THE UNIVERSITY
of EDINBURGH

Investigating the Natural History of Liver Disease in Type 2 Diabetes and Predicting the Risk of its Progression to Advanced Disease

Sheila Marion Grecian

THE UNIVERSITY *of* EDINBURGH

Doctor of Medicine

The University of Edinburgh

2021

Declaration

I declare that the content of this thesis is my own and that all contributions and collaborations have been explicitly acknowledged in the text. No material presented in this thesis has been submitted to any other university or for any other degree.

Sheila Marion Grecian

19.01.2021

Abstract

Introduction

In people with Type 2 diabetes; chronic liver disease, particularly non-alcoholic fatty liver disease (NAFLD), is more common and has an increased risk of progression to cirrhosis and hepatocellular carcinoma. European guidelines (European Association for the Study of the Liver, European Association for the Study of Diabetes and European Association for the Study of Obesity) recommend screening for NAFLD in Type 2 diabetes yet both the natural history of liver disease in Type 2 diabetes and the factors associated with higher risk of progression to clinically significant disease are still incompletely understood. Further, it is thought that the recommended generic NAFLD risk prediction tools may perform sub-optimally in people with Type 2 diabetes.

Aims

This study aimed to use a community cohort of over one thousand older people with Type 2 diabetes followed for 11 years to:

1. Define the absolute and relative cohort incidence of liver disease to date.
2. Determine whether current non-invasive fibrosis risk prediction tools reliably identified incident cirrhosis and hepatocellular carcinoma in people with Type 2 diabetes.
3. Determine whether the addition of baseline biomarkers to existing fibrosis risk prediction tools improved their ability to predict incident cirrhosis and hepatocellular carcinoma.
4. Identify whether potential non-invasive tests for non-alcoholic fatty liver disease (those identifying steatosis, serum liver enzymes, markers of fibrosis) are associated with incident cirrhosis, hepatocellular carcinoma or all-cause mortality.

Methods

The Edinburgh Type 2 Diabetes Study recruited men and women with Type 2 diabetes (n=1,066, aged 60–75 at baseline) in 2006. Liver markers were measured at baseline and year 1; steatosis and fibrosis markers were calculated according to independently

published formulae. During follow-up, cases of cirrhosis and HCC were identified. Logistic regression (odds ratio) was used to determine associations between markers and outcomes, with competing risks regression used for sensitivity analyses. The predictive ability of tests was assessed using sensitivity, specificity, positive predictive value, negative predictive value, false positive and false negative rates.

Results

Over 11 years 43/1059 participants with no baseline cirrhosis or hepatocellular carcinoma developed incident liver disease. The 11-year incidence of liver cirrhosis was 3.92 per 1000 person years and of hepatocellular carcinoma 1.28 per 1000 person years (whole population rates). 58% of those with cirrhosis had clinical complications of varices, ascites or hepatic encephalopathy.

Existing non-invasive NAFLD fibrosis risk-stratification tools (AST:ALT ratio, AST: platelet ratio index (APRI), Enhanced Liver Fibrosis panel (ELF), Fibrosis 4 index (FIB-4), NAFLD Fibrosis Score (NFS)) were significantly associated (Odds Ratios, $p<0.05$) with incident cirrhosis and hepatocellular carcinoma but their ability to accurately identify those who developed incident disease was poor with low positive predictive values (5-46%) and high false negative and false positive rates (up to 60% and 77% respectively). When fibrosis risk scores were used in conjunction with the European algorithm, predictive performance was modestly improved.

Among the risk-stratification scores tested, FIB-4 and APRI performed best. Of additional biomarkers assessed, hyaluronic acid, gamma-glutamyl transferase, glycated haemoglobin, C-reactive protein, interleukin-6, body mass index and deprivation index were each, individually, significantly associated with future cirrhosis or hepatocellular carcinoma after adjustment for age, sex and existing components of the base models. However, only the addition of hyaluronic acid (cut-point ≥50μg/L) to FIB-4 (cut-point ≥1.3) reduced the number of people falsely identified as ‘high-risk’ by ~50% whilst retaining a false negative rate of ≤25%.

Serum liver enzymes, the Fatty Liver Index, hepatic steatosis on ultrasound, FIB-4 and FIB-4 with hyaluronic acid all had false positive or false negative rates of >20% or >35% respectively for the identification of cirrhosis or HCC. A raised Fatty Liver Index was statistically associated with mortality (hazard ratio 1.45 (1.13-1.87)) but all

tests showed high false positive and false negative rates (>20% or >75% respectively) for mortality.

Conclusions

The increased incidence of cirrhosis and hepatocellular carcinoma in people with Type 2 diabetes were confirmed, with NAFLD the predominant aetiology. Markers of fibrosis were associated with incident cirrhosis and hepatocellular carcinoma but no non-invasive risk prediction tools reliably identified participants at increased risk of incident disease. The addition of hyaluronic acid to FIB-4 showed promise by reducing the proportion of people inappropriately identified as ‘high-risk’ but no combination of tests examined, provided a ‘good balance’ between false positive and negative rates in the identification of risk for cirrhosis, HCC or mortality.

These results need to be validated in independent cohorts but suggest that the evidence does not exist for formal liver disease screening in people with Type 2 diabetes and presently the only option for non-invasive liver disease surveillance is to use tests with a relatively low false positive rate in order to identify a proportion of those likely to develop incident cirrhosis and HCC.

Lay Summary

Non-alcoholic liver disease is the leading cause of liver disease affecting 1 in 4 adults worldwide, and occurs more often in people with obesity or Type 2 diabetes. Build-up of fat in the liver often causes inflammation and damage that can get worse and lead to fibrosis (scarring), cirrhosis (severe scarring), liver failure and liver cancer. In advanced disease people have symptoms and may die without a liver transplant. Therefore, it is not surprising that screening for liver disease in people with Type 2 diabetes has been recommended by international organisations even though there are big gaps in the detailed understanding needed to design reliable approaches.

The Edinburgh Type 2 Diabetes Study is a unique resource for studying this issue. It included 1066 men and women with Type 2 diabetes and aged 60-75 at baseline who were followed for 11 years. Detailed clinical measurements at the start were linked with the development of new liver problems and death over the follow up period.

This study found that the number of people who developed serious liver disease and its complications was higher than in the general population. But, unfortunately, the currently recommended screening pathways did not reliably identify the people who went on to develop advanced liver disease. These tests too often missed people at risk (false negatives) and wrongly identified others as being at high risk (false positives). We showed that adding a blood test measurement to one of the risk prediction tools (measurement of a chemical called hyaluronic acid to the Fibrosis-4 index) reduced the false identification of people as ‘high-risk’ (for developing cirrhosis or liver cancer). However, no test provided a ‘good balance’ between accurately identifying those who developed severe liver disease and those who did not.

Further research is needed if a reliable method of screening for risk of serious liver disease in people with Type 2 diabetes is to be developed.

Acknowledgements

I would first like to thank my supervisors, Professor Mark Strachan, Professor Jackie Price and Dr Nicola Zammitt. They have provided continuing support and guidance. Their patience, advice and enthusiasm have been invaluable.

This research would not have been possible without the efforts of every person who has been involved in the Edinburgh Type 2 Diabetes Study. Particular mention goes to those involved with the design of the study (Prof Jackie Price, Prof Mark Strachan, Prof Rebecca Reynolds, Prof Brian Frier, Prof Peter Hayes, Prof Jonathan Fallowfield, Prof Neil Guha, Dr Stephen Glancy). My work in particular would not have been possible without the efforts of Dr Rachel Williamson and Dr Joanne Morling who were instrumental in the data collection and analysis during the earlier years of the liver sub study. We indebted to every participant in the study.

Of course, this research would not have been possible without the funding of the study from the following grants. 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 sub-study was supported by a grant from Pfizer (Unrestricted Investigator Led Grant). I am very grateful to have been supported by NHS Lothian for the funding of my time in research.

I would like to thank Dr Stela McLachlan, Dr Rachel Foster and Dr Anniek Sluiman who worked alongside me at ET2DS, for their guidance, patience in the face of many questions, and company. I wish to say a particular thank you to Prof Jonathan Fallowfield who has provided invaluable hepatology and broader research advice; I have been so grateful for his expertise and patience. Thanks also go to Dr Patrick Kearns, Dr Anna Price, Dr Tom Drake and the team at Healthy R for their assistance in my statistical education, with a particular mention to Patrick with his assistance with the competing risk regression analysis.

Finally, thank you to my family, without whom I would not be in the position of writing a thesis. Thank you to my parents for their unending encouragement and support. Thank you to my husband Rob for your love, support, encouragement and patience. Thank you to my daughters, Iona (5) and Ailsa (1), and our dog, Holly for giving us so much enjoyment, your boundless love and for keeping me grounded.

List of Publications and Presentations

Publications

Grecian S, Fallowfield J, McLachlan S, Hayes P, Guha N, Morling J, Glancy S, Williamson R, Reynolds R, Frier B, Zammitt N, Price J, Strachan M. The association of non-alcoholic fatty liver disease screening tests with incident advanced chronic liver disease and mortality in people with Type 2 diabetes. *Diabetic Medicine* (Under Review)

Grecian S, McLachlan S, Fallowfield J, Hayes P, Guha N, Morling J, Glancy S, Williamson R, Reynolds R, Frier B, Zammitt N, Price J, Strachan M. Addition of hyaluronic acid to the FIB-4 liver fibrosis score improves prediction of incident cirrhosis and hepatocellular carcinoma in Type 2 diabetes: The Edinburgh Type 2 Diabetes Study. *Obesity Science & Practice* 2021 https://doi.org/10.1002/osp4.484

Forster R, Sandoval Garcia E, Sluiman A, Grecian S, McLachlan S, MacGillivray T, Strachan M, Price J. Retinal venular tortuosity and fractal dimension predict incident retinopathy in adults with type 2 diabetes: The Edinburgh Type 2 Diabetes Study. *Diabetologia* 2021 May;64(5):1103-1112.

Grecian S, McLachlan S, Fallowfield J, Kearns P, Hayes P, Guha N, Morling J, Glancy S, Williamson R, Reynolds R, Frier B, Zammitt N, Price J, Strachan M. Non-invasive risk scores do not reliably identify future cirrhosis or hepatocellular carcinoma in Type 2 diabetes: The Edinburgh Type 2 Diabetes Study. *Liver International* 2020 Sep;40(9):2252-2262

Carr A, Sluiman A, Grecian S, Forster R, McLachlan S, Strachan M, Price J. Depression as a risk factor for dementia in older people with type 2 diabetes and the mediating effect of inflammation. *Diabetologia* 2020 https://doi.org/10.1007/s00125-020-05301-6

Presentations

Grecian S, McLachlan S, Fallowfield J, Guha N, Morling J, Williamson R, Reynolds R, Zammitt N, Price J, Strachan M. Serum hyaluronic acid improves the performance of the Fibrosis-4 Index in the identification of advanced liver disease in Type 2 diabetes. The Edinburgh Type 2 Diabetes Study. Accepted *Diabetes UK Professional Conference* 2020 (cancelled due to COVID-19) (oral presentation)

Grecian S, McLachlan S, Sluiman A, Forster R, Williamson R, Reynolds R, Frier B, Zammitt N, Price J, Strachan M. Hepatic steatosis does not predict 10 year incident cirrhosis, hepatocellular cancer or mortality in people with Type 2 diabetes. *Diabetes UK Professional Conference* 2019, Liverpool. (oral presentation, Co-winner of YDEF Travel Award, Diabetes UK 2019)

Grecian S, McLachlan S, Sluiman A, Forster R, Williamson R, Reynolds R, Frier B, Zammitt N, Price J, Strachan M. Current European clinical practice guidelines for the management of non-alcoholic fatty liver disease: Is the diagnostic algorithm useful in people with Type 2 diabetes? *Diabetes UK Professional Conference* 2019, Liverpool. (poster presentation)

Grecian S, McLachlan S, Sluiman A, Forster R, Williamson R, Reynolds R, Frier B, Zammitt N, Price J, Strachan M. Frequency of Cirrhosis and Hepatocellular Carcinoma in Older People with Type 2 Diabetes. *Scottish Society of Physicians* 2018, Edinburgh. (poster presentation)

List of Abbreviations

AASLD	The American Association for the Study of Liver Diseases
AGEs	advanced glycosylation end products
AIC	Aikaike Information Criterion
ALD	Alcoholic Liver Disease
ALP	Alkaline Phosphatase
ALT	Alanine Aminotransferase
APRI	AST to platelet ratio index
AST	Aspartate Transaminase
AUROC	Area Under the Receiver Operating Characteristic curve
BMI	Body Mass Index
Bp	Blood Pressure
CI	Confidence Interval
CRP	C-reactive Protein
DAG	di-acylglycerol
EASL-EASD-EASO	European Association for the Study of the Liver, European Association for the Study of Diabetes and European Association for the Study of Obesity
ELF	Enhanced Liver Fibrosis Score
ER	Endoplasmic reticulum
ET2DS	Edinburgh Type 2 Diabetes Study
FFA	Free fatty acids
FIB-4	Fibrosis-4 Score
FLI	Fatty Liver Index
GGT/ γGT	Gamma Glutamyltransferase
GLP-1	Glucagon-like Peptide-1

HA	Hyaluronic Acid
HbA1c	Glycated haemoglobin
HCC	Hepatocellular carcinoma
HOMA-IR	Homeostatic Model Assessment of Insulin Resistance
HSC	hepatic stellate cells
HR	Hazard Ratio
IKKβ	Inhibitor of nuclear factor kappa-B kinase subunit beta
IL-6	Interleukin 6
JNK-1	c-jun N terminal protein kinase 1
MI	Myocardial Infarction
MRS	Magnetic Resonance Imaging (spectroscopy)
NAFL	Non-alcoholic fatty liver
NAFLD	Non-alcoholic Fatty Liver Disease
NASH	Non-alcoholic steatohepatitis
NF-κB	Nuclear factor kappa-B
NFS	NAFLD Fibrosis Score
NPV	Negative predictive value
OR	Odds Ratio
PKC	Protein kinase C
PPARγ	Peroxisome proliferator-activated receptor gamma
PPV	Positive predictive value
ROS	Reactive Oxygen Species
RR	Relative risk
SGLT-2	Sodium-glucose co-transporter 2
T2DM	Type 2 Diabetes Mellitus
TNFα	Tumour Necrosis Factor alpha
USS	Ultrasound

VIF	Variance inflation factor
VLDL	Very low density lipoprotein

List of Figures

List of Tables

Contents

Chapter 1 Introduction

1.1 Non-alcoholic Fatty Liver Disease (NAFLD)- definition and diagnosis

1.1.1 Definition

NAFLD is defined as the presence of excessive fat (steatosis) in the liver when no secondary cause of hepatic steatosis can be found. NAFLD is the liver manifestation of the metabolic syndrome (a cluster of conditions including abdominal obesity, impaired glucose regulation or diabetes, hypertension, hypercholesterolaemia and hypertriglyceridaemia- all of which are associated with increased cardiovascular risk). The NAFLD spectrum encompasses: isolated steatosis (non-alcoholic fatty liver (NAFL)), steatosis with inflammation, ballooning or both (non-alcoholic steato-hepatitis (NASH)), and NAFLD associated with fibrosis, cirrhosis (irreversible liver scarring) and/or hepatocellular carcinoma (HCC, a primary liver tumour that usually develops in the setting of chronic liver disease).

1.1.2 Clinical presentation and implications

Most people with NAFLD are asymptomatic and are diagnosed during investigation of raised serum liver enzymes or by incidental detection of hepatic steatosis on abdominal imaging. People with NAFLD, particularly NASH, may complain of fatigue or right upper abdominal discomfort. On examination, hepatomegaly may be found although this is highly variable at all stages, and indeed even if a person presents with late-stage disease and cirrhosis, the cirrhotic liver may indeed be small. People who have developed cirrhosis as a result of NAFLD may present with weight loss, fatigue, stigmata of chronic liver disease (for example spider naevi, palmar erythema, clubbing, gynaecomastia, splenomegaly, caput medusae) or with signs or symptoms of decompensated cirrhosis (for example variceal bleeding, ascites, hepatic encephalopathy). HCC is often asymptomatic but those affected may develop wide ranging symptoms from new decompensation of existing cirrhosis, jaundice, abdominal pain, weight loss, a palpable mass, bleeding from the tumour, fever, lymphadenopathy, symptoms from metastatic disease and paraneoplastic syndromes including hypoglycaemia, hypercalcemia due to PTHrP, diarrhoea and various non-

specific cutaneous features. Without liver transplant, the life expectancy of someone with decompensated cirrhosis is around 6 months- 2 years while median 5-year survival for HCC is 15% (https://www.cancerresearchuk.org/).

1.1.3 Diagnosis

Diagnosis of NAFLD is made by detecting hepatic steatosis (defined as >5% liver fat, by ultrasound (USS), MRI or biopsy) where secondary causes of steatosis have been excluded from the medical history or liver screen tests (e.g. heavy alcohol intake (defined as alcohol intake >14 units/ week (female) or >21 units/ week (male) or a history of previous or current alcohol excess), hepatitis B or C infection; use of steatosis associated drugs (for example amiodarone, methotrexate, tamoxifen, valproate, anti-retroviral agents); parenteral nutrition; haemochromatosis; liver auto-immune disease; Wilson's disease; alpha1-antitrypsin deficiency; lipodystrophy or inborn errors of metabolism; and fatty liver associated with pregnancy). However, it is important to note that there is increasing awareness of a substantial population who have both NAFLD and alcoholic liver disease; they are thought to have a more aggressive disease course than people with either condition in isolation. [1]

NASH is a histological diagnosis that can be made only on liver biopsy. Fibrosis, cirrhosis and HCC are diagnosed by imaging and/or biopsy. Surrogate, non-invasive markers of fibrosis are also used and are discussed further below. If clinical, laboratory and radiological data strongly suggest the presence of cirrhosis, confirmatory biopsy is often not required. People with NAFLD may have elevated serum liver enzymes, though this is not a consistent finding. Laboratory abnormalities in people with advanced, fibrotic, disease include raised serum bilirubin or serum liver enzymes, prolonged prothrombin time, hypoalbuminaemia, hyponatraemia and thrombocytopaenia. Alpha-fetoprotein can be raised in the context of HCC but levels do not correlate well with clinical features and not all tumours secrete this protein.

1.2 The population importance of NAFLD

Liver disease is now documented as the second most significant contributor to years of working life lost (the first being ischaemic heart disease). [2] Contrary to most diseases in the UK, mortality rates from liver disease have risen 400% since 1970. [3] Whilst the majority of this impact is attributable to alcohol related liver disease, NAFLD and HCC are increasing contributors. [2] The future impact of NAFLD is concerning as it is clear that the numbers of people who are obese are rising, with a worrying rising trend in young people. [2,3] Despite the proportion of patients with NAFLD who progress to end stage liver disease being small, the vast population of patients with NAFLD means that a large proportion of liver related death is attributable to NAFLD, and NAFLD is now the second most common indication for liver transplant. [2-6] In addition, cardiovascular disease remains a significant cause of mortality in patients with NAFLD. [6]

1.3 The prevalence and incidence of NAFLD

NAFLD is thought to be the most common liver disease in the Western world. In the European population the median prevalence of NAFLD (any stage) is 25-26% (results calculated including paediatric populations, studies from 2000- present). [6,7] Another meta-analysis estimates a global adult prevalence (with imaging diagnostics) of 25.24% (95% CI 22.10-28.65%) though in this analysis studies exclusively examining high risk groups were not included. [6-8] The prevalence of NAFLD is increased in 'high risk' groups including people with Type 2 diabetes (T2DM), dyslipidaemia and obesity. In people with T2DM population prevalence (by USS assessment) is reported between 40-70%.[9-12] In the whole population among those with a body mass index (BMI) >30, a prevalence of >30% is seen on imaging studies whilst prevalence (biopsy diagnosis) in a morbidly obese bariatric surgery population can exceed 90%.[13-16] It is important to note that precise prevalence of NAFLD in the general population has been difficult to define due to the use of many different diagnostic criteria (including USS, liver enzymes, magnetic resonance spectroscopy imaging (MRS) and non-invasive tools developed using combinations of biomarkers (such as the fatty liver index (FLI) and liver biopsy). Of the non-invasive tools there is no gold-standard test increasing the difficulty of estimation. [17]

There is a paucity of data with regards to the incidence of NAFLD. One meta-analysis reports an incidence of 28/1000 person years (95% confidence interval 19.34-40.57) in the Western population and others report incidence rates of 52.34/1000 person years in the Asian population. [7,8,18,19] One English study reports an incidence of 29/100,000 person years but this was based on ICD-10 diagnosis and as such is likely to be a significant underestimate. [7] Similarly, incidence is seen to be increased in high risk populations. El-Serag et al studied the incidence of NAFLD in a large cohort reporting significantly higher incidence of NAFLD in subjects with T2DM compared to those without (18.13 vs 9.55 per 10000 person years $p<0.0001$). [20] Similar results have been seen in a biobank study from China, with hazard ratio (HR) for developing NAFLD over 10 years being 1.76 (95% CI 1.47-2.16) in the context of T2DM compared to the general population. [21]

Data on the population prevalence of the non-alcoholic steatohepatitis (NASH) and fibrosis stages is sparse; mainly due to the gold standard diagnostic test for diagnosis being liver biopsy, and as an invasive test, unsuitable for use in population screening. For NASH, the only population data is in people who have had biopsies for a non-

clinical reason (for example as a living donor) where the proportion of people with NASH was seen to be 6%.[7] In the bariatric surgery population, NASH is found in 25-70% people who agreed to liver biopsy at the time of surgery. [15,16,22]

Although there is no routine population registration of new diagnoses of cirrhosis in the UK, a comprehensive search of routinely collected electronic health care data from registries in England estimated the incidence of cirrhosis in England to be 0.37 per 1000 person years, with approximately 25% of those thought to be related to NAFLD (giving a likely incidence rate of NAFLD cirrhosis as 0.07 per 1000 person years). [23] For HCC, one meta-analysis quotes an incidence in defined NAFLD cohorts of 0.44 per 1000 person years (95% CI 0.29-0.66). [8] On a population level, data on liver cancer as a whole tend to be presented rather than HCC alone, although HCC contributes around 90% of liver cancer cases (www.britishlivertrust.org.uk). In Scotland, the incidence of liver cancer is 0.12 per 1000 person years (www.isdscotland.org). A recent USA paper looked at 296707 people with NAFLD (29.9% diabetes), followed for mean 9 years. They identified an incidence of HCC of 0.21 per 1000 person years. Importantly, whilst incidence of HCC was greater in those with cirrhosis, 20% of people with HCC did not have cirrhosis. [24] Rates of disease progression are discussed further in section 1.6.

