Prognostic non-invasive biomarkers for all-cause mortality in non-alcoholic fatty liver disease: A systematic review and meta-analysis

Nicole Cianci, Mohsan Subhani, Trevor Hill, Amardeep Khanna, Dong Zheng, Abhishek Sheth, Colin Crooks, Guruprasad P Aithal

BACKGROUND
Non-alcoholic fatty liver disease (NAFLD) represents a growing public health concern, with patients having higher risk of morbidity and mortality. It has a considerably high prevalence in the general population, estimated 20%-40% in Europe, and is asymptomatic until late in the disease course. It is therefore important to identify and validate tools that predict hard outcomes such as mortality for use in clinical practice in risk-stratifying NAFLD patients.

AIM
To evaluate available evidence on the use of non-invasive test(s) as prognostic factors for mortality in NAFLD.

METHODS
We performed electronic searches of Medline and EMBASE (Ovid) until 7th
January 2021 of studies in NAFLD populations. Prognostic markers included serum biomarkers, non-invasive scoring systems, and non-invasive imaging. The population included all spectrums of disease severity, including NAFLD and non-alcoholic steatohepatitis (NASH). Outcomes included all-cause, and cardiovascular mortality. All non-invasive tests were synthesised in a narrative systematic review. Finally, we conducted a meta-analysis of non-invasive scoring systems for predicting all-cause and cardiovascular mortality, calculating pooled hazard ratios and 95% confidence (STATA 16.1).

RESULTS

Database searches identified 2850 studies – 24 were included. 16 studies reported non-invasive scoring systems, 10 studies reported individual biomarkers, and 1 study reported imaging modalities. 4 studies on non-invasive scoring systems (6324 participants) had data available for inclusion in the meta-analysis. The non-invasive scoring system that performed best at predicting all-cause mortality was NAFLD fibrosis score (NFS) \([pHR \, 3.07 \, (1.62-5.83)]\), followed by fibrosis-4 index \([pHR \, 3.06 \, (1.54-6.07)]\), BARD \([pHR \, 2.87 \, (1.27-6.46)]\), and AST to platelet ratio index \([pHR \, 1.90 \, (1.32-2.73)]\). NFS was also prognostic of cardiovascular-related mortality \([pHR \, 3.09 \, (1.78-5.34)]\).

CONCLUSION

This study reaffirms that non-invasive scoring systems, especially NFS, are reliable prognostic markers of all-cause mortality and cardiovascular mortality in NAFLD patients. These findings can inform clinical practice in risk stratifying NAFLD patients.

Key Words: Nonalcoholic fatty liver disease; Nonalcoholic steatohepatitis; Biomarkers; Non-invasive; Prognosis; Mortality

©The Author(s) 2022. Published by Baishideng Publishing Group Inc. All rights reserved.

Core Tip: Non-alcoholic fatty liver disease (NAFLD) represents a growing public health concern, with an estimated prevalence in European general populations of 20%-40% and epidemiological projections of significant future increase in prevalence. NAFLD patients are at increased risk of morbidity and mortality, so it’s important to validate non-invasive prognostic markers for predicting mortality in NAFLD. This systematic review highlighted several non-invasive prognostic markers including biomarkers and imaging modalities. This meta-analysis showed that NAFLD fibrosis score is a useful prognostic marker for all-cause and cardiovascular mortality, which can be implemented in clinical practice to risk stratify and target high risk NAFLD patients.

Citation: Cianci N, Subhani M, Hill T, Khanna A, Zheng D, Sheth A, Crooks C, Aithal GP. Prognostic non-invasive biomarkers for all-cause mortality in non-alcoholic fatty liver disease: A systematic review and meta-analysis. World J Hepatol 2022; 14(5): 1025-1037

URL: https://www.wjgnet.com/1948-5182/full/v14/i5/1025.htm

DOI: https://dx.doi.org/10.4254/wjh.v14.i5.1025

INTRODUCTION

Non-alcoholic fatty liver disease (NAFLD) represents a growing public health concern, with an estimated prevalence in European general populations of 20%-40% in various studies[1], and epidemiological projections of significant future increase in prevalence[2] owing to its bidirectional association with other growing metabolic diseases such as type 2 diabetes mellitus (T2DM) and dyslipidaemia[3]. Epidemiological studies have shown that NAFLD patients have higher risk of morbidity, and mortality from all causes, and specifically from cardiovascular and liver-related causes[4-6]. Hence, it is important to identify and validate tools that predict hard outcomes such as mortality for use in clinical practice.

NAFLD is characterised by excessive hepatic fat accumulation, defined by the presence of steatosis in > 5% of hepatocytes. This can be ascertained histologically, by liver biopsy, or radiologically, including by magnetic resonance imaging (MRI) or proton magnetic resonance spectroscopy (1H-MRS). NAFLD encompasses a wide spectrum of liver disease ranging from non-alcoholic fatty liver (NAFL) to non-alcoholic steatohepatitis (NASH). NASH includes liver fibrosis, cirrhosis and hepatocellular carcinoma (HCC), which carries a significantly worse clinical prognosis[7]. Liver biopsy is the gold standard for histological assessment of NAFLD. Fibrosis, quantified by liver biopsy, has been validated as being an
important prognostic measure of disease-related outcomes and mortality\cite{8,9}. However, liver biopsy has well-established limitations including its risk and availability, making it a less than ideal tool for widespread and repeated use in clinical practice and clinical research.

International guidelines now recommend the use of ultrasound, non-invasive biomarkers and scoring systems as reliable and validated tools to diagnose NAFLD\cite{7}. There are no guidelines advocating use of non-invasive tests for prognostic purposes in NAFLD. Increasing the evidence available on the prognostic value of non-invasive tests for mortality in NAFLD would facilitate their inclusion in international guidelines, which would facilitate risk stratification of NAFLD patients and more intensive targeting of higher risk groups.

