Reviewer's report

Title: Non-invasive Score Identifies Ultrasound-diagnosed Non-Alcoholic Fatty Liver Disease and Predicts Mortality in United States

Version: 1 Date: 26 June 2014

Reviewer: Vincent Wong

Reviewer's report:

Cheung and colleagues reported a post-hoc analysis of the Third National Health and Nutrition Examination Survey (NHANES III). The authors have two main aims. First, they tested the accuracy of several fatty liver scores based on clinical and biochemical parameters against abdominal ultrasonography. In this part, the liver fat score (LFS) was found to be the most accurate. Second, they tested the use of LFS to predict overall and disease-specific mortality.

As non-alcoholic fatty liver disease (NAFLD) is currently the most common chronic liver disease worldwide, the study question is potentially important. The sample size and follow-up duration are also respectable. However, there are a few weaknesses that may limit the validity of the findings.

Major points:
1. Abdominal ultrasonography is an imperfect reference standard. It is not sensitive enough for mild hepatic steatosis. Therefore, the so-called mild fatty liver by abdominal ultrasonography may not be mild. In this study, only patients with moderate to severe fatty liver by abdominal ultrasonography were considered to have fatty liver, while those with mild fatty liver were classified as not having fatty liver. This might introduce bias. The authors should provide data to show that patients with mild fatty liver by abdominal ultrasonography did not have increased mortality as compared to those without fatty liver.

2. The accuracy of LFS was compared against abdominal ultrasonography. It does not necessarily mean that it is superior to the other scores in predicting clinical outcomes.

3. Although the sample size appears reasonable, it should be noted that only 17 patients died of liver causes during follow-up. Most patients died of malignancy and cardiovascular disease instead. Therefore, the role of a liver-based score is dubious. For example, previous studies have shown that a number of biomarkers independently predict cardiovascular events, but when elucidated in full most did not improve the prediction over existing models such as the Framingham Risk Score. Unless the authors can show that LFS significantly improves risk prediction, the findings have limited clinical utility.

Minor points:
1. The positive and negative predictive values of the scores should also be
reported.

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

None declared.