenhagen, Denmark). Stata was used to pool statistical indexes and draw statistical graphs, such as the pooled forest graph of positive likelihood ratio (PLR), negative likelihood ratio (NLR), and diagnostic odds ratio (DOR) with corresponding 95% confidence intervals (CI), and the area under hierarchical summary receiver operating characteristic curve (HSROC). RevMan software was used to assess the methodological quality of the eligible studies.

The inconsistency index ($I^2$) and Cochrane Q statistic were used to estimate the heterogeneity across the included studies. If the Cochrane Q statistic value was $p<0.1$ or $I^2<50\%$, we used a fixed-effect model. Otherwise, a random effects model was selected [14]. Furthermore, groups were divided into subgroups based on the heterogeneity between studies.

**Publication bias**

Deek’s funnel plot asymmetry test was used to test the potential publication bias. It was conducted by a regression of diagnostic log odds ratio (lnDOR) versus the inverse of the square root ($1/\sqrt{ESS}$) and weighted by effective sample size. A $p$-value $<0.10$ for the slope coefficient indicated significant asymmetry [15].

**Results**

**Literature search results and characteristics of the studies**

Based on the predefined search strategy, the initial literature search yielded 960 papers. The flow chart of the study selection process is detailed in Figure 1. After screening, 9 original papers (Table I) fulfilled the inclusion criteria and were ultimately included in our final data for the meta-analysis.

**Study characteristics**

All 9 papers [13,16-23] were prospective diagnostic cohort studies. A total of 1046 FLLs from 968 patients were identified, 679 (64.9%) lesions being malignant.

Methodology quality assessment using the QUADAS-2 questionnaire

According to the methodological assessment of the QUADAS-2 checklist by Revman 5.0.3, the items of patient selection, index test, and reference standard indicated good quality are detailed in figure 2. The items of flow and timing were of low quality. Three studies performed histology for final confirmation using fine needle aspiration cytology [17], targeted biopsies of lesions [21], and surgical specimens [23]. Other studies [13,16,18-20,22]