Several non-invasive scores for liver fibrosis that combine serum tests, clinical and demographic data have been developed and validated to stratify levels of liver fibrosis. These include the NAFLD fibrosis score (NFS), fibrosis-4 index (FIB-4), and Enhanced Liver Fibrosis score (ELF). The association between biopsy-proven liver fibrosis and more advanced disease, with poor outcomes of morbidity and mortality has been well-described. In particular, a recent systematic-review and meta-analysis of studies on patients with NAFLD or NASH found that fibrosis stage measured by liver biopsy had an unadjusted increased risk of all-cause mortality, liver-related mortality, liver transplantation, and liver-related events\cite{9}. Given that increasing degree of biopsy-proven liver fibrosis is associated with mortality, it is plausible that non-invasive measures of liver fibrosis might find the same association. Indeed, studies have shown that non-invasive scoring systems can predict important clinical outcomes such as mortality. One systematic review and meta-analysis of 5 studies of NFS and all-cause mortality has successfully validated the association\cite{10}. A second systematic review and meta-analysis has evaluated the association between NFS, and a further two scoring systems, FIB-4 and AST to platelet ratio index (APRI), and mortality. This also confirmed the association between NFS and mortality, in a meta-analysis of 5 studies, but did not find an association between APRI or FIB-4 and mortality\cite{11}. Finally, a recent systematic review and meta-analysis of non-invasive scoring systems (NFS, APRI, FIB-4, and BARD) and histological scoring systems associated with important clinical outcomes, including liver disease decompensation and mortality, once again validated the prognostic ability of NFS in relation to mortality but rejected that of the remaining scoring systems\cite{12}. These studies have not evaluated the ability of scoring systems to predict cause-specific mortality, namely cardiovascular or liver-related mortality; two of the commonest causes of death in NAFLD patients.

The pathophysiology of NAFLD is multifactorial and multisystem. In addition to genetic predisposition, there is close relation to endocrine and metabolic dysfunctions\cite{13-16}. This suggests several factors are at play in disease progression and poor outcomes, and several biomarkers not necessarily specific to liver function may be useful in outcome prediction. Overall, in clinical practice, different non-invasive markers can be combined to achieve a series of clinical uses. This includes, in primary care, identifying those individuals who have risk factors for metabolic syndrome and insulin resistance and are therefore at higher risk of having NAFLD. EASL guidelines presently recommend that individuals with insulin resistance or features of metabolic syndrome should undergo diagnostic procedures for NAFLD. The individual markers quoted are waist circumference, arterial pressure, serum triacylglycerols, fasting glucose and HDL cholesterol. Guidelines also recommend individuals with obesity, which can be defined as BMI > 30, as well as persistently raised liver enzymes should be screened for NAFLD. It is plausible such markers could also have a prognostic implication in patients with NAFLD. Some studies have successfully addressed this issue, linking the presence of T2DM and insulin resistance, as well as renal impairment, with poor outcomes, including mortality, in NAFLD patients\cite{17, 18}.

Certain genetic polymorphisms have also been implicated in susceptibility to NAFLD and, indeed, more severe disease. The clinical application of this is currently limited, though guidelines do mention genotyping may be considered in select cases and in clinical studies. An example of a potential future clinical application of genotyping in NAFLD is a clinical study in 152 children that developed a risk score based on 4 genetic polymorphisms that predicts presence of NASH\cite{19}.

EASL guidelines advocate the use of ultrasonography as a first line diagnostic test for hepatic steatosis, seeing as, despite its limited ability to detect low grade steatosis, it is reliable in identifying moderate and severe steatosis. It is preferred in clinical practice to MRI due to lower cost and better availability. Non-invasive imaging modalities including ultrasound, elastography and MRI have been shown in individual studies to play an important prognostic role for clinically significant outcomes such as mortality\cite{20,21}. To date, no systematic review has evaluated the prognostic use of these non-invasive modalities in NAFLD. The main aim of this systematic review was to evaluate available evidence on the use of any non-invasive test, including serum biomarkers, non-invasive scoring systems, and imaging modalities, in predicting all-cause mortality, and disease-specific mortality, in NAFLD. We aimed to validate one or a combination of measures that can be used as prognostic factors for mortality in NAFLD.
MATERIALS AND METHODS

The systematic review was conducted following "Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)" and was registered with the International Prospective Register of Systematic Reviews (PROSPERO) and protocol for this systematic review was published on the PROSPERO website. Registration number: CRD4202021207.

Search strategy

We searched MEDLINE via OvidSP (January 1946 to 7th January 2021) and EMBASE via OvidSP (January 1947 to 7th January 2021). A citation search of key included studies using Google Scholar and examining references was also conducted. We restricted our search to human only and English language studies.

A detailed search strategy using both indexing languages (MeSH, EmTree) and free text key words was developed in consultation with a university librarian. We manually searched the reference lists of all included papers for other relevant primary articles. Appendix 1 displays the final search strategy used for one of the database searches. A senior librarian from University of Nottingham was consulted to finalise the search strategy. A combination of the terms below was used for the final search strategy (PICOTS tool).

Population or Condition of interest: “Non-alcoholic fatty liver or non-alcoholic steatohepatitis”, “NASH or NAFLD or NAFL”, “fatty liver”, “Liver fibrosis”, “cirrhosis”, “Liver Disease”.

Prognostic Factors: “Fibroscan”, “Transient Elastography”, “magnetic resonance elastography”, “Elasticity Imaging Techniques”, “Liver Biomarkers”, “The Enhanced Liver Function test (ELF)”, “Hepascore”, “BARD score”, “NFS”, “Fibrometer NAFLD”, “FibroTest”, “FIB-4”, “APRI”, “FLI”, “HSI”, “SteatoTest”, “LAP”, “ION”, “NAFLD-LFS”, “Liver Function tests”, “Liver function”, “Liver enzymes”, “Liver test”, “blood test”, “blood marker” “serum marker”, “non-invasive” “ALT”, “AST”, “GGT”, “AST/ALT ratio”, “platelets”, “triglycerides”, “HBAlc”, “plasma glucose”, “fasting plasma glucose”, “insulin”, “cholesterol”.

