Dear Editors and Reviewer,

We are very grateful the thorough review given our review article "Noninvasive evaluation for nonalcoholic fatty liver disease, nonalcoholic steatohepatitis, and associated fibrosis" (Invited Review, No. 45715). According to the criticisms from reviewer, we have addressed the questions raised in comments and made corrections in the revised manuscript. Our point-by-point responses to the reviewers' comments are as follows. All of the revisions are highlighted in red.

Thanks for your consideration of our review for publication.

Sincerely yours,

Hongliang Li, M.D., Ph.D.

Department of Cardiology, Renmin Hospital of Wuhan University

Institute of Model Animal of Wuhan University

No. 115 Donghu Road, Wuchang District, Wuhan 430071, Hubei, China

Tel/Fax: 86-27-68759302

E-mail: lihl@whu.edu.cn

Responses to Reviewer #1's comments:

1. **What is not sufficiently recognized is that MRI already is the standard method for quantitative liver fat content assessment. Novel MRI applications for the study of elastography and diffusion as indications of fibrosis and cirrhosis are only mentioned very briefly, whereas the more outdated methods such as US and CT get too much focus.**

   Response: Thanks for your review and consideration. We have further discussed the MRI diagnosing NAFL on page 10-11.**The changes as follows:** Magnetic resonance imaging (MRI) determines steatosis by signal intensity differences on opposed-phase or fat saturation magnetic resonance imaging. Magnetic resonance imaging -derived proton density fat fraction (MRI-PDFF) is a robust, noninvasive MRI-based methods for assessing hepatic steatosis. It uses MRI-visible protons that combine with fat in the liver to quantify steatosis by dividing all protons in the liver. Tang et al found that MRI-PDFF was significantly associated with the histological steatosis grade according to the NASH-CRN grade ($\rho = 0.69$, $P < 0.001$), independent of age, sex, other NASH parameters and NASH diagnosis. The robust correlation was confirmed in several studies. Tang et al. also reported AUROC values of 0.99 for any grade of steatosis versus grade 0, 0.83 for grade 2 or higher versus grade 1 or lower, and 0.89 for grade 3 versus grade 2 or lower. In addition, MRI-PDFF is superior to other imaging tools for the assessment of hepatic steatosis, and its performance is not affected by obesity. MRI-PDFF is also regarded as a robust noninvasive method to monitor the treatment effect; this aspect will be described in detail below. Proton magnetic resonance spectroscopy （^{1}H-MRS） is another MR-base techniques that directly measures the chemical compositions of the liver. It is usually used in clinical studies of NAFLD representing biopsy for measurement of intrahepatocellular lipid (IHCL) through calculating PDFF. ^{1}H-MRS was reported a high correlation with biopsy in steatosis assessment and had a sensitivity of 80%, sensitivity of 80% for diagnosis of liver fat content ≥5%. ^{1}H-MRS was reported a good accuracy to

detect small amounts of liver fat. Nasr et al found ^{1}H-MRS had a specificity of 100% and sensitivity of 79% with PDFF cut-off value (3%), a specificity of 94% and sensitivity of 87% with PDFF cut-off value (2%). Although recognized as the most accurate noninvasive method to assess PDFF quantitatively, MRS is limited to its device- and operator- dependency, complexity, and potentially errors.Complex-based chemical shift imaging-based MRI(CSE- MRI) is emerging as a promising method for noninvasively quantifying PDFF, which could quantitatively assess liver fat content with a refined pulse sequence corrects for T1 bias,T2* decay, and spectral complexity of fat. It exhibits a high corrections with MRS- PDFF (r^2 = 0.985 for 1.5 T MR systems, r^2 = 0.991 for 3.0 T MR systems). MR diffusion weighted imaging (DWI) measures motion of water protons diffusing and tissue perfusing, it is regarded another promising tool for assessing liver fat content, while it exert poor performance for detecting steatosis in comparison with MRS and dual echo in phase and out of phase imaging (DEI). Therefore, more studies are needed to evaluate the performance of DWI in the future.

2. **Some language editing is needed.**

   Response: Thanks for the suggestion.We have checked the grammar and spelling problems again in our manuscript. In addition, we have sent our manuscript to "Nature Research Editing Service" for English language editing, and the certification of language editing has been uploaded in the the F6Publishing system.