**Table I. Characteristics of the included studies.**

<table>
<thead>
<tr>
<th>Author/Years</th>
<th>Country</th>
<th>Design</th>
<th>N of Patients</th>
<th>N of lesions (N of malignant lesions)</th>
<th>Inclusion criteria – nodule size (mm)</th>
<th>Diagnosis criteria</th>
<th>Reference standard</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
<th>Se (%)</th>
<th>Sp (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cho 2010 [16]</td>
<td>Korea</td>
<td>Prospective</td>
<td>51</td>
<td>60(43)</td>
<td>NA</td>
<td>2m/s</td>
<td>HP or CI</td>
<td>26</td>
<td>3</td>
<td>17</td>
<td>14</td>
<td>61</td>
<td>81</td>
</tr>
<tr>
<td>Kapoor 2010 [17]</td>
<td>India</td>
<td>Prospective</td>
<td>42</td>
<td>42(27)</td>
<td>NA</td>
<td>2.5m/s</td>
<td>HP</td>
<td>24</td>
<td>3</td>
<td>3</td>
<td>12</td>
<td>88</td>
<td>83</td>
</tr>
<tr>
<td>Shuang 2011 [18]</td>
<td>China</td>
<td>Prospective</td>
<td>116</td>
<td>128(68)</td>
<td>≥10</td>
<td>2.22m/s</td>
<td>HP or CI</td>
<td>61</td>
<td>3</td>
<td>7</td>
<td>57</td>
<td>89.7</td>
<td>95</td>
</tr>
<tr>
<td>Yu 2011 [19]</td>
<td>Canada</td>
<td>Prospective</td>
<td>89</td>
<td>105(41)</td>
<td>≥10mm</td>
<td>1.9m/s</td>
<td>HP or CI</td>
<td>28</td>
<td>20</td>
<td>13</td>
<td>44</td>
<td>68</td>
<td>69</td>
</tr>
<tr>
<td>Kim 2013 [20]</td>
<td>Korea</td>
<td>Prospective</td>
<td>74</td>
<td>101(93)</td>
<td>≥18.2mm</td>
<td>2.73m/s</td>
<td>HP or CI</td>
<td>90</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>96.4</td>
<td>65.8</td>
</tr>
<tr>
<td>Hana 2013 [21]</td>
<td>Korea</td>
<td>Prospective</td>
<td>47</td>
<td>47(39)</td>
<td>≥20mm</td>
<td>1.8 m/s</td>
<td>HP or CI</td>
<td>28</td>
<td>2</td>
<td>11</td>
<td>6</td>
<td>71.8</td>
<td>75</td>
</tr>
<tr>
<td>Zhang 2014 [22]</td>
<td>China</td>
<td>Prospective</td>
<td>156</td>
<td>170(112)</td>
<td>≥10mm</td>
<td>2.16m/s</td>
<td>HP or CI</td>
<td>91</td>
<td>15</td>
<td>21</td>
<td>43</td>
<td>81.3</td>
<td>74.1</td>
</tr>
<tr>
<td>Guo 2015 [23]</td>
<td>China</td>
<td>Prospective</td>
<td>134</td>
<td>134(55)</td>
<td>≥10mm</td>
<td>2.13m/s</td>
<td>HP or CI</td>
<td>46</td>
<td>17</td>
<td>9</td>
<td>62</td>
<td>83.3</td>
<td>77.9</td>
</tr>
<tr>
<td>Lu 2015 [23]</td>
<td>China</td>
<td>Prospective</td>
<td>259</td>
<td>259(201)</td>
<td>NA</td>
<td>13Kpa</td>
<td>HP or CI</td>
<td>157</td>
<td>10</td>
<td>44</td>
<td>49</td>
<td>78</td>
<td>83</td>
</tr>
</tbody>
</table>

relied on histological or clinical confirmation. Guo [13] did not describe the excluded patients.

Data synthesis and analysis

The pooled sensitivity, specificity, PLR, NLR, and DOR of SWE in differentiating malignant and benign liver lesions were 82.2% (95% CI: 73.4–88.5), 80.2% (95% CI: 73.3–85.7), 4.159 (95% CI: 2.899–5.966), 0.222 (95% CI: 0.140–0.352), and 18.749 (95% CI: 8.746–40.195), respectively (fig 3). The area under the HSROC curve was 87% (95% CI: 84–90) (fig 4). The Fagan nomogram shows the impact of SWE as a screening tool for malignant FLLs when the pretest probability is 20%, 60%, 80% (fig 5).

The data were pooled in stata 12.0 using the bivariate midas model. As shown in figure 3, significant heterogeneity in pooling the sensitivity ($I^2=80.16\%$, $p<0.01$) and specificity ($I^2=56.09\%$, $p<0.02$) was detected.

Subgroup analysis

The threshold effect test showed that the Spearman correlation coefficient was 0.033, $p=0.932$, suggesting that factors other than threshold effect might lead to heterogeneity among included studies. Based on the heterogeneity and the threshold effect between studies, the subgroups were established for further analysis. The dis-
crepancy between tumor size, SWV records, the type or model of US also required analysis. The meta-regression analysis yielded \( p \)-values for tumor size, SWV records, the type or model of US of 0.61, 0.64, 0.53, 0.15, 0.81, respectively, suggesting that the tumor size (≥20 mm or <20 mm, exclude not mentioned size), SWV records (≥2.50 m/s or <2.50 m/s), SWV records (≥2.00 m/s or <2.00 m/s), type or model of US (SWV or SWE) did not influence the heterogeneity.

**Publication bias**

The Deeks’ funnel plot for testing publication bias showed that the studies were distributed symmetrically with \( p=0.848 \), indicating no clear evidence of publication bias (fig 6). Regression analysis of \( \ln \text{DOR} \) against \( 1/\text{effective sample size (ESS)}^{1/2} \) showed no obvious small-study bias in this meta-analysis.

