Received: 10 September 2019 | Revised: 6 June 2020 | Accepted: 30 June 2020

DOI: 10.1111/liv.14590

ORIGINAL ARTICLE

Liver INTERNATIONAL WILEY

Non-invasive risk scores do not reliably identify future cirrhosis or hepatocellular carcinoma in Type 2 diabetes: The Edinburgh type 2 diabetes study

Sheila M. Grecian[1] | Stela McLachlan[1] | Jonathan A. Fallowfield[2] | Patrick K. A. Kearns[3] | Peter C. Hayes[2] | Neil I. Guha[4] | Joanne R. Morling[1] | Stephen Glancy[5] | Rachel M. Williamson[6] | Rebecca M. Reynolds[7] | Brian M. Frier[7] | Nicola N. Zammitt[8] | Jackie F. Price[1] | Mark W. J. Strachan[6]

[1]Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, UK

[2]Centre for Inflammation Research, Queen's Medical Research Institute, University of Edinburgh, UK

[3]Centre for Clinical Brain Sciences, University of Edinburgh, UK

[4]NIHR Nottingham Biomedical Research Centre, Nottingham University Hospitals NHS Trust, University of Nottingham, Nottingham, UK

[5]FRCR, Department of Radiology, Western General Hospital, Edinburgh, UK

[6]Western General Hospital, Edinburgh, UK

[7]University/BHF Centre for Cardiovascular Science, Queen's Medical Research Institute, University of Edinburgh, UK

[8]Royal Infirmary of Edinburgh, Edinburgh, UK

Correspondence
Sheila M Grecian, c/o Type 2 Diabetes Research Study Office, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, UK.
Email: sheila.grecian@ed.ac.uk

Present address
Joanne R. Morling, Division of Epidemiology and Public Health, University of Nottingham
Rachel M. Williamson, Borders General Hospital, Melrose, UK

Abstract

Background: The incidence of cirrhosis and hepatocellular carcinoma (HCC) is increased in Type 2 diabetes, primarily secondary to non-alcoholic fatty liver disease (NAFLD). European guidelines recommend screening for NAFLD in Type 2 diabetes. American guidelines, while not advocating a screening protocol, suggest using non-invasive markers of fibrosis for risk-stratification and guiding onward referral.

Aims: To test the ability of individual fibrosis scores and the European screening algorithm to predict 11-year incident cirrhosis/HCC in an asymptomatic community cohort of older people with Type 2 diabetes.

Methods: The Edinburgh Type 2 Diabetes Study investigated men and women with Type 2 diabetes (n = 1066, aged 60–75 at baseline). Liver markers were measured at baseline and year 1; steatosis and fibrosis markers were calculated according to independently published calculations. During 11 years of follow-up, cases of cirrhosis and HCC were identified.

Results: Forty-three out of 1059 participants with no baseline cirrhosis/HCC developed incident disease. All scores were significantly associated with incident liver disease by odds ratio ($P < .05$). The ability of the risk-stratification tools to accurately identify those who developed incident cirrhosis/HCC was poor with low-positive predictive values (5-46%) and high false-negative and -positive rates (up to 60% and 77%) respectively. When fibrosis risk scores were used in conjunction with the European algorithm, they performed modestly better than when applied in isolation.

Abbreviations: APRI, AST to platelet ratio index; EASL-EASD-EASO, European Association for the Study of the Liver, the European Association for the Study of Diabetes and the European Association for the Study of Obesity; ELF, the enhanced liver fibrosis panel; ET2DS, the edinburgh type 2 diabetes study; FIB-4, fibrosis 4 index; HCC, hepatocellular carcinoma; NAFLD, non-alcoholic fatty liver disease; NFS, NALFD fibrosis score; ALT, alanine aminotransferase; AST, aspartate aminotransferase; γGT, gamma-glutamyltransferase; AASLD, The American Association for the Study of Liver Diseases; AIC, aikaike information criterion; OR, odds ratio; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; HR, hazard ratio.

Funding Information
The Edinburgh Type 2 Diabetes Study was funded by a grant from the U.K. Medical Research Council (Project Grant G0500877) and the Chief Scientist Office of Scotland (Programme Support Grant CZQ/1/38), and the liver substudy was supported by a grant from Pfizer (Unrestricted Investigator Led Grant). The study sponsor was not involved in the design of the study; the collection, analysis and interpretation of data; writing the report; or the decision to submit the report for publication.

Handling Editor: Salvatore Petta

Conclusions: In a cohort with a moderately low incidence of cirrhosis/HCC, existing risk scores did not reliably identify participants at high risk. Better prediction models for cirrhosis/HCC in people with Type 2 diabetes are required.

KEYWORDS
cirrhosis, hepatocellular carcinoma, non-alcoholic fatty liver disease, risk prediction, screening, type 2 Diabetes

1 | INTRODUCTION

People with Type 2 diabetes have a higher incidence of cirrhosis and hepatocellular carcinoma (HCC) than the general population.[1-3] The commonest cause of liver disease in Type 2 diabetes is non-alcoholic fatty liver disease (NAFLD) with estimates of prevalence from 40% to 70%.[4-7]

It would be valuable to identify those at high risk of developing cirrhosis/HCC because NAFLD (at the pre-cirrhotic stage) is potentially reversible by weight loss, and it would direct screening and early treatment for varices and HCC, while promoting intensive management of increased cardiovascular risk.[8,9]

A significant problem in creating appropriate risk assessment tools for NAFLD is that no consistent risk factors for progressive disease have been identified. Cohort studies report variable results and in meta-analyses the only consistent factor predicting progressive disease is histological identification of liver fibrosis.[10,11] However, liver biopsy is an invasive procedure, with a complication rate that is not acceptable for population screening. Several groups have developed non-invasive risk scoring models to identify those with fibrosis (including the Fibrosis 4 Index (FIB-4), the NALFD Fibrosis Score (NFS), AST:ALT ratio, the AST to Platelet Ratio Index (APRI) and the Enhanced Liver Fibrosis test (ELF)).[12-16] These scores have been validated in cohorts with NAFLD. However, subsequent studies have shown variable performance with the strength of association with incident cirrhosis, HCC, the need for liver transplantation and death varying significantly between cohorts.[17-21] Most of these studies have been small and only included people under secondary care hepatology services. In addition, when applied to specific groups, literature based cut-offs result in very variable proportions of populations being classed as 'high risk' with poor agreement between the top 5% of the distribution of risk scores.[22,23]

Consensus guidelines on the management of NAFLD, published by the European Association for the Study of the Liver, the European Association for the Study of Diabetes and the European Association for the Study of Obesity (EASL-EASD-EASO) recommend screening for NAFLD as part of routine care in Type 2 diabetes.[8] These guidelines suggest a screening algorithm that advises referral for specialist hepatology assessment if there is evidence of steatosis *and* non-invasive markers suggest medium or high risk of fibrosis; *or* if there is a raised alanine aminotransferase (ALT), aspartate aminotransferase (AST) or gamma-glutamyltransferase (γGT) (Figure 1). The American Association for the Study of Liver Diseases (AASLD), while not recommending a specific screening algorithm, states that there should be 'a high index of suspicion for NAFLD and NASH in Type 2 diabetes'.[9] The AASLD suggests the use of existing liver fibrosis risk scores or assessment methodologies (such as the FIB-4, NFS or transient elastography) to assess at-risk patients.[9]

One study of the EASL-EASD-EASO referral algorithm reported that around one third of people routinely attending a diabetes clinic would fulfil the criteria for hepatology referral; the incidence of subsequent cirrhosis and HCC in that cohort was not reported.[24] It is possible that the ability of the non-invasive tests to accurately identify incident disease may be affected by low event rates in community populations. Moreover, it has been suggested that current risk scores may be less accurate in people with Type 2 diabetes than in those without.[25] There remains significant uncertainty about the utility of these screening methods in Type 2 diabetes.

