Round 1

Dear Editor,

Please find enclosed the edited manuscript in Word format (file name: 86696-Revised manuscript.docx).

Title: Diagnostic role of transient elastography in patients with autoimmune liver diseases: A systematic review and meta-analysis

Author: Hong Chen, Yue Shen, Sheng-Di Wu, Qin Zhu, Cheng-Zhao Weng, Jun Zhang, Mei-Xia Wang, and Wei Jiang

Name of Journal: *World Journal of Gastroenterology*

ESPS Manuscript NO: 86696

We appreciate your constructive and insightful comments and those from the reviewer. The key points emphasized provide new insight for us to improve our meta–analysis. Revisions in the manuscript have been highlighted in yellow and strikethrough.

The manuscript has been improved according to the suggestions of reviewers:

1 Format has been updated

2 Before our initial submission, we used language editing services provided by the biomedical editing companies you recommended, but we did not meet the publication requirement (Grade A). We have sent our revised manuscript to another professional English language editing company to polish the manuscript further so that the revised manuscript will meet the publication requirement (Grade A). We provide a new language certificate along with the manuscript.

3 Point–to–point replies to the concerns that the reviewer raised have been made:

Reviewer #1

Major comments

#1. P16, lines 14-17. “TE had excellent accuracy, with summary AUROC

values of 0.84, 0.88, and 0.90 for SF, AF, and cirrhosis, respectively, in AIH patients and 0.93, 0.93, and 0.91, respectively, in PBC patients." This description is inconsistent with the description from P7, line 22 to P8, line 2, "If the summary AUROC value was above 0.90, the method was considered to have excellent accuracy, while less than 0.80 was considered to have poor accuracy [18]."

Reply: We have added additional explanations that If the summary AUROC value was between 0.80 and 0.90, the method was considered to have moderate accuracy. The description of AIH patients is not right, so we have modified it to TE had a moderate to excellent accuracy with 0.84, 0.88 and 0.90, respectively, in AIH patients.

#2. P16, line 22 to P17, line 3. "Moreover, our results showed that TE had a higher specificity and relatively low sensitivity in the diagnosis of AILDs, implying that TE was a better noninvasive method for ruling in than for ruling out." This description is not correct. In AF of PBC, the sensitivity (0.91) is higher than the specificity (0.82) when the cutoff values are 9.6-10.7.

Reply: We have deleted this inappropriate description.

Minor comments

#1. Table 1. The number of study had better correspond to the reference number of study in supplementary table 2.

#2. Table 3, 4 and 5. There should be lines between AIH, PBC and PSC as in Table 2.

Reply: We have revised the order of the reference in supplementary table 2 to make its serial numbers and the number of studies in Table 1 match. The format of these tables has been updated.

Reviewer #2

Comments

1. In the supplementary Table 2, please arrange these references according to the reference No in Table 1.

Reply: This question is consistent with that of reviewer 1, which we have answered before.

2. In 3.1 Characteristics of the included studies and patients, please give the references for those reports included AIH and PBC, or PSC and PBC.

Reply: We have added these three references.

3. In Table 2, the modalities and cut-off values are confusing in related to SF, AF and cirrhosis.

Reply: Because these cutoff values have only been defined in a single population using ROC curves to maximize sensitivity and specificity and not applied to a validation cohort. So the cutoff values vary form different studies. Several studies set the cutoff value of significant fibrosis in autoimmune hepatitis at 5.8 KPa[1], while another study set it at 10.05 Kpa[2]. Therefore, we pooled the data with similar cutoff value to form an interval. The same modalities and cut-off values can be seen in other study[3].

References:

[1] **Hartl J**, Denzer U, Ehlken H, Zenouzi R, Peiseler M, Sebode M, Hübener S, Pannicke N, Weiler-Normann C, Quaas A, Lohse AW, Schramm C. Transient elastography in autoimmune hepatitis: Timing determines the impact of inflammation and fibrosis. *J Hepatol.* 2016; **65**: 769-775. [PMID: 27238753 DOI: 10.1016/j.jhep.2016.05.023]

[2] **Anastasiou EO**, Buchter M, Baba AH, Korth J, Canbay A, Gerken G, Kahraman A. Performance and Utility of Transient Elastography and Non-Invasive Markers of Liver Fibrosis in Patients with Autoimmune Hepatitis: A Single Centre Experience. *Hepat Mon.* 2016; **16**: e40737. [PMID: 28070199 DOI: 10.5812/hepatmon.40737]

[3] **Xiao GQ**, Zhu SX, Xiao X, Yan L, Yang JY, Wu G. Comparison of

laboratory tests, ultrasound, or magnetic resonance elastography to detect fibrosis in patients with nonalcoholic fatty liver disease: A meta‐analysis. *Hepatology*. 2017; **66**: 1486-1501. [PMID: 28586172 DOI: 10.1002/hep.29302]

4. In Figures, the markers, authors, titles, and CI are too small to be seen.

Reply: We redrew the drawing and adjusted the font size to make it more clear to be seen.