1.4 The association between NAFLD and T2DM

It is thought that there is a bidirectional relationship between NAFLD and T2DM. Both are associated with an increased prevalence of the other in epidemiological cohorts and the presence of both is associated with worse patient outcome that the presence of either alone.

1.4.1 Evidence that the presence of NAFLD is associated with increasing prevalence, incidence and worsened outcomes of T2DM

The prevalence of T2DM is consistently higher in cohorts with NAFLD than in matched cohorts without. In two studies, patients identified with NAFLD (MR SPECT diagnosis) had a much greater prevalence of diabetes, impaired glucose regulation and adipose tissue insulin resistance compared to obese controls without NAFLD (85% vs 30% in one study). [25,26] In addition, even in patients without T2DM, the presence of NAFLD is associated with increased HbA1c within the normal range, and the level of steatosis on biopsy has been seen to correlate with increasingly impaired glucose regulation. [27,28]

In addition, rates of incident T2DM have been shown to be higher in people with NAFLD compared to those without. Several cohort studies have investigated the effect of NAFLD on the development of incident diabetes. Of these, the vast majority have found NAFLD to be a significant risk factor for the development of T2DM (Table 1). The consistency of finding through these cohort studies suggests a true association. This is supported by a meta-analysis where a HR of 2.22 (95% CI 1.84-2.6) for incident T2DM was found in the presence of imaging diagnosed NAFLD. [29] However, it is important to note that there is significant inter-study variation in the quantification of the risk attributed to NAFLD. This is most likely attributable to the wide variation in study methodology including differing diagnostic assessment of NAFLD (though most studies use ultrasonography), the lack of differentiation in most studies between NAFLD severity, different lengths of follow up, variation in diagnostic criteria for T2DM and varying and frequently incomplete confounder adjustments. [30] In addition, all except two studies were confined to Asian populations, so caution should be used when extrapolating to other populations. Nonetheless, evidence from these cohort studies supports the hypothesis that NAFLD may contribute to or be closely associated with the development of T2DM. Furthermore, 3 studies assessed incident diabetes in comparison to severity of liver disease. Two found increasing

incidence of T2DM in participants whose liver steatosis was more significant. [31,32] In the only study looking at participants with biopsy proven NAFLD, participants with NASH had a threefold increased risk of developing T2DM over mean 13 year follow up when compared with those with simple steatosis. [33] Interestingly, two studies looked at a subgroup of participants where liver steatosis had regressed during the period of follow-up. Importantly, in these subgroups, the rate of incident T2DM was lower than in the population where steatosis had persisted suggesting a potential causal association although this is unproven. [34,35]

Finally, there is epidemiological evidence to support an association between severity of NAFLD and worsening glycaemic control, macrovascular and microvascular complications of diabetes. [36] The Valpollicella Heart Diabetes Study looked at 2103 patients with T2DM, free of cardiovascular disease or viral hepatitis. There was an association between USS diagnosed liver steatosis and renal disease (odds ratio (OR) 1.87 95% CI 1.3-4.1), and diabetic retinopathy (OR 1.75 95% CI 1.1-3.7) compared to those with T2DM and no NAFLD, even after correcting for age, sex, BMI, waist circumference, diabetes duration, glycated haemoglobin (HbA1c), lipids, hypertension, smoking status and medication use. [37] Macrovascular risk was also increased. [37] Two studies examined the link between liver steatosis and glycaemic control. One which examined inter-individual variation in insulin requirements of patients taking once daily insulin and metformin, showed that hepatic fat correlated with daily insulin dose and ability of intravenous insulin to suppress endogenous glucose production. [38] Another, examining a cohort of 300 patients with impaired glucose regulation, found the presence of liver fat on MRS imaging was a significant predictor for failure of lifestyle management for glycaemic control. [39] Additionally, one retrospective study of 337 people with diabetes showed increased mortality if there was a concurrent diagnosis of NAFLD (HR 2.2 95% CI 1.1-4.2). [40]

Putative pathogenic mechanisms to explain how hepatic steatosis can exacerbate T2DM have been put forward. The link between hepatic steatosis and insulin resistance is seen in its most pronounced form in lipodystrophy patients where lipids accumulate in ectopic tissue (including liver) due to the lack of adipose tissue. Interestingly, if adipose tissue is transplanted into lipodystrophic mice, permitting lipid

Table 1. Studies Investigating Incident Diabetes in the Context of NAFLD

Study	Study Type, Country (n)	Modality used to diagnose NAFLD	Mean Follow-up (years)	Criteria for Diagnosis of Diabetes	Diabetes Diagnosis at baseline (%)	Main Findings	Adjustments for confounders
Okamoto et al. 2003 [41]	Retrospective cohort, Japan (840)	Ultrasound	10	Fasting glucose ≥7.8mmol/l/ HbA1c ≥6.5%/ drug treatment	0	Steatosis not associated with incident T2DM (HR 1.83 95% CI 0.95-3.51)	age, alcohol intake, BMI, family history of T2DM, fasting glucose, HbA1c, sex
Ekstedt et al. 2006 [33]	Retrospective cohort, Sweden (129)	Biopsy	14	2h 75g OGTT ≥11.1mmol/l/ treatment (diet or drug)	8.5	Number with T2DM increased from 8.5% to 58%. Further 20% had developed impaired glucose tolerance.	
Fan et al. 2007 [42]	Case-control, China (358 cases, 788 control)	Ultrasound	7	Fasting glucose ≥7.8mmol/l/ clinical history/ drug treatment	14.0 (NAFLD) 2.8 (control)	Steatosis associated with incident T2DM (OR 4.6 (95% CI 3.0-7.1))	age, BMI, sex
Shibata et al 2007 [43]	Retrospective cohort, Japan (3189)	Ultrasound	4	Fasting glucose ≥7.0mmol/l/ 2hour 75g OGTT ≥11.1 mmol/l	0	Steatosis associated with incident T2DM (HR 5.5 (95% CI 3.6-8.5))	age, BMI

Study	Study Type, Country (n)	Modality used to diagnose NAFLD	Mean Follow-up (years)	Criteria for Diagnosis of Diabetes	Diabetes Diagnosis at baseline (%)	Main Findings	Adjustments for confounders
Kim et al. 2008 [31]	Retrospective cohort, S Korea (5372)	Ultrasound	5	Fasting glucose ≥7.0mmol/l/ clinical history/ drug treatment	0	Steatosis associated with incident T2DM (RR 1.51 (95% CI 1.04-2.2))	age, ALT, BMI, family history of T2DM, fasting glucose, lipids, sex, smoking, triglycerides
Yamada et al. 2010 [44]	Retrospective cohort, Japan (12,375)	Ultrasound	5	Fasting glucose ≥7.0mmol/l/ clinical history/ drug treatment	0	Steatosis associated with incident impaired fasting glucose or T2DM in men (OR 1.91 (95% CI 1.56-2.34)) and women (OR 2.15 (95% CI 1.53-3.01))	age, alcohol, BMI, bp, family history of T2DM, sex, smoking
Bae et al. 2011 [45]	Retrospective cohort, S Korea (7849)	Ultrasound	5	Fasting glucose ≥7.8mmol/l/ HbA1c ≥6.5%/ drug treatment	0	Steatosis associated with incident T2DM if pre-existing impaired fasting glucose (HR 1.33 (95% CI 1.07-1.66))	age, alcohol, BMI, bp, impaired fasting glucose, lipids, physical activity, sex, smoking
Sung et al. 2012 [46]	Retrospective cohort, S Korea (12,853)	Ultrasound	5	Fasting glucose ≥7.0mmol/l/ clinical history/ drug treatment	0	Steatosis associated with incident T2DM (OR 2.42 (95% CI 1.74-3.36))	age, alcohol, ALT, BMI, educational status, HOMA-IR, lipids, physical activity, sex, smoking

No text

Study	Study Type, Country (n)	Modality used to diagnose NAFLD	Mean Follow-up (years)	Criteria for Diagnosis of Diabetes	Diabetes Diagnosis at baseline (%)	Main Findings	Adjustments for confounders
Park et al. 2013 [32]	Prospective cohort, S Korea (25,232 Male only)	Ultrasound	5	Fasting glucose ≥7.0mmol/ HbA1c ≥6.5%/ clinical history	0	Steatosis associated incident T2DM- moderate/severe steatosis (HR 1.73 (95% CI 1.0-3.0))	age, bp, creatinine, CRP, family history of T2DM, HOMA-IR, lipids, metabolic syndrome, physical exercise, waist circumference
Kasturiaratne et al. 2013 [47]	Prospectie cohort, Sri Lanka (2276)	Ultrasound	3	Fasting glucose ≥7.0mmol/l/ clinical history/ drug treatment	0	Steatosis associated with incident T2DM (HR 1.64 (95% CI 1.2-2.2))	age, ALT, BMI, bp, family history of T2DM, impaired fasting glucose, lipids, sex, waist circumference
Chang et al. 2013 [48]	Prospective cohort, S Korea (38,291)	Ultrasound	5	Fasting glucose ≥7.0mmol/ HbA1c ≥6.5%/ drug treatment	0	NFS associated with incident T2DM- high NFS (HR 4.74 (95% CI 3.7-6.1))	age, alcohol, CRP, exercise, family history of T2DM, HOMA-IR, lipids, sex, smoking
Choi et al. 2013 [49]	Retrospective cohort, S Korea (7849)	Ultrasound and liver function tests	5	Fasting glucose ≥7.8mmol/l/ HbA1c ≥6.5%/ drug treatment	0	Steatosis and ALT associated with incident T2DM (HR 1.64 (95% CI 1.3-2.1))	age, alcohol, BMI, bp, exercise, impaired fasting glucose, lipids, sex, smoking
Sung et al. 2013 [34]	Retrospective cohort, S Korea (13,218)	Ultrasound	5	Fasting glucose ≥7.0mmol/l/ clinical history/ drug treatment	0	Steatosis associated with incident T2DM- new (OR 2.49 (95% CI 1.5-4.1)), worsening (OR 6.13 (95% CI 2.6-14.7))	age, alcohol, BMI, bp, exercise, glucose, insulin, lipids, liver function tests, sex, smoking

Study	Study Type, Country (n)	Modality used to diagnose NAFLD	Mean Follow-up (years)	Criteria for Diagnosis of Diabetes	Diabetes Diagnosis at baseline (%)	Main Findings	Adjustments for confounders
Ming et al. 2015 [50]	Retrospective cohort, China (508)	Ultrasound	5	Fasting glucose ≥7.0mmol/l/ 2hour 75g OGTT ≥11.1 mmol/l/ drug treatment	0	Steatosis associated with incident T2DM (HR 4.46 (95% CI 1.9-10.7))	age, alcohol, BMI, bp, education, exercise, family history of T2DM, fasting glucose, lipids, oral glucose tolerance test, sex, smoking,
Yamazaki et al. 2015 [35]	Retrospective cohort, Japan (3074)	Ultrasound	11	Fasting glucose ≥7.0mmoll/ hba1c ≥6.5%/ clinical history/ drug treatment	0	Steatosis associated with incident T2DM (OR 2.37 (95% CI 1.6-3.5))	age, BMI, bp, exercise, family history of T2DM, impaired fasting glucose, lipids, sex
Li et al. 2015 [51]	Retrospective cohort, China (4736)	Ultrasound	4	Fasting glucose ≥7.0mmol/l/ clinical history/ drug treatment	0	Steatosis associated with incident T2DM (HR 3.37 (95% CI 2.4-4.3))	age, ALT, bp, creatinine, lipids, sex, uric acid
Shah et al. 2015 [52]	Prospective cohort, US (3153)	CT	9	Fasting glucose ≥7.0mmol/l/ clinical history/ drug treatment	0	Steatosis associated with incident T2DM (HR 2.06 (95% CI 1.5-2.8))	age, BMI, bp, CRP, ethnicity, exercise, family history of T2DM, glucose, sex, waist circumference

Study	Study Type, Country (n)	Modality used to diagnose NAFLD	Mean Follow-up (years)	Criteria for Diagnosis of Diabetes	Diabetes Diagnosis at baseline (%)	Main Findings	Adjustments for confounders
Fukuda et al. 2016 [53]	Retrospective cohort, Japan (4629)	Ultrasound	13	Fasting glucose ≥7.0mmoll/ hba1c ≥6.5%/ drug treatment	0	Steatosis associated with incident T2DM – normal BMI (HR 3.59 95%CI 2.14-5.76), raised BMI HR 6.77 (95%CI 5.17-8.91)	age, alcohol, exercise, family history of T2DM, HbA1c, sex, smoking

accumulation in transplanted adipocytes as opposed to liver, insulin resistance improves, suggesting a causal link. [54,55] Furthermore, in high fat diet rat models,hepatic steatosis can be seen at day three before the development of obesity, while insulin resistance post-dates hepatic steatosis but pre-dates obesity. [30,55,56] Of mechanisms postulated, the most consistent theme is a role for hepatic di-acylglycerol (DAG) and protein kinase C (PKC)ε. It is thought that hepatic lipid leads to an increase in hepatic DAG which activates hepatic PKCε. This is associated with decreased activation of the insulin receptor and subsequent insulin resistance. Supporting this are studies in rat and mouse models where if PKCε expression is inhibited or knocked-out, decreased hepatic insulin resistance is seen in the high fat diet model. [56,57] Likewise, in a study that looked at reversal of insulin resistance, that reversal was associated with reduced DAG and PKCε expression. [58] However, one study where DAG was overexpressed in mice lead to triglyceride accumulation but no increased insulin resistance, suggesting that the mechanism is multifactorial. [59] This can be extrapolated to humans where biopsy samples in obese participants show that hepatic DAG and PKCε activation (as measured by mass spectrometry) were the strongest predictors of insulin resistance (as measured by HOMA-IR). [60,61] Other suggested mechanisms have included the possibility that liver inflammation and inflammatory pathways may inhibit phosphorylation of the insulin receptor leading to hepatic insulin resistance. Rat studies have shown that activation of nuclear factor kappa B (NF-κB) can lead to insulin resistance and knocking-out inhibitor of nuclear factor kappa-B kinase subunit beta (IKK-β) can lead to diminished insulin resistance. [62,63] However, inconsistent results have been obtained in other studies. [30,64] Alternatively, it has been proposed that fetuin-B, a hepatokine, may be implicated. Fetuin-B is differentially secreted in NAFLD compared to control individuals, and increased secretion is seen in patients with NAFLD and T2DM. Mechanistically, in vitro studies have shown that fetuin-B can impair insulin action, and in vivo mouse experiments have indicated that expression can lead to impaired glucose tolerance, with silencing of fetuin-B leading to improved glycaemic profile. [65] Lastly, it is known that ceramides can be increased in NAFLD and be deposited in hepatocyte cell membranes, interrupting the insulin receptor. Certainly, two rodent models have shown that inhibiting ceramide synthesis can improve glucose tolerance in obese models. [55,66,67]

1.4.2 Evidence that the presence of T2DM is associated with increasing prevalence of and progression of disease in NAFLD

It is widely acknowledged that the prevalence of NAFLD is increased in the T2DM population. Overall European population median prevalence is 25-26%, rising to 40-70% in T2DM populations. [6,9-11,37,68] Furthermore, studies have shown increasing prevalence of not just steatosis but NASH, fibrosis, cirrhosis and NAFLD related HCC in people with T2DM compared to those without. In biopsy samples taken at the time of gastric bypass surgery the odds of finding NASH were 128 times greater (95% CI 5.2-13137.0) and severe fibrosis 75 times greater (95% CI 4.5-123.7) if the patient had T2DM compared to no T2DM. [16] Similarly, biopsies undertaken in an unselected NAFLD cohort found the prevalence of cirrhosis to be 25.5% in participants with diabetes, compared to 10.2% in those without (p=0.04) and another study has shown increased prevalence of cirrhosis in those with T2DM compared to people without diabetes. [69,70]

Many cohort studies have investigated potential risk factors for progression of NAFLD (Table 3). The influence of baseline factors on the speed of and risk of disease progression is discussed in section 1.7. From these studies, approximately 50% found diabetes at baseline to be associated with more frequent progression of disease. In addition, one large biobank study and one large community study have shown strong associations between progression of NAFLD to incident cirrhosis and HCC and the presence of diabetes at baseline. [21,71] One meta-analysis showed a statistically significant increased risk of HCC in those with diabetes compared to those without (RR 2.31, 95%CI 1.87-2.84), a finding replicated in a large community prospective cohort study where the presence of T2DM was associated with a HR or 2.96-7.52 of developing a diagnosis of HCC, interestingly with duration of diabetes being associated with increasing risk. [72,73] Furthermore, one recent multicentre international study of 200 individuals with Childs-Pugh A cirrhosis showed that those who also had T2DM had increased rates of decompensated liver disease (6.6/100 person years compared to 4.2/100 person years p<0.01) and HCC (3.1/100 person years compared to 1.2/100 person years p<0.01). [74]

Several epidemiological studies have investigated mortality in NAFLD cohorts. In these studies, diabetes has consistently been found to be a risk factor for mortality. One, looking at people with cirrhosis quoted a rate of 4.9/100 person years in those

with T2DM compared to 3.0 per 100 person years in those without ($p<0.01$) . [69,74-79] Similarly, another study using population data has shown that T2DM is associated with an increased risk of both hospital admission for liver disease and death from liver disease. [80]

The pathogenic mechanisms which determine the link between T2DM and NAFLD are not fully understood. Possible pathways by which insulin resistance and diabetes contribute to the development of hepatic steatosis are discussed in the pathogenesis section of this thesis (section 1.5). Supporting evidence from mouse models suggests a causal rather than simply associative role. Importantly, Lo et al. showed that if mice fed a high fat diet were compared to those fed a high fat diet and rendered diabetic with streptozotocin, those with diabetes developed much more significant hepatic fibrosis. [81] In addition, Guimaraes et al. showed in 2010 in in-vitro experiments that hepatic stellate cells (HSC) express receptors for advanced glycosylation end products (AGEs) and that AGEs can stimulate HSC production of reactive oxygen species (ROS). [82] Another group, also using in-vitro methodology have shown that raised glucose can induce HSC proliferation and activation. [83] This suggests a mechanism by which patients with T2DM (in whom AGEs accumulate) may develop more severe liver inflammation and fibrosis than those without.

1.5 The pathogenesis of NAFLD

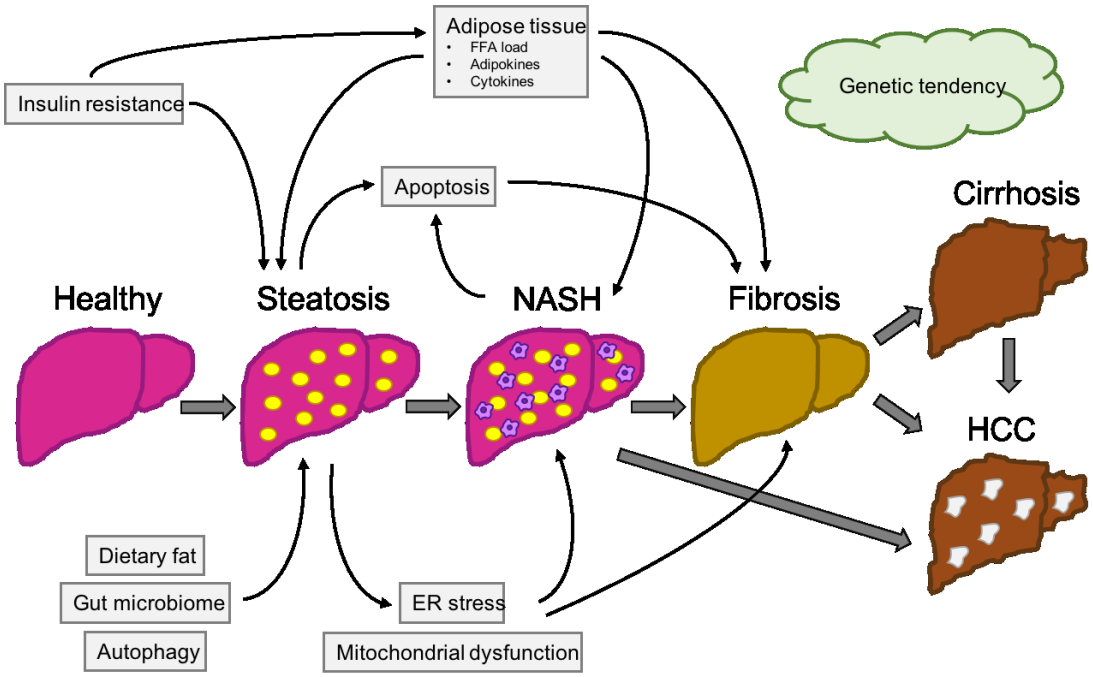
The pathophysiology of NAFLD has not been fully elucidated. NAFLD was first identified when it was noted that there was a consistent association between the metabolic syndrome (defined as ≥3 of raised waist circumference, raised triglycerides, low HDL cholesterol, raised blood pressure and evidence of impaired glucose regulation or T2DM) and the development of 'cryptogenic cirrhosis'. [84,85] NAFLD is now widely accepted as the 'liver component' of the metabolic syndrome.