Lay summary

The incidence of end-stage liver disease is increased in people with Type 2 diabetes, primarily because of non-alcoholic fatty liver disease. Screening for liver disease in Type 2 diabetes is recommended. This study showed that, in a cohort of older people with Type 2 diabetes, current recommended screening pathways did not reliably identify those at risk of developing end-stage liver disease.

2 | AIMS

We aimed to assess the ability of individual fibrosis scores and of the EASL-EASD-EASO screening algorithm to predict 11-year incident

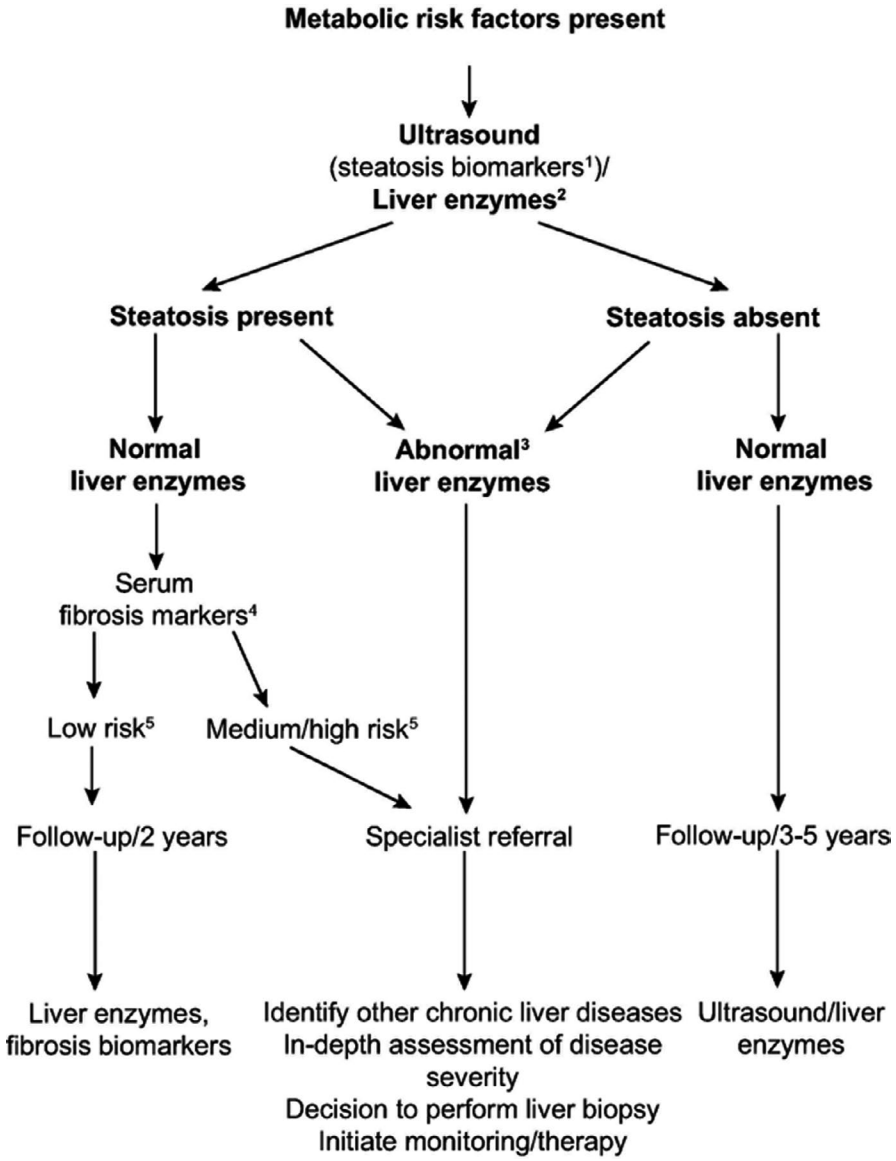

FIGURE 1 EASL-EASD-EASO algorithm (with permission of the initial publisher). Diagnostic flow-chart to assess and monitor disease severity in the presence of suspected NAFLD and metabolic risk factors. [1]Steatosis biomarkers: Fatty Liver Index, Steato Test, NAFLD Fat score (see Tables). [2]Liver tests: ALT AST, γGT. [3]Any increase in (FibroTest, FibroMeter, ELF). [5]Low risk: indicative of no/mild fibrosis; Medium/high risk: indicative of significant fibrosis or cirrhosis

cirrhosis and/or HCC in an asymptomatic community cohort of older people with Type 2 diabetes.

3 | METHODS

3.1 | The Edinburgh type 2 diabetes study

The Edinburgh Type 2 Diabetes Study (ET2DS) is a population-based prospective cohort study, designed to investigate the progression of complications in people with Type 2 diabetes. The full methods have been described previously.[26] In summary, in 2006/07 participants aged 60-74 with Type 2 diabetes were randomly selected (in age and sex bands) from the Lothian Diabetes Register (a database of almost 30 000 patients with diabetes living in Lothian, Scotland, UK, managed in both primary and secondary care). Invitations to participate were sent to 5454 people, of whom 1066 (20%) attended baseline assessment. These people have been shown to be representative of all those invited and thus of the target population.[26] All who attended the baseline clinic were invited to re-attend a clinical and liver assessment at year 1 and 4. A total of 939 attended the year 1 clinic (of the original baseline cohort, deceased n = 15, unable to contact n = 19, unable to attend n = 93) and 831 at year 4 (of the baseline cohort, deceased n = 88, unsuitable for clinical reasons n = 26, unable to contact n = 23, unable to attend n = 98). The characteristics of the cohort who attended the year 1 clinic were similar to the whole cohort at baseline.[6] All 1066 participants were followed up for outcome assessment to death (320 participants throughout the study) or end of follow-up.

3.2 | Data collection—baseline biomarker assessment

Research clinics were undertaken at the Wellcome Trust Clinical Research Facility, Western General Hospital, Edinburgh, UK.

Standardized operating procedures were used for every aspect of data collection as previously detailed.[26] ALT, AST, γGT, platelets and triglycerides were measured on fasting venous samples at the baseline research clinic and were analysed using a Vitros Fusion chemistry system (Ortho Clinical Diagnostics, Bucks, UK). The Enhanced Liver Fibrosis test (ELF) was measured on fasting venous blood samples from the year 1 clinic and was analysed using the ADVIA Centaur immunoassay system (Siemens Healthcare Diagnostics Inc., New York, USA) at the iQur laboratory (London, UK). Ultrasound was undertaken at the year 1 clinic following a 4-hour fast (Sonoline Elegra Ultrasound Imaging System (Siemens Medical Systems Inc., Washington, USA)). Ultrasounds were graded for hepatic steatosis using established criteria (0=normal liver, 1=indeterminate, 2=mild steatosis, 3=severe steatosis) and validated by three different graders and ^{1}H MRI spectroscopy in a subset, as previously described.[27] This showed a median fat fraction in those with 'severe' steatosis of 19.4% (interquartile range 12.9-27.5), compared to 4.1% (interquartile range 3.1-8.5) in those with 'indeterminate'/ 'mild' steatosis and 4.2% (interquartile range 1.2-5.7) in those with 'no steatosis'. As a result of this validation which showed significant overlap between grade 0-2 steatosis, only those with grade 3 steatosis on ultrasound assessment were deemed to have 'definite steatosis'. Individuals with an ultrasound grading of 0-2 were considered to have 'no definite steatosis'.