3. **A paper about the latter method was published in WJG in 2010 (not cited).**

   Response: Thanks for your remind.We have cited the article as the reference 59 (Springer F et al. *World J Gastroenterol* 2010 ;16:1560-1566)

4. **Mention in the title that this is a review paper. Now it sounds as a patient study.**

Response: Thank you very much for your advice. With the increasing number of individuals with diabetes and obesity, nonalcoholic fatty liver disease (NAFLD) is becoming increasingly prevalent, affecting more than one-quarter of adults in the world. Although recognized as the gold standard, biopsy is limited by its sampling bias, poor acceptability and severe complications, such as mortality, bleeding, and pain. Therefore, This review discussed the current noninvasive methods for assessing NAFLD in adults, including steatosis, NASH and NAFLD-related fibrosis, and explores the advantages and disadvantages of measurement tools. In order to avoiding misunderstanding, we change the title of of manuscript. The new title of our review is as following: Noninvasive evaluation of nonalcoholic fatty liver disease: current evidence and practice

5. **Introduction. State somewhere what the normal liver fat content is and that steatosis is fat > 5%. Include references.**

   Response: Thanks for introducing excellent review to us. We have stated Normal hepatic fat content is commonly defined when macroscopic steatosis in liver histology is less than 5% of hepatocytes. We have added relevant references（reference 32-34）in this part.

6. **Page 5, line 1: replace "urgent" by "preferable".**

   Response: Thanks for the suggestion. We have replaced "urgent" by "preferable" on Page 6.

7. **Upgrade the sections about MRI with details and figures elating to novel techniques.**

   Response: Thank you very much for your advice. We have further discussed the MRI diagnosing NAFL on page 10-11. We also discuss the other MRI techniques assessing liver fat content , such as Complex-based chemical shift imaging-based MR and MR diffusion weighted imaging in detail. **The changes as follows:** Magnetic resonance imaging (MRI) determines steatosis by signal intensity

differences on opposed-phase or fat saturation magnetic resonance imaging. Magnetic resonance imaging -derived proton density fat fraction (MRI-PDFF) is a robust, noninvasive MRI-based methods for assessing hepatic steatosis. It uses MRI-visible protons that combine with fat in the liver to quantify steatosis by dividing all protons in the liver. Tang et al found that MRI-PDFF was significantly associated with the histological steatosis grade according to the NASH-CRN grade ($\rho = 0.69$, $P < 0.001$), independent of age, sex, other NASH parameters and NASH diagnosis. The robust correlation was confirmed in several studies. Tang et al. also reported AUROC values of 0.99 for any grade of steatosis versus grade 0, 0.83 for grade 2 or higher versus grade 1 or lower, and 0.89 for grade 3 versus grade 2 or lower. In addition, MRI-PDFF is superior to other imaging tools for the assessment of hepatic steatosis, and its performance is not affected by obesity. MRI-PDFF is also regarded as a robust noninvasive method to monitor the treatment effect; this aspect will be described in detail below. Proton magnetic resonance spectroscopy（^{1}H-MRS）is another MR-base techniques that directly measures the chemical compositions of the liver. It is usually used in clinical studies of NAFLD representing biopsy for measurement of intrahepatocellular lipid (IHCL) through calculating PDFF. ^{1}H-MRS was reported a high correlation with biopsy in steatosis assessment and had a sensitivity of 80%, sensitivity of 80% for diagnosis of liver fat content ≥5%. ^{1}H-MRS was reported a good accuracy to detect small amounts of liver fat. Nasr et al found ^{1}H-MRS had a specificity of 100% and sensitivity of 79% with PDFF cut-off value (3%), a specificity of 94% and sensitivity of 87% with PDFF cut-off value (2%). Although recognized as the most accurate noninvasive method to assess PDFF quantitatively, MRS is limited to its device- and operator-dependency, complexity, and potentially errors.Complex-based chemical shift imaging-based MRI(CSE- MRI) is emerging as a promising method for noninvasively quantifying PDFF, which could quantitatively assess liver fat content with a refined pulse sequence corrects for T1 bias,T2* decay, and spectral complexity of fat. It exhibits a high corrections with MRS- PDFF ($r^2 = 0.985$ for

1.5 T MR systems, r^2 = 0.991 for 3.0 T MR systems). MR diffusion weighted imaging (DWI) measures motion of water protons diffusing and tissue perfusing, it is regarded another promising tool for assessing liver fat content, while it exert poor performance for detecting steatosis in comparison with MRS and dual echo in phase and out of phase imaging (DEI). Therefore, more studies are needed to evaluate the performance of DWI in the future.