Outcomes: “mortality”, “death”, “all-cause mortality”, “cardiovascular”, “liver-related”, “extrahepatic”, “malignancy”, “cancer”, “diabetes”.

For this review NAFLD was defined as “excessive hepatic fat accumulation in the liver, as characterised by the presence of steatosis in more than 5% of hepatocytes. The term NAFLD encompasses all spectrum of liver disease including non-alcoholic fatty liver (NAFL), non-alcoholic steatohepatitis (NASH), various stages of liver fibrosis, and cirrhosis”.

Inclusion and exclusion criteria

For the purposes of our study, we included all adult patients with confirmed NAFLD according to agreed international diagnostic criteria (including imaging or biopsy based) of any spectrum, including NASH[7]. We included any non-invasive biomarker used in predicting disease specific or all-cause mortality. These included but were not limited to serum biomarkers, imaging, and combined scoring systems. The study outcome was defined as disease specific and all-cause mortality. We included studies conducted in any setting i.e., primary, or secondary care. We included any observational study (retrospective, prospective, cohort, and case-control studies), and interventional study.

We excluded studies that did not diagnose NAFLD in their study population according to international guidelines; or did not exclude other causes of chronic liver disease. We excluded studies using invasive markers such as liver biopsy and studies reporting non-quantifiable markers such as presence of a comorbidity and studies reporting outcomes other than mortality, or combined outcomes that included mortality. We also excluded cross-sectional studies and systematic reviews and meta-analyses (although these were used to aid manual reference searching). Finally, we excluded studies only available as abstracts if the full paper could not be obtained.

Study selection

Initial title and abstract screening was done using Rayyan. QCRI and Microsoft Excel (2016). Article screening was done by two independent reviewers (NC and MS, and conflicts resolved by AS). Following shortlisting, three independent reviewers identified studies for inclusion by full text review (MS, AS and NC). Where the two reviewers disagreed about eligibility of studies for inclusion, disagreement was resolved through discussion. A third reviewer (GPA) was available to resolve any further disagreements but was not required. Reasons for exclusion of ineligible studies were recorded and the selection process recoded in a PRISMA flow diagram and “Characteristics of Excluded Studies” table.

Data extraction

A modified data extraction form was created using Cochrane CHARMS checklist as a guide. For each study included in the systematic review, two reviewers (NC and DZ) extracted the data using a standardised template. Extracted data was checked by a third reviewer (AK) prior to meta-analysis. Data collected included study characteristics (study design, length of follow-up, method of NAFLD
diagnosis, number of participants, comorbidities, NAFLD severity), index test features (type of biomarker, cut off thresholds, how they were determined and context of use), study outcomes, mortality data (including relative risk, hazard ratio, and/or any reported outcome measure), statistical analysis and adjustment methods.

**Risk of bias assessment**
Two reviewers independently assessed the quality of individual studies using the Quality in Prognostic Studies (QUIPS) tool, and determined risk of bias rating (high, moderate, low) for each study based on information presented in the published study.

**Data synthesis and statistical analysis**
We aimed to create a systematic narrative review, as we anticipated there would be considerable diversity in the design, index tests, and outcomes used in individual studies. We sought to include detailed tables and figures to display characteristics and findings (including numerical outcomes) of included studies, as well as bias ratings.

We assessed feasibility and appropriateness of conducting meta-analysis for the primary outcome, the prognostic effect of individual non-invasive tests for mortality in NAFLD. We considered conducting a meta-analysis of adjusted and unadjusted prognostic effect estimate of individual non-invasive tests by pooling any accepted measure of all-cause or disease-specific mortality of included studies. Meta-analysis was performed where 2 or more studies reported the same outcome measure for mortality for a given non-invasive test, having used equivalent cut-off values and statistical methods and in a similar study population. We excluded studies that had overlapping study population due to data duplication.

We performed a meta-analysis of multivariable adjusted hazard ratios and 95% confidence intervals of individual studies, calculated the pooled hazards ratio and 95% confidence interval with p value for overall effect.

The between-studies heterogeneity was measured by the Q test and Higgins's inconsistency index ($I^2$). In line with the Cochrane Handbook[23], we interpreted heterogeneity values of 0%-40% as low heterogeneity, 30%-60% as representing moderate heterogeneity, 50%-90% as representing substantial heterogeneity and 75%-100% as representing considerable heterogeneity. The p value for Cochrane’s Q statistic was also calculated to evaluate the statistical significance of the heterogeneity, considering a $P < 0.05$ as statistically significant.

Where the heterogeneity was statistically significant the results of the random effects analysis (DerSimonian-Laird method) are reported. Where heterogeneity was not a concern the results of the fixed effects analysis (inverse variance method) are reported.

Data was analysed using STATA (version 16.1).

**RESULTS**
The study selection process is summarized in a PRISMA flow diagram (Figure 1). Initial searches identified a total of 2850 records, which were narrowed down to 1725 records after exclusion of duplicates. One hundred and forty-five records were selected for full-text review. References and reasons for exclusion of full-text articles are detailed in Appendix 1. After full-text review, 24 articles were judged to meet our inclusion and exclusion criteria.

**Systematic review**
**Individual serum and imaging prognostic markers:** Overall, 10 studies reported individual biomarkers (2 studies on bilirubin, 2 studies on HbA1C, 2 studies on albumin, 1 study on Apolipoprotein A1, Haptoglobin, GGT, platelets, serum ferritin, prothrombin time, TSH, serum vitamin E and Vitamin E:Cholesterol, serum Vitamin D, and PNPLA3 genotype). One study reported an imaging-based prognostic marker. The individual characteristics, outcomes, and conclusions of these studies are summarized in Supplementary Table 1.