Participants underwent full diagnostic liver screen (including Hepatitis B and C serology, liver autoantibody titres, alpha-foeto protein, ferritin) and history to assess alcohol status, medication use and past medical history. Any participant with routine liver enzyme tests above the laboratory upper limit of normal (ALT >50 U/L, AST >45 U/L, γGT >55 U/L, alkaline phosphatase >125 U/L), AST:ALT ratio >1, hyaluronic acid >100 μ g/L (in the absence of known joint disease), positive liver autoantibodies, ferritin >1000 ng/mL, alpha-foeto protein >6 ng/mL, positive hepatitis B or C serology, spleen diameter >13 cm, platelets <150 × 10^9/L in the absence of known haematological cause, or suspected cirrhosis on ultrasound was referred for specialist hepatology review.

Steatosis and Fibrosis scores were calculated and cut-off levels used as per published literature.

- AST to platelet ratio index (APRI) was calculated as: ((AST(U/L)/Upper limit normal) /platelets(×10^9/L)) ×100. Cut-point low to medium/high risk of fibrosis >0.5.[12]
- AST: ALT ratio was calculated as: AST(U/L)/ALT(U/L). cut-point ≥0.8[16].
- Fibrosis-4 index (FIB-4) was calculated as ((age(years) ×AST(U/L))/(plt(×10^9/L)xsqrt ALT(U/L))). Cut point low-medium risk ≥1.3 and medium-high risk >2.67.[13,17,28]
- NAFLD Fibrosis Score (NFS) was calculated as: 1.675+(0.037×age(years))+(0.094×BMI(kg/m^2))+(1.13×IFG/diabetes (yes=1, no=0))+(0.99× (AST(U/L)/ALT(U/L))- (0.013xplatelet count(×10^9/L))-(0.66×albumin (g/dL)). Cut-point for low-medium risk ≥-1.455, medium-high risk >0.676.[14]
- Fatty liver index (FLI) was calculated as: $e^y/(1+e^y)$×100 where y = 0.953 × ln(triglycerides, mg/dl) + 0.139 × BMI, kg/m^2 + 0.718 × ln (γGT, U/L) + 0.053 × waist circumference, cm − 15.745).[29]

The EASL-EASD-EASO referral decision algorithm (Figure 1) was used.[8]

3.3 | Data collection—identification of liver disease

Possible *prevalent* liver disease was identified through a patient clinical history questionnaire at the baseline clinic. Possible cases were confirmed if a clinician diagnosis was recorded in primary or secondary care medical records.

Incident cirrhosis and HCC cases were identified and corroborated using multiple sources of information: retrospective review of all participants' secondary care medical notes (TrakCare, InterSystems Corp., Cambridge, USA), patient and GP questionnaires provided at year 4 and year 10 follow-up, ISD (Information Services Division, NHS Scotland) discharge summary coding of hospital admissions and death coding (data from year 0-8). Cases were confirmed if a clinician diagnosis was recorded in secondary care medical notes. Participants were identified as having 'screen-detected' cirrhosis/HCC if they were referred to hepatology as a result of year 1 or 4 investigation and remained under hepatology follow-up until definitive diagnosis was made. Prevalence and 10-year incidence data from the Year 1 cohort have previously been reported; these data include only those individuals who attended for the year 1 visit, by contrast with this study which has reported data from the entire cohort.[30]

3.4 | Data analysis

Data were analysed using R (R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.). Logistic regression was used to identify the strength of association between baseline prediction scores and incident cirrhosis/HCC in our cohort. Complete case analysis was undertaken; <5% data were missing for any variable with the exception of ELF and ultrasound measurement (calculated at year 1 attendance; (n = 681 for ELF, n = 933 for ultrasound). Aikaike Information Criterion (AIC) and C-statistic were used to assess performance of the regression models. C-statistic assesses discrimination (the ability for a model to correctly identify those in two different groups). In logistic regression it is calculated as a comparison between the odds of each individual having the outcome based on the model variables and the actual outcome achieved and examines if the model performs better than chance; a value of >0.8 considered to be good. AIC assesses overall model performance using a combination of discrimination and calibration (the ability of the model to rank increased risk appropriately); it has no scale, but lower values suggest improved performance. Because of our mixed population of screen-detected and clinician-diagnosed outcomes, possibly skewing our time-to-event data as

those who were screen-detected were often diagnosed at a pre-symptomatic stage, our primary analysis (logistic regression) does not include a time component. We additionally ran a sensitivity analysis using competing risks regression to assess whether there was a significant impact of the competing risk of non-liver death on model performance. The Bayesian Information Criterion (BIC) was used to assess model performance for the competing risks regression, with lower values suggesting improved performance. Performance was additionally assessed through calculation of sensitivity, specificity, positive predictive value, negative predictive value, false-positive and false-negative level.

3.5 | Ethics

Ethical permission for the study was granted by Lothian Medical Research Ethics Committee (REC reference 16/SS/0098). All participants gave written informed consent.

4 | RESULTS

4.1 | Subject characteristics at baseline

Participants were aged 60-74 years (mean 67.9), 51.3% male. Mean duration of Type 2 diabetes was 8 years, mean HbA1c 57 mmol/mol (7.4%) and mean BMI 31.4 kg/m^2. Alcohol intake was above recommended limits in 19.9% and 14.4% were current smokers (Table 1). Seven people had prevalent cirrhosis/HCC.

TABLE 1 Baseline characteristics of the study population

Baseline characteristic		ET2DS population (n = 1066)
Age (years)		67.9 (4.2)
Sex (male)		547 (51.3)
Scottish Index of Multiple Deprivation quintile	1 (most deprived)	12 (11.9)
	2	208 (19.5)
	3	188 (17.6)
	4	194 (18.2)
	5 (least deprived)	349 (32.7)
Duration Type 2 diabetes (years)		8.1 (6.5)
HbA1c (%)		7.4 (1.1)
HbA1c (mmol/mol)		57.0 (12.0)
BMI (kg/m^2)		31.4 (5.7)
Smoker (current)		154 (14.4)
Alcohol (excess)[a]		207 (19.9)

Note: Values are mean (SD) or n (%).

[a]Defined as females >14 units/week, males >21 units/week or patient disclosed history of a current or prior alcohol problem.

4.2 | Incident cirrhosis/HCC

Out of 1059 people without cirrhosis/HCC at baseline, 43 developed this outcome over 11 years of follow-up (11-year incidence 4.1%) (Figure 2). Twenty-three cases were 'screen-detected' from year 1 clinic, and eight from year 4 clinic. Twelve cases were diagnosed following clinical referral. Range of time to diagnosis overlapped between the 'screen-detected' group (163-2251 days) and the 'clinician-detected' group (920-3977 days). Out of the 43 people identified with cirrhosis/HCC, 37 cases were attributed to NAFLD, NAFLD with alcohol above the recommended limit as a cofactor, or mixed aetiology NAFLD and alpha-1-antitrypsin deficiency; 30 developed cirrhosis, nine both cirrhosis and HCC and four HCC. Of those with cirrhosis, 58% developed varices, ascites and/or encephalopathy. This equated to an 11-year incidence of 3.7% (3.66/ 1000 person years) (cirrhosis) and 1.2% (1.31/ 1000 person years) (HCC).