As previously mentioned, NAFLD comprises a spectrum of disease extending from isolated steatosis (NAFL) to end-stage cirrhosis. As a disease entity NAFL can progress to NASH, fibrosis and cirrhosis but this pathway is not always direct, for example it is thought that NASH can develop in the absence of NAFL. In addition, fibrosis, NASH and NAFL are all, to an extent, reversible, as evidenced in natural history studies (section 1.6). Lastly the precise mechanisms by which NAFL, NASH and fibrosis develop are not fully confirmed. Putative models of NAFLD pathogenesis have been suggested and will be discussed. It is likely that NAFLD is a multifactorial disease with many components contributing to progression or regression of disease (Figure 1).

In 1998, Day and James proposed the '2 hit hypothesis' for NAFLD pathogenesis. [86] In this model, the accumulation of liver triglyceride was considered to be a 'first hit' that sensitised the liver to 'second hits' such as oxidative stress, lipid peroxidation, and mitochondrial dysfunction. It was postulated that is was to be the second hit that led to hepatocyte inflammation, damage and fibrosis. Subsequent studies have suggested that this hypothesis may not hold.

Firstly, some studies have shown that triglyceride accumulation may not necessarily predispose to liver inflammation, damage and fibrosis. Triglyceride accumulation in itself could well be protective and it is probably the production of lipotoxins as part of dysfunctional hepatic lipid metabolism that leads to hepatic damage. [87,88] One study inhibited an enzyme in the final common pathway of liver triglyceride synthesis in mice fed a high fat diet. Although less steatosis was evident, these mice developed worsening hepatic inflammation and fibrosis in comparison to controls. [89] This phenomenon has not been consistently replicated. [90] Secondly, studies have shown an association between the level of steatosis and the chances of progression to NASH

Figure 1. Factors Associated with the Development of NAFLD and progression to Cirrhosis and HCC



and liver fibrosis suggesting that liver steatosis, or a factor associated with it, is contributing to hepatocyte damage and it is not the innocent bystander suggested in the initial 2 hit hypothesis. [88,91]

NAFLD is now generally considered to be a multifactorial disease process whereby multiple parallel 'hits' including insulin resistance, disrupted lipid metabolism, lipotoxicity, oxidative stress, mitochondrial dysfunction, endoplasmic reticulum (ER) stress, autophagy, gastrointestinal endotoxins, alcohol, altered cytokine and adipokine signalling and genetic predisposition lead to hepatocyte inflammation, damage and progressive liver disease, hepatic fibrosis, cirrhosis and HCC (Figure 1). [92]

I will discuss in turn proposed contributory factors to the pathogenesis of NAFLD.

1.5.1 The development of steatosis in NAFLD

Hepatic steatosis is a key histopathological feature of NAFLD. [93] Its presence is a reflection of disordered hepatic lipid homeostasis. Hepatic fat is normally stored in the form of triglycerides. The liver derives lipid from 3 sources- triglyceride remnant from dietary lipid not absorbed into chylomicrons, circulating free fatty acids (FFA) derived from adipose tissue lipolysis and de novo lipogenesis which is primarily regulated by

insulin and glucose at a transcriptional level. One isotope study showed a contribution from all 3 sources in NAFLD with 59% hepatic triglyceride derived from serum, 26% from de novo lipogenesis and 15% from dietary sources. [94] This is especially notable with the increased proportion derived from de novo lipogenesis, thought to contribute to only 5% hepatic fat in people without NAFLD. [93]

Lipids can be processed by hepatocytes via several pathways. They can be used in phospholipid synthesis or re-esterified to create triglycerides and then very low density lipoprotein (VLDL) which can then be secreted from hepatocytes. Alternatively, they can undergo β-oxidation in mitochondria, oxidation in peroxisomes or microsomal oxidation in the ER (via p450) leading to adenosine triphosphate production, but at the expense of the production of reactive oxygen species (ROS). It is postulated that hepatic steatosis develops when VLDL secretion and oxidation of FFA are unable to keep up with the hepatic lipid overload and it is excessive FFA that are esterified into triglycerides and stored in lipid droplets. [93] Fitting with this, increased rather than decreased VLDL secretion is seen in NAFLD. [87] Interestingly, one study showed increased VLDL secretion in NAFL, but decreased in NASH suggesting that failure of VLDL secretion may contribute to the progression of disease from NAFL to NASH. [95]

Influences on the development of hepatic steatosis- insulin resistance

Insulin resistance is seen in most patients with NAFLD. [96,97] In euglycaemic clamp studies, insulin resistance correlates with intrahepatic triglyceride concentration. [26,98,99] It is thought to contribute to the development of hepatic steatosis via three mechanisms. Firstly, in health insulin acts on the liver to inhibit hepatic glucose production, stimulate hepatic glucose uptake and promote hepatic de novo lipogenesis. [93] In the context of insulin resistance, although the inhibitory effects are diminished, the stimulatory effect on de novo lipogenesis seems to be retained due to postulated differential enzymatic pathways. [100] In addition, hyperinsulinaemia and hyperglycaemia are thought to stimulate further hepatic de novo lipogenesis via stimulation of the carbohydrate and sterol response element binding protein transcription factors. [87,101,102] Secondly, in the context of insulin resistance, FFA uptake, metabolism and lipolysis in adipose tissue is altered, leading to increased FFA release and delivery to the liver. [103-106] Thirdly, insulin resistance in skeletal muscle leads to decreased muscle glucose uptake, leading to the increased delivery

of glucose to the liver. [93,98,99] The likely significant contribution of the above three mechanisms to the development of hepatic steatosis has led to the consensus that insulin resistance may be a primary pathological determinant of NAFLD.

Influences on the development of hepatic steatosis- dysfunctional adipose tissue

It is accepted that there is a correlation between increased visceral fat and NAFLD. It has been shown that rising visceral fat measurements correlate with the extent of hepatic inflammation and fibrosis, even when corrected for insulin resistance. [107] Excessive lipid storage in adipose tissue leads to adipose tissue hypertrophy and hyperplasia. Whilst thought to be initially protective as it stores excess FFA, it is also thought to impair normal adipose tissue function and stimulates altered gene expression, promoting the production of cytokines akin to those produced by macrophages. [87] FFA and lipid by-products (such as diacylglycerol and ceramide) stimulate the induction of multiple inflammatory pathways, including Ikκ-β and subsequently NF- κβ and c-jun N terminal protein kinase 1 (JNK1) signalling which lead to the production of multiple inflammatory mediators including TNFα, interleukin-6 (IL-6) and interleukin-1beta. These inflammatory mediators can interrupt insulin signalling, contributing to insulin resistance and then hepatic steatosis. [87,92] Additionally, they likely exert a direct effect on the hepatocyte in NAFLD; this will be discussed later in this review. Supporting this mechanism is the knowledge that IL-6 and TNFα are known to be increased in fat cells, increase in obesity and decrease with weight loss. [92,108,109] In addition, in murine models, JNK knockout mice fed with a high fat diet do not develop hepatic insulin resistance. [110-112]

The secretion of adipokines from adipocytes is known to be altered in obesity and may play a role in the development of steatosis. Adiponectin is normally secreted by adipose tissue and acts in the hepatocyte to stimulate FFA oxidation. [101] Systemic levels of adiponectin are reduced in obesity and NAFLD. [113] Studies have shown that adiponectin levels inversely correlate with liver steatosis, and in some studies hepatic inflammation and fibrosis. [113,114] Interestingly, in the ob/ob mouse model, the administration of adiponectin can alleviate NAFLD. [115] Leptin is also thought to play a role. Leptin deficient ob/ob mice are known to develop hepatic steatosis and it has been shown that leptin administration can reduce triglyceride levels. [93,116] However,

most people with NAFLD have increased systemic levels of leptin so its exact role is unclear, although it is postulated that leptin resistance may contribute to the picture.[93]

Influences on the development of hepatic steatosis- dietary intake of fructose, gastrointestinal microbiota and gastrointestinal tract permeability

Dietary intake of fructose has been shown in many studies to be associated with the development of NAFLD steatosis in humans.[117] Fructose seems to be associated with a tendency towards the development of liver steatosis and inflammation above and beyond what would be expected from its calorific content. Animal studies have also shown increased liver steatosis when given a fructose rich diet compared to those without, despite equal total energy intake.[118] Similarly, mice who have no fructokinase enzyme are more resistant to the development of liver steatosis.[119] This is thought to be mediated firstly by liver metabolism of fructose which can lead to decreased ATP and increased uric acid, both of which are thought to be related to the development of hepatic steatosis and inflammation.[117] Secondly, the metabolism of fructose in the small intestine can lead to increased gut permeability, the effects of which will be discussed in the following paragraph.[117]

Obesity is associated with increased intestinal mucosal permeability, and NAFLD is associated with an altered composition of gut microbiota, most consistently with the decreased presence of bacteroides. [101,120-122] The role of the gut microbiome in the metabolic syndrome has been illustrated in mouse studies where it has been demonstrated that germ free mice develop less total body fat on the same diet; and when exposed to a gut microbiota from obese mice, normal mice develop a 60% increase in body fat and insulin resistance. [123] In addition mice fed with a high fat diet and treated with antibiotics show reduced hepatic triglyceride accumulation. [124] The mechanism by which altered microbiota and gastrointestinal tract permeability can impact on the development of steatosis is not fully defined. It is postulated that these changes can lead to increased serum ethanol levels (with increased abundance of ethanol producing bacteria in NAFLD microbiomes) and decreased choline (as it is known that specific microbial enzymes are required for choline conversion to trimethylamine and subsequent absorption). [93,125] Ethanol is well known to contribute to hepatic steatosis and hepatocyte damage in alcoholic liver disease, and choline deficient diets are commonly used to generate NAFLD models in rodents. [93] Adding weight to this argument are studies that show raised serum ethanol levels in NASH

compared to healthy control and obese participants without NASH, and increased expression of ethanol metabolism genes in NASH patients compared with controls.[125,126]

Influences on the development of hepatic steatosis- genetics

NAFLD prevalence varies by ethnicity, suggesting a potential genetic influence in NAFLD development. [127] Supporting this, twin studies have shown correlation between the development of NAFLD in monozygotic but not dizygotic twins. [128-130]

Furthermore, genome wide association studies have identified several candidate gene associations with NAFLD. [131] It is beyond the remit of this review to discuss all candidate genes, but the best defined, the rs738409 minor allele variant of PNPLA3 and rs58542926 minor allele variant of the TM6SF2 gene, merit mention. These gene variants are the only two to be identified as to be associated with NAFLD in two or more studies. [132]

Genome wide association studies of participants with MRS, biopsy or CT diagnosed NAFLD have all identified an association with the rs738409 SNP of the PNPLA3 gene. [131,133,134] Further studies have also found the rs738409 SNP to be associated with the histological severity of NAFLD, even after adjustment for metabolic risk factors. [135-137] In addition, studies including a meta-analysis have shown rs738409 to be independently associated in multivariate analysis as an independent predictor of HCC occurrence in both NAFLD and ALD. [132,138,139]

The rs58542926 polymorphism in the TM6SF2 gene has been shown to be significantly associated with NAFLD, and additionally has been identified to be independently associated with increased fibrosis in studies comparing studies with biopsy proven NASH and fibrosis to healthy control participants. [132,135,140-142]

PNPLA3 codes for adiponutrin which is expressed in hepatocytes, adipocytes and hepatic stellate cells. It is thought to play a role in lipid droplet remodelling and VLDL secretion and in vitro affects phospholipase and triglyceride lipase. [93] Although mice studies with PNPLA3 knock-out have not been consistent, in vitro studies have suggested that presence of the RS738409 is associated with lipid accumulation. [93,137,143,144] Similarly, TM6SF2 is thought to play a role in VLDL secretion, thus adding

to evidence that impaired VLDL secretion may be a key factor in the development of NAFLD. [93,137]

Lastly, one recent large population wide genetics study looking at 46544 people has shown an association with a splice variant of the HSD17B13 gene, which encodes for a lipid droplet protein, is associated with decreased prevalence of NAFLD (as diagnosed by evidence of electronic hospital patient record), suggesting that genetic factors may confer protection from as well as predisposition to NAFLD. [145]

Influences on the development of hepatic steatosis- autophagy

Autophagy is the process by which lysosomes degrade intracellular proteins, organelles and lipids. In health it is thought to function to recycle cellular constituents, maintain cellular homeostasis and to provide energy in the context of starvation. [93] Mouse studies have shown that the inhibition of autophagy can lead to increase triglyceride storage in hepatocytes, and induction of autophagy can decrease lipid storage. [111] In addition, a pilot immunohistochemical study has shown a reduction in autophagy marker LC3 with increased steatosis. [146] Thus defective autophagy is postulated to contribute to the development of hepatic steatosis, although its exact function in the development of NAFLD is not fully understood.

1.5.2 The development of liver inflammation

NASH is a histological diagnosis characterised by lobular inflammation and hepatocellular ballooning. The progression of NAFL to NASH, fibrosis and cirrhosis is thought to be secondary to hepatocellular damage caused by a combination of factors resulting in oxidative stress and inflammation. [147]

Liver inflammation- the role of oxidative stress and mitochondrial dysfunction

Whether the presence of NAFL contributes to the development of NASH or NAFLD fibrosis has been controversial. However, it is increasingly thought that, whilst triglyceride accumulation in itself may not be hepatotoxic, and in fact may be a compensatory protective mechanism in response to triglyceride overload, that the presence of excess fat in hepatocytes can lead to the accumulation of fatty acid metabolites that act as lipotoxins, resulting in liver injury. [88,148] It is thought that, in the presence of excessive FFAs, increased mitochondrial oxidation leads to the increased production of ROS. In addition, in FFA excess, peroxisomes and microsomes are

recruited to oxidise FFA, a process with leads to proportionally greater ROS production than mitochondrial oxidation. [93] When ROS production exceeds the hepatocyte's antioxidant capacity this leads to intracellular oxidative stress, nuclear and mitochondrial DNA damage, phospholipid membrane disruption, release of pro-inflammatory cytokines and upregulation of apoptosis. [104] Immunochemical studies show evidence of oxidative stress and DNA damage and it is known anti-oxidant genes are upregulated in NAFLD, presumably as a compensatory mechanism. [93,149] Furthermore, there is well documented evidence of mitochondrial disruption in NASH and NAFLD fibrosis but not in NAFL suggesting that this contributes to the progression of disease. [85,96,150]

Liver inflammation- the role of ER stress

ER stress is thought to contribute to hepatocyte inflammation and stress in the context of NAFLD. ER stress, putatively caused by hyperlipidaemia and hyperinsulinaemia, affects normal translation leading to an 'unfolded protein response' whereby there is halting of normal translation, increased protein degradation and activation of inflammatory cascades including those that can aggravate insulin resistance. [92,93,101,151] Human histological studies have shown markers of the unfolded protein response in NASH and correlation of damage with histological severity of disease. [152]

Liver inflammation- the creation of an inflammatory milieu

In addition to intra-hepatocyte inflammation, distant inflammatory cascades are thought to contribute. Adipokines and the inflammatory cytokines produced by dysfunctional adipocytes are through to play a role, with adiponectin known to have an systemic anti-inflammatory effect and leptin has been shown to stimulate the release of pro-inflammatory cytokines from hepatocytes. [92,153] Secondly, changes in the gut microbiome may affect inflammation- with one study showing that the administration of a pro-biotic was associated with decreased inflammation and fibrosis compared to controls in a high fat diet rat model. [154,155]

Overall, increased expression of IKK-β, NF-κβ, TNFα and IL-6 are seen in mouse models. [63] In human studies increased serum IL-6 and TNFα are seen, and increased cytokine gene expression is seen in patients with NASH compared to obese control, with TNFα correlating with disease severity. [63,156-158] Conversely, inhibiting cytokine production in high fat diet and methionine-choline deficient mouse models attenuates

hepatic inflammation and insulin resistance. [153,159] Both intra-hepatic inflammation, increased hepatic FFA and systemic inflammation are thought to promote the pro-inflammmatory phenotype of the Kuppfer cell (resident liver macrophage) in NASH. Kuppfer cells in health act to remove pathogens (normally gut derived) and injured hepatocytes, but if overactivated can result in harmful inflammation. [87,93]

1.5.3 Progression to fibrosis, cirrhosis and hepatocellular carcinoma

Hepatic fibrosis (of which cirrhosis is the most advanced stage) is the final common pathway of pathogenesis for most chronic liver disease; and presence of fibrosis in the context of NAFLD or NASH confers a worse prognosis. [93,103] It is defined as a wound healing response which is characterised by the excessive deposition of collagen and extracellular matrix which leads to the formation of scar tissue. [93]

In health, the liver repairs damage through the duplication of mature hepatocytes. In chronic oxidative stress and injury, this is not seen and instead recruitment of hepatic progenitor cells is seen. [103] Additionally, HSCs are activated. These normally play a role in extracellular matrix homeostasis but are activated by free fatty acids, ROS and inflammatory cytokines, injured hepatocytes, hepatic progenitor cells and gut derived peptides to a myofibroblast-like phenotype. [93] It is important to note that in brief hepatic injury these mechanisms work to repair hepatic tissue, however in the context of persistent insult they seem to fail, and activation of HSCs leads to acquisition of fibrogenic potential. [93,103] Immunohistochemical and immunofluorescence studies have shown increased hepatic progenitor cells in human and murine models of NAFLD, and correlate with the extent of NASH and fibrosis. [160,161] Hepatic stellate cell activation has additionally been documented in humans in correlation with hepatic inflammation. [162] In murine models, inhibiting transforming growth factor beta cytokine expression can arrest fibrosis. [163,164]

In addition, the sustained inflammation and aberrant regeneration that are seen in NASH, fibrosis and cirrhosis lead to genetic and epigenetic events in the hepatocytes, the development of dysplastic nodules, preneoplastic lesions and hepatocellular carcinoma. [165]

1.6 The natural history of NAFLD

Better understanding of the natural history of NAFLD is needed to interpret the clinical relevance of specific findings and to develop potential interventional strategies. Presently the natural history of NAFLD is poorly understood and it is difficult to determine an individual's risk of developing cirrhosis when diagnosed with NAFLD. However, it is becoming increasingly clear that the longstanding viewpoint that NAFLD simple steatosis was a benign condition that did not progress to clinically significant disease is incorrect; simple steatosis, as well as NASH and NAFLD fibrosis can all progress to cirrhosis and HCC. [3-5,166-168] In addition, the presence of NAFLD is associated with an increased mortality from cardiovascular disease and all cause malignancy, in fact non-liver-related mortality is significantly more common that liver-related mortality for people with NAFLD. [17]

When examining the specific progression of NAFLD, the best available human data is found in studies looking at patients who have undergone sequential liver biopsies. Many studies have examined such data, discussed in more detail in further sections (Table 3). They report progression of NAFLD to increasing fibrosis, inflammation and cirrhosis in 25-53% subjects with meta-analyses reporting progression in 36%.[33,48,169-180] Importantly those people with steatosis only on initial biopsy were also susceptible to disease progression, though meta-analysis suggests that rate of disease progression on meta-analysis is more rapid in those with NASH on initial biopsy. [180] However in addition these studies also indicate the potential of disease regression with 15-30% participants seeing improved disease on repeat biopsy (20-22% on systematic review and meta-analysis). [33,169,171,173,174,177-180] It is important to note the limitations of these studies. The numbers of patients studied in all these studies is small, with a maximum number of patients in any one study being 132, and only a total of 411 patients in meta-analysis. Duration of follow-up, likely to be a key factor in disease progression, is very variable ranging from one to twenty two years. Additionally in most studies patients were recruited after at least the first, if not the repeat, biopsy. Thus, there will be a significant selection bias within these cohorts for those patients whose condition merited specialty review and assessment at the start of analysis and may well not reflect the natural history of a community population with undiagnosed NAFLD. Contributory to this, as participants were reviewed at specialist clinics they are likely to have been provided with undocumented lifestyle modification therapy, as would be routine clinical practice, which may impact on the natural history

progression seen. Lastly, as with any biopsy study, the potential sampling error for liver biopsy must be acknowledged.

Additionally, some cohort studies have examined clinical outcomes in NAFLD cohorts. Four studies document progression to cirrhosis. Adams et al. reported 9/103 of their cohort developing cirrhosis over the study period (range 0.7-21 years), Ekstedt et al reported 5.4% progression to cirrhosis with complications over mean 13 year follow-up, Dam-Larsen et al. reported a 1.2% period prevalence of cirrhosis over 20 year follow-up and Sebastiani et al. reported the incidence of varices in 10 and ascites in 13 of their 148 participants over a median 5 year follow-up. [33,171,181,182] Of note, only one study specifically identified NASH at baseline, all others were NAFLD at any pre-cirrhotic stage. Four studies also report on the development of HCC. Hashimoto et al. report an incidence of HCC in 11/359 over 5 years in their NASH cohort while Ascha et al. noted a yearly cumulative incidence of 2.6% in people with NAFLD cirrhosis. [183,184] Ekstedt et al. and Sebastiani et al. document progression to HCC in 3/129 and 1/148 of their unselected NAFLD cohorts. [33,185] Importantly, HCC can develop in pre-cirrhotic livers in the context of NAFLD, with one study identifying 50% people developing HCC in the context of NAFLD developing it in a non-cirrhotic liver. [186] Importantly, a population study using routinely collected data has shown that those who develop HCC in the context of NAFLD have a shorter survival time and more advanced tumour stage than those who develop HCC secondary to alternative pathology. [187] Interestingly, a similar study comparing outcome to those with HCC secondary to hepatitis C found no mortality difference when patients were matched but acknowledged that HCC in the context of NAFLD is often identified at a late stage.[186]

It has also been shown that NAFLD is associated with an increased risk of cardiovascular disease and malignancy. In one large cohort of people with T2DM (134,368 of whom 1,452 had NAFLD), the presence of NAFLD was linked to increased cardiovascular disease (HR 1.7 (95% CI 1.52-1.9), malignancy (HCC and other) and mortality (HR 1.6 (95%CI 1.4-1.83). [188] Seven cohort studies have looked at mortality in NAFLD populations. [33,76,181,189-192] Of these, 5 took an initial cohort of biopsy proven NAFLD, while 2 followed forward cohorts of patients with liver steatosis on USS imaging. There are obvious dichotomies between the diagnostic certainty of the biopsy cohort versus the more community-relevant nature of the USS cohorts. These studies showed that the main causes of mortality in this population were from

cardiovascular disease and malignancy, with Rafiq et al. documenting additionally an increase in liver related mortality, but only in their NASH cohort. Ekstedt et al. also demonstrated an increase in mortality in a NASH population compared to a reference population. [33] Additionally, supporting this Chang et al. showed an association between the presence of NAFLD on USS and increased coronary artery calcification score on CT (OR 1.1 (95% CI 1.05-1.16)). [193]

Thus, there is increasing evidence for NAFLD being a progressive condition that has an impact on population morbidity and mortality. However, the natural history is also dynamic and complex with varying rates of progression and evidence that regression can also occur. The precise timescale and nature of progression, and the reasons for the inter-study discrepancies in findings remaining unclear.