5. In supplementary Table 4, the first 3 lines were not data of ARFI.

6. In the discussion section, the first paragraph is unnecessary. Please start with "TE had excellent accuracy,".

Reply: # 5 and #6 We have deleted.

7. Both fibrosis and inflammation may have a significant impact on liver stiffness. Autoimmune liver disease is characterized by persistent liver inflammation. This report suggests that the AUROC of TE in AILD is as good as those in hepatitis C, and better than hepatitis B. Was there any published data that may support this point?

Reply: Firstly, Thank you for your kind remarks and constructive comments. Indeed, inflammation can influence the liver stiffness value. We think this description "the AUROC of TE in AILD is as good as those in hepatitis C, and better than hepatitis B" is not correct. Although Afdhal et al.[1] demonstrate that in patients with HBV or HCV infection the AUROC values of FibroScan for diagnosis of SF and cirrhosis are 0.89 and 0.92, respectively. The AUROC of TE to diagnose different liver diseases is similar. However, it is irrational to directly compare chronic liver disease from different etiology. We do not find any published data that support this point.

References:

[1] **Afdhal NH**, Bacon BR, Patel K, Lawitz EJ, Gordon SC, Nelson DR, ,

Challies TL, Nasser I, Garg J, Wei LJ, McHutchison JG. Accuracy of fibroscan, compared with histology, in analysis of liver fibrosis in patients with hepatitis B or C: a United States multicenter study. *Clin Gastroenterol Hepatol.* 2015; **13**: 772-779. [PMID: 25528010 DOI: 10.1016/j.cgh.2014.12.014]

8. Are there any differences in the cut-off values between AIH and PBC? In addition, how about the differences in cut-off values between AILD and other diseases?

Reply: Our results show there are similar cutoff values between AIH and PBC. We have mentioned that several previous studies have demonstrated that inflammation in the liver (reflected by elevated ALT levels) and extrahepatic cholestasis (reflected by total bilirubin) may influence the stiffness value. However, our results of subgroup analysis indicate they have no significant effect on diagnostic accuracy. However, due to the limited number of studies, further investigation is needed to confirm the results. The final setting of the cutoff value is affected by some factors, such as the prevalence of fibrosis, etiology, race, and so on. Our results suggest the appropriate cutoff values for staging advanced fibrosis and cirrhosis ranged from 9.6 to 10.7 and 14.4 to 16.9 KPa for PBC patients. If the ALT level is normal, the cutoff value for advanced fibrosis in hepatitis B is 9 KPa, while the ALT level is elevated, and the cutoff value is 12 KPa[1]. So the cutoff varies in different etiology.

[1] EASL-ALEH Clinical Practice Guidelines: Non-invasive tests for evaluation of liver disease severity and prognosis. *J Hepatol.* 2015; **63**: 237-264. [PMID: 25911335 DOI: 10.1016/j.jhep.2015.04.006]

Reviewer #3

Comments

1. a very important pint in the inclusion criteria is the characteristics of enrolled patients. Were they studied at diagnosis or after starting treatment? This is a very important issue because the LS value might be affected by

inflammatory infiltrate that when is significant in liver parenchima may produce higher stiffness values.

Reply: Thank you for your kind remarks and constructive comments. Considering the influence of inflammation on diagnostic accuracy, we take into account the inclusion criteria of patients when extracting data. However, some studies did not mention whether patients had received treatment before inclusion. So we conducted a subgroup analysis of treatment conditions, which showed that diagnostic accuracy for staging liver fibrosis was comparable between pretreatment and posttreatment in patients with both PBC and AIH. But the data were limited. Further large studies are needed to validate this conclusion.

2. liver fibrosis assessment: this point also deserve a comment and should be discussed since METAVIR is originally developed for chronic viral diseases and not for autoimmune liver diseases. The authors should discuss whether in your opinion the METAVIR criteria could be considered as reliable histological assessment for AILD.