4.3 | Performance of fibrosis risk scores in predicting incident cirrhosis/HCC

Table 2 describes the association of existing fibrosis risk scores, using published cut-points, with the development of cirrhosis/HCC. All risk scores revealed a significant relationship by odds ratio (OR) with incident cirrhosis/HCC ($P < .05$). Confidence intervals for the OR were wide. The score with the highest C-statistic was APRI (cut-point >0.5), that with the lowest AIC was ELF (cut-point ≥10.51).

The ability of the risk scores to correctly identify people who developed cirrhosis/HCC was variable (sensitivity 33-93%, specificity 22-98%, PPV 5-46%) The NPV's for all scores were 97-99%, probably because the outcome (cirrhosis/HCC) was relatively rare. All except two scores had false-negative rates >20% (with some 60%). For example using FIB-4 (cut-point >2.67) or AST:ALT (cut-point ≥0.8), 24 out of 40 people who developed cirrhosis/HCC were wrongly classified as 'low risk'. For scores with false-negative rates <20%, the false-positive rates were very high (41-78%); indicating that if a score was used where a false negative was less likely, a significant proportion of the population who would not develop cirrhosis/HCC would be classified as 'high risk'. For example using NFS (cut-point ≥1.455), 806 people would have been classified as 'high risk' (of whom only 37 developed cirrhosis/HCC). Using APRI (the score with the best performing C-statistic), 19/40 people who developed incident cirrhosis/HCC would have been classified as 'low risk'; 78 people would have been classified as 'high risk' and referred, of whom 21 developed cirrhosis/HCC.

4.4 | Performance of the EASL-EASD-EASO algorithm in predicting incident cirrhosis/HCC

Table 2 describes how the EASL-EASD-EASO algorithm outcome was associated with the development of cirrhosis/HCC in our cohort.

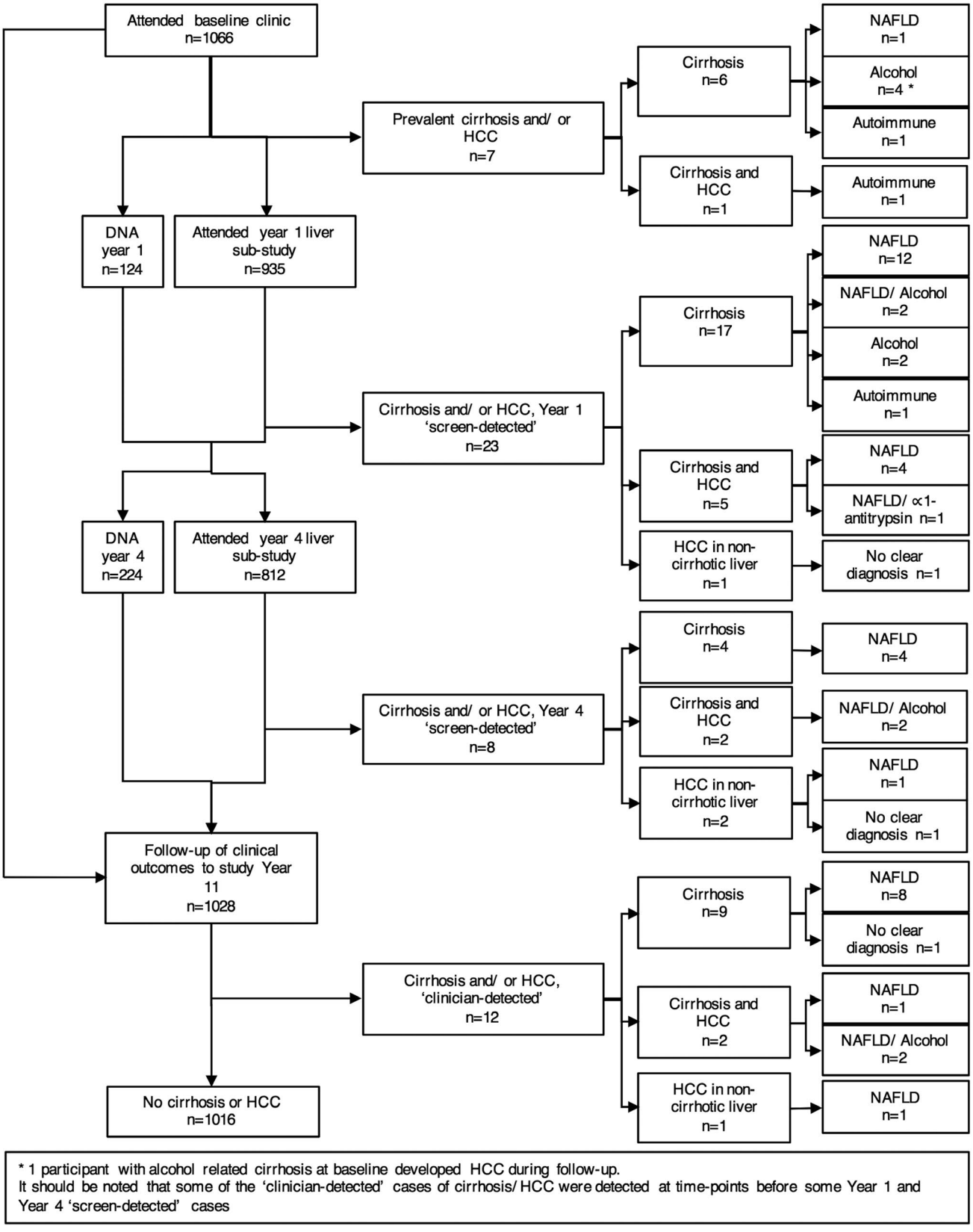

FIGURE 2 Cirrhosis and HCC events at baseline and through 11-year follow-up

Different steatosis and fibrosis scores were used within this algorithm to see whether combinations of different scores within the algorithm affected algorithm performance. No significant difference was observed in how well the algorithm 'advise to refer' outcome associated with incident cirrhosis/HCC based on the marker of steatosis used (Table 2). Irrespective of the fibrosis score used in the algorithm,

TABLE 2 Performance of Fibrosis scores in prediction of 11-year incident cirrhosis/HCC