1.7 Risk factors that determine the rate of disease progression in NAFLD

A key interest in the NAFLD literature is the identification of factors that may be associated with disease progression. Identification of risk factors that accurately predict disease progression could enable targeted screening and surveillance of 'at risk' populations, in addition to forming potential therapeutic targets.

Both cross-sectional studies and cohort studies have been undertaken to try to identify risk factors for disease progression. Firstly, we will discuss the data findings from cohort studies.

1.7.1 Evidence from cohort studies

The literature search for cohort studies looking at disease progression was undertaken using a systematic search protocol. The Medline and Embase databases were searched. The search strategy involved a MeSH terms search for ('fatty liver OR non-alcoholic fatty liver disease') AND ('cohort studies OR follow-up studies OR longitudinal studies OR prospective studies OR retrospective studies') AND ('disease progression' OR 'prognosis' OR 'risk factors'). Results were confined to 'English' and 'Human' and studies from 1998-2020. After results were combined and deduplicated the search returned 1302 papers. Abstracts and full articles were reviewed. To be included the studies had to fulfil the following criteria that they were a cohort study, examining progression of NAFLD, examined risk factors for progression, participants were not on disease modifying treatment, participants were followed up for at least 1 year and participants had not undergone liver transplant. After abstract review, 39 papers were selected that fulfilled these criteria, and 32 of those were selected for analysis following full text review. In addition, two systematic reviews/ meta-analyses and their associated references were reviewed. [179,180] From these, two additional relevant cohort studies were identified and included in analysis. [174,194]

As described above, a total of 34 cohort studies were identified and have been included in this assessment of the literature (Table 3). Diagnosis was mostly defined by biopsy showing NAFLD. Four used a baseline USS diagnosis of NAFLD. [189,195-197] Two used ISD clinical diagnosis and one used a positive SteatoTest or FibroTest.[71,176,198] Other liver pathology was excluded.

Participants were followed up for a wide variety of durations, with range 1-41 years. Outcomes assessed in most studies were primarily either change in histology on repeat biopsy or mortality. Eleven studies each looked at the development of clinical cirrhosis or HCC. [33,71,171,181-184,198-201] Diabetes prevalence within the cohorts varied significantly.

The results of these studies have been highly variable, both with regards to rate of progression (section 1.6) and additionally in factors found to be associated with disease progression (Table 2). Importantly, time to follow-up varied significantly within and between studies, but only three reported on the effect of time with regards to disease progression, and two of these reported a link between follow-up time and progression of disease, so this must be taken into account. Although studies excluded participants based on internationally agreed alcohol excess cut-offs, only one study looked into the relationship between disease progression and any lifetime alcohol intake, finding an association between any lifetime alcohol intake and risk of progression. [184]

There are benefits and limitations of these cohort studies. These are the only studies that have looked at disease progression over time in humans. In addition, some studies have followed up patients for >20 years, a substantial follow-up period. However, all but five studies examine populations who were referred to secondary care for concerns regarding liver disease. Thus, in many studies there is a selection bias towards patients who are likely to have more severe disease. This is introduced again in retrospective biopsy studies where the repeat biopsy was undertaken for clinical concern, thus again selecting for patients with a potentially clinically more severe disease. There are also limitations with diagnostic techniques used. Whilst biopsy is the gold standard diagnostic test for NAFLD, NAFLD histopathology is known not to be consistent throughout the liver and so sampling error can be significant. In those studies not using biopsy but instead using non-invasive imaging or markers, there is concern that they may not reliably diagnose all patients, and it is not possible to distinguish between simple steatosis and NASH with non-invasive diagnostics. Lastly, follow up between studies is variable with regard to follow up time, routine care (including potentially disease modifying lifestyle advice), and end points assessed. All these may affect end points and the ability to compare predictive factors.

Table 2. Factors identified in cohort studies associated with disease progression in NAFLD

Baseline Factor		Number studies - factor associated	Number studies- factor not associated
Age		11	9
Gender		1	13
Race		1	4
BMI		3	14
Increasing Weight		1	0
Smoking		2	5
Diagnosed Diabetes		10	10
Hypertension		1	11
Hyperlipidaemia		2	12
Metabolic syndrome		1	3
Histology	↑ fibrosis	10	2
	Steatosis	1	9
	↑ inflammation	4	5 no association
	↓ inflammation	1	
NAFLD scores	Fibrotest/ steatotest	1	0
	FIB-4	5	2
	NFS	3	2
	APRI/ BARD	2	2
Liver Biochemistry	AST	2 high 1 low	7
	ALT	3	10
	AST:ALT	3	3
	ALP	2	4
	Bilirubin	2	4
	GGT	0	6
	Albumin	5	4
	PT/ INR	2	3
	Platelets	5	3
	Ferritin	1	5
Diabetes Biochemistry	HbA1c	1	4
	HOMA-IR	1	4
	Random/fasting glucose	1	3
Other: TNFα, adiponectin, IL-6, CRP, leptin, CK-18, IgA, IgG, uric acid		0	2 for all

FIB-4 Fibrosis-4 score, **NFS** NAFLD fibrosis score, **APRI** AST to platelet ratio index, **BARD** BARD fibrosis score, **PT** prothrombin time, **INR** international normalised ratio, **TNFα** tumour necrosis factor alpha, **IL-6** interleukin 6, **CRP** C reactive protein, **CK-18** cytokeratin 18 fragment, **IgA** immunoglobulin A, **IgG** immunoglobulin G

There is one systematic review and one meta-analysis in the literature, both of which analyse a subsection of these cohort studies. [179,180] Argo et al. primarily analysed older studies, not included in this analysis. They identified inflammation on biopsy and age to be the only factors significantly associated with disease progression. Singh et al. looked at 11 studies, totalling 2145.5 person-year follow-up. They noted an increased proportion of their cohort developing progressive disease if NASH was present at baseline compared to NAFL, though there was evidence of progressive disease in both cohorts. On meta-analysis they identified hypertension and a low AST:ALT ratio to be the only features associated with progressive disease. Note was made of the limitations of the cohort studies, discussed above, and in addition concern was noted that most studies included were at least moderate risk of bias. [180]

The Edinburgh Type 2 Diabetes cohort study (ET2DS) examined development of clinically significant liver disease in a T2DM population. [202] Although it is not included above because it investigated all-cause liver disease, 14/15 incident cases of clinically significant liver disease were at least partially attributed to NAFLD. Baseline factors assessed were age, gender, index of multiple deprivation, duration of diabetes, fasting Glucose, HbA1c, antihyperglycaemic treatment, BMI, cholesterol, triglycerides, IL-6, TNFalpha, CRP, ALT, AST, AST:ALT, GGT, CK-18, % steatosis on USS, APRI, ELF, FIB-4, HA, NFS and platelets. Of these, SIMD, insulin use, BMI, IL-6, TNFalpha, CRP, ALT, AST, AST:ALT, GGT, CK-18, APRI, ELF, FIB4, HA, NFS were associated with development of clinically significant liver disease over 4-6 year follow up.

1.7.2 Evidence from cross-sectional studies

In addition to the cohort studies discussed above, several cross-sectional studies have been undertaken to identify factors that are associated with cohorts with more significant disease at a single point in time. [14,16,190,203-211] As with the cohort studies, the numbers of participants in each study was small, and there is wide ranging discrepancy with regards to which factors, if any, associate with more advanced disease. Four studies did examine findings when patients had liver biopsy at the time of bariatric surgery. [14-16,210] Whilst this population is not necessarily representative of the general population, it is interesting to examine the results in an asymptomatic population. Even in this small subsection of studies there was no inter-study correlation between examined baseline factors and disease severity, although three out of four studies found T2DM to be associated with more severe disease. However, an appreciable concern with cross-sectional data in this context is the fact that factors

that may be associated with the presence of more advanced disease at a single time point may well not be those which determine or predict the rate or significance of progression of the disease.

1.7.3 Genetic association with disease progression

Although none of the cohort or cross-sectional studies mentioned above examined genetic factors linked to disease progression, there is increasing interest in the role of genetic factors in the progression of NAFLD. As discussed in the pathogenesis section (section 1.5), there is evidence to suggest that two genes in particular, namely the RS738409 SNP of the PNPLA3 gene and RS 58542926 of the TM6SF2 gene, may play a role in the development of NAFLD and are associated with a more advanced disease phenotype. [132,134-137,140] One systematic review has shown an association between the RS738409 SNP of the PNPLA3 gene and progressive fibrosis. [138] In addition, in one study looking exclusively at patients with T2DM, evidence of fibrosis on fibrotest in a non-selected population of people with T2DM was associated with increasing prevalence of SNP RS 738409 PNPLA3. [212]

1.7.4 Summary

Thus, in summary, although many potential associations with NAFLD disease progression in humans have been identified no clear cut, dominant baseline factors have yet emerged.

Table 3. Cohort Studies Assessing Risk Factors for Progressive NAFLD. (note not all factors assessed in every study, only those reported documented

Reference	Type of study/ Country	Diabetes at baseline (%)	No. participants	Baseline diagnosis	Length of follow up (years)	Outcome	Baseline factors associated with progressive disease	Baseline factors not associated with progressive disease
Matteoni et al. 1999 [194]	retrospective cohort/ USA	39-60	98	Biopsy NAFLD	average 8.3y +/- 5.4	Mortality	fibrosis, steatonecrosis	-
Harrison et al. 2003 [169]	retrospective cohort/ USA	41	22	Biopsy NAFLD	mean 5.7 (range 1.4-15.7)	Increase in biopsy fibrosis score	-	age, albumin, ALT, AST:ALT, bilirubin, BMI, bp, diabetes, ethnicity, gender, lipids
Fassio et al. 2004 [170]	prospective cohort/ Argentina	36	22	Biopsy NASH	mean 4.3 (range 3.0-14.3)	Increase in biopsy fibrosis score	obesity	age, albumin, ALT, AST:ALT, diabetes, gender, inflammation on biopsy, lipids, prothrombin time, steatosis
Adams et al. 2005 [171]	retrospective cohort/ USA	42	103	Biopsy NAFLD	mean 3.2 (range 0.7-21.3)	Increase in biopsy fibrosis score, Cirrhosis	BMI, low fibrosis stage	age, ALT, AST, AST:ALT, bp, diabetes, ferritin, gender, glucose, HOMA-IR, lipids, metabolic syndrome, NASH, obesity, platelets, prothrombin time, weight gain
Hui et al. 2005 [172]	prospective cohort/ Hong Kong	24	17	Biopsy NAFLD	mean 6.1 (range 3.8-8.0)	Increase in biopsy fibrosis score	-	Inflammation biopsy, steatosis

Reference	Type of study/ Country	Diabetes at baseline (%)	No. participants	Baseline diagnosis	Length of follow up (years)	Outcome	Baseline factors associated with progressive disease	Baseline factors not associated with progressive disease
Ekstedt et al. 2006 [33]	prospective cohort/ Sweden	9	129	Biopsy NAFLD	mean 13.7 (sd +/- 1.3)	Increase in biopsy fibrosis score, Mortality, Cirrhosis, HCC	periportal fibrosis	age, AST, ALT, ALP, bilirubin, BMI, bp, diabetes, gender, GGT, glucose, HOMA-IR, iron studies, lipids, metabolic syndrome, prothrombin time, TSH
Dam-Larsen et al. 2009 [181]	retrospective cohort/ Denmark	unknown	170 (82 return clinic)	Biopsy NAFLD	20.7 (range 0.1-27.9)	Mortality, Cirrhosis, HCC	albumin	ALT, AST, BMI, fibrosis, gender, GGT
Hashimoto et al. 2009 [183]	prospective cohort and case-control/ Japan	46	382	Biopsy NASH	mean 3.4 (range 0.5-15.1)	Mortality, HCC	age, albumin, ALP, ALT, AST, diabetes, fibrosis, hyaluronic acid, lipids, necroinflammation, platelets, prothrombin time, steatosis	-
Rafiq et al. 2009 [76]	retrospective cohort/ USA	29	173	Biopsy NAFLD	median (no NASH) 13.0, (NASH) 10.5 (range 5.0-28.5)	Mortality	age, albumin, ALP, NASH, T2DM	-

Reference	Type of study/ Country	Diabetes at baseline (%)	No. participants	Baseline diagnosis	Length of follow up (years)	Outcome	Baseline factors associated with progressive disease	Baseline factors not associated with progressive disease
Sorrentino et al. 2010 [173]	prospective cohort/ Italy	21	132	Biopsy NAFLD	median 6.4 (range 5.0-8.3)	Increase in biopsy fibrosis score	bp, HOMA-IR, intralobular fibronectin	-
Ascha et al. 2010 [184]	prospective cohort/ USA	72	195	Biopsy NAFLD Cirrhosis meeting transplant criteria	median 3.2y (IQR 1.7-5.7)	HCC	age, any lifetime alcohol, BMI	Diabetes, ethnicity, sex, smoking
Wong et al. 2010 [174]	prospective cohort/ Hong Kong	50	52	Biopsy NAFLD	3.0	Increase in biopsy fibrosis score	-	adiponectin, age, ALT, BMI, bp, CK-18, diabetes, gender, HOMA-IR, IL-6, leptin, lipids, metabolic syndrome, TNFα, waist circumference
Kim et al. 2013 [189]	prospective cohort/ USA	8	4083	USS NAFLD steatosis	median 14.5 (range 0.03-18.1)	Mortality	FIB-4, NFS	-
Pais et al. 2013 [175]	retrospective cohort/ France	35	70	Biopsy NAFLD	mean 3.7 (range 1.0-12.0)	Increase in biopsy fibrosis score	age, BMI, fibrosis, inflammation on biopsy, steatosis	bp, HOMA-IR, lipids

Reference	Type of study/ Country	Diabetes at baseline (%)	No. participants	Baseline diagnosis	Length of follow up (years)	Outcome	Baseline factors associated with progressive disease	Baseline factors not associated with progressive disease
Stepanova et al. 2013 [190]	retrospective cohort/ USA	26	289	Biopsy NAFLD	mean 12.5, maximum 28.5	Mortality	age, NASH, T2DM	Ethnicity, gender, lipids, obesity
Xun et al. 2014 [195]	retrospective cohort/ China	9.4	180	USS NAFLD steatosis	median 6.6 (range 0.5-14.8)	Mortality	NFS	albumin, ALT, ALP, APRI, AST, BARD, BMI, bp, cardiovascular disease, bilirubin, FIB-4, GGT, glucose, lipids, platelets, T2DM
Chan et al. 2015 [213]	prospective cohort/ Malaysia	54.3	35	Biopsy NAFLD	mean 6.4 (sd +/- 0.8)	Increase in biopsy fibrosis score	-	BMI
Perazzo et al. 2014 [176]	prospective cohort/ France	39	2312	Steatotest and Fibrotest	median 12 (range 5.0-15.0)	Mortality, Increase in Fibrotest Score	fibrotest, HbA1c, steatotest	HbA1c, insulin
Sebastiani et al. 2015 [185]	retrospective cohort/ Canada	33.1	148	Biopsy NASH	median 5.0 (IQR 3.0-8.0)	Mortality, Decompensated cirrhosis, HCC	albumin, APRI, bilirubin, diabetes, FIB-4, fibrosis, NFS, platelets,	age, BMI, bp, cholesterol, metabolic syndrome, sex, steatosis

Reference	Type of study/ Country	Diabetes at baseline (%)	No. participants	Baseline diagnosis	Length of follow up (years)	Outcome	Baseline factors associated with progressive disease	Baseline factors not associated with progressive disease
McPherson et al. 2015 [177]	retrospective cohort/ UK	48	108	Biopsy NAFLD	median 6.6 (range 1.3-22.6)	Increase in biopsy fibrosis score	AST, AST:ALT, FIB-4, platelets	Age, ALT, BMI, diabetes, ferritin, fibrosis stage on biopsy, gender, GGT, IgA, IgG, NASH, NFS, steatosis
Ekstedt et al. 2015 [192]	prospective cohort/ Sweden	14	229	Biopsy NAFLD	mean 26.4 (range 6.0-33.0)	Mortality	fibrosis	NAFLD activity score
Angulo et al. 2015 [77]	retrospective cohort/ Multicentre	37.5	619	Biopsy NAFLD	median 12.6 (range 0.3-35.1)	Mortality, Transplant	age, diabetes, fibrosis, smoking	NAFLD activity score, NASH
Pelusi et al. 2016 [178]	retrospective cohort/ Italy	25	118	Biopsy NAFLD	median 3.0 (IQR 2.0-6.4)	Increase in biopsy fibrosis score	NASH, T2DM, time between biopsy	age, BMI, bp, ferritin, liver function tests, lipids, non-invasive fibrosis scores, platelets, sex
Hagstrom et al. 2016 [214]	prospective cohort/ Sweden	18	222	Biopsy NAFLD	median 15.6 (range 0.5-34.2)	Mortality	ferritin	-
Lee et al. 2017 [198]	retrospective cohort/ Taiwan	37	18080	ISD code NAFLD	median 6.3 (IQR 3.0-10.0)	HCC	age, ALT	bp, diabetes, gout, lipids, sex

Reference	Type of study/ Country	Diabetes at baseline (%)	No. participants	Baseline diagnosis	Length of follow up (years)	Outcome	Baseline factors associated with progressive disease	Baseline factors not associated with progressive disease
Akuta et al 2018 [199]	retrospective cohort/ Japan	31.3	402	Biopsy NAFLD	Median 4.2 (range 0.0-41.4)	HCC	age, fibrosis, platelets,	ALT, AST, BMI, CRP, bp, ferritin, gender, GGT, HbA1c, lipids, steatosis, T2DM, uric acid
Vilar-Gomez et al. 2018 [200]	prospective cohort/ Spain, Australia, Hong Kong, Cuba	67	458	Biopsy NAFLD Fibrosis or cirrhosis	Mean 5.5 (range 2.7-8.2)	Mortality, Decompensated cirrhosis, Transplant	age, albumin, AST:ALT, bilirubin, cholesterol, cirrhosis rather than fibrosis on biopsy, INR, platelets, sex, smoking, T2DM	BMI, bp, ethnicity
Castro et al. 2019 [215]	retrospective cohort/ Brazil	(or glucose intolerance) 76.9	39	Biopsy NAFLD Fibrosis	10	Increase in fibrosis by transient elastography	-	NASH
Kleiner et al. 2019 [216]	prospective cohort/ USA	33.4	446	Biopsy NAFLD	4.9 (sd +/- 2.8)	Increase in biopsy fibrosis score	AST, ALT, ethnicity, fibrosis, metabolic syndrome, time between biopsies	age, HOMA-IR, smoking
Caruso et al. 2019 [196]	prospective cohort/ Italy	15.8	457	Ultrasound NAFLD	Median 12	Mortality	-	Steatosis

Reference	Type of study/ Country	Diabetes at baseline (%)	No. participants	Baseline diagnosis	Length of follow up (years)	Outcome	Baseline factors associated with progressive disease	Baseline factors not associated with progressive disease
Kim et al. 2019 [197]	prospective cohort/ S Korea	4.9	40700	Ultrasound NAFLD with low APRI	Median 6 (IQR 3.9-10.0)	Change in APRI	increasing weight	decreasing weight
Onnerhag et al. 2019 [201]	retrospective cohort/ Sweden	22.2	144	Biopsy NAFLD	Median 17.7 (IQR 12.1-25.7)	Mortality, Decompensated cirrhosis, HCC	APRI, BARD, FIB-4, NFS	-
Alexander et al. 2019 [71]	retrospective cohort/ UK, Netherlands, Italy, Spain	19.8	136703	ICD code NAFLD/NASH	Mean 3.3 (IQR 1.8-5.3)	Cirrhosis, HCC	Age, diabetes, FIB-4, NASH	BMI, bp, smoking

Sd standard deviation, **IQR** interquartile range, **TNFα** tumour necrosis factor alpha, **FIB-4** Fibrosis-4 score, **NFS** NAFLD fibrosis score, **APRI** AST to platelet ratio index, **BARD** Bard fibrosis score

1.8 Risk prediction in NAFLD

1.8.1 The need for risk prediction tools in NAFLD

As NAFLD progresses from NAFL/ NASH and fibrosis to cirrhosis, decompensated cirrhosis and HCC; there is a prolonged often asymptomatic course with symptoms often not developing until cirrhosis or HCC is established. However, pre-cirrhotic disease is potentially reversible, and it is additionally helpful to identify those with pre-cirrhotic NAFLD to enable optimal management of metabolic disease and cardiovascular risk; and consider screening for HCC in high risk groups. As discussed, T2DM is associated with an increased prevalence of NAFLD, and increased risk of progression to cirrhosis and HCC. People with T2DM are thus considered a high risk group for the development of NAFLD and proactive screening for pre-cirrhotic NAFLD is advised in joint European guidelines (European Association for the Study of the Liver, European Association for the Study of Diabetes, European Association for the Study of Obesity, EASL-EASD-EASO). [217]

The gold standard tool for identifying NASH and fibrosis is liver biopsy. However, liver biopsy is an invasive procedure and so inappropriate for use when considering population screening. Transient elastography (Fibroscan®), a non-invasive USS-based imaging method to assess liver fibrosis, has been shown to perform reasonably well in identifying those with NAFLD related fibrosis but when considering population screening approaches, would be difficult to scale up on a population screening basis. [211,218] Current interest thus lies in the identification of blood biomarker based risk prediction tools to identify those who would benefit from further investigation and appropriate treatment.