8. **There is a lot of uncited literature on DWI and MRE in fibrosis and cirrhosis as a function of disease stage. Please include this in order to get an more up to date review paper.**

   Response: Thanks for constructive insights. We have added the reference 153-156 in the section of MRE diagnosing NAFLD related fibrosis.

Responses to Reviewer #2's comments:

1. **As the authors mentioned in the abstract, are there any suggestions of effective algorithms for evaluating the stage of NAFLD by using non-invasive tools?**

   Response: Thanks for your remind. In order to convey the message of clinicians clearly, we added the section of "Clinical implication" at the end of the parts of "diagnosis of NAFLD", "diagnosis of NASH", and "diagnosis of NAFLD related fibrosis". We think it could provide effective information for readers about non-invasive tools for NAFLD. In addition, we provide a effective clinical algorithms for evaluating advanced fibrosis in the Figure 1.

2. **Are there any previous reports about the mortality rate of liver biopsies?**

   Response: One study uses hospital episode statistics collected by the National Health Service in England from 1998 to 2005 of elective percutaneous liver biopsies, it reported death within 7 days directly related

to liver biopsy occurred, at most, every 1 in 10,000 biopsies in patients investigated for liver disease or abnormal liver function test results[1].

Reference:

[1] West J, Card TR. Reduced mortality rates following elective percutaneous liver biopsy. Gastroenterology 2010;139(4):1230-1237.[PMID: 20547160 DOI: 10.1053/j.gastro.2010.06.015]

3. **What is the prevalence of NAFLD in Asian countries?**

Response: In a recent meta-analysis, it reported the prevalence of 27.37% for Asian countries[1]. In addition, another studies reported the pooled prevalence of 20.09% in mainland of China[2]. We have depicted the prevalence of Asian in the introduction. **The changes as follows:** In Asian, the prevalence of NAFLD has reach to 27.37%, and 20.09% in China.

Reference:

[1] Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. *Hepatology* 2016;64:73-84.[PMID: 26707365 DOI: 10.1002/hep.28431]

[2] Li Z, Xue J, Chen P, Chen L, Yan S, Liu L. Prevalence of nonalcoholic fatty liver disease in mainland of China: a meta-analysis of published studies. *J Gastroenterol Hepatol* 2014 ;29(1):42-51.[PMID: 24219010 DOI: 10.1111/jgh.12428]

4. **Is there a difference in the prevalence of NAFLD between the developed and developing countries?**

Response: We think there is a difference in the prevalence of NAFLD between the developed and developing countries. Two studies conducted in developing countries, Sudan, Nigeria and Iran reported a prevalence of NAFLD with 20%, 8.67%, 15.6% respectively [1-3]. The prevalence in those countries is significantly lower than 24.13% in America [4]. We have added these date in the

introduction. **The changes as follows:** In the United States, the prevalence of NAFLD in adults is 24.13%, and it is forecasted to be 33.5% in 2030, and NAFLD cases will reach 100.9 million in the general population.In some developing countries, such as Sudan, Nigeria and Iran, the prevalence of NAFLD is about 8.7%-20%.

Reference:

[1] Almobarak AO, Barakat S, Khalifa MH, Elhoweris MH, Elhassan TM, Ahmed MH. Non alcoholic fatty liver disease (NAFLD) in a Sudanese population: what is the prevalence and risk factors? *Arab J Gastroenterol* 2014;15:12-15.[PMID: 24630507 DOI: 10.1016/j.ajg.2014.01.008]

[2] Onyekwere CA, Ogbera AO, Balogun BO. Non-alcoholic fatty liver disease and the metabolic syndrome in an urban hospital serving an African community. *Ann Hepatol* 2011;10:119-124.[PMID:21502672]

[3] Eshraghian A, Dabbaghmanesh MH, Eshraghian H, Fattahi MR, Omrani GR. Nonalcoholic fatty liver disease in a cluster of Iranian population:thyroid status and metabolic risk factors. Arch Iran Med 2013;16:584-589. [PMID: 24093139 DOI: 0131610/AIM.007]

[4]Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. *Hepatology* 2016;64:73-84. [PMID:26707365 DOI: 10.1002/hep.28431]

5. **Abstract and Introduction: "NALFD" should be corrected to "NAFLD".**

Response: We apologize for the mistake of this abbreviation. We have corrected this mistake in our manuscript.