In summary, out of the 10 studies reviewed, 9 found a statistically significant association between a serum marker and risk of all-cause mortality in NAFLD patients. Only 3 reported a statistically significant association between a serum marker and cardiovascular mortality. Ferritin was found in 1 study to be a prognostic marker of all-cause mortality in NAFLD. Bilirubin has in 2 studies been found to be a marker for all-cause mortality. One study found HbA1C to be predictive of all-cause and cardiovascular mortality in NAFLD and non-NAFLD participants; and one further study found HbA1C was prognostic of all-cause and liver-related mortality in NAFLD participants. Vitamin D level has been investigated in 1 study in relation to its prognostic ability for mortality and was found to only be prognostic of Alzheimer’s disease mortality but not of all-cause, or other cause-specific mortality. Low TSH has been found in 1 study to be prognostic of all-cause and cardiovascular mortality in NAFLD but not in non-NAFLD population. Low platelet count, low albumin, and high GGT was prognostic for all-cause and liver-related mortality in 1 study. One study found that low ApoA1 and high haptoglobin...
were prognostic of all-cause mortality. One study found that both serum vitamin E and lipid-corrected vitamin E were negatively associated with all-cause mortality only in non-diabetic NAFLD participants, but not in pre-diabetic or diabetic NAFLD participants. One study found that the homozygous PNPLA3 I148M (rs738409) GG genotype showed an increase in all-cause mortality in the general population and NAFLD population.

Only one study of non-invasive imaging modalities met our inclusion criteria. This was a study of 2245 NAFLD participants in China and France, which found that LSM > 12 kPa (≤ 12 as reference) was prognostic for all-cause mortality, with a HR 2.85 (1.65 - 4.92).

Non-invasive scoring systems: Overall, 16 studies reported scoring systems (11 studies on NFS, 5 studies on FIB-4, 7 studies on APRI, 2 studies on BARD, 1 study on FibroTest, SteatoTest-2, NashTest-2, renal impairment, ASCVD, Hepascore). The studies reporting prognostic value for mortality of non-invasive scoring systems are summarised in Supplementary Table 2.

All 11 studies that included NFS found that it performed well in predicting all-cause mortality. Two further studies investigated cardiovascular mortality, one of which found NFS had a prognostic value, and a further study found that NFS is a prognostic marker for cerebro-cardiovascular mortality. One study investigated NFS and liver-related mortality and found that it was not a prognostic marker.

All 5 studies that included FIB-4 found that it was a prognostic marker for all-cause mortality. Two studies investigated the prognostic value of FIB-4 and liver-related mortality and one had positive results. One study found that FIB-4 was prognostic of cardiovascular mortality and one further study found that FIB-4 was prognostic of cerebro-cardiovascular mortality.

APRI was investigated for its prognostic value for all-cause mortality in 3 studies, with positive results in 2 studies. APRI was also investigated in 1 further study for liver-related mortality and 1 study for cardiovascular mortality, both of which had negative results. BARD was found to be a prognostic marker for all-cause mortality in 2 studies. FibroTest was found to be a prognostic marker for all-cause and cardiovascular mortality, but not liver-related mortality, in 1 study. NashTest-2 and SteatoTest-2 were found not to be prognostic markers for all-cause, liver-related, or cardiovascular mortality. Renal impairment (measured by eGFR or albumin-creatinine ratio) was found to be a prognostic marker for both all-cause and cardiovascular mortality in 1 study. ASCVD score was also found to be a prognostic marker for all-cause and cardiovascular mortality in 1 study. Hepascore was found to be a prognostic marker for all-cause mortality in 1 study.

**Meta-analysis**

Non-invasive scoring systems were the only prognostic markers reported in two or more studies, to enable pooling of results via meta-analysis. A total of 4 studies were included in data analysis. The study characteristics for the 4 studies are summarized in Table 1.

The 4 included studies recruited a total of 6324 NAFLD participants. Two studies included participants with NAFLD diagnosed by liver biopsy, 1 study included participants with NAFLD diagnosed by imaging presence of hepatic steatosis, and 1 study included participants with NAFLD diagnosed by USFLI score ≥ 30.

All were retrospective cohort studies. Study participants with NAFLD had a weighted mean age of 48.2 years, 36.9% had hypertension, 14.3% had diabetes, and were overweight with a mean BMI of 30.4. The 4 studies included 2 with cohorts from the United States, 1 from Sweden, and 1 multinational cohort (United States, United Kingdom, Australia, Italy, Thailand, Iceland).
Table 1 Characteristics of studies included in the meta-analysis

| Study characteristics | Le et al[18] | Kim et al[29] | Hagstrom et al[30] | Angulo et al[31] |
|-----------------------|-------------|--------------|-------------------|-----------------|
| Year                  | 2019        | 2013         | 2019              | 2013            |
| Country (setting)     | United States (community care) | United States (community care) | Sweden (secondary care) | United States, United Kingdom, Australia, Thailand, Italy, Iceland (secondary care) |
| Study design          | Retrospective cohort study | Retrospective cohort study | Retrospective cohort study | Retrospective cohort study |
| NAFLD diagnostic method | USFLI¹ | Ultrasound² | Liver biopsy³ | Liver biopsy³ |
| Number of NAFLD participants | 4680 | 4079 | 646 | 320 |
| Average length of follow-up (yr) | Not specified, however follow-up period was 1999-2016 | 14.5 (median) | 19.9 (mean) | 8.7 (median) |
| Proportion of males (%) | 56.3 | 50.9 | 62 | 43 |
| Mean age (yr)         | 52.6        | 46.2         | 50              | 52             |
| Caucasians (%)        | 74.8        | 75.8         | Not specified   | 92             |
| Mean BMI (kg/m²)      | 34.3        | 29.05        | 28              | 33             |
| Cardiovascular disease (%) | 13.3 | 7.1 | Not specified | Not specified |
| Hypertension (%)      | 52.3        | 32.4         | 30              | 47.5           |
| Diabetes (%)          | 24.4        | 9.5          | 14              | 36.2           |
| Smoking (%)           | 45          | 55.2         | Not specified   | Not specified  |
| Deaths (n)            | 683         | 778          | 214             | 22             |
| Low NFS⁴ (%)          | 32.4        | 67.5         | 76.2            | 39.1           |
| Intermediate NFS⁴ (%) | 51.7        | 28.2         | 5.1             | 37.5           |
| High NFS⁴ (%)         | 17.3        | 4.2          | 2.3             | 23.4           |