1.8.2 Existing NAFLD fibrosis prediction tools/ models.

A risk prediction tool is designed to combine factors at baseline which are associated with a future outcome. These factors may be causal or simply associated and predictive. Liver enzyme (AST, ALT, ALP, GGT) levels are commonly used in models. The enzymes AST and ALT are found in liver cells (though AST is not truly liver specific) and play a role in amino acid catabolism; in the context of liver injury they are released from liver cells leading to a rise in serum levels.[219] ALP and GGT are involved in bile production and amino acid metabolism respectively; a rise in serum levels of these enzymes is generally considered a marker of cholestasis but can also

be associated with more general liver injury.[219] Bilirubin is metabolised in the liver and so raised levels are indicative of impaired liver function. The liver is key to the synthesis of albumin and platelets so levels of both can decrease due to liver dysfunction, and splenomegaly secondary to portal hypertension in cirrhosis can further reduce platelet levels through increased destruction. None of these blood tests is truly liver specific because levels can be altered by other systemic diseases and treatments (for example cardiac and skeletal muscle damage, haemolysis and the use of statins)). Many prediction scores also use non-hepatic factors associated with likelihood of disease such as age, obesity and the presence of diabetes each of which is known to be associated with the metabolic syndrome and, in turn, NAFLD. As already discussed, however, few individual factors have been shown in isolation consistently to predict progressive disease in NAFLD and so there is much interest in developing models using multiple factors that can identify those at risk. Existing tools were initially designed to identify fibrosis at a point in time, rather than being designed to predict incident cirrhosis, HCC or death. However, as fibrosis is known to predict outcome, they have subsequently been used as risk-prediction surrogates with variable success. There are many existing potential tools. Those used more frequently, and those used within this thesis, will be discussed individually.

AST to platelet ratio index (APRI)

The APRI is derived from the AST level and platelet count. Initially described by Wai et al. in a Hepatitis C cohort, APRI was reported as having an area under the ROC curve (AUROC) of 0.8 for significant fibrosis compared to biopsy, accurately predicting fibrosis in 51% participants using their defined cut-points of ≤0.5 for no fibrosis and >1.5 for definite fibrosis. [220] Similar results have been shown in a NAFLD population. [221] Other cohorts have used different APRI cut-offs, such as a single cut-off of 1, but without significant improvement in performance. [221]

AST to ALT ratio (AST:ALT)

The AST: ALT was initially developed to distinguish NAFLD from ALD. This is because, in alcohol vitamin B6 is often decreased. B6 is required for synthesis of AST and ALT, but mostly for ALT leading to a disproportionate effect on ALT. However, it has since been seen as a marker for liver injury and a level of >0.8 or >1 is seen as an indicator of fibrosis, with one study quoting a negative predictive value (NPV) of >90% with a cut-point of 0.8 in a NAFLD cohort. [221,222]

Enhanced Liver Fibrosis (ELF) test

The ELF test comprises measurement of hyaluronic acid, amino terminal type III procollagen peptide and tissue inhibitor of metalloproteinase 1. These are constituents of the matrix and mediators of matrix remodelling seen in liver fibrosis. Levels have been shown to correlate with liver fibrosis on liver biopsy. [223] Validated by Guha et al. it has reported ability to distinguish severe fibrosis (compared with biopsy) with an AUROC of 0.9 for severe fibrosis and 0.82 for moderate fibrosis in a secondary care population with NAFLD and elevated liver enzymes. [224] Furthermore, in a hospital clinic population with mixed aetiology liver disease, at 6 year follow up 14/16 of those who developed decompensated cirrhosis or HCC had a baseline ELF >9.8 (though 73 participants had high baseline ELF). [225] The proportion of participants with diabetes mellitus is not described. Various cut-points for the score have been used in studies; UK National Institute of Clinical Excellence guidelines recommend a cut-point of ≥10.51 to identify those at high risk of fibrosis. [226]

Fibrosis-4 (FIB-4) score

The FIB-4 score comprises age, AST, ALT and platelets. Initially developed in a cohort with Hepatitis C and HIV co-infection, it was reported to have an AUROC of 0.77 for differentiating significant fibrosis compared to biopsy. [227] In a cohort with Hepatitis C mono-infection, similarly severe fibrosis was identified with an AUROC of 0.85. [228] It has subsequently been validated in NAFLD cohorts, although cut-points vary between studies in Hepatitis C and NAFLD. NAFLD studies suggest a FIB4 <1.3 would suggest low risk of fibrosis and FIB4 >2.67 high risk, with studies suggesting a NPV of 90% for FIB 4 <1.3 in comparison to biopsy. [221,229]

NAFLD Fibrosis Score (NFS)

The NFS was developed in a hospital cohort of participants with biopsy diagnosed NAFLD. Comprising age, hyperglycaemia, BMI, platelets, albumin and AST:ALT, it reported an AUROC 0.82-0.88 for identification of fibrosis (compared to biopsy), with high positive predictive value (PPV) and NPV with cut-points of <-1.45 to identify those at low risk and >0.676 to identify those at high risk of fibrosis. [230] 1/3 of this initial population had diabetes. Subsequently, one study has looked at 12-year mortality and liver outcomes against baseline NFS in a hospital population with NAFLD, 16% of whom had diabetes. In those with NFS >-1.5, there was increased incidence of

decompensated cirrhosis, HCC and increased mortality during follow-up (though total numbers who developed liver events were low with only 6/302 experiencing an event).[231]

Comparison of non-invasive biomarker performance

Several studies have compared the performances of these tools. Some have compared performance for the identification of fibrosis at the time of a simultaneous biopsy. In one secondary care NAFLD population NFS and APRI were compared suggesting respective AUROCs 0.88 and 0.87, and misclassification rates 14% and 16%. 24% of the population had diabetes. [232] Other studies have similarly shown that tools often perform similarly though correlation with outcome in differing populations is very variable, as is best performing model. One (20-30% participants with diabetes) showed FIB-4 as having the highest AUROC 0.8 (NFS 0.77, AST:ALT 0.72, APRI 0.72); another (50% participants with diabetes) also found FIB-4 to have the highest AUROC of 0.86 (AST:ALT 0.83, NFS 0.81, APRI 0.67). [221,229] However, even with the AUROCs in these studies, sensitivity was calculated at 50% and whilst NPV was >90% PPV was mostly <50%. A further study compared ELF, NFS and FIB-4 undertaken at the time of biopsy (30% with diabetes). The AUROC for ELF was superior to that of FIB-4 or NFS but all were >0/8 for F≥3 fibrosis. [233]

Other studies have looked at the use of non-invasive fibrosis tools in their association with incident cirrhosis, decompensated cirrhosis, HCC or progressive disease. One 320 participant cohort with NAFLD (at any stage), found an increase in incident decompensated cirrhosis, HCC and liver transplant in people with NFS ≥-1.455, APRI >0.5 or FIB4 ≥1.3 over 9-year follow-up. [234] Another study found an increased HR for decompensated cirrhosis, HCC or death at 5-year follow up in NASH participants with high risk APRI, FIB4 or NFS scores. [182] However, another study looking at 300 participants with mixed liver disease reported that, at 6-year follow up, of the 16 who developed cirrhosis or HCC, 14/16 had an ELF ≥9.8 at baseline but only 6/16 had an APRI >1.5 and only 4/16 had a FIB-4 >3.25. [225] One study has looked at a large community population with NAFLD (any stage). In it, FIB-4 ≥1.3 was associated with an increased incidence of cirrhosis and HCC over mean 3.3-year follow-up. [71] Two studies have examined paired biopsy results, identifying rising markers (APRI, AST:ALT, FIB-4, NFS – not all used in both studies) to be associated with worsening fibrosis on interval biopsy. [177,235]

Lastly, studies have assessed association with mortality. In the NHANES study, participants with an USS diagnosis of NAFLD steatosis were followed up for 14.5 years. High risk scores for NFS and APRI, but not FIB-4, were associated with an increased HR for mortality, with almost all dying from cardiovascular disease. [189] A further two cohorts with USS diagnosis of NAFLD steatosis with around 6-year follow-up showed an association with raised NFS but not other scores (APRI, AST:ALT, FIB-4, not all assessed in both) and mortality. [191,231] However, a 320 participant cohort with NAFLD (any stage) found increased mortality rates in those with NFS ≥-1.455, APRI >0.5 or FIB-4 ≥1.3 over 9-year follow-up with another 18.8-year cohort finding high APRI, FIB-4 and NFS associated with mortality. [201,234]

Therefore, whereas there is consistent evidence of an association between existing risk prediction tools and outcomes, it is inconsistent and variable in strength of association depending on the cohort to which it is applied. Additionally, there are acknowledged concerns about the application of existing tools to use as screening tools in the general population. Most models have been developed in populations who are under secondary care services for NAFLD, thus testing on a limited population with likely symptomatic and higher-risk disease more likely to progress and so results may not extrapolate to the general population. It is acknowledged that in community populations with or without diabetes, that fibrosis scores perform less well with a large diagnostic grey zone, and possibly the inability to classify up to 1/3 population. [236] In addition, many of the biomarkers used in the tools can be influenced by non-liver disease (for example hyaluronic acid can be raised in joint disease, platelets can be affected by haematological disease).

People with T2DM are thought to be at high risk for progressive disease in NAFLD, and guidelines advise screening this population for NAFLD.[217] However few tools have been specifically validated in cohorts of people with diabetes, and there is increasing evidence that these tools perform less well in people with diabetes. A recent study of 284 hepatology clinic patients of whom 53% had T2DM looked at median 51 month outcomes. [237] Whilst it showed that T2DM conferred an increased risk of death/transplantation (HR 3.4(1.2-9.1)), decompensated cirrhosis (HR 4.7 (2-11.3)) and HCC (HR 2.9 (1.2-7.3)); it showed that the accuracy of FIB-4 and APRI in predicting outcome was reduced in those with T2DM ($p<0.005$). Additionally, whilst no participant without diabetes and with a low score developed decompensated disease, 21% those with T2DM and low score developed decompensated cirrhosis

and 27% developed HCC. A recent study has shown poorer correlation between FIB-4 scores and elastography values (a validated imaging method of identifying hepatic fibrosis) in people with diabetes than those without. [238] Furthermore, other studies have shown, that in populations with T2DM, the estimation of the percentage of the population to have fibrosis based on risk prediction model classification is extremely variable, and the agreement of the top 5% of the risk prediction model scores is poor.[239,240] The reasons for these discrepancies are not fully elucidated. It has been shown that the measurement of AST and ALT may correlate less well with liver pathology in people with diabetes and this may affect the performance of scores based on these biomarkers. [241] Specifically, AST and ALT have been shown to be associated with increased insulin resistance in humans, and higher levels of AST and ALT are seen in mice models of diabetes although whether this is related to worsening NAFLD and metabolic syndrome, or is related to insulin dependent metabolism aside from NAFLD is unclear. [242,243]

1.9 Potential treatment of NAFLD

Ultimately, the aim of better understanding the natural history of NAFLD is to help identify targets for treatment to prevent disease progression, or even encourage disease regression. The identification of a reliably effective treatment is eagerly anticipated given the dramatically increasing prevalence of both NAFLD steatosis and progressive NAFLD with complicated liver disease. However, although multiple targets for treatment have been identified, definition of a specific candidate has proved elusive. [244] Within the limits of this section the broad areas of research interest and potentially positive candidates will be discussed, and ongoing difficulties in the search for treatments will be outlined.

European guidelines recommend that every person with NAFLD should have full assessment for other components of the metabolic syndrome, optimisation of cardiovascular risk prevention (for example statin treatment, management of hypertension, management of co-existing diabetes), lifestyle advice regarding weight loss, and non-invasive monitoring for the development of fibrosis. [217] People with NAFLD are recommended to abstain from alcohol, as the 'safe' alcohol limit in the context of NAFLD is unknown. Immunisation for Hepatitis B should be considered. People with NAFLD associated cirrhosis are screened regularly for variceal disease, liver function and HCC. Management of individual complications of cirrhosis will not be discussed in detail here. People with decompensated cirrhosis should be considered for transplant.

1.9.1 Weight loss

Weight loss has been the only treatment identified to consistently reduce disease progression and also improve not only steatosis, but additionally inflammation and fibrosis. Prospective studies have shown that even modest weight loss of ≤5kg can be an independent predictor of disease progression. [19,245] One study of 293 individuals examining interval liver biopsies showed that weight loss as a result of lifestyle change of ≥3% improved steatosis, ≥5% improved inflammation, and ≥10% improved fibrosis. [246] In a cohort of people with T2DM, it has been shown that moderate caloric restriction and increased physical activity with weekly support resulted in increased weight loss, and associated decline in hepatic steatosis compared to control. [247] As a result of these findings, guidelines have recommended lifestyle management with target weight loss of 5-10%.[7,217] However it is known that weight loss through lifestyle change is challenging, and historically only 10-20% of people are able to lose ≥10%

body weight over 1-2 year period, with maintenance even more difficult although results in a different group of people with T2DM (DiRECT) are slightly more promising. [248,249]

The role of weight loss through bariatric surgery has been assessed. One key study looked at 109 people with biopsy-proven NASH who underwent bariatric surgery. [250] At one year post surgery 85% people had improvement in NASH and fibrosis reduced in 34%. Importantly, those who showed less improvement tended to have more advanced initial disease, and to have lost less weight post-surgery. This suggests the link is related to weight loss rather than another consequence of surgery though this is not conclusive.

1.9.2 Pharmacological therapy

Attempts have been and are being made to identify targets in NAFLD pathogenesis that would be amenable to pharmacological therapy. They include those that may alter fatty acid metabolism (for example peroxisome proliferator-activated receptor (PPAR)α/δ/γ modulators), antioxidants such as Vitamin E, agents that modulate glucose regulation (such as PPARγ agonists, glucagon like peptide (GLP)-1 receptor agonists and sodium-glucose co-transporter-2 (SGLT-2) inhibitors), apoptosis inhibitors, anti-inflammatory and anti-fibrotic agents (such as Galectin 3 inhibitors, lipopolysaccharide antagonists, farsenoid X receptor agonists) and the intestinal microbiome (such as probiotics which may target an endotoxin linked to NASH). [244]

This next paragraph will focus in on agents already used in the context of T2DM. PPARγ modulators, such as pioglitazone, are thought to act by reducing inappropriate fat storage, improving insulin sensitivity and upregulating adiponectin. These have been shown to improve steatosis and inflammation but not fibrosis, and in addition primary trial outcomes have never been reached. [147,251] Furthermore, pioglitazone is associated with complications including osteopaenia, fluid retention and a potential risk of bladder cancer and so is not without risk. A small study of 84 people with NAFLD and T2DM showed that the combined treatment with the SGLT-2 inhibitor dapagliflozin and omega-3 resulted in reduced MRI determined liver proton density fat fraction and improvement in ALT, AST, GGT and cytokeratin-18 at 12 weeks. [252] Treatment with dapagliflozin and dapagliflozin and omega-3 resulted in weight loss, with only the combination affecting liver MRI proton density fat fraction. A recent study of 320 people with biopsy confirmed NASH and liver fibrosis showed that taking a

GLP-1 agonist, semaglutide, resulted in a significantly higher percentage of people experience NASH resolution compared to placebo over 72 week follow-up. However, there was no significant improvement in fibrosis and the significant weight decrease in the semaglutide group was not corrected for. [253]

Of all other treatments in development, vitamin E has been shown to have the most profound anti-steatohepatitic effect, but it does not reduce fibrosis. [251] Whilst it is not within the remit of this thesis to discuss all agents individually, in summary no pharmacological agent has yet been found that improves fibrosis without significant adverse events. [147]

It is important to acknowledge the challenges faced in the search for a NAFLD pharmacological therapeutic agent. Firstly, the complex and dynamic natural history makes trial endpoint interpretation difficult because of the variable placebo response rate of up to 34%.[244] Associated with this, knowledge of the natural history of NAFLD, key pathogenic drivers and thus identification of those at risk of progressive disease is incomplete. [244] Secondly, the ability to assess therapeutic response is hindered by the lack of validation of non-invasive predictive biomarkers to identify disease response, and even biopsy results, the 'gold standard' but invasive investigation, must be interpreted in the context of known sampling variability, variability in processing, reading strategies and inter-observer agreement for key features of NASH. [147,244] In addition, primary endpoints are highly variable between trials. Thirdly, there is inadequate knowledge of in vitro and animal models to mimic disease and test new targets. [244]

In conclusion, the search for appropriate therapeutic agents in NAFLD still presents many ongoing challenges and uncertainties. Weight loss whether through lifestyle change or bariatric surgery does seem to be effective and should be the cornerstone of clinical management. Developing other therapeutic strategies in NAFLD is dependent, however, on more accurately identifying those truly at 'high risk' for the development of complicated disease and understanding which non-invasive markers reliably predict disease progression or remission.

1.10 Summary and Aims

There is a substantial body of evidence demonstrating the clinical significance of NAFLD, both from an individual patient and population perspective. Nonetheless, the disease burden in community populations is poorly understood. In addition, the evidence that the concurrent presence of NAFLD and T2DM worsens prognosis suggests that there may be merit in active screening of high risk populations to enable the implementation of potential NAFLD treatment, cardiovascular risk prevention and identification and treatment of complications. [36] In particular, the increasing awareness that up to 50% of HCC in NAFLD can occur in those without cirrhosis enhances the need to develop improved surveillance strategies to improve outcome (which is currently poor at 18% 5 year survival). [165,186] Consequently, UK defined key goals for the development of liver treatment include strengthening the detection of early disease and the use of new diagnostic pathways to identify people with NAFLD.[3]

However, there is no consensus on prognostic indicators and whilst EASL-EASD-EASO guidelines suggest possible pathways of care for patients with NAFLD but there is ongoing doubt about their applicability to the T2DM population. [217,254]

Accordingly, there is a need for better understanding of the natural history and incidence of NAFLD in people with diabetes, and to discern those predictive factors that can potentially identify those at high risk of developing complicated disease and be utilised in screening strategies.

This thesis will address the following specific aims:

1. Define the absolute and relative cohort incidence of liver disease to date in the ET2DS cohort
2. Determine whether current non-invasive fibrosis risk prediction tools reliably identify incident cirrhosis and hepatocellular carcinoma in a community cohort of older people with T2DM
3. Determine whether the addition of other biomarkers to existing fibrosis risk prediction tools improve their performance in predicting incident cirrhosis and hepatocellular carcinoma in a community cohort of older people with T2DM
4. Identify whether potential non-invasive screening tests for NAFLD (those identifying steatosis, serum liver enzymes, markers of fibrosis) associated with incident cirrhosis, hepatocellular carcinoma and mortality in people with T2DM

Chapter 2 Methods

This Chapter will discuss general methods used in the study. Methods used for individual analyses will be discussed in specific chapters.

2.1 The study population: The Edinburgh Type 2 Diabetes Study

The ET2DS is a population based prospective cohort study. It was designed to examine cognitive function and factors associated with cognitive decline in a population of older people with T2DM. At year 1 a second assessment arm was initiated to examine the prevalence of liver disease in this cohort, to study the natural history of liver disease progression and to identify factors associated with this progression in this cohort.

Participants were selected from over 20,000 patients with T2DM on the Lothian Diabetes Register. The Lothian Diabetes Register (now incorporated into SCI-Diabetes) is a computerised database established in 2001. It contains details of people with diabetes (as defined by WHO criteria) living in Lothian, Scotland, UK. It includes both those cared for in primary care and in hospital clinics. Work by the Lothian Diabetes Services Advisory Group demonstrated that this database contains almost everyone diagnosed with diabetes in the area (Sarah Wild, personal communication). [255]

Participants were selected by gender and 5-year age bands from a computer-randomised list. Inclusion criteria were age 60-74 on 01/08/2006 and a diagnosis of T2DM. Exclusion criteria were non-English speaking (as fluent English was required for cognitive testing), visual acuity worse than 6/36 at distance, or unable to read large print text (as this was required for the testing), who were unwilling or unable to give informed consent or physically unable to complete the clinical or cognitive examination.

Criteria to confirm the presence of diabetes were: currently having treatment with oral antihyperglycaemic medication and/or insulin; or currently treated with dietary modification alone and an HbA1c>6.5%. If recruits were treated with dietary modification alone and had an HbA1c ≤6.5%, their medical records were reviewed by a consultant Diabetologist (Mark Strachan) to ensure that the diagnosis of diabetes was robust. The diagnosis of T2DM was additionally reviewed if the participant had

commenced insulin within 1 year of diagnosis, if they had reported a history of pancreatic surgery at the research clinic or if the participant was treated with insulin and had been diagnosed aged <35 years. Participants in whom it was not possible to confirm a diagnosis of T2DM were excluded.

A total of 5,454 invitations to participate were sent out between 20th June 2006 and 1st June 2007; 3,286 people replied, of whom 1,252 were interested in the study. Of these 1,077 attended the baseline clinic and 1,066 were included in the study (unable to complete tests n=4, did not fulfil study criteria for the diagnosis of T2DM n=7). Every effort was made to ensure participants attended the research clinics. This included multiple attempts to contact the participant to arrange appointments, providing a choice of clinic dates, paid travel or provision of transport to clinic and reminder calls prior to the clinic attendance.