Fibrosis score (cut-point value used)	OR (95% CI)	Sens (%, 95% CI)	Spec (%, 95% CI)	PPV (%, 95% CI)	NPV (%, 95% CI)	False $^{+}$ve n (%)	False $^{-}$ve n (%)
Individual scores							
ELF (≥10.51)[a]	30.4 (11.3-83.5)***	39 (22-59)	98 (97-99)	46 (26-67)	97 (96-98)	13 (2)	17 (61)
APRI (>0.5)	23.7 (11.5-49.8)***	53 (36-68)	94 (93-96)	27 (18-39)	98 (97-99)	57 (6)	19 (48)
AST:ALT (≥0.8)	3.9 (2.1-7.6)***	41 (26-58)	26 (23-29)	2 (1-14)	92 (88-95)	261 (26)	24 (59)
NFS (≥-1.455)	3.7 (1.3-15.6)*	93 (80-98)	22 (20-25)	5 (3-6)	99 (96-100)	769 (78)	3 (8)
NFS (>0.676)	8.2 (4.2-15.9)***	45 (29-62)	91 (89-93)	17 (10-26)	98 (96-98)	88 (9)	22 (56)
FIB 4 (≥1.3)	8.1 (3.7-20.6)***	82 (67-93)	59 (56-63)	8 (5-11)	99 (98-100)	402 (41)	7 (18)
FIB 4 (>2.67)	39.5 (17.4-91.9)***	40 (25-57)	98 (97-99)	46 (29-63)	98 (96-98)	19 (2)	24 (60)
Scores used within EASL-EASD-EASO algorithm, uss steatosis as steatosis marker							
ELF (≥10.51)[a]	10.1 (4.6-25.6)***	82 (66-92)	66 (63-69)	10 (7-14)	99 (97-99)	276 (34)	7 (18)
APRI (>0.5)[a]	13.0 (6.1-31.0)***	80 (64-91)	73 (70-75)	10 (7-14)	99 (98-100)	274 (27)	8 (20)
AST:ALT (≥0.8)[a]	7.7 (3.7-16.7)***	80 (65-91)	63 (60-66)	8 (6-11)	99 (98-99)	367 (37)	8 (20)
NFS (≥-1.455)[a]	6.0 (2.5-17.6)***	88 (73-96)	44 (41-47)	6 (4-9)	99 (97-100)	527 (56)	5 (13)
FIB 4 (≥1.3)[a]	11.7 (5.0-34.5)***	88 (73-96)	60 (57-64)	8 (6-11)	99 (98-100)	384 (40)	5 (13)
Scores used within EASL-EASD-EASO algorithm, fatty liver index as steatosis marker							
ELF (≥10.51)[a]	9.1 (4.3-21.7)***	79 (64-91)	67 (64-70)	10 (7-14)	99 (97-99)	277 (33)	8 (21)
APRI (>0.5)	12.6 (5.9-30.1)***	80 (64-91)	72 (69-75)	10 (7-14)	99 (98-100)	281 (28)	8 (20)
AST:ALT (≥0.8)	10.0 (5.0-20.0)***	76 (60-88)	73 (70-76)	10 (7-14)	99 (98-99)	273 (27)	10 (24)
NFS (≥-1.455)	5.6 (2.2-19.0)**	90 (76-97)	36 (33-39)	5 (4-7)	99 (97-100)	635 (64)	4 (10)
FIB 4 (≥1.3)	13.7 (5.4-46.2)***	90 (76-97)	58 (55-61)	8 (6-11)	99 (98-100)	422 (42)	4 (10)

Note: AIC and C statistic documented in Supporting Information 1.

Abbreviations: OR, odds ratio (age and sex adjusted); sens, sensitivity; spec, specificity; PPV, positive predictive value; NPV, negative predictive value; ELF, enhanced liver fibrosis panel; NFS, NAFLD fibrosis score; APRI, AST:Platelet ratio index; FIB 4, fibrosis 4 index; EASL-EASD-EASO, European Association for the Study of the Liver, the European Association for the Study of Diabetes and the European Association for the Study of Obesity algorithm (Figure 1).

[a] ELF and ultrasound measured at year 1 only- so calculated 10 not 11-year incident cirrhosis/HCC.

*$P < .05$, **$P < .01$, ***$P < .001$.

people categorized as requiring referral were significantly more likely to develop cirrhosis/HCC (OR's range 0.1-13.7 with wide CIs, all $P < .05$). When used within the algorithm, the fibrosis score that resulted in the greatest ability to discriminate and appropriately associate algorithm 'advise to refer' outcome with cirrhosis/HCC was APRI, based on a C-statistic of 0.82. AIC was lowest when ELF was used within the algorithm, though APRI provided not dissimilar AIC performance.

The algorithm, regardless of steatosis marker or fibrosis score inserted, performed variably in how the 'advise to refer' outcome associated with incident cirrhosis/HCC (sensitivity 79%-90%, specificity 36%-73%). PPV was low (5-10%) indicating that a 'advise to refer' outcome was not a good predictor of incident cirrhosis/HCC. NPV was high at 99% but may again reflect the relative rarity of the outcome. False-negative rates were lower when using algorithm compared to fibrosis score alone, but were still 10-20%, which would have resulted in 4-8/40 who developed cirrhosis/HCC being classified as 'low risk'. False-positive rates ranged from 27-64%, with higher false-positive rates seen in using risk score combinations with lower false-negative rates. This again demonstrates that if scores are chosen that reduce the number who were at true risk of cirrhosis/HCC to being classified as 'low risk', a very large number of people who are not at risk of developing cirrhosis/HCC over 11 years would be advised to be referred to hepatology. For example using NFS (cut-off ≥1.455), 671 people would obtain a 'advise to refer' outcome, of whom only 36 developed cirrhosis/HCC. Using APRI, the model with the highest C-statistic, 8/40 people who developed incident cirrhosis/HCC would have been classified as 'low risk,' whereas 306 (using ultrasound steatosis as the steatosis marker) and 313 (using the FLI steatosis score as steatosis marker) would have been classified as 'high risk' and referral advised, with only 32 of those developing incident cirrhosis/HCC.

4.5 | Sensitivity analysis

Two sensitivity analyses were undertaken. The first demonstrates that there is no improvement in test performance when an outcome of 'presence of varices, ascites or encephalopathy in the context of cirrhosis or HCC' was used (Table 3). The second excluded all those with definite non-NAFLD disease (n = 3) and showed similar results

to those presented for the whole cohort with mixed aetiology disease above (Supporting Information 2).

Additionally, analysis was re-run using competing risks regression methodology with the competing risk being non-liver death. Results were similar to those obtained from logistic regression methodology with all risk scores showing a significant association with the development of cirrhosis/HCC and APRI providing the best improvement from null model by BIC (Supporting Information 3).

5 | DISCUSSION

In this study cohort of older people with Type 2 diabetes, during 11 years of follow-up a moderate rate of incident cirrhosis (3.66 per 1000 person years) and HCC (1.31 per 1000 person years) was identified. These are substantially higher than reported population rates (0.36-0.54 per 1000 person-years for cirrhosis; 0.41-0.58 per 1000 person years for 'liver cancer') (www.isdscotland.org) [31]. However, despite these findings (and consistent with other studies showing that Type 2 diabetes is a risk factor for the development of cirrhosis/HCC), the performance of existing non-invasive risk stratification tools in identifying those at risk of developing disease was poor.