¹USFLI ≥ 30 and exclusion of liver disease of other aetiology.
²Ultrasound hepatic steatosis and exclusion of liver disease of other aetiology.
³Liver biopsy confirmed non-alcoholic fatty liver disease (NAFLD) and exclusion of liver disease of other aetiology.
⁴The same NAFLD fibrosis score (NFS) cut-offs have been used in all studies: High NFS (> 0.676), intermediate NFS (-1.455-0.676), low NFS (<-1.455).
NAFLD: Non-alcoholic fatty liver disease; BMI: Body mass index; NFS: NAFLD fibrosis score.

The scoring system used, outcomes and main conclusions for the 4 included studies are summarised in Supplementary Table 3. Risk of bias for included studies was assessed by using the QUIPS tool. Overall, studies included in the meta-analysis were low – moderate risk of bias (Supplementary Table 4).

After considering the feasibility of pooling results for analysis, we were able to conduct a meta-analysis of several non-invasive scoring systems and all-cause mortality, and a meta-analysis of NFS and cardiovascular-related mortality.

Overall, our analysis found that non-invasive scoring systems can predict all-cause mortality in patients with NAFLD, with higher scores having a higher pooled hazard ratio for mortality than intermediate scores, when compared with low scores (Table 2). The non-invasive scoring system that performed best at predicting all-cause mortality when comparing “high” with “low” scores was NFS, pHR 3.07 (1.62-5.83), followed by FIB-4, pHR 3.06 (1.54-6.07), BARD, pHR 2.87 (1.27-6.46), and finally APRI, pHR 1.90 (1.32-2.73). These pHR were all statistically significant (P < 0.05). When comparing “intermediate” and “low” scores, NFS, FIB-4, and BARD, but not APRI, were also statistically significant for prognosis of mortality, with intermediate scores showing a higher pHR for all-cause mortality, though the individual values were lower than the respective values for “high” score in each scoring system. The forest plots for NFS are shown in Figure 2 and for FIB-4 in Figure 3. The forest plots for the remaining analyses can be found in Supplementary Figures 1 and 2.

We also report that NFS was associated with cardiovascular-related mortality, with higher scores being prognostic of higher mortality risk (Figure 2). “High” NFS had a pHR of 3.09 (1.78-5.34), and “intermediate” NFS had a pHR of 2.12 (1.41-3.17).
We report heterogeneity using $F$ and Cochrane’s $Q$ statistic and found that several of the analyses reported moderate to considerable heterogeneity. However, the significance of this is uncertain due to the small number of studies included in the analyses. Due to the small number of studies involved in each analysis, we were unable to investigate sources of heterogeneity via statistical methods such as sensitivity analyses, sub-group analyses and/or meta-regression. For the same reason we did not assess publication bias via the usual methods such as funnel plot as this was uninformative with such a low number of included studies.

**DISCUSSION**

This systematic review identified a substantial number of individual observational studies reporting several non-invasive markers of prognostic value for mortality in NAFLD, including individual blood markers, imaging modalities, and non-invasive scoring systems. Only non-invasive scoring systems were examined in a sufficient number of studies to enable meta-analysis of the results. Our analysis reaffirms previous evidence; with higher scores in non-invasive scoring systems, there is a stepwise prognostic value for all-cause mortality. NFS appears to be the most reliable among the non-invasive markers, with highest pHR and greatest number of included studies and patients. Another non-invasive marker with very similar performance in predicting all-cause mortality was FIB-4. The pHR, confidence intervals, and heterogeneity levels of FIB-4 and NFS with all-cause mortality were indeed very similar. This can likely be attributed to all 4 of the individual components of the FIB-4 score (age, AST, platelets, and ALT) being part of the NFS (which in addition contains BMI, impaired fasting glucose or diabetes, etc.).
and albumin). It is encouraging to find a scoring system with fewer components seems to have a similar performance, as it may be easier to implement in clinical practice. However, our study found only 3 studies, with a total of 5045 NAFLD patients, that evaluated the prognostic performance of FIB-4. This is significantly less than the 9725 NAFLD patients included in the analysis of NFS and all-cause mortality. Further epidemiological studies are warranted to enable a head-to-head comparison of NFS and FIB-4 performance to help develop clinical guidelines on the best non-invasive scoring system to use in clinical practice.
Further, for the first time in the literature, our study reports that NFS has a prognostic value for cardiovascular-related mortality in patients with NAFLD. Although only 2 studies were included in this meta-analysis, they included a large number of participants and events, with 8759 NAFLD patients, and 1461 deaths. The main cause of death in NAFLD patients is cardiovascular disease[24], and our findings suggest it is possible to predict those NAFLD patients at higher risk of cardiovascular death such that more intensive clinical care can be provided to modify cardiovascular risk.