2.1.1 Representativeness of Data

Non-identifiable data gathered from the Lothian Diabetes Register was able to confirm that the baseline study participants (1,066) were representative to the randomly selected 5,454 patients in age, HbA1c, duration of T2DM, proportion requiring insulin treatment and total cholesterol. It was thus considered that the study participants were largely representative of the target population (Table 4). [256] It is important to note that there was a significant difference in the proportion of men the study invited compared to the target population. This was due to the selection of participants in age range brackets.

2.1.2 Power and sample size

The study aimed to recruit 1,000 subjects which would allow 90% power at the 2-sided 5% significance level to detect a Pearson correlation coefficient of ≥0.10 between a continuous outcome measure and predictor variable; and estimated to allow for the detection of any risk factor that contributed 1% or more to the variance in the outcome for observed associations. [255] Using a post-hoc power calculation, the study had 82% power to detect a 3% difference in rates of incident cirrhosis and HCC with a putative baseline rate of 1%. The study had 89% power to detect a 10% difference in all-cause mortality rates with a baseline rate of 29% and 88% power to detect a 7.5% difference in cardiovascular mortality rates with a baseline rate of 12% (2-tailed, probability of a type 1 error 0.05).

Table 4. Representativeness table. Baseline characteristics of ET2DS study population compared with non-responders (adapted from Marioni et al. 2010) [256]

		ET2DS (n = 1066)		Non-responders (n = 4386*)	
Age		67.9	(4.2)	67.9	(4.4)
Sex - Male		547	(51.3%)	1839	(41.9%)
Systolic blood pressure (mmHg)		133.3	(16.4)	137.2	(18.2)
Total cholesterol (mmol/L)		4.3	(0.9)	4.2	(1.0)
HbA1c (%Hb)		7.4	(1.1)	7.4	(1.4)
HbA1c (mmol/mol)		57	(12.4)	57	(14.9)
Diabetes Treatment (Insulin)		186	(17.4%)	704	(16.1%)
Duration of diabetes - years	Up to 5 years	516	(48.4%)	2135	(48.7%)
	5 years or more	550	(51.6%)	2251	(51.3%)
Scottish Index of Multiple Deprivation	1 (most deprived)	127	(11.9%)	736	(16.8%)
	2	208	(19.5%)	1134	(25.9%)
	3	188	(17.6%)	820	(18.7%)
	4	194	(18.2%)	782	(17.8%)
	5 (least deprived)	349	(32.7%)	897	(20.5%)

Values are mean (sd) or n (%); *4388 was actual number of non-responders but two subjects from the Lothian Diabetes Register did not have any data so were discarded from analyses
HbA1c glycated haemoglobin, **mmHg** millimetres of mercury, **mmol/L** millimol/litre

2.1.3 Ethics and consent

Ethics permission for the study was granted by Lothian Medical Research Ethics Committee (REC reference 16/SS/0098). The Lothian Diabetes Services Advisory Group and the Caldicott guardian for NHS Lothian provided permission for the use of the Lothian Diabetes register (now SCI-diabetes) in patient selection. Participants renewed written informed consent at all clinic visits.

2.2 Data Collection

2.2.1 Study Phases

Assessment took place primarily at baseline, year 1, year 4 and year 10. Assessment included attendance at dedicated research clinics, collection of biochemical samples, physical examination, cognitive testing, completion of self-assessment questionnaires by participants and GPs, data linkage to the information services division in Scotland

(ISD) (www.isdscotland.org), National Records Scotland and SCI-diabetes, and assessment of hospital electronic patient records (Table 5).

All participants who were alive and who had consented to be contacted about additional studies were invited to return to Year 1, 4 and 10 assessment. At year 1, 1,054 were invited (died n=2, declined n=5, medically unsuitable for contact n=3, withdrawn from further contact n=2). 940 attended (died n=13, unable to contact n=19, unable to attend for health reasons n=23, unable to attend for other reasons n=38, did not attend appointment n=21). Of these 940, one participant was unable to complete the research assessment. At year 4, 974 were invited (died n=81, withdrawn from contact n=11). 830 attended (unable to contact n=15, declined to attend n=100, withdrew from further contact n=30). At year 10, 845 were invited of whom 581 were able to complete assessment (died n=84, unable to attend n=112, declined appointment n=44, unable to contact n=24).

2.2.2 Research Clinics

All research clinics took place at the Wellcome Trust Clinical Research Facility, Western General Hospital, Edinburgh, UK. Standardised operating procedures were used for every aspect of data collection. Detailed assessment of modifiable risk factors and cognitive testing was undertaken and has been documented previously.[255] Individuals with clinically significant findings during testing were referred to an appropriate clinician for follow up.

2.2.3 Questionnaires

Questionnaires were completed by participants at the time of the clinic appointment. These recorded demographic data, data regarding diabetes history, other past medical history including a history of liver or joint disease, medication use, alcohol and smoking history, and chest pain and claudication scales. If participants were unable to attend clinic at Year 4 and 10, modified questionnaires to obtain demographic data, diabetes history, updated medical history and medication use were sent to the participant's GP.

Table 5. Data Collection Undertaken throughout Study

Baseline	Year 1	Year 4	Year 6	Year 10	Data Collected
✓	✓	✓		✓	General questionnaire: Demographics, Past Medical History, current medications, alcohol consumption, smoking
✓	✓	✓		✓	Diabetes questionnaire: Diabetes history, diabetes medication history, history of hypoglycaemia
	✓	✓		✓	Liver questionnaire: liver disease history, joint disease history, hepatotoxic medication use
✓		✓		✓	Cardiovascular questionnaire: chest pain and claudication scales
✓	✓	✓		✓	Physical examination (including BMI, bp, waist circumference)
✓		✓		✓	Cognitive testing
✓	✓	✓		✓	Fasting venous blood sample (including HbA1c, cholesterol, AST, ALT, ALP, GGT, bilirubin, albumin, platelets, HA, TNFα, IL-6, CRP, triglycerides and blood for DNA at baseline, ELF at year 1)
	✓	✓			Liver Ultrasound
		✓			Transient Elasstography
✓		✓			ECG
✓		✓	✓	✓	Data linkage with ISD Scotland
				✓	Data from SCI-diabetes
		✓		✓	Data linkage with electronic secondary care record
				✓	Death Certificate data from National Records Scotland

BMI body mass index **bp** blood pressure **AST** aspartate aminotransferase **ALT** alanine aminotransferase **ALP** alkaline phosphatase **GGT** gamma glutamyltransferase **HA** hyaluronic acid **TNFα** tumour necrosis factor alpha **IL-6** interleukin 6 **CRP** C-reactive protein **ELF** Enhanced Liver Fibrosis Score

2.2.4 Routine data collection

Data was collected for all participants from general and acute inpatient (excluding psychiatric and obstetric wards) discharge records and death certification using record linkage to the SMR01 scheme at NHS National Services Scotland, Information Services Division (www.isdscotland.org). This data linkage was undertaken at baseline (data acquired 1981-2007), with follow up linkage up to 2011 and 2015. This was used to supplement and confirm self-reported history. In addition, death data was collected in 2018 from National Records Scotland.

Selected data held on the Lothian Diabetes Register (SCI-diabetes) was extracted to provide routinely recorded HbA1c data and data on medication prescription during the period of the study.

Secondary care records from Lothian Hospitals were interrogated to supplement and confirm liver, cardiovascular, dementia and death events. To undertake this, the TRAK electronic patient record was reviewed (TrakCare, InterSystems Corp., Cambridge, USA) at year 4 and 10.

2.2.5 Variable collection

Discussed below in detail are the variables collected relevant to this project. Full details of data collection can be found in previous publications. [255] Unless otherwise specified, blood samples were analysed using a Vitros Fusion chemistry system (Ortho Clinical Diagnostics, Bucks, UK) at the Western General Hospital (Edinburgh, UK).

Demographics: Date of birth and sex was obtained from self-report questionnaire and confirmed against clinical records. Ethnicity was obtained from self-report questionnaire. Socio-economic status was measured using the Scottish Index of Multiple Deprivation (SIMD) 2006 converted from patient home postcodes at baseline (see http://openscotland.gov.uk/Topics/Statistics/SIMD/FAQs#lookups) and defined as quintiles (1 most deprived, 5 least deprived).

Diabetes Background: Duration of Diabetes was calculated based on date of diagnosis on self-report questionnaire. Hba1c at baseline was measured on venous blood samples. For those people who did not have an HbA1c taken due to failure of venepuncture at baseline clinic, baseline HbA1c was obtained from routine data

download of care records if possible, with an HbA1c within 6 months of the baseline clinic appointment accepted as a valid value.

Lifestyle Factors: Smoking and alcohol were assessed on self-report questionnaire. For alcohol, average weekly alcohol intake was determined using two questions adapted from the AUDIT-C screening tool: “How often did you have a drink containing alcohol in the past year? Consider a “drink” to be a can or bottle of beer, a glass of wine, or one cocktail or a measure of spirits (like scotch, gin or vodka)” [257]. (A drink was considered to be the equivalent of one and a half units of alcohol); and “How many drinks did you have on a typical day when you were drinking in the last year?”. Alcohol excess was defined according to established criteria as alcohol intake >14 units/ week (female) or >21 units/ week (male) or subject self-report of current or previous alcohol excess.

Metabolic Factors: Height was measured (to nearest mm) standing without shoes. Weight was measured (to nearest 0.1kg) without outdoor clothing or shoes using SECA 761 electronic weighing scales. BMI was calculated as height (m)/ (weight (kg))2. Waist and hip circumference were measured (to nearest 0.5cm) using a non-expandable tape measure. Waist circumference was measured at the level midway between the lower rib margin and the iliac crest, with the subject standing with their feet 30 cm apart and with their hands by their sides, during exhalation. The average of 2 readings was taken. Total cholesterol and triglycerides were collected from fasting venous blood samples.

Inflammatory markers: CRP, IL-6 and TNFα were assessed from venous blood samples. CRP was measured using an immunonephelometric assay, and IL-6 and TNF-α were measured using the ELISA system (R&D Systems, Oxon, UK), Glasgow Royal Infirmary, UK.

Non-invasive liver markers: ALT, AST, ALP, GGT, bilirubin, albumin, platelets and HA were measured on fasting venous blood samples at baseline. ELF was measured on fasting venous blood samples at 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). HA was measured using a radiometric assay (Pharmacia, Uppsala, Sweden).

Hepatic steatosis was determined by USS measurement following a 4 hour fast at the year 1 clinic (Sonoline Elegra Ultrasound Imaging System (Siemens Medical Systems Inc, Washington, USA), software version 6, with a 3.5 MHz transducer). The technique has previously been described in detail. [258] Briefly, a single sonographer, blinded to clinical history, undertook all scanning and graded using established criteria (normal, indeterminate steatosis (possible slight increase in echogenicity or slightly impaired visualization of the diaphragm/ intrahepatic vessels/ difficult to grade as a result of diseased or absent R kidney), mild steatosis (definite increased echogenicity and/ or definite impaired visualisation of intrahepatic vessels and diaphragm, no/ little evidence of focal fatty sparing), severe steatosis (marked increase echogenicity and or poor or no visualization of the diaphragm and intrahepatic vessels, with or without focal fatty sparing). To confirm results and assess for inter- and intra-observer variation, a subset of the USS (71) were re-graded (still images only) by 2 additional independent graders (a consultant radiologist and a medical trainee in radiology) within 1 month of the USS. ≥2 months later, all 3 graders regraded all 71 USS by still image assessment. Graders were blinded to the other graders results and clinical and lab results of the participants. No significant inter or intra-observer variability was seen. In addition, a subset (50 participants) underwent 1H MRI spectroscopy (the gold standard non-invasive tool for assessing hepatic steatosis) to assess the validity of USS as a technique to identify steatosis. This has been described in detail previously [258]. In brief, 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 graded ‘normal’, ‘indeterminate’ and ‘mild’ steatosis, only those with severe steatosis on USS assessment were deemed to have ‘definite steatosis’. Individuals with any other USS grading were considered to have ‘no definite steatosis’.

Calculated variables: A wide range of markers of fibrosis were measured and calculated as per original cited research. Documented cut-off levels were utilised.

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(x109/L)) x100. Cut-point low to medium/high risk of fibrosis >0.5.[220]
- AST: ALT ratio was calculated as: AST(U/L)/ALT(U/L). cut-point ≥0.8. [221]
- Fibrosis-4 index (FIB-4) was calculated as ((age(years)xAST(U/L))/(plt(x109/L)x√ALT(U/L))). Cut point low-medium risk ≥1.3 and medium-high risk >2.67. [227,229,234]
- NAFLD Fibrosis Score (NFS) was calculated as: 1.675+(0.037xage(years))+(0.094xBMI(kg/m2))+(1.13xIFG/diabetes (yes=1, no=0))+(0.99x(AST(U/L)/ALT(U/L))- (0.013xplatelet count(×109/L))-(0.66xalbumin (g/dL)). Cut-point for low-medium risk ≥-1.455, medium-high risk >0.676. [230]
- Fatty liver index (FLI) was calculated as: ey/(1+ey)x100 where y=0.953 x ln(triglycerides, mg/dl) + 0.139 x BMI, kg/m2 + 0.718 x ln (GGT, U/L) + 0.053 x waist circumference, cm – 15.745). [259]
- The EASL-EASD-EASO referral decision algorithm (Figure 2) was used. [217]

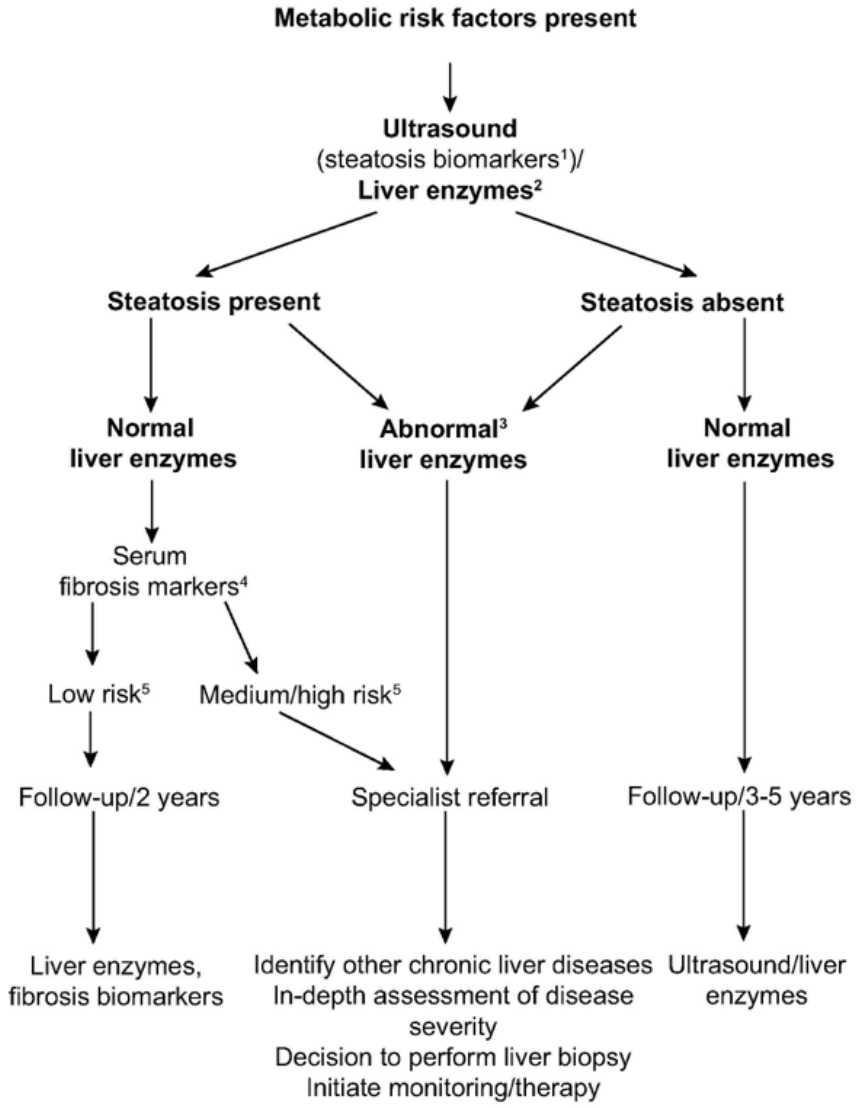
Outcomes- development of cirrhosis, HCC, varices, ascites, encephalopathy and hepato-renal syndrome: The presence of cirrhosis, HCC and other cirrhosis complications (varices, ascites, encephalopathy or hepato-renal syndrome) were determined, with data collection mechanisms discussed below.

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 liver disease was 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. Cases were confirmed if a clinician diagnosis was recorded in secondary care medical notes. All cases identified through data linkage were able to be confirmed in secondary care

Figure 2. EASL-EASD-EASO Algorithm[217]

Permission to reproduce (Appendix 1)



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, SteatoTest, NAFLD Fat score (see Tables). [2]Liver tests: ALT AST, γGT. [3]Any increase in ALT, AST or γGT. [4]Serum fibrosis scores: NAFLD Fibrosis Score, FIB-4, Commercial tests (FibroTest, FibroMeter, ELF). [5]Low risk: indicative of no/mild fibrosis; Medium/high risk: indicative of significant fibrosis or cirrhosis.

medical notes. Date of diagnosis was determined as the date of the first clinical documentation of the diagnosis in the medical notes (clinic letter or inpatient stay discharge summary).

ICD codes analysed and correlated with medical records were: B18 (chronic viral hepatitis), B19 (unspecified viral hepatitis), C22.0 (liver cell carcinoma), C22.7 (other specified carcinomas of liver), C22.9 (malignant neoplasm of liver, unspecified), K70 (alcoholic liver disease), K71 (toxic liver disease), K72.0 (acute and subacute hepatic failure), K72.1 (chronic hepatic failure), K72.9 (hepatic failure, unspecified), K73 (chronic hepatitis not elsewhere classified), K74.0 (hepatic fibrosis), K74.1 (hepatic sclerosis), K74.2 (hepatic fibrosis with hepatic sclerosis), K74.3 (PBC), K74.4 (secondary biliary cirrhosis), K74.5 (biliary cirrhosis, unspecified), K74.6 (other and unspecified cirrhosis of the liver), K76.0 (fatty liver not elsewhere classified), K76.1 (chronic passive congestion of the liver), K76.2 (central haemorrhage necrosis of liver), K76.3 (infarction of liver), K76.4 (peliosis hepatitis), K76.5 (hepatic vena-occlusive disease), K76.6 (portal hypertension), K76.7 (hepatorenal syndrome), K76.8 (other specified diseased of liver), K76.9 (liver disease, unspecified), K77.0 (liver disorders in infectious and parasitic diseases), K92.0 (haematemesis), K92.1 (Melaena), K92.2 (GI haemorrhage, unspecified), R16.0 (hepatomegaly not elsewhere classified), R16.1 (splenomegaly not elsewhere classified), R16.2 (hepatomegaly with splenomegaly not elsewhere classified), R17 (unspecified jaundice), R18 (ascites), R58 (haemorrhage, not elsewhere classified).

It should be noted that some incident cirrhosis/ HCC was identified following referral after year 1 and year 4 clinic screening. Referral criteria for hepatology review were any of: routine liver enzyme tests above the laboratory upper limit of normal (ALT >50 U/L, AST >45 U/L, GGT >55 U/L, ALP >125 U/L); AST:ALT ratio >1; positive autoantibodies (anti-nuclear antibody, anti-smooth muscle antibody, anti-mitochondrial antibody), ferritin >1000ng/ml, positive hepatitis B or C serology, hyaluronic acid >100 microg/L (in the absence of known joint disease), spleen >13cm (in the absence of known haematological cause), platelets <150 x109/L (in the absence of known haematological cause), suspected cirrhosis on USS or alpha-feto protein >6ng/l. 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 (Figure 3).

Outcomes- Cause of Death: Cause of death was identified through data linkage to death coding in ISD data (collected up to year 8 follow-up). For those where ISD data was not available, cause of death was identified from death certification in secondary care medical records or the analysis of death certificates at National Records Scotland.

Death data was reviewed manually to ascertain primary cause of death. Cardiovascular Death was determined to be a primary cause of death termed as any of fatal myocardial infarction, fatal cerebrovascular accident, other fatal ischaemic heart disease, other fatal cerebrovascular disease (note this did not include vascular dementia), fatal other cardiovascular disease. Death from cirrhosis or HCC was determined to be a primary cause of death as cirrhosis, HCC or as a direct complication of these.

2.3 Data Entry

All data obtained was entered onto a master Microsoft Access database (Microsoft Access 2003/2010, Microsoft Corporation, Washington, USA). At baseline, the majority of data from paper records was double entered and discrepancies resolved by reference to the original documentation. At follow up years, at least 10% of the data from paper records was double entered.

2.4 Personal contribution to data collection

My personal contribution was to the following parts of the data collection and entry for the year 10 follow-up of the study:

- An interrogation of participants' electronic secondary care records from NHS Lothian hospitals, Lothian, UK (TrakCare, InterSystems Corp., Cambridge, USA) was undertaken. The records of all 1066 participants in the study were searched for liver outcomes (at any point in the study), dementia and cardiovascular outcomes (from the point at which records had previously been interrogated (2010 and 2014 respectively)), and death records. Specifically for liver outcomes; evidence of a diagnosis of cirrhosis, hepatocellular carcinoma, cirrhosis related variceal disease, ascites, subacute bacterial peritonitis, encephalopathy or hepato-renal syndrome, attendance at liver clinic or enrolment into an HCC or varices surveillance programme was recorded.
- With other team members, a data download from the SCI-Diabetes database to obtain retinopathy screening, blood pressure, HbA1c, lipid and medication prescription data was organised.
- The National Records Scotland database was searched to identify additional death certificate details for those participants who have died but for whom the study did not have cause of death details. Through this search, out of over 300 participants who have passed away, the study now has cause of death for all but 20.
- I contributed to sending self-assessment questionnaires to participants and GPs (for those where the study was unable to contact the participant), the completion of the database data entry for the 10 year follow up clinics, the double data entry process for the 10 year follow up clinic data and reviewing and ensuring correct coding of conditions and causes of death in the database to facilitate future database searching.