Responses to Reviewer #3's comments:

1. **The article is overall well comprehensive but several relevant reference are**

not included: World J Hepatol. 2015 Nov 18;7(26):2664-75. Diabetes Res Clin Pract. 2018 Oct;144:144-152.

Response: Thanks for the suggestion. We have discussed the cardiovascular risk markers and clinical endpoints of NAFLD, and emerging tools in imaging and urinary biomarkers for NAFLD. In addition, we have added these two articles in our manuscript, reference 170-171.

2. **Fig. 1 is probably overambitious and difficult to follow. Try to simplify it.**

   Response: Thanks for constructive insights. We have make some minor modifications in the Figure 1. Here, we manage the individuals with advanced fibrosis through three steps, identifying patients with suspected NAFLD, diagnosing advanced fibrosis with combining NFS or FIB-4 with fibroscan, providing different suggestions for different risks of advanced fibrosis. It may be effective in clinical practice to reduce unnecessary biopsies.

Responses to Reviewer #4's comments:

1. **The different methods are listed without any critical interpretation of the advantage or limitation of the test. Indicating the AUROC, PPV and NPV is not sufficient to convey the message of what a clinician should rely on most important, what we still lack.**

   Response: Thanks for your remind. In our manuscript, we comprehensively introduce the serum biomarkers, imaging, and biomarker panels for diagnosing NAFLD, NASH, NAFLD related fibrosis. We discuss the diagnostic performance of these tools with sensitivity, specificity, AUROC, PPV and NPV. Considering that these information could not sufficiently guide clinical practice for clinicians, we have added the section of "Clinical implication" at the end of the parts of "diagnosis of NAFLD", "diagnosis of NASH", and "diagnosis of NAFLD related fibrosis".It may be helpful for clinical practice. In addition, we discussed the

advantage or limitation of the tests for diagnosing NAFLD and NAFLD related fibrosis in Table 1-3.

2. **The inclusion of scores based on serum values with methods based on imaging makes the aim of the review too broad. It will be nice and appropriate having the Ms limited to methods base on serum-based tests.**

Response: Thanks for introducing excellent review to us. First, noninvasive methods for NAFLD are an attractive field. Taking account of limitations of biopsy, more and more studies try to investigate effective noninvasive methods, including serum biomarker, biomarker panels, and imaging. Therefore, we discuss serum biomarker, biomarker panels, and imaging tools in our review. Second, based on the fact that ultrasound is first-line methods for diagnosing NALFD and combining serum biomarkers or clinical rules with imaging tools to diagnose fibrosis could reduce unnecessary diagnostic liver biopsies, it is necessary to discuss imaging tools for diagnosing NAFLD in our manuscript. Finally, based on discussing the role of serum biomarker, biomarker panels, and imaging in diagnosing fibrosis extensively, we could come up with a clinical algorithms consisting of imaging and nonimaging biomarkers for detecting advanced fibrosis. We also rewrote the section of serum biomarker of NASH, we hope it could provide more information for serum biomarker diagnosing NASH.

The changes as follows:

Serum biomarkers

Cytokeratin-18 (CK18)

Cytokeratin-18 (CK18), an intermediate filament protein, is one of the most studied biomarkers for the diagnosis of NASH. It is cleaved during the period of cell death, containing CK-18 M30 and CK-18 M6[]. A meta-analysis of 25 studies reported that M30 and M65 had pooled AUROCs of 0.82 and 0.80, while the pooled sensitivity and specificity were 75% and 77%, and 71% and 77%,

respectively. Therefore, CK18 is commonly used with other serum biomarkers to diagnose NASH. Anty et al. found that combining metabolic syndrome, ALT and CK18 in a morbidly obese population could achieve an AUROC of 0.88 compared with CK18 alone, with an AUROC of 0.74. Grigorescu et al. reported that the triple combination of adiponectin, CK18 and IL-6 achieved an AUROC of 0.90, a specificity of 85.7%, and a sensitivity of 84.5%. However, the results should be further verified in future studies. In addition, some studies have examined the difference in the accuracy of CK18 in assessing NASH with different stages of fibrosis. Huang et al. found an AUROC of 0.93 for NASH with fibrosis stage 3-4 and 0.63-0.78 for NASH with fibrosis stage 0-2, which may indicate that CK18 can predict the disease severity in NASH patients.