A significant association was demonstrated between all NAFLD fibrosis risk scores, and the EASL-EASD-EASO algorithm 'advise to refer' outcome, and incident cirrhosis/HCC by OR. However, confidence intervals of the OR were wide. The model that yielded the highest C-statistic, both in isolation, and as part of the EASL-EASD-EASO algorithm, suggesting best discriminatory ability, was APRI with a cut point of >0.5. However, this score in isolation would have resulted in 47.5% (19/40) people who developed cirrhosis/HCC being classified as 'low risk'. Using APRI within the EASL-EASD-EASO algorithm, 20% (8/40) people who developed cirrhosis/HCC would have been classified as 'low risk' (received a 'do not refer' outcome), whereas 29% (306 or 313 individuals using ultrasound steatosis or FLI respectively) would have been classified as 'high risk' (receiving a 'advise to refer' outcome), with only 32 of those developing cirrhosis/HCC over 11 years. Using any model, significant numbers of people would have been classified as 'high risk' who did not develop cirrhosis/HCC over 11 years, whereas a large proportion of those who developed cirrhosis/HCC would have been classified inappropriately as 'low risk'. It is important to note that many of the risk

TABLE 3 Performance of Fibrosis scores in prediction of 11-year incident cirrhosis-related varices, ascites, encephalopathy or HCC

Fibrosis score (cut-point value used)	OR (95% CI)	Sens (%, 95% CI)	Spec (%, 95% CI)	PPV (%, 95% CI)	NPV (%, 95% CI)	False $^{+}$ve n (%)	False $^{-}$ve n (%)
Individual scores							
ELF (≥10.51)[a]	25.41 (8.45-76.40)***	38 (18-62)	98 (96-99)	33 (16-55)	98 (97-99)	16 (2)	13 (62)
APRI (>0.5)	31.30 (13.57-75.97)***	61 (41-78)	94 (92-95)	22 (13-33)	94 (92-95)	61 (6)	11 (39)
AST:ALT (≥0.8)	5.50 (2.55-12.57)***	34 (18-54)	26 (23-29)	1 (1-2)	93 (90-96)	754 (74)	19 (66)
NFS (≥-1.455)	4.05 (1.18-25.44)	93 (76-99)	22 (20-25)	3 (2-5)	99 (97-100)	780 (78)	2 (7)
NFS (>0.676)	8.48 (3.85-18.47)***	46 (28-66)	91 (89-92)	12 (7-20)	98 (97-99)	93 (9)	15 (54)
FIB 4 (≥1.3)	15.67 (5.28-67.35)***	89 (72-98)	59 (56-62)	6 (4-8)	99 (99-100)	410 (41)	3 (11)
FIB4 (>2.67)	59.99 (24.00-157.79)***	50 (31-69)	98 (97-99)	40 (24-58)	99 (98-99)	21 (2)	14 (50)
Scores used within EASL-EASD-EASO algorithm, uss steatosis as steatosis marker							
ELF (≥10.51)[a]	6.32 (2.75-16.36)***	75 (55-89)	65 (62-69)	7 (4-10)	99 (97-99)	286 (35)	7 (25)
APRI (>0.5)[a]	10.74 (4.51-29.76)***	79 (59-92)	72 (69-75)	7 (5-11)	99 (98-100)	284 (28)	6 (21)
AST:ALT (≥0.8)[a]	5.45 (2.41-13.97)***	76 (56-90)	62 (59-65)	6 (3-8)	99 (98-100)	378 (38)	7 (24)
NFS (≥-1.455)[a]	3.70 (1.51-11.14)**	82 (63-94)	44 (40-47)	4 (3-6)	99 (97-100)	539 (56)	5 (18)
FIB 4 (≥1.3)[a]	7.34 (2.98-22.07)***	82 (63-94)	60 (57-63)	5 (4-8)	99 (98-100)	396 (40)	5 (18)
Scores used within EASL-EASD-EASO algorithm, fatty liver index as steatosis marker							
ELF (≥10.51)[a]	6.53 (2.84-16.89)***	75 (55-89)	66 (63-69)	7 (4-10)	99 (97-100)	287 (34)	7 (25)
APRI (>0.5)	10.47 (4.40-29.01)***	79 (59-92)	71 (68-74)	7 (4-10)	99 (98-100)	291 (29)	6 (21)
AST:ALT (≥0.8)	7.52 (3.07-22.54)***	83 (64-94)	60 (57-63)	6 (4-8)	99 (98-100)	407 (40)	15 (17)
NFS (≥-1.455)	4.96 (1.71-21.00)**	89 (72-98)	36 (33-39)	4 (2-5)	99 (98-100)	646 (64)	3 (11)
FIB 4 (≥1.3)	12.23 (4.23-51.78)***	89 (72-98)	57 (54-60)	5 (4-8)	99 (98-100)	433 (43)	3 (11)

Abbreviations: OR, odds ratio (age and sex adjusted); sens, sensitivity; spec, specificity; PPV, positive predictive value; NPV, negative predictive value; ELF, enhanced liver fibrosis panel; NFS NAFLD, fibrosis score; APRI, AST:Platelet ratio index; FIB 4, fibrosis 4 index, EASL-EASD-EASO, European Association for the Study of the Liver, the European Association for the Study of Diabetes and the European Association for the Study of Obesity algorithm (Figure 1).

[a] ELF and ultrasound measured at year 1 only- so calculated 10 not 11-year incident cirrhosis/HCC.

*$P < .05$, **$P < .01$, ***$P < .001$.

scores were designed to identify advanced fibrosis as opposed to cirrhosis/HCC. However, given the time span of follow-up we would have expected those with advanced fibrosis to progress to cirrhosis over 11 years and there thus to be a correlation. In addition, a significant proportion of our population underwent ultrasound at year 1. All abnormal ultrasounds were followed up and those diagnosed with fibrosis at year 1 progressed to cirrhosis over the period of the study.

ET2DS is a study of moderate size that has reviewed long-term liver outcomes in individuals with Type 2 diabetes who were asymptomatic of liver disease at baseline. Almost all other studies have examined outcomes in people recruited from secondary care hepatology clinics, with known NAFLD and a higher likelihood of cirrhosis/HCC. Although the ET2DS studied a cohort at higher risk of cirrhosis/HCC than the general population, the absolute probability of cirrhosis/HCC was moderately low. Therefore, validated risk scores and a European consensus algorithm have been tested in a cohort where the pre-test probability is low; in contrast to previous studies. However, this represents precisely the scenario in which European guidelines recommend screening for liver disease. Participants in the ET2DS were well characterized at baseline allowing accurate documentation of baseline risk factors, and have been followed longitudinally and extensively using multiple sources of information.

There are limitations to our study. ET2DS is a single centre study, undertaken in people aged 60-75 years, of predominantly Caucasian origin (98.3%). Whilst this was a representative sample of people with Type 2 diabetes in the population sampled (Lothian, Scotland, UK), care should be taken in extrapolating to other populations. All-cause cirrhosis and HCC was investigated. While aetiology was predominantly NAFLD, individuals with advanced liver disease because of other causes or with known co-factors (eg alcohol above the NAFLD threshold) were also included. Determining the precise aetiology of cirrhosis/HCC can often be difficult in a real-world setting. It is likely that some individuals had liver disease where both alcohol and obesity contributed, therefore including individuals with all-causes of liver disease seemed more clinically relevant. A sensitivity analysis excluding the three participants who had definite non-NAFLD disease did not reveal significantly different results (Supporting Information 2). Medication exposure data were not analysed, so any modifying effect will not have been detected.