Reply: Although METAVIR, including two parts (the grade and stage), is originally developed for chronic hepatitis C[1], it is now used to stage liver fibrosis in various chronic liver diseases other than viral liver disease. Viral hepatitis (B, C and D) and autoimmune hepatitis (AIH) both belong to the chronic necroinflammatory diseases, the grade of them is considered to be the degree of inflammation and hepatocellular injury, which can gradually contribute to fibrosis[2]. This activity score is generated by combining the degree of piecemeal necrosis (PMN) and lobular necrosis (LN) in the liver specimen. PMN can be often seen in viral hepatitis and AIH. So I think the METAVIR criteria is also suitable for AIH.

Primary biliary cholangitis (PBC) and primary sclerosing cholangitis (PSC) belong to chronic cholestatic diseases. Bile ducts loss and interface hepatitis occurs as the disease progresses, gradually progressing to fibrosis. Ludwing

proposed four stages, including stage 1 (portal), stage 2 (periportal), stage 3 (septal) and stage 4 (cirrhosis), this algorithm for staging is similar to Metavir[3-4].

References:

[1] **Bedossa P**, Poynard T. An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group. *Hepatology.* 1996; **24**: 289–293. [PMID: 8690394 DOI: 10.1002/hep.510240201]

[2] **Goodman ZD**. Grading and staging systems for inflammation and fibrosis in chronic liver diseases. *J Hepatol.* 2007; **47**: 598–607. [PMID: 17692984 DOI: 10.1016/j.jhep.2007.07.006]

[3] **Ludwig J**, Dickson ER, McDonald GS. Staging of chronic nonsuppurative destructive cholangitis (syndrome of primary biliary cirrhosis). *Virchows Arch A.* 1978; **379**: 103–112. [DOI: 10.1007/BF00432479]

[4] **Ludwig J**, Barham SS, LaRusso NF, Elveback LR, Wiesner RH, McCall JT. Morphologic features of chronic hepatitis associated with primary sclerosing cholangitis and chronic ulcerative colitis. *Hepatology.* 1981;**1**:632–640. [DOI: 10.1002/hep.1840010612]

3. AIH selected studies: among the 22 selected AIH studies, there are differences in AIH diagnostic criteria that were used since the diagnostic scoring systems (1999 original revised AIH score vs 2008 Simplified score) and this could be introduce unintentional bias in enrolled population. Therefore, in my opinion, the authors should recall the two different diagnostic scoring systems and their differences as well described in a comprehensive review (Diagnosis and therapy of autoimmune hepatitis. Mini Rev Med Chem. 2009 Jun;9(7):847-60. doi: 10.2174/138955709788452676.) that highlighted that the two scoring systems are not interchangeable, and each may be useful in certain clinical situations. In particular, the original scoring system has greater value in diagnosing patients with few or atypical features of AIH, especially in patients with cryptogenic or autoantibody-negative

chronic hepatitis, while the simplified scoring system is more useful to exclude the diagnosis in patients with etiologically distinctive disease who have concurrent immune manifestations. Importantly, for the diagnostic purpose, the diagnostic accuracy of the simplified AIH score has been validated in real-life, clinical practice, as previously reported (Validation of simplified diagnostic criteria for autoimmune hepatitis in Italian patients. Hepatology. 2009 May;49(5):1782-3;).

Reply: Firstly, Thank you for your kind remarks and constructive comments. This question truly deserves to be discussed. Therefore, we re-extracted information on the diagnostic criteria of each article about AIH. Moreover, we conducted a subgroup analysis on it. Regrettably, several studies did not mention which diagnostic criteria they chose. Based on the data available, our result shows that diagnostic accuracy was comparable between IAIHG 2008 and IAIHG 1999 to diagnose AIH. But the data, after all, were limited. Further large studies are needed to compare the diagnostic accuracy between IAIHG 2008 and IAIHG 1999.

Thank you again for your kind remarks and guidance. We appreciate your patience and we hope that the revisions accompanied by the response letter will make our manuscript qualified for publication in the *World Journal of Gastroenterology*.

We look forward to hearing from you.