2.5 Data analysis

Data was 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/).

2.5.1 Missing Data

There are three generally accepted approaches to dealing with missing data. Firstly, if it is considered that in those people who had missing data, that data is missing at random, then it is reasonable to use only those with complete data sets in the analysis. However, if the data is not missing at random then there is a risk of introducing systematic bias. [260,261] Additionally, if a large subsection of the research population has missing data, then excluding them will reduce the power of the analysis. Secondly, it is possible to use incomplete data but add in a random effects term to any model created, in recognition of the fact that the missing data may contribute an important but unknown effect. Thirdly, multiple imputation can be utilised. In this method, missing items are assigned a generated value and then incomplete records can be 'topped up' in this way and included in analysis alongside cases with full data sets. Analysis can be repeated multiple times with alternative imputed values to aim to ensure that the imputed values are not substantially affecting the results. This enables a fuller analysis of the whole cohort on one hand, however any imputed value is not a true reading and may in itself bias the results. [260,261]

For this study, initial paper records have been reviewed for all missing data by at least 2 investigators to ensure that the data-set is as complete as possible, and that all missing data is truly missing.

For this analysis, most explanatory variables have <5% missing data. In previous analysis of liver data undertaken in our cohort, assessment was made of the whole data set. If <5% of the participants had missing data for a particular variable and whether it was missing was presumed random, then analysis was undertaken on an available case analysis and participants with missing data were excluded. However, not all patients were available to return to follow up for the year 1 follow-up and thus variables collected at this point (relevant to this analysis are ELF, USS assessment of hepatic steatosis) have >5% missing data. In addition, the variables with >5% missing data are key variables where imputation would be difficult to determine (an average value would be difficult to obtain based on other variables as often they are

not linked). Multiple imputation analysis additionally was attempted for ELF in previous work undertaken by this group and results were not significantly different from those obtained from available case analysis. [240]

The decision was thus made that analysis would be undertaken on an available case analysis basis, understanding the potential limitations of power this approach might present.

An assessment was undertaken assessing the baseline characteristics of the whole population compared to the available case populations for each analysis, which showed these were broadly similar

Table 6).

2.5.2 Outliers

Outliers were identified using the Tukey boxplot method, looking at any value greater than 1.5 times the interquartile range from the median value. All values in this category were manually checked for clinical plausibility, with extreme outliers checked with paper records to exclude the possibility of transcription error. As a result of these checks, no items were excluded.

Table 6. Baseline characteristics of the population for the analyses

		Total population (minus those missing in each variable) (n=1066)	**Number missing in each variable (from total population)**	**Total population without cirrhosis or HCC at baseline (n=1059)**	**Complete cases for steatosis analysis (n=933)**	**Complete cases for analyses including ELF (n=681)**	**Complete cases for model development analysis (n=999)**
Used in Question:		1	-	3, 5	2	3	4
Age		67.9 (4.2)	0	67.9 (4.2)	67.9 (4.2)	67.8 (4.2)	67.9 (4.2)
Sex (male)		547 (51.3)	0	544 (51.4)	487 (52.1)	358 (52.6)	515 (51.6)
Scottish Index of Multiple Deprivation quintile	1 (most deprived)	127 (11.9)		125 (11.8)	107 (11.5)	81 (11.9)	119 (11.9)
	2	208 (19.5)		206 (19.5)	176 (18.9)	128 (18.8)	196 (19.6)
	3	188 (17.6)		187 (17.7)	161 (17.3)	116 (17.0)	178 (17.8)
	4	194 (18.2)		193 (18.2)	169 (18.1)	117 (17.2)	180 (18.0)
	5 (least deprived)	349 (32.7)		348 (32.9)	320 (34.3)	239 (35.1)	326 (32.6)
Duration T2DM (years)		8.1 (6.5)	13	8.1 (6.5)	8.0 (6.4)	7.8 (6.3)	8.0 (6.5)
HbA1c (%)		7.4 (1.1)	9	7.4 (1.1)	7.4 (1.1)	7.4 (1.1)	7.4 (1.1)
HbA1c (mmol/mol)		57 (12)	9	57 (12)	57 (12)	57 (12)	57 (12)
BMI (kg/m^2)		31.4 (5.7)	1	31.4 (5.7)	31.3 (5.7)	31.2 (5.7)	31.3 (5.6)
Waist-Hip Ratio		0.97 (0.1)	5	0.97 (0.1)	0.96 (0.1)	0.96 (0.1)	0.96 (0.1)
Smoker (current)		154 (14.4)	0	153 (14.4)	122 (13.1)	89 (13.1)	143 (14.3)
Alcohol (excess)[a]		207 (19.9)	27	204 (19.8)	187 (20.0)	139 (20.1)	195 (19.5)
Cholesterol (mmol/L)		4.3 (0.9)	9	4.3 (0.9)	4.3 (0.9)	4.3 (0.9)	4.3 (0.9)
ALT (U/L)		43.2 (14.3)	9	43.2 (14.3)	43.5 (14.4)	44.0 (14.1)	43.4 (14.4)
AST (U/L)		31.0 (10.5)	11	31.0 (10.4)	31.1 (10.3)	31.1 (9.8)	31.1 (10.4)
ALP (U/L)		91.8 (27.4)	9	91.7 (27.3)	90.9 (27.6)	90.9 (27.2)	91.6 (27.5)

	Total population (minus those missing in each variable) (n=1066)	Number missing in each variable (from total population)	Total population without cirrhosis or HCC at baseline (n=1059)	Complete cases for steatosis analysis (n=933)	Complete cases for analyses including ELF (n=681)	Complete cases for model development analysis (n=999)
GGT (U/L)	29.9 (42.3)	11	29.4 (40.3)	29.1 (40.0)	28.6 (39.3)	29.7 (40.9)
Bilirubin (μmol//L)	10.1 (5.0)	9	10.0 (4.7)	10.1 (4.9)	10.2 (5.0)	10.0 (4.7)
Albumin (g/L)	44.8 (3.3)	11	44.8 (3.3)	44.8 (3.2)	44.9 (3.2)	44.8 (3.3)
Platelets (10^9/L)	258.0 (69.9)	21	258.7 (69.3)	257.9 (68.7)	260.9 (70.4)	258.5 (69.8)
Hyaluronic Acid (ng/ml)	57.9 (58.9)	9	56.1 (46.6)	56.1 (46.8)	55.9 (48.6)	56.4 (47.1)
TNF-$\propto$ (pg/ml)	1.4 (1.5)	3	1.4 (1.5)	1.4 (1.6)	1.3 (1.1)	1.4 (1.6)
IL-6 (pg/ml)	3.9 (3.5)	2	3.9 (3.5)	3.8 (3.4)	3.7 (3.3)	3.9 (3.4)
CRP (mg/L)	3.9 (6.0)	24	3.9 (6.0)	3.6 (5.6)	3.6 (5.7)	3.8 (6.0)

Values are mean(sd) or n(%)

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

HbA1c glycated haemoglobin, **BMI** body mass index, **ALT** alanine aminotransferase, **AST** aspartate aminotransferase, **ALP** alkaline phosphatase, **GGT** gamma glutamyltransferase, **TNF-$\propto$** tumour necrosis factor-alpha, **IL-6** Interleukin-6, **CRP** c-reactive protein, **mmol/L** milimol per litre, **kg/m^2** kilograms per square metre, **U/L** international units per litre, **μmol//L** micromol per litre, **g/L** grams per litre, **pg/ml** pico-grams per mililitre, **mg/L** milligrams per litre, **ng/ml** nanograms per mililitre

Chapter 3 Non-invasive risk scores do not reliably identify future cirrhosis or hepatocellular carcinoma in Type 2 diabetes: The Edinburgh Type 2 Diabetes Study

This section was published in Liver International under the same title by Sheila M Grecian (SMG), Stela McLachlan (SM), Jonathan A Fallowfield (JF), Patrick KA Kearns (PK), Peter C Hayes (PH), Indra Neil Guha (NG), Joanne R Morling (JM), Stephen Glancy (SG), Rachel M Williamson (RW), Rebecca M Reynolds (RR), Brian M Frier (BF), Nicola N Zammitt (NZ), Jackie F Price (JP) and Mark WJ Strachan (MS). [262] SMG wrote the manuscript. JP was principal investigator of the ET2DS, designed the study, analysed and interpreted the data. MS was lead investigator of the ET2DS liver sub-study, designed the study, analysed and interpreted the data. RR, BF, PH, JF, RW, NG and SG contributed to study design. SMG, SM, RW, JM and PK contributed to data collection, analysis and interpretation. All authors contributed to revision and final approval of the article.

There is increased incidence of cirrhosis and HCC in people with T2DM, with the primary aetiology being NAFLD (section 1.4). However, although it has been suggested that it may be beneficial to screen high risk population groups for NAFLD in order that advice can be targeted about lifestyle modification to reduce rate of disease progression or reverse disease progression, cardiovascular risk management optimised and complications identified and treated promptly, there is no consensus on which tools work best to predict clinically significant disease. [217] In addition, it has been shown that existing risk prediction tools tend to perform worse in populations with diabetes. [237] In this study we investigated the incidence of cirrhosis and HCC in a community cohort of older people with T2DM. Furthermore, we assessed the ability of existing non-invasive risk prediction tools to identify incident liver disease in our community population with T2DM.

Please note- formulae used for the calculation of incidence and the predictive ability of risk prediction tools can be found in appendix 2.

Please note- all tables for this section sit at the end of the chapter text.

3.1 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=1,066, 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: 43/1059 participants with no baseline cirrhosis/HCC developed incident disease. All scores were significantly associated with incident liver disease by odds ratio ($p<0.05$). 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.

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.

3.2 Introduction

People with Type 2 diabetes have a higher incidence of cirrhosis and hepatocellular carcinoma (HCC) than the general population. [20,21,75] The commonest cause of liver disease in Type 2 diabetes is non-alcoholic fatty liver disease (NAFLD) with estimates of prevalence from 40-70%.[9-11,263]

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. [7,217]

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. [179,180] 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)). [220,221,224,227,230] 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. [185,189,225,231,234] 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.[202,264]

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. [217] 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'. [7] 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. [7]

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. [254] 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. [237] There remains significant uncertainty about the utility of these screening methods in Type 2 diabetes.

3.3 Aims

We aimed to assess the ability of individual fibrosis scores and of the EASL-EASD-EASO screening algorithm to predict 11-year incident cirrhosis and/or HCC in an asymptomatic community cohort of older people with Type 2 diabetes.

3.4 Methods

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. [255] 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. [255] 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. [9] All 1066 participants were followed up for outcome assessment to death (320 participants throughout the study) or end of follow-up.

Data Collection - Baseline biomarker assessment

Research clinics were undertaken at the Wellcome Trust Clinical Research Facility, Western General Hospital, Edinburgh, UK. Standardised Operating Procedures were used for every aspect of data collection as previously detailed. [255] 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 1H MRI spectroscopy in a subset, as previously described. [258] 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-feto 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 >100µg/L (in the absence of known joint disease), positive liver autoantibodies, ferritin >1000ng/mL, alpha-feto protein >6ng/mL, positive hepatitis B or C serology, spleen diameter >13cm, platelets <150x109/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(x109/L)) x100. Cut-point low to medium/high risk of fibrosis >0.5. [220]
- AST: ALT ratio was calculated as: AST(U/L)/ALT(U/L). cut-point ≥0.8 [221].
- Fibrosis-4 index (FIB-4) was calculated as ((age(years)xAST(U/L))/(plt(x109/L)xsqrt ALT(U/L))). Cut point low-medium risk ≥1.3 and medium-high risk >2.67. [227,229,234]
- NAFLD Fibrosis Score (NFS) was calculated as: 1.675+(0.037xage(years))+(0.094xBMI(kg/m2))+(1.13xIFG/diabetes (yes=1, no=0))+(0.99x(AST(U/L)/ALT(U/L))- (0.013xplatelet count(×109/L))-

(0.66xalbumin (g/dL)). Cut-point for low-medium risk ≥-1.455, medium-high risk >0.676. [230]

- Fatty liver index (FLI) was calculated as: ey/(1+ey)x100 where y=0.953 x ln(triglycerides, mg/dl) + 0.139 x BMI, kg/m2 + 0.718 x ln (γGT, U/L) + 0.053 x waist circumference, cm – 15.745). [259]
- The EASL-EASD-EASO referral decision algorithm (Figure 2) was used. [217]

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 the present study which has reported data from the entire cohort. [265]

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 was 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. Due to 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.

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.

3.5 Results

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 57mmol/mol (7.4%) and mean BMI 31.4 kg/m2. Alcohol intake was above recommended limits in 19.9% and 14.4% were current smokers (Table 7). Seven people had prevalent cirrhosis/HCC.

Incident cirrhosis/HCC

Of 1059 people without cirrhosis/HCC at baseline, 43 developed this outcome over 11 years of follow-up (11-year incidence 4.1%) (

Figure 3). Twenty-three cases were 'screen-detected' from year 1 clinic, and 8 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). 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, 9 both cirrhosis and HCC and 4 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).

Performance of fibrosis risk scores in predicting incident cirrhosis/HCC

Table 8 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 ($p<0.05$). 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.

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

Table 8 describes how the EASL-EASD-EASO algorithm outcome was associated with the development of cirrhosis/HCC in our cohort. 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 8). Irrespective of the fibrosis score used in the algorithm, people categorised as requiring referral were significantly more likely to develop cirrhosis/HCC (OR's range 0.1-13.7 with wide CIs, all $p < 0.05$). 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' while 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.

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 9). 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 (Table 10).

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 (T

Table 11- Table 13).

3.6 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) [23]. 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), while 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, while 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 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 characterised 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 due to other causes or with known co-factors (e.g. 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 3 participants who had definite non-NAFLD disease did not reveal significantly different results (Table 10). Medication exposure data was 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. Due to 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 (Table 12- Table 13) 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). [185,231,234] 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). [221,229,232] 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). [237] 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. By 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 is 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. [237] 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. [21,75] Risk prediction scores and international guidelines have attempted to provide non-invasive methods of assessing risk of incident disease in this high-risk population. The present

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.

3.7 Additional Information

Further studies have been published that relate to this work since this paper was submitted for publication. Of particular note are two studies. The first used the database from the National Health and Nutrition Examination Survey (NHANES) in the US and undertook a cross-sectional analysis of 2940 adults with T2DM, applying the EASL-EASD-EASO algorithm to them. Using FLI as the steatosis marker, around 40% of participants with T2DM in that study would have been identified as requiring referral to hepatology services. [266] There is no associated data on outcome in these participants. Secondly, a study has examined the association between FIB-4 and transient elastography (an imaging method of identifying hepatic fibrosis), showing a poorer association between FIB-4 score and elastography result in those with T2DM than those without. [238] These two studies are consistent with the findings of this study which highlight that risk-stratification tools perform sub-optimally in people with T2DM and work to improve risk-stratification methods is needed.

Table 7. 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)†		207	(19.9)

Values are mean (sd) or n (%)

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

Table 8. Performance of fibrosis scores in prediction of 11-year incident cirrhosis/HCC

Fibrosis score (cut-point value used)	OR (95% CI)	AIC	C-Stat.	Sens (%, 95% CI)	Spec (%, 95% CI)	PPV (%, 95% CI)	NPV (%, 95% CI)	False ⁺ve n (%)	False ⁻ve n (%)
INDIVIDUAL SCORES									
ELF (≥10.51) †	30.4 (11.3-83.5)***	198.6	0.67	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)***	276.0	0.80	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)***	334.0	0.70	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)*	336.1	0.65	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)***	309.7	0.73	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)***	311.6	0.75	82 (67-93)	59 (56-63)	8 (5-11)	99 (98-100)	402 (41)	7 (18)
FIB4 (>2.67)	39.5 (17.4-91.9)***	277.1	0.76	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) †	10.1 (4.6-25.6)***	275.7	0.79	82 (66-92)	66 (63-69)	10 (7-14)	99 (97-99)	276 (34)	7 (18)
APRI (>0.5) †	13.0 (6.1-31.0)***	291.6	0.82	80 (64-91)	73 (70-75)	10 (7-14)	99 (98-100)	274 (27)	8 (20)
AST:ALT (≥0.8) †	0.1 (0.1-0.3)***	315.7	0.77	80 (65-91)	63 (60-66)	8 (6-11)	99 (98-99)	367 (37)	8 (20)
NFS (≥-1.455) †	6.0 (2.5-17.6)***	319.1	0.72	88 (73-96)	44 (41-47)	6 (4-9)	99 (97-100)	527 (56)	5 (13)
FIB 4 (≥1.3) †	11.7 (5.0-34.5)***	300.3	0.79	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) †	9.1 (4.3-21.7)***	285.9	0.78	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)***	292.8	0.82	80 (64-91)	72 (69-75)	10 (7-14)	99 (98-100)	281 (28)	8 (20)
AST:ALT (≥0.8)	0.1 (0.1-0.2)***	306.4	0.79	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)**	327.0	0.70	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)***	301.1	0.79	90 (76-97)	58 (55-61)	8 (6-11)	99 (98-100)	422 (42)	4 (10)

† ELF and ultrasound measured at year 1 only- so calculated 10 not 11-year incident cirrhosis/HCC
* p<0.05, ** p<0.01, ***p<0.001
OR odds ratio (age and sex adjusted), **C-Stat** C-statistic, **AIC** akaike information criterion, **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 2)

Table 9. 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) †	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) †	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) †	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) †	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) †	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) †	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) †	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)

† ELF and ultrasound measured at year 1 only- so calculated 10 not 11-year incident cirrhosis/HCC
* p<0.05, ** p<0.01, ***p<0.001
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

Table 10. Sensitivity Analysis- Performance of the fibrosis scores in the prediction of 11 year incident cirrhosis/HCC (excluding those with definite non-NAFLD disease (n=3))

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) †	35.44 (12.90-100.18)***	44 (24-65)	98 (97-99)	46 (26-67)	99 (96-99)	13 (2)	14 (56)
APRI (>0.5)	24.80 (11.80-53.60)***	54 (37-71)	94 (93-96)	26 (17-37)	98 (97-99)	57 (6)	17 (46)
AST:ALT (≥0.8)	3.36 (1.73-6.60)***	45 (29-62)	26 (23-29)	2 (1-4)	93 (89-95)	747 (74)	21 (55)
NFS (≥-1.455)	5.28 (1.58-32.83)*	95 (82-99)	22 (20-25)	4 (3-6)	99 (97-100)	769 (78)	2 (5)
NFS (>0.676)	8.48 (4.22-16.88)***	46 (29-63)	91 (89-93)	16 (10-25)	98 (97-99)	88 (9)	20 (54)
FIB 4 (≥1.3)	8.86 (3.82-24.22)***	84 (68-94)	59 (56-62)	7 (5-10)	99 (98-100)	402 (41)	6 (16)
FIB4 (>2.67)	40.72 (17.48-97.37)***	41 (25-58)	98 (97-99)	44 (27-62)	98 (97-99)	19 (2)	22 (59)
SCORES USED WITHIN EASL-EASD-EASO ALGORITHM, USS STEATOSIS AS STEATOSIS MARKER							
ELF (≥10.51) †	11.26 (4.88-30.65)***	83 (66-93)	66 (63-69)	10 (6-13)	99 (98-100)	276 (34)	6 (17)
APRI (>0.5) †	14.18 (6.40-36.03)***	81 (65-92)	73 (70-75)	10 (7-14)	99 (98-100)	274 (27)	7 (19)
AST:ALT (≥0.8) †	8.36 (3.83-20.99)***	82 (66-92)	63 (60-66)	8 (5-11)	99 (98-100)	367 (37)	7 (18)
NFS (≥-1.455) †	7.13 (2.79-24.16)***	89 (75-97)	44 (41-47)	6 (4-8)	99 (98-100)	527 (56)	4 (11)
FIB 4 (≥1.3) †	13.98 (5.47-47.34)***	89 (75-97)	60 (57-64)	8 (6-11)	99 (98-100)	384 (40)	4 (11)
SCORES USED WITHIN EASL-EASD-EASO ALGORITHM, FATTY LIVER INDEX AS STEATOSIS MARKER							
ELF (≥10.51) †	9.89 (4.47-25.08)***	81 (64-92)	67 (64-70)	9 (6-13)	99 (97-100)	277 (33)	7 (19)
APRI (>0.5)	13.81 (6.23-35.07)***	81 (65-92)	72 (69-75)	10 (7-13)	99 (98-100)	281 (28)	7 (19)
AST:ALT (≥0.8)	7.38 (3.39-18.49)***	82 (66-92)	61 (57-64)	7 (5-10)	99 (98-100)	397 (39)	7 (18)
NFS (≥-1.455)	5.22 (2.04-17.68)**	89 (75-97)	36 (33-39)	5 (3-7)	99 (97-100)	635 (64)	4 (11)
FIB 4 (≥1.3)	12.70 (4.97-43.01)***	89 (75-97)	58 (55-61)	7 (5-10)	99 (98-100)	422 (42)	4 (11)

† ELF and ultrasound measured at year 1 only- so calculated 10 not 11-year incident cirrhosis/HCC
* p<0.05, ** p<0.01, ***p<0.001
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

Table 11. Performance of fibrosis scores in prediction of 11-year incident cirrhosis/HCC, competing risks regression analysis with non-liver death as the competing risk