**Discussions**

In 1991, Ophir [24] developed a method for quantitative imaging of strain and elastic modulus distributions in soft tissues and first used the term “elastography”. This technique has currently been developed for use with potential clinical imaging modalities and is widely used
in ultrasound (US) diagnosis of diseases of superficial organs [25-28]. These studies have demonstrated that this modality may be useful for differentiating malignant from benign breast masses or other superficial organs. The use of elastography to diagnose liver disease has limitations because the high-frequency transducer cannot penetrate deep tissue and rib shadowing must be minimized. SWE is a new method of US elastography that provides both conventional images as well as numerical measurements of tissue elasticity. Unlike the traditional compression elastography, a short-duration (0.03–0.4 ms), high-intensity acoustic “pushing pulse” is transmitted to generate a internal tissue excitation (1–20 mm) in the region of interest (ROI) by the transducer, and is followed by a series of diagnostic intensity pulses, which are used to track the displacement of the tissue caused by the pushing pulse. In that way, it is possible to quantify the stiffness of deep tissues. Because the velocity of the shear wave depends on the tissue stiffness, SWE technology can be applied to evaluate the deep tissue stiffness.

Liver cancer has a poor prognosis, the average survival being only 6–9 months [29]. Therefore, early diagnosis of liver cancer is essential to improve patient outcome. FLLs have commonly been found on routine US examination. However, because US is associated with a high rate of false-positive and false-negative results, its diagnostic accuracy for liver cancer is poor [30,31]. With SWE, differences in stiffness between lesions can now be detected, thus offering a new modality for the differential diagnosis of FLLs.

Based on the predefined search strategy, 9 original papers fulfilled the inclusion criteria and were included in our final data. Different reference standards and cut-off values were used in the different diagnostic test accuracy studies (DTAs). The inconsistency index ($I^2$) test of the 9 studies showed notable heterogeneity. We used the bi-variate midas model of stata 12.0 to pool statistical data. This model retains the two-dimensional characteristic of the original data and takes into consideration the negative correlation between sensitivity and specificity. Because significant heterogeneity was present, we used a meta-regression analysis to evaluate the factors that could be the source of the heterogeneity. Although the subgroups of patients and the study designs were closely examined, none were found to affect SWV imaging accuracy.

The Fagan nomogram plot analysis was used to explore the clinical utility of SWV. Our results show that with a pre-test probability at 60%, SWE had an 86% probability of correctly diagnosing FLLs following a “positive” measurement. The Deeks’ funnel plot testing result indicated no clear evidence of publication bias ($p=0.848$).

In our meta-analysis, the pooled sensitivity, specificity, PLR, NLR, and DOR were calculated using the bi-variate mixed-effects regression model developed by van Houwelingen and modified for evaluating diagnostic test data [32,33]. The area under the ROC curve (AUROC) was between 0.5 and 1.0. An AUROC $>0.8$ indicates a good test [34]. The area under the HSROC curve was 87% (95%CI: 84–90) in our study. These results support the significant value of SWE in differentiating malignant from benign FLLs.

Certain limitations in this meta-analysis should be taken into consideration. First, differences in the sections of FLLs, for example the depth of the lesions, could affect the elasticity value. Second, because of the limited number of studies, we included studies that used different ultrasound equipment; third, due to DTA itself, there was significant heterogeneity, which we believe primarily affected the cut-off value. These limitations should be addressed with future multi-center studies that involve large samples sizes.

**Conclusions**

In conclusion, our meta-analysis demonstrated that SWE is a useful technique that is valuable in differentiating malignant and benign FLLs. Combined with US, SWE improves the diagnostic accuracy of FLLs.

**Acknowledgements:** This project was supported by Commission of  Scientific and Technology of Shenzhen (No.JCYJ20150403101028181).

**Conflicts of interest:** None

**References**