An important limitation of this meta-analysis is that few studies were included, leading to high heterogeneity. From the baseline study characteristics (Table 1), one can infer there were differences in study design and population which may explain this heterogeneity, namely differences in the setting and country of study, the method of NAFLD diagnosis, age, sex, and prevalence of different comorbidities. The high levels of heterogeneity ultimately limits the generalisability of the results of our meta-analysis. Our definition of NAFLD included all spectrums of disease, and, in the inclusion criteria for the population included in our study, we sought to evaluate both NAFL and NASH, however very few studies included subgroups comprising NASH. Indeed the only studies who did were studies where NAFLD was diagnosed via liver biopsy (Supplementary Table 1). This is likely to be due to the fact that, currently, international and national clinical guidelines recommend for NASH to be diagnosed histologically by liver biopsy, so studies where NAFLD was diagnosed by imaging and non-invasive scores would not be able to include NASH as a subgroup. In addition there aren’t robust, validated non-invasive markers to identify NASH independent of fibrosis. So, it is unsurprising that in our systematic review we weren’t able to identify any relationship between NASH and mortality.

Cohort studies have consistently shown association of fibrosis stage in NAFLD with overall and disease specific mortality[9,25,26]. The algorithms that we have identified include parameters such as age, BMI and type 2 diabetes which are well recognised risk factors for cardiovascular and all-cause mortality. Therefore, it is understandable that particular biomarkers are also associated with all-cause mortality.

Previous systematic reviews and meta-analyses report non-invasive scoring systems are prognostic of all-cause mortality in NAFLD[11,12]. These authors also report NFS is the best non-invasive tool for prognosis and risk-stratification of all-cause mortality in NAFLD patients. However, their analysis included studies using different NFS cut-off values, which may make the results less precise, and seem to have missed out the largest observational study of NFS and mortality to date[18]. Despite including less studies, our meta-analysis stringently scrutinised individual studies, ensuring the study population, non-invasive biomarker cut-offs, and multivariable adjustment were equivalent and therefore comparable. For instance, the study by Golabi et al[21], which was included in the Liu et al[12] analysis, reports HRs for all-cause mortality in “low” and “high” NFS groups, however they do not state what their reference group was. Further, their study population was extracted from NHANES III (1988-1994) data, which is the same population used in another of the studies included in their analysis. This would introduce bias due to data duplication.

The main limitations of our study were derived from the design and reporting of primary included studies. Several individual studies reported non-invasive markers having a prognostic use for mortality in NAFLD, however these were not replicated in different studies to enable a meta-analysis. It is well-recognised NAFLD is associated with extra-hepatic disease, and commonest causes of death include diabetes-related and extra-hepatic cancers, as well as cardiovascular disease. It is unsurprising that studies found markers including HbA1C[27], renal impairment[18] and ferritin[28] demonstrated good prognostic value for mortality in NAFLD when adjusting for other variables. Further studies aiming to better characterise prognostic markers for disease-specific mortality in NAFLD are warranted, to enable a more targeted approach for risk stratification and reduction in mortality of NAFLD patients. Future studies should also consider the prognostic role of imaging-based tests. One prospective observational study of 2245 participants found liver stiffness measurement using transient elastography had very good performance in identifying patients at predicting overall survival and liver events[20]. Transient elastography is a non-invasive, increasingly widespread test that may in future prove to be a useful complement to non-invasive biomarkers and liver biopsy in risk-stratifying NAFLD patients.

Biopsy-proven liver fibrosis has been well-described as being prognostic for mortality in NAFLD, with higher stages of fibrosis being prognostic of higher rates of mortality[9]. Our study adds to the available literature supporting NFS as a simple, non-invasive marker for biopsy-proven fibrosis that has a growing body of evidence suggesting it as a useful surrogate marker to predict important clinical outcomes. Further studies assessing whether a reduction in NFS value then translates to a reduction in mortality are crucial in establishing the use of NFS in clinical practice to improve outcomes in NAFLD patients.

CONCLUSION

In conclusion, our study reaffirms non-invasive scoring systems, especially NFS, is a reliable prognostic marker of all-cause mortality in NAFLD patients. We further report NFS can be used specifically to predict cardiovascular-related mortality, and our systematic review has highlighted several other non-invasive biomarkers having prognostic value.
invasive prognostic markers for mortality in NAFLD. These findings can be applied to clinical practice to stratify patients needing further investigation such as liver biopsy, closer follow-up such as referral to specialist liver services, and more intense treatment including addressing metabolic risk factors. With the increasing prevalence of NAFLD in the global population and general strain on healthcare systems, the ability to stratify NAFLD patients according to the risk of adverse outcomes can have a crucial role on clinical practice and help guide future research in NAFLD.

ARTICLE HIGHLIGHTS

Research background
Non-alcoholic fatty liver disease (NAFLD) represents a growing public health concern, highly prevalent in the general population, and with wide range of disease severity and prognosis.

Research motivation
Some NAFLD patients are at increased risk of morbidity and mortality, so it’s important to validate non-invasive prognostic markers for predicting mortality in these patients, to guide risk stratification and more intense clinical focus on high risk patients.

Research objectives
The aim of this systematic review and meta-analysis was to evaluate available evidence on the use of non-invasive test(s) as prognostic factors for mortality in NAFLD.

Research methods
The authors performed electronic searches of Medline and EMBASE (Ovid) until 7th January 2021 of studies in NAFLD populations. We conducted a meta-analysis of non-invasive scoring systems for predicting all-cause and cardiovascular mortality, calculating pooled hazard ratios and 95% confidence (STATA 16.1).

Research results
The authors identified multiple individual non-invasive biomarkers and imaging modality that had a prognostic value in NAFLD patients. Non-invasive scoring systems were the only marker to have been studied in a sufficient number of studies to permit meta-analysis. The non-invasive scoring system that performed best at predicting all-cause mortality was NAFLD fibrosis score (NFS) [pHR 3.07 (1.62-5.83)], followed by fibrosis-4 index (FIB-4) [pHR 3.06 (1.54-6.07)], BARD [pHR 2.87 (1.27-6.46)], and AST to platelet ratio index [pHR 1.90 (1.32-2.73)]. NFS was also prognostic of cardiovascular-related mortality [pHR 3.09 (1.78-5.34)].