Inflammatory markers

CXCL10 is a proinflammatory cytokine involved in diabetes and obesity[76]. In a previous study, CXCL10 exhibited a moderate accuracy for differentiating NASH from simple steatosis (AUROC, 0.68) and non-NASH (AUROC, 0.77). Tumor necrosis factor-α(TNF-α) and interleukin-8 (IL-8) were common inflammatory markers, which also exhibits moderate performance with sensitivity and specificity of 72% and 76%, 65% and 68%, respectively. However, when combining these two markers with pyroglutamate, the panel could achieve a sensitivity of 91%, specificity of 87%.

Adipocytokines and hormones

Fibroblast growth factor 21 (FGF21) secreted by the liver is another potential biomarker for NASH. One study reported that FGF21 had an AUROC of 0.62, and the two cutoffs of 126 and 578 pg/ml had >90% sensitivity and specificity for diagnosing NASH, but the positive predictive value (PPV) and NPV of FGF21 were moderate (0.59-0.78) and low (0.49-0.60), respectively. To improve the PPV and NPV, FGF21 was combined with CK-18, which improved the PPV to 82%

and the NPV to 74%. Adiponectin was reported decreased in NASH patients, which had an AUROC of 0.71 for diagnosing NASH. However, the AUROC could reach to 0.90 when adiponectin was combined with CK-18 M65, interleukin-8.Other adipocytokines, such as leptin, resistin may be potentially markers for diagnosing NASH, while they are needed to be further validated in more groups.

Other serum biomarkers

Serum iron is a common protein associated with oxygen radicals, which contribute to necroinflammation and fibrosis, two important parameters of NAFLD. Serum iron was high in individuals with NASH compared with those with simple steatosis. In a Japanese study, serum ferritin exhibited a moderate performance for distinguishing NASH from simple steatosis (AUROC, 0.73). Another study of 619 biopsy-proven NAFLD patients constructed a scoring system that combined serum ferritin with type IV collagen 7S and fasting insulin, which could be used to predict NASH with an AUROC of 0.78-0.85.

3. **Even more dispersive is the section dealing with the use of tests in assessing treatment or progression of the disorders.**

Response: Thank you very much for your advice. Non-invasive methods not only could be applied in diagnosing NAFLD, but also could be used to track disease processes and monitor treatment effects. Therefore, we also discuss the role of non-invasive method in tracking disease processes and monitoring treatment effects. We hope to help readers learn the role noninvasive methods for NAFLD more comprehensively. In addition, there are not enough studies to investigate the role of non-invasive method in tracking disease processes and monitoring treatment effects; we hardly discuss this part according to the order of serum biomarker, biomarker panels, and imaging tools. However, we hope this part of our manuscript enlighten people to pay more attentions to investigate the role of

non-invasive evaluation in disease processes and treatment effects, thus contributing to disease progression and development of therapy.

Responses to Reviewer #5's comments:

1. **Some grammatical errors should be corrected: a. "……gamma-glutamyl transferase γ …." b. "…degrees of hepatic steatosis43[39]." c. "elastoghy" d. "imagin" e. "steatosis73, 74[69,70]." f. "cirrhosis110, 111[105,106]." g. "30 kg/m(2)[124]."**

   Response: We apologize for these mistakes. We have corrected these grammatical errors in our manuscript, the corrected part are marked by the red highlight.

Responses to Reviewer #6's comments:

1. **This is a timely review of the validity of several non-invasive methods in NAFLD. It is concise and well-written. I have found only one typographical error stating NAFL as NFAL.**

   Response: We apologize for the mistake of this abbreviation. We have corrected this mistake in our manuscript.