The main outcome was cirrhosis/HCC. It is possible that some participants developed cirrhosis/HCC during follow-up, but were asymptomatic or did not seek medical advice for symptoms. These individuals would not have been identified as research screening for cirrhosis/HCC was not repeated at 11-year follow-up. A substantial proportion of the diagnoses were made after hepatology referral following year 1 and year 4 screening investigations. This has two implications. Firstly, as the natural history of NAFLD progression is very prolonged, it is possible that those who were diagnosed following referral from screening had cirrhosis/HCC at baseline and had prevalent rather than incident disease. However, the range of time from year 1 clinic to diagnosis overlaps significantly in the 'screen-detected' and 'clinician-detected' groups. Moreover, several of those who were 'screen-detected' were not identified with cirrhosis/HCC on initial hepatology review, but follow-up was continued because of concern regarding 'high-risk' features and they were diagnosed with cirrhosis/HCC several years later. Therefore, we defined prevalent disease as that which was clinically apparent at baseline. Secondly, the screening process may have led to an earlier diagnosis of cirrhosis/HCC in some people who may have died from other causes before cirrhosis/HCC was clinically apparent, inflating incidence rates. However, 58% of those identified with cirrhosis developed varices, ascites and/or encephalopathy and 23% developed HCC, so while investigation may have advanced the time of diagnosis, many would likely have presented during the period of follow-up. While all other biomarkers were measured at baseline, ELF and liver ultrasound were undertaken at the year 1 clinic, so analyses using these markers have examined slightly different 'baseline' time points. However, the performance of the EASL-EASD-EASO algorithm did not differ when using Fatty Liver Index (measured at baseline) and ultrasound as the steatosis marker, so with respect to the steatosis assessment, it is unlikely that this had a material effect on the present results. Because of limitations in the time to diagnosis data, both a logistic regression analysis and a competing risks regression approach (as a sensitivity analysis) were used. The former analysis has the disadvantage of not taking into account deaths during follow-up, whereas in the latter approach, time-to-event discrepancies may also introduce bias. Results of the competing risks regression were similar to the logistic regression assessment (Supporting Information 3) suggesting that neither the proportion of non-liver death in our population nor the mixed screen-detected and clinician-detected events substantially affected results.

Several studies have compared non-invasive markers of fibrosis to clinical outcome in NAFLD, mostly undertaken in populations of people under secondary care hepatology clinic follow-up, who had an initial liver biopsy. Three studies (median 5-12 year follow-up) showed increased hazard ratios (HR) for all-cause mortality, decompensated cirrhosis, rates of HCC or liver transplant in those with raised NFS, APRI or FIB-4 scores (16-36% participants had diabetes).[17,19,20] None reported sensitivity, specificity, PPV or NPV. Three studies compared non-invasive scores with severity of fibrosis on biopsy at the time of testing and showed strong associations between NFS, APRI, FIB-4 and AST: ALT ratio, and biopsy with an area under the receiving curve of 0.7-0.88, depending on score used (19-50% participants had diabetes).[16,28,32] However, all described decreasing specificity with increasing sensitivity, for risk score cut-points used. A recent study examined median 4 year outcomes in a cohort of 284 participants under hepatology clinic follow-up for NAFLD (>80% biopsy confirmed, 53% had diabetes).[25] As expected in a hepatology clinic population, rates of cirrhosis/HCC were high (9.2% liver-related death or transplant, 14.8% decompensated cirrhosis, 9.9% HCC). A diagnosis of diabetes conferred an increased HR of developing a liver outcome (death/transplantation HR 3.4 (95% confidence interval 1.2-9.1), decompensated cirrhosis HR 4.7 (2-11.3) and HCC HR 2.9

(1.2-7.3)). However, NFS, APRI or FIB-4 scores in the people with diabetes were substantially less good (by C-statistic comparison) at predicting outcome than in the individuals without diabetes. In those with diabetes, 21% of those with a 'low-risk' NFS, 15% with 'low-risk FIB-4' and 15% 'low-risk' APRI developed decompensated cirrhosis, and 27% with 'low-risk' scores developed HCC. In contrast, in individuals without diabetes, no participant with a 'low-risk' fibrosis score developed decompensated cirrhosis or HCC during follow-up. Therefore, the results of our study confirm what is reported in previous publications; that non-invasive risk scores do associate with outcome, but false-positive and -negative levels are high.

Current risk prediction scoring fails to identify a significant proportion of people with Type 2 diabetes who develop incident cirrhosis/HCC. Our population representative approach implies that general use of current risk scores and algorithms in people with Type 2 diabetes will result in unnecessary additional referral and investigation in large numbers of people who will not develop incident cirrhosis/HCC over 11 years. This has significant resource implications for hepatology services. Our study importantly examines outcomes from an unselected community population, for which these screening algorithms are advocated.

It remains unclear why the fibrosis risk scores perform better in people without diabetes than those with diabetes. It is possible that there are confounders influencing the biomarkers used in the non-invasive scores that are affected by diabetes. For example it has been described that measurements of AST and ALT in mouse models are affected by hyperglycaemia.[25] Future research is required to identify improved methods of predicting incident cirrhosis/HCC in this high-risk population, possibly through combining existing risk scores, examining whether serial monitoring is a more effective screening strategy or investigating novel or alternative biomarkers.

Type 2 diabetes is associated with an increased rate of cirrhosis/HCC.[2,3] Risk prediction scores and international guidelines have attempted to provide non-invasive methods of assessing risk of incident disease in this high-risk population. This study shows only modest performance of these risk scores and screening algorithm. Use would lead to significant pressure on hepatology services from high referral rates coupled with increased patient anxiety generated by false-positive results. Furthermore, the risk scores fail to identify a significant proportion of the population that are potentially vulnerable to incident disease. Future work to improve prediction methods in this population is necessary.

ACKNOWLEDGEMENTS

S.M.G. wrote the manuscript. J.F.P. was principal investigator of the ETDS, designed the study, analysed and interpreted the data. M.W.J.S., was lead investigator of the liver substudy of the ET2DS, designed the study, analysed and interpreted the data. R.M.R., B.M.F., P.C.H, J.A.F, R.M.W, N.G. and S.G. contributed to study design. S.M.G., S.M., R.M.W., J.M. and P.K. contributed to data collection, analysis and interpretation.

CONFLICT OF INTEREST

J.A.F. has received consultancy fees from Ferring Pharmaceuticals, Macrophage Pharma, Galecto Biotech, Caldan Therapeutics, Cypralis Ltd, NorthSea Therapeutics, Gilde Healthcare, Guidepoint, Techspert.ioand research grant funding from GlaxoSmithKline, Intercept Pharmaceuticals and Novartis for work unconnected with that reported in this article. J.M. reported salary funding from Diabetes UK for the ET2DS. R.M.W. reported salary funding by a research grant from Pfizer for 2 years. B.M.F. has participated in a Speakers' Bureau for: Lilly, Novo Nordisk, MSD, Sanofi, Roche, Abbott and Boehringer Ingelheim, and has served on advisory boards for Lilly, Novo Nordisk, Locemia Solutions and Zucara. M.W.J.S has received consultancy fees from Servier and Novo Nordisk and speaking fees from Napp and Eli Lilly. J.R.M. reported salary funding during the project from Diabetes UK and reports salary funding (not associated with this project) from MRC. No other potential conflicts of interest relevant to this article were reported.