Research conclusions
This study reaffirms that non-invasive scoring systems, especially NFS, are reliable prognostic markers of all-cause mortality and cardiovascular mortality in NAFLD patients. Further, we have identified multiple individual biomarkers and imaging modalities that have prognostic value.

Research perspectives
NFS and FIB-4 may be of value in clinical practice in risk-stratification of NAFLD patients with highest risk of mortality. Several other individual serum and imaging markers identified by this systematic review could be studied further to evaluate and validate their prognostic ability.

ACKNOWLEDGEMENTS

The authors are very grateful to Alison Ashmore, senior librarian at University of Nottingham, who helped finalizing the search strategy.

FOOTNOTES

Author contributions: Cianci N contributed to the literature search, article screening, data extraction, data analysis and interpretation, manuscript writing; Subhani M contributed to the literature search, article screening, data extraction, proof reading manuscript; Hill T contributed to the data analysis and interpretation, proof reading manuscript; Khanna A contributed to the article screening, data extraction, proof reading manuscript; Zheng D contributed to the data extraction, proof reading manuscript; Sheth A contributed to the article screening, data extraction, proofreading of final manuscript; Crooks C contributed to the project supervisor, study design, review of
progress, manuscript review; and Aithal G contributed to the project supervisor, study design, review of progress, manuscript writing and review.

Conflict-of-interest statement: No conflicts of interest are declared.

PRISMA 2009 Checklist statement: The authors have read the PRISMA 2009 Checklist, and the manuscript was prepared and revised according to the PRISMA 2009 Checklist.

Open-Access: This article is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution NonCommercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: https://creativecommons.org/Licenses/by-nc/4.0/

Country/Territory of origin: United Kingdom

ORCID number: Nicolle Cianci 0000-0002-7547-6359; Mohsan Subhani 0000-0001-8739-7263; Trevor Hill 0000-0002-0277-4988; Amardeep Khanna 0000-0001-7714-5409; Dong Zheng 0000-0001-8610-5003; Aashishek Sheth 0000-0002-1322-4631; Colin Crooks 0000-0002-6794-6621; Guruprasad P Aithal 0000-0003-9294-4830.

S-Editor: Ma YJ
L-Editor: A
P-Editor: Ma YJ

REFERENCES

1 Bluchier M, Leleu H, Peck-Radosavljevic M, Vallà DC, Roudot-Thoraval F. The burden of liver disease in Europe: a review of available epidemiological data. J Hepatol 2013; 58: 593-608 [PMID: 23419824 DOI: 10.1016/j.jhep.2012.12.005]

2 Shingina A, DeWitt PE, Dodge JL, Biggins SW, Gralla J, Sprague D, Bambha K. Future Trends in Demand for Liver Transplant: Birth Cohort Effects Among Patients With NASH and HCC. Transplantation 2019; 103: 140-148 [PMID: 30451739 DOI: 10.1097/TP.0000000000002497]

3 Masuoka HC, Chalasani N. Nonalcoholic fatty liver disease: an emerging threat to obese and diabetic individuals. Ann NY Acad Sci 2013; 1281: 106-122 [PMID: 23363012 DOI: 10.1111/nyas.12016]

4 Paik JM, Henry L, De Avila L, Younossi E, Racila A, Younossi ZM. Mortality Related to Nonalcoholic Fatty Liver Disease Is Increasing in the United States. Hepat Commun 2019; 3: 1459-1471 [PMID: 31701070 DOI: 10.1002/hep4.1419]

5 Söderberg C, Stål P, Askling J, Glaumann H, Lindberg G, Marmur J, Hultcrantz R. Decreased survival of subjects with elevated liver function tests during a 28-year follow-up. Hepatology 2010; 51: 595-602 [PMID: 20014114 DOI: 10.1002/hep.23314]

6 Adams LA, Lynp JF, St Sauver J, Sanderson SO, Lindor KD, Feldstein A, Angulo P. The natural history of nonalcoholic fatty liver disease: a population-based cohort study. Gastroenterology 2005; 129: 113-121 [PMID: 16012941 DOI: 10.1053/j.gastro.2005.04.014]

7 European Association for the Study of the Liver (EASL); European Association for the Study of Diabetes (EASD); European Association for the Study of Obesity (EASO). EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. J Hepatol 2016; 64: 1388-1402 [PMID: 27062661 DOI: 10.1016/j.jhep.2015.11.004]

8 Ekstedt M, Hagström H, Nasr P, Fredrikson M, Stål P, Kechagias S, Hultcrantz R. Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. Hepatology 2015; 61: 1547-1554 [PMID: 25125077 DOI: 10.1002/hep.27368]

9 Taylor RS, Taylor RJ, Bayliss S, Hagström H, Nasr P, Schattenberg JM, Ishigami M, Toyoa H, Wai-Sun Wong V, Peleg N, Shlomai A, Sebastiani G, Seko Y, Bhala N, Younossi ZM, Anstee QM, McPherson S, Newsome PN. Association Between Fibrosis Stage and Outcomes of Patients With Nonalcoholic Fatty Liver Disease: A Systematic Review and Meta-Analysis. Gastroenterology 2020; 158: 1611-1625.e12 [PMID: 32027911 DOI: 10.1053/j.gastro.2020.01.043]

10 Jaruvongvanich V, Wijarnpreecha K, Ungrasert P. The utility of NAFLD fibrosis score for prediction of mortality among patients with nonalcoholic fatty liver disease: A systematic review and meta-analysis of cohort study. Clin Res Hepatol Gastroenterol 2020; 41: 629-634 [PMID: 28716600 DOI: 10.1016/j.clinre.2017.03.010]