Sincerely yours,

Wei Jiang
Dept. of Gastroenterology and Hepatology
Zhongshan Hospital (Xiamen), Fudan University
Xiamen 361015, China
Email address: jiang.wei@zs-hospital.sh.cn

Tel: +86(592)3569-767

Round 2

Reviewer:

The PI addressed most of the questions. In question 7 of 2nd reviewer, I agree that most of the reports suggest that elastography showed similar accuracy between hepatitis B and C. However, many of them included a small number of patients with chronic hepatitis B. In a meta analysis from Friedrich-Rust M et al. JVH 2012, the AUROC of HBV was lower in HBV than other etiologies. Similar findings can be seen from Hsu TH et al. JMU2019 which also included AIH. The reason for poor performance could be due to unstable inflammation of HBV. 1. Friedrich-Rust M, Nierhoff J, Lupsor M, Sporea I, Fierbinteanu-Braticevici C, Strobel D, Takahashi H, Yoneda M, Suda T, Zeuzem S, Herrmann E. Performance of Acoustic Radiation Force Impulse imaging for the staging of liver fibrosis: a pooled meta-analysis. J Viral Hepat. 2012 Feb;19(2):e212-9. doi: 10.1111/j.1365-2893.2011.01537.x. Epub 2011 Oct 30. PMID: 22239521. 2. Hsu TH, Tsui PH, Yu WT, Huang SF, Tai J, Wan YL, Tai DI. Cutoff Values of Acoustic Radiation Force Impulse Two-Location Measurements in Different Etiologies of Liver Fibrosis. J Med Ultrasound. 2019 May 17;27(3):130-134. doi: 10.4103/JMU.JMU_7_19. PMID: 31867175; PMCID: PMC6905267.

Answer: Firstly, we appreciate your constructive and insightful comments. We think this description "the AUROC of TE in AILD is as good as those in hepatitis C, and better than hepatitis B" is deserved to be discussed. The references (Friedrich-Rust et al[1] and Hsu et al[2]) mentioned by the reviewer are about Acoustic Radiation Force Impulse (ARFI) but not about TE. However, TE and ARFI are different noninvasive imaging methods to detect fibrosis. Since TE is mainly based on one-dimensional ultrasound technology, while ARFI can acquire the grayscale images of the liver, and can dynamically display the two-dimensional acoustic images of the liver in real time. Second,

we agree that the LS value might be affected by inflammatory infiltrate since many studies proposed that hepatic inflammation has been identified as a potential confounder that may lead to false positive LS values in different liver diseases[3-6]. Hsu et al[2] found the AUROC of HBV was lower in HBV than other etiologies, but in this meta-analysis, the ALT and AST levels for patients with AILD (94.41 and 90.94 U/L) are higher than that for patients with CHB (50.17 and 60.34 U/L). Moreover, in our study, we also considered the influence of inflammation, so we conducted a subgroup analysis of treatment conditions, which showed that diagnostic accuracy for staging liver fibrosis was comparable between pretreatment and posttreatment in patients with both PBC and AIH. But the data were limited. Therefore, we think further studies with large number of AILD patients are needed to validate the influence of inflammation on diagnostic accuracy of TE.

1. Friedrich-Rust M, Nierhoff J, Lupsor M, Sporea I, Fierbinteanu-Braticevici C, Strobel D, Takahashi H, Yoneda M, Suda T, Zeuzem S, Herrmann E. Performance of Acoustic Radiation Force Impulse imaging for the staging of liver fibrosis: a pooled meta-analysis. J Viral Hepat 2012, 19:e212-e219.

2. Hsu TH, Tsui PH, Yu WT, Huang SF, Tai J, Wan YL, Tai DI. Cutoff Values of Acoustic Radiation Force Impulse Two-Location Measurements in Different Etiologies of Liver Fibrosis. J Med Ultrasound 2019, 27: 130-134.

3. Tapper EB, Cohen EB, Patel K, Bacon B, Gordon S, Lawitz E, et al. Levels of alanine aminotransferase confound use of transient elastography to diagnose fibrosis in patients with chronic hepatitis C virus infection. Clin Gastroenterol Hepatol 2012; 10: 932-937.

4. Romanque P, Stickel F, Dufour JF. Disproportionally high results of transient elastography in patients with autoimmune hepatitis. Liver Int 2008: 1177-1178.

5. Sagir A, Erhardt A, Schmitt M, Häussinger D. Transient elastography is unreliable for detection of cirrhosis in patients with acute liver damage. Hepatology 2008; 47: 592-595.

6. Arena U, Vizzutti F, Corti G, et al. Acute viral hepatitis increases liver stiffness values measured by transient elastography. Hepatology 2008; 47: 380-384.

Thank you again for your kind remarks and guidance. We appreciate your patience and we hope that the revisions accompanied by the response letter will make our manuscript qualified for publication in the World Journal of Gastroenterology. We look forward to hearing from you.

Sincerely yours,

Wei Jiang

Dept. of Gastroenterology and Hepatology Zhongshan Hospital (Xiamen), Fudan University Xiamen 361015, China

Email address: jiang.wei@zs-hospital.sh.cn

Tel: +86(592)3569-767