AUTHORSHIP

All authors contributed to revision and final approval of the article. The authors thank all the staff and participants involved in the Edinburgh Type 2 Diabetes Study

ORCID

Sheila M. Grecian https://orcid.org/0000-0002-1600-559X
Jonathan A. Fallowfield https://orcid.org/0000-0002-5741-1471

REFERENCES

1. El-serag HB, Tran T, Everhart JE. Diabetes increases the risk of chronic liver disease and hepatocellular carcinoma. *Gastroenterology*. 2004;126(2):460–468.
2. Campbell PT, Newton CC, Patel AV, Jacobs EJ, Gapstur SM. Diabetes and cause-specific mortality in a prospective cohort of one million U.S. adults. *Diabetes Care*. 2012;35(9):1835–1844.
3. Pang Y, Kartsonaki C, Turnbull I, et al. Plasma glucose, and incidence of fatty liver, cirrhosis, and liver cancer: A prospective study of 0.5 million people. *Hepatology. Wiley-Blackwell*. 2018;68(4):1308–1318.
4. Targher G, Bertolini L, Rodella S, et al. Non-alcoholic fatty liver disease is independently associated with an increased prevalence of chronic kidney disease and proliferative/laser-treated retinopathy in type 2 diabetic patients. *Diabetologia*. 2008;51(3):444–450.
5. Leite NC, Salles GF, Araujo ALE, Villela-Nogueira CA, Cardoso CRL. Prevalence and associated factors of non-alcoholic fatty liver disease in patients with type-2 diabetes mellitus. *Liver Int. Blackwell Publishing Ltd*. 2009;29(1):113–119.
6. Williamson RM, Price JF, Glancy S, et al. Prevalence of and risk factors for hepatic steatosis and nonalcoholic Fatty liver disease in people with type 2 diabetes: the Edinburgh Type 2 Diabetes Study. *Diabetes Care. American Diabetes Association*. 2011;34(5):1139–1144.
7. Yi M, Chen R-P, Yang R, Chen H. Increased prevalence and risk of non-alcoholic fatty liver disease in overweight and obese patients with Type 2 diabetes in South China. *Diabet Med*. 2017;34(4):505–513.
8. 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 Hepatology*. European Association for the Study of the Liver; 2016;64:1388–402.
9. Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of nonalcoholic fatty liver disease: Practice guidance from the American Association for the Study of Liver Diseases. *Hepatology*. 2017;55(1 Suppl):2005.
10. Argo CK, Northup PG, Al-Osaimi AMS, Caldwell SH. Systematic review of risk factors for fibrosis progression in non-alcoholic steatohepatitis. *J Hepatol*. 2009;51(2):371–379.
11. Singh S, Allen AM, Wang Z, Prokop LJ, Murad MH, Loomba R. Fibrosis progression in nonalcoholic fatty liver vs nonalcoholic steatohepatitis: a systematic review and meta-analysis of paired-biopsy studies. *Clin Gastroenterol Hepatol*. 2015;13(4):643–54. e1–9–quize39–40.
12. Wai C-T, Greenson JK, Fontana RJ, et al. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis. *C. Hepatology. W.B. Saunders*. 2003;38(2):518–526.
13. Sterling RK, Lissen E, Clumeck N, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology*. 2006;43(6):1317–1325.
14. Angulo P, Hui JM, Marchesini G, et al. The NAFLD fibrosis score: A noninvasive system that identifies liver fibrosis in patients with NAFLD. *Hepatology*. 2007;45(4):846–854.
15. Guha IN, Parkes J, Roderick P, et al. Noninvasive markers of fibrosis in nonalcoholic fatty liver disease: Validating the European Liver Fibrosis Panel and exploring simple markers. *Hepatology*, 2008;47(2):455–460.
16. McPherson S, Stewart SF, Henderson E, Burt AD, Day CP. Simple non-invasive fibrosis scoring systems can reliably exclude advanced fibrosis in patients with non-alcoholic fatty liver disease. *Gut. BMJ Publishing Group*. 2010;59(9):1265–1269.
17. Angulo P, Bugianesi E, Bjornsson ES, et al. Simple noninvasive systems predict long-term outcomes of patients with nonalcoholic fatty liver disease. *Gastroenterology*. 2013;145(4):782–784.
18. Kim D, Kim WR, Kim HJ, Therneau TM. Association between noninvasive fibrosis markers and mortality among adults with nonalcoholic fatty liver disease in the United States. *Hepatology*. 2013;57(4):1357–1365.
19. Treeprasertsuk S. NAFLD fibrosis score: A prognostic predictor for mortality and liver complications among NAFLD patients. *World J Gastroenterol*. 2013;19(8):1219.
20. Sebastiani G, Alshaalan R, Wong P, et al. Prognostic value of non-invasive fibrosis and steatosis tools, hepatic venous pressure gradient (HVPG) and histology in nonalcoholic steatohepatitis. *PLOS ONE*. 2015;10(6):e0128774.
21. Irvine KM, Wockner LF, Shanker M, et al. The Enhanced liver fibrosis score is associated with clinical outcomes and disease progression in patients with chronic liver disease. *Liver Int. Wiley/Blackwell*. 2016;36(3):370–377.
22. Morling JR, Fallowfield JA, Guha IN, et al. Using non-invasive biomarkers to identify hepatic fibrosis in people with type 2 diabetes mellitus: the Edinburgh type 2 diabetes study. *J Hepatol*. 2014;60(2):384–391.
23. Morling JR, Fallowfield JA, Guha IN, et al. Clinically significant chronic liver disease in people with Type 2 diabetes: the Edinburgh Type 2 Diabetes Study. *QJM*. 2016;109(4):249–256.
24. Sberna AL, Bouillet B, Rouland A, et al. European association for the study of the liver (EASL), European Association for the Study of Diabetes (EASD) and European Association for the Study of Obesity (EASO) clinical practice recommendations for the management of non-alcoholic fatty liver disease. *Diabet Med*. 2018;35(3):368–375.
25. Bertot LC, Jeffrey GP, de Boer B, et al. Diabetes impacts prediction of cirrhosis and prognosis by non-invasive fibrosis models in non-alcoholic fatty liver disease. *Liver Int*. 2018;25(59):2188.
26. Price JF, Reynolds RM, Mitchell RJ, Williamson RM, Fowkes FGR, Deary IJ, et al. he Edinburgh Type 2 Diabetes Study: study protocol. *BMC Endocrine Dis*. 2008;8(1):18.
27. Williamson RM, Perry E, Glancy S, et al. The use of ultrasound to diagnose hepatic steatosis in type 2 diabetes: intra- and interobserver variability and comparison with magnetic resonance spectroscopy. *Clin Radiol*. 2011;66(5):434–439.
28. Shah AG, Lydecker A, Murray K, et al. Comparison of noninvasive markers of fibrosis in patients with nonalcoholic fatty liver disease. *Clin Gastroenterol Hepatol*. 2009;7(10):1104–1112.
29. Bedogni G, Bellentani S, Miglioli L, et al. The Fatty Liver Index: a simple and accurate predictor of hepatic steatosis in the general population. *BMC Gastroenterology*. 2006;6(1):33.
30. Grecian SM, McLachlan S, Fallowfield JA, et al. Hepatic steatosis does not predict 10-year incident cirrhosis, hepatocellular cancer or mortality in people with Type 2 diabetes. The Edinburgh Type 2 Diabetes Study. *Diabet Med*. 2019;18;36(Supplement 1):19.
31. Ratib S, West J, Crooks CJ, Fleming KM. Diagnosis of liver cirrhosis in England, a cohort study, 1998–2009: A comparison with cancer. *Am J Gastroenterol*. 2014;109(2):190–198.
32. Calès P, Lainé F, Boursier J, et al. Comparison of blood tests for liver fibrosis specific or not to NAFLD. *J Hepatol*. 2009;50(1):165–173.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Grecian SM, McLachlan S, Fallowfield JA, et al. Non-invasive risk scores do not reliably identify future cirrhosis or hepatocellular carcinoma in Type 2 diabetes: The Edinburgh type 2 diabetes study. *Liver Int*. 2020;00:1–11. https://doi.org/10.1111/liv.14590