11 Salomone F, Micek A, Gods J. Simple Scores of Fibrosis and Mortality in Patients with NAFLD: A Systematic Review with Meta-Analysis. J Clin Med 2018; 7 [PMID: 30111756 DOI: 10.3390/jcm7080219]

12 Liu CH, Ampsurero J, Pavlides M, Wong VW, Fan JG, Bai L, Li H, Wu DB, Zhou LY, Du LY, Yang TK, Jiang W, Shi Y, Gil-Gómez A, Zhang WT, Liang JX, Romero-Gómez M, Tang H. Simple non-invasive scoring systems and histological scores in predicting mortality in patients with non-alcoholic fatty liver disease: A systematic review and meta-analysis. J Gastroenterol Hepatol 2021; 36: 1754-1768 [PMID: 33569851 DOI: 10.1111/jgh.15431]

13 Valentí L, Al-Serri A, Daly AK, Galmozzi E, Rametta R, Dongiovanni P, Nobili V, Mozzi E, Roviaro G, Vanni E, Bugianesi E, Maggioli M, Facranzani AL, Fargion S, Day CP. Homozgyosity for the patatin-like phospholipase-3/adiponutrin I-I48M polymorphism influences liver fibrosis in patients with nonalcoholic fatty liver disease. Hepatology 2010; 51: 1209-1217 [PMID: 20373368 DOI: 10.1002/hep.23622]
10.1053/j.gastro.2013.06.057
Graham S, Morley M. [The true significance of foot care]. Servir 1985; 33: 204-207 [PMID: 3931256 DOI: 10.1053/gastro.2013.06.057]

Le MH, Yoo YH, Henry L, Nguyen MH. Nonalcoholic Fatty Liver Disease and Renal Function Impairment: A Cross-Sectional Population-Based Study on Its Relationship From 1999 to 2016. Hepatol Commun 2019; 3: 1334-1346 [PMID: 31922492 DOI: 10.1002/hep.41048]

Nobili V, Donati B, Panera N, Vongskulayanan A, Alisi A, Dallapiccola B, Valenti L. A 4-polyphenol risk score predicts steatohepatitis in children with nonalcoholic fatty liver disease. J Pediatr Gastroenterol Nutr 2014; 58: 632-636 [PMID: 24345846 DOI: 10.1097/MPG.0000000000000279]

Shili-Masmoudi S, Wong GL, Hiriart JB, Liu K, Chermak F, Shu SS, Foucher J, Tse YK, Bernard PH, Yip TC, Merrerouche W, Chan HL, Wong VW, de Ledinghen V. Liver stiffness measurement predicts long-term survival and complications in non-alcoholic fatty liver disease. Liver Int 2020; 40: 581-589 [PMID: 31749300 DOI: 10.1111/liv.14301]

Golabi P, Stepanova M, Pham HT, Cable R, Rafiq N, Bush H, Gogoll T, Younossi ZM. Non-alcoholic steatofibrosis (NASF) can independently predict mortality in patients with non-alcoholic fatty liver disease (NAFLD). BMJ Open Gastroenterol 2018; 5: e000198 [PMID: 29607054 DOI: 10.1136/bmjgast-2018-000198]

Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009; 6: e1000097 [PMID: 19621072 DOI: 10.1371/journal.pmed.1000097]

Higgins JPT, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions (updated February 2021). Cochrane; 2021 [DOI: 10.1002/14651858.ed000142]

Mantovani A, Scarletti E, Mosca A, Alisi A, Byrne CD, Targher G. Complications, morbidity and mortality of nonalcoholic fatty liver disease. Metabolism 2020; 111S: 154170 [PMID: 32060558 DOI: 10.1016/j.metabol.2020.154170]

Hagström H, Nasr P, Ekstedt M, Hammar U, Stål P, Hultcrantz R, Kecharias S. Fibrosis stage but not NASH predicts mortality and time to development of severe liver disease in biopsy-proven NAFLD. J Hepatol 2017; 67: 1265-1273 [PMID: 28803953 DOI: 10.1016/j.jhep.2017.07.027]

Buzzetti E, Hall A, Ekstedt M, Manuqueria R, Guerrero Misas M, Covelli C, Leandro G, Luong T, Kecharias S, Manesis EK, Pinzani M, Dhillon AP, Younossi ZM. Non-alcoholic steatofibrosis (NASF) can independently predict mortality in patients with non-alcoholic fatty liver disease. J Hepatol 2017; 67: 1265-1273 [PMID: 28803953 DOI: 10.1016/j.jhep.2017.07.027]

Paik JM, Deshpande R, Golabi P, Younossi I, Henry L, Younossi ZM. The impact of modifiable risk factors on the long-term outcomes of non-alcoholic fatty liver disease. Aliment Pharmacol Ther 2020; 51: 291-304 [PMID: 31782543 DOI: 10.1111/apt.15580]

Hagström H, Nasr P, Bottai M, Ekstedt M, Kecharias S, Hultcrantz R, Stål P. Elevated serum ferritin is associated with increased mortality in non-alcoholic fatty liver disease after 16 years of follow-up. Liver Int 2016; 36: 1688-1695 [PMID: 27064133 DOI: 10.1111/liv.13144]

Ataman AV. [Tissue respiration characteristics of arterial and venous vessels]. Fiziol Zh 1987; 33: 84-88 [PMID: 3622816 DOI: 10.1002/hep.26156]

Hagström H, Nasr P, Ekstedt M, Stål P, Hultcrantz R, Kecharias S. Accuracy of Noninvasive Scoring Systems in Assessing Risk of Death and Liver-Related Endpoints in Patients With Nonalcoholic Fatty Liver Disease. Clin Gastroenterol Hepatol 2019; 17: 1148-1156.e4 [PMID: 30471458 DOI: 10.1016/j.cgh.2018.11.030]