Fibrosis score (cut-point value used)	CRR Hazard (95% CI)		BIC
INDIVIDUAL SCORES			
Null Model		n/a	348.95
ELF (≥10.51) †	24.84	(9.98-61.83)***	310.64
APRI (>0.5)	18.94	(9.96-35.99)***	313.66
AST:ALT (≥0.8)	3.85	(2.00-7.40)***	342.88
NFS (≥-1.455)	3.65	(1.14-11.63)*	353.06
NFS (>0.676)	7.59	(4.06-14.19)***	331.11
FIB 4 (≥1.3)	7.77	(3.48-17.31)***	332.79
FIB4 (>2.67)	30.52	(15.18-61.35)***	309.52
SCORES USED WITHIN EASL-EASD-EASO ALGORITHM, USS STEATOSIS AS STEATOSIS MARKER			
Null model		n/a	348.95
ELF (≥10.51) †	9.56	(4.13-22.12)***	333.82
APRI (>0.5) †	12.09	(5.54-26.35)***	329.59
AST:ALT (≥0.8) †	7.38	(3.41-15.97)***	338.60
NFS (≥-1.455) †	5.70	(2.28-14.78)***	348.16
FIB 4 (≥1.3) †	11.11	(4.31-28.65)***	337.66
SCORES USED WITHIN EASL-EASD-EASO ALGORITHM, FATTY LIVER INDEX AS STEATOSIS MARKER			
Null model		n/a	348.86
ELF (≥10.51) †	8.60	(3.87-19.11)***	333.96
APRI (>0.5)	11.77	(5.39-25.72)***	330.31
AST:ALT (≥0.8)	7.72	(3.41-17.46)***	333.03
NFS (≥-1.455)	5.48	(1.93-15.56)**	346.70
FIB 4 (≥1.3)	13.02	(4.58-37.07)***	331.15

† ELF and ultrasound measured at year 1 only- so calculated 10 not 11-year incident cirrhosis/HCC

* p<0.05, ** p<0.01, ***p<0.001

CRR Hazard Exponentiated coefficient of the subdistribution hazard model (Fine and Gray), adjusted for age and sex**, BIC** Bayesian Information Criterion, **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 2)

Note- all available case analysis used for calculation of the CRR hazard, complete case only analysis used for calculation of BIC

Table 12. Sensitivity Analysis- Performance of the fibrosis scores in the prediction of 11-year incident Cirrhosis/HCC (excluding those with definite non-NAFLD disease (n=3)), competing risks regression analysis with non-liver death as the competing risk

Fibrosis score (cut-point value used)	CRR Hazard (95% CI)	BIC
INDIVIDUAL SCORES		
Null Model	n/a	348.78
ELF (≥10.51) †	29.04 (11.17-75.49)***	310.54
APRI (>0.5)	20.61 (10.59-40.10)***	313.57
AST:ALT (≥0.8)	3.33 (1.71-6.51)***	342.78
NFS (≥-1.455)	5.21 (1.27-21.37)*	352.92
NFS (>0.676)	7.85 (4.12-14.98)***	331.01
FIB 4 (≥1.3)	8.55 (3.60-20.28)***	332.75
FIB4 (>2.67)	31.76 (15.45-65.29)***	309.41
SCORES USED WITHIN EASL-EASD-EASO ALGORITHM, USS STEATOSIS AS STEATOSIS MARKER		
Null model	n/a	348.78
ELF (≥10.51) †	10.61 (4.29-26.28)***	333.75
APRI (>0.5) †	13.22 (5.74-30.46)***	329.53
AST:ALT (≥0.8) †	8.03 (3.52-18.34)***	338.53
NFS (≥-1.455) †	6.91 (2.44-19.61)***	348.00
FIB 4 (≥1.3) †	13.24 (4.65-37.73)***	337.60
SCORES USED WITHIN EASL-EASD-EASO ALGORITHM, FATTY LIVER INDEX AS STEATOSIS MARKER		
Null model	n/a	348.69
ELF (≥10.51) †	9.34 (3.98-21.97)***	333.89
APRI (>0.5)	12.89 (5.59-29.74)***	330.25
AST:ALT (≥0.8)	7.12 (3.12-16.25)***	332.99
NFS (≥-1.455)	5.10 (1.78-14.58)***	346.64
FIB 4 (≥1.3)	12.11 (4.23-34.70)***	331.11

† ELF and ultrasound measured at year 1 only- so calculated 10 not 11-year incident cirrhosis/HCC

* p<0.05, ** p<0.01, ***p<0.001

CRR Hazard Exponentiated coefficient of the subdistribution hazard model (Fine and Gray), adjusted for age and sex**, BIC** Bayesian Information Criterion, **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 2)

Note- all available case analysis used for calculation of the CRR hazard, complete case only analysis used for calculation of BIC

Table 13. Sensitivity Analysis- Performance of fibrosis scores in prediction of 11-year incident cirrhosis related varices, ascites, encephalopathy or HCC, competing risks regression analysis with non-liver death as the competing risk

Fibrosis score (cut-point value used)	CRR Hazard (95% CI)	BIC
INDIVIDUAL SCORES		
Null Model	n/a	258.30
ELF (≥10.51) †	24.46 (8.11-73.77)***	233.20
APRI (>0.5)	26.31 (12.08-57.32)***	231.13
AST:ALT (≥0.8)	5.45 (2.45-12.15)***	252.95
NFS (≥-1.455)	4.02 (0.96-16.82)	262.65
NFS (>0.676)	8.04 (3.79-17.10)***	246.64
FIB 4 (≥1.3)	15.07 (4.55-49.92)***	243.71
FIB4 (>2.67)	46.02 (20.89-101.38)***	226.42
SCORES USED WITHIN EASL-EASD-EASO ALGORITHM, USS STEATOSIS AS STEATOSIS MARKER		
Null model	n/a	258.30
ELF (≥10.51) †	6.18 (2.58-14.80)***	254.92
APRI (>0.5) †	10.39 (4.20-25.71)***	251.67
AST:ALT (≥0.8) †	5.36 (2.29-12.56)***	257.07
NFS (≥-1.455) †	3.66 (1.40-9.57)**	262.05
FIB 4 (≥1.3) †	7.16 (2.68-19.12)***	255.91
SCORES USED WITHIN EASL-EASD-EASO ALGORITHM, FATTY LIVER INDEX AS STEATOSIS MARKER		
Null model	n/a	258.24
ELF (≥10.51) †	6.38 (2.66-15.33)***	254.99
APRI (>0.5)	10.13 (4.09-25.14)***	252.10
AST:ALT (≥0.8)	7.37 (2.80-19.44)***	252.31
NFS (≥-1.455)	4.92 (1.48-16.38)*	260.40
FIB 4 (≥1.3)	11.91 (3.58-39.62)***	250.28

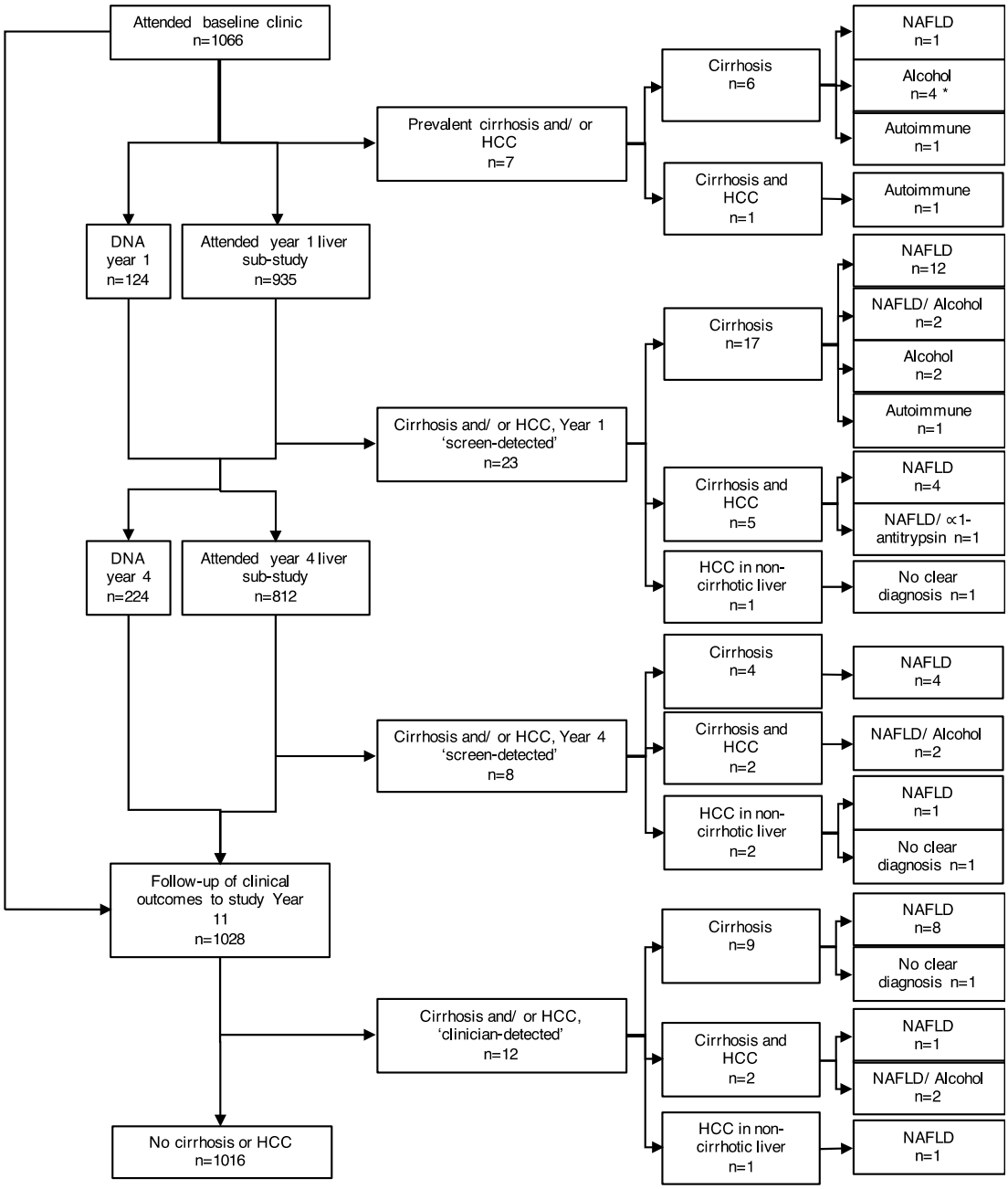
† ELF and ultrasound measured at year 1 only- so calculated 10 not 11-year incident cirrhosis/HCC

* p<0.05, ** p<0.01, ***p<0.001

CRR Hazard Exponentiated coefficient of the subdistribution hazard model (Fine and Gray), adjusted for age and sex**, BIC** Bayesian Information Criterion, **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 2)

Note- all available case analysis used for calculation of the CRR hazard, complete case only analysis used for calculation of BIC

Figure 3. Cirrhosis and HCC events at baseline and through 11-year follow-up



* 1 participant with alcohol related cirrhosis at baseline developed HCC during follow-up.
It should be noted that some of the 'clinician-detected' cases of cirrhosis/ HCC were detected at time-points before some Year 1 and Year 4 'screen-detected' cases

Chapter 4 Addition of hyaluronic acid to the FIB-4 liver fibrosis score improves prediction of incident cirrhosis and hepatocellular carcinoma in Type 2 diabetes: The Edinburgh Type 2 Diabetes Study

This section has been submitted and is published in Obesity Science & Practice under the same title by Sheila M Grecian (SMG), Stela McLachlan (SM), Jonathan A Fallowfield (JF), Peter C Hayes (PH), Indra Neil Guha (NG), Joanne R Morling (JM), Stephen Glancy (SG), Rachel M Williamson (RW), Rebecca M Reynolds (RR), Brian M Frier (BF), Nicola N Zammitt (NZ), Jackie F Price (JP) and Mark WJ Strachan (MS). SMG wrote the manuscript. JP was principal investigator of the ET2DS, designed the study, analysed and interpreted the data. MS was lead investigator of the ET2DS liver sub-study, designed the study, analysed and interpreted the data. RR, BF, PH, JF, RW, NG and SG contributed to study design. SMG, SM, RW and JM contributed to data collection, analysis and interpretation. All authors contributed to revision and final approval of the article.

In summary, in the previous chapter it was discussed that whilst it is considered beneficial to screen for NAFLD in the context of T2DM, existing NAFLD fibrosis risk prediction tools perform less well. In this study we investigated whether combining additional biomarkers with existing risk prediction tools could improve the associative and predictive ability of these tools in the identification of incident cirrhosis and HCC in a cohort of people with diabetes.

Further information regarding how potential biomarkers were identified, and additional detail on validation of the models built can be found in appendix 3. Formulae used for predictive ability calculations can be found in appendix 2.

Please note- all tables for this section sit at the end of the chapter text.

4.1 Abstract

Background: Type 2 diabetes is associated with increased risk of progression to cirrhosis and hepatocellular carcinoma (HCC) in people with chronic liver diseases, particularly non-alcoholic fatty liver disease (NAFLD). However, the absolute risk of progression is low so it is crucial to accurately identify patients who would benefit most from hepatology referral and intensified management. Current risk-stratification tools are sub-optimal and perform worse in people with diabetes.

Aims: To determine whether the addition of complementary biomarker(s) to current NAFLD risk-stratification tools in people with Type 2 diabetes could improve the identification of people who are at increased risk of developing incident cirrhosis or HCC.

Methods: The Edinburgh Type 2 diabetes Study (ET2DS) is a cohort study of men and women with Type 2 diabetes (n=1066, age 60-75 at baseline). Cases of cirrhosis and HCC were identified over 11-years of follow-up. Biomarkers were measured at baseline and year one and association with incident disease assessed using logistic regression.

Results: Of existing risk-stratification scores tested, the Fibrosis-4 (FIB-4) index and the AST:platelet ratio index (APRI) performed best in our cohort. Addition of hyaluronic acid (cut-point ≥50μg/L) to FIB-4 (cut-point ≥1.3) maintained a false negative rate ≤25% and reduced the number of people incorrectly identified as 'high-risk' for incident disease by ~50%.

Conclusions: The addition of hyaluronic acid to FIB-4 reduced the proportion of people inappropriately identified as 'high-risk' for development of cirrhosis/HCC in a community population of otherwise asymptomatic people with Type 2 diabetes. These findings require validation in independent cohorts.

4.2 Introduction

Non-alcoholic fatty liver disease (NAFLD) is recognised as the liver component of the metabolic syndrome, a cluster of conditions including abdominal obesity, impaired glucose regulation or diabetes, hypertension, hypercholesterolaemia and hypertriglyceridaemia, which are associated with increased cardiovascular risk. [84] With rising population levels of obesity, prevalence of NAFLD is rising and 25% of people globally may be affected. [8] Type 2 diabetes is associated with a further increased prevalence of NAFLD, the prevalence of NAFLD steatosis being 40-70%.[9-11,263] Furthermore, people with Type 2 diabetes have a higher incidence of, and risk of progression to, cirrhosis and hepatocellular carcinoma (HCC). [20,21,71,75,265]

In Type 2 diabetes, identifying those at increased risk of developing cirrhosis/HCC is important to prompt intensified lifestyle interventions, enhanced monitoring of disease progression and timely initiation of surveillance for varices and HCC. Screening for NAFLD in Type 2 diabetes is advocated in European guidelines (European Association for the Study of the Liver, European Association for the Study of Diabetes, European Association for the Study of Obesity (EASL-EASD-EASO)). [217]

Liver biopsy is the gold standard test for staging NAFLD, with histological fibrosis the most important factor predictive of disease progression in meta-analyses. [179,180] However, biopsy is not suitable for population screening as it is an invasive procedure with a risk of serious complications. Consequently, interest in the identification of non-invasive markers that predict those at risk of disease progression has increased. Many scores have been developed and validated in NAFLD, including the Fibrosis-4 Index (FIB-4), NAFLD Fibrosis Score (NFS), aspartate aminotransferase (AST):alanine aminotransferase (ALT) ratio, AST to Platelet Ratio Index (APRI) and Enhanced Liver Fibrosis (ELF) test). [220,221,224,227,230] While these were initially developed to identify liver fibrosis at the time of testing, their ability to predict incident cirrhosis and HCC has also been validated. [71,185,225,231,234]

The performance of these scores varies between research cohorts. [185,189,225,231,234] Typically, study populations have consisted of patients attending hepatology secondary care services and there is much less evidence to support their utility in community populations. Furthermore, these scores perform less well in people with Type 2 diabetes, with one study reporting that, over 4 years follow-up, 15% of people with diabetes with a ‘low-risk’ FIB-4 score developed decompensated cirrhosis, and

17% developed HCC; by contrast, in individuals without diabetes, no participant with a 'low-risk' score developed decompensated cirrhosis or HCC. [237] This group has reported that use of current risk-stratification tools would have resulted in large numbers of people who did not develop cirrhosis/HCC over 11-years follow-up being classified as 'high-risk' (41% with FIB-4), while a significant proportion (18% with FIB-4) who did develop cirrhosis/HCC were classified as 'low-risk' at baseline. [262]

This study hypothesised that the addition of a complementary biomarker(s) could improve the performance of current risk-stratification tools for the accurate identification of people with Type 2 diabetes who are at increased risk of developing cirrhosis or HCC.

4.3 Methods

The Edinburgh Type 2 Diabetes Study

The Edinburgh Type 2 Diabetes Study (ET2DS) is a community-based prospective cohort study of older people with Type 2 diabetes. Full methods have been described previously. [255] Briefly, 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 all people with Type 2 diabetes living in Lothian, Scotland), and were subsequently found to be largely representative of this sampling population. [256]

Invitations to participate were sent to 5454 people, of whom 1066 (20%) attended baseline clinic. All 1066 were invited to re-attend a clinical and liver assessment after 1 and 4 years. 939 attended the year 1 clinic (deceased n=15, unable to contact n=19, unable to attend n=93) and 831 the year 4 clinic (deceased n=88, unsuitable for clinical reasons n=26, uncontactable n=23, unable n=98). All 1066 participants were followed up for outcomes until death (320 participants) or end of follow-up in 2018.

Data collection- baseline biomarker assessment

Assessments were undertaken at dedicated research clinics at the Wellcome Trust Clinical Research Facility, Western General Hospital, Edinburgh, UK by specially trained research staff using Standard Operating Procedures. [255] Fasting venous blood samples were collected at baseline. Glycated haemoglobin (HbA1c), ALT, AST, alkaline phosphatase (ALP), gamma-glutamyltransferase (γGT), albumin, bilirubin and platelets were analysed using a Vitros Fusion chemistry system (Ortho Clinical Diagnostics, Bucks, UK) at the Western General Hospital. C-reactive protein (CRP) was measured using an immunonephelometric assay; interleukin-6 (IL-6) and tumour necrosis factor-alpha (TNFα) were measured using ELISA (R&D Systems, Oxon, UK), Glasgow Royal Infirmary, UK. Hyaluronic acid was measured using a radiometric assay (Pharmacia, Uppsala, Sweden). The Enhanced Liver Fibrosis test (ELF) was measured on fasting venous blood samples from the year-1 clinic and analysed using the ADVIA Centaur immunoassay system (Siemens Healthcare Diagnostics Inc., New York, USA) at the iQur laboratory (London, UK).

Participants attending the year-1 research clinics underwent a full diagnostic liver screen if serum liver enzymes or abdominal ultrasound was abnormal (including Hepatitis B and C serology, liver autoantibody titres, alpha-fetoprotein, ferritin) and all completed standard questions on alcohol consumption (AUDIT-C questionnaire), 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 (ALP) >125 U/L), AST:ALT ratio >1, hyaluronic acid >100µg/L (in the absence of known joint disease), positive liver autoantibodies, ferritin >1000ng/mL, alpha-feto protein >6ng/mL, positive hepatitis B or C serology, spleen diameter >13cm, platelets <150x109/L (in the absence of known haematological cause), or suspected cirrhosis on ultrasound, was referred for specialist hepatology review.

Fibrosis scores were calculated and cut-point levels used as per published work.

- AST to Platelet Ratio Index (APRI) was calculated as: ((AST(U/L)/Upper limit normal)/platelets(x109/L))x100. [220]
- AST:ALT ratio was calculated as: AST(U/L)/ALT(U/L). [221]
- Fibrosis-4 index (FIB-4) was calculated as ((age(years)xAST(U/L))/(plt(x109/L)xsqrt ALT(U/L))). [227,229,234,267]
- NAFLD Fibrosis Score (NFS) was calculated as: 1.675+(0.037xage(years))+(0.094xBMI(kg/m2))+(1.13xIFG/diabetes (yes=1, no=0))+(0.99x(AST(U/L)/ALT(U/L))- (0.013xplatelet count(×109/L))-(0.66xalbumin (g/dL)). [230]

Data Collection - Identification of liver disease

Possible prevalent liver disease was identified by self-completion questionnaire at baseline with subsequent confirmation if a clinician diagnosis was recorded in primary or secondary care medical records. Incident cirrhosis/HCC was identified and confirmed using multiple data sources, including review of all participants' hospital medical notes (TrakCare, InterSystems Corp., Cambridge, USA) at 11-year follow-up; responses recorded in patient and GP questionnaires sent at year 4 and year 10 follow-up; hospital discharge data (diagnosis and death codes) collated by ISD (Information Services Division, NHS Scotland) and collected at year 8 follow-up). All confirmed cases required clinician diagnosis in secondary care medical notes. Participants were identified as having 'screen-detected' cirrhosis/HCC if they were

referred to hepatology following year 1 or 4 research clinic investigation and remained under hepatology follow-up until definitive diagnosis was made. People with prevalent cirrhosis or HCC at baseline were excluded from analysis on incident disease.

Data Analysis

Data were analysed with R (R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.) using complete-case analysis. <5% of data was missing for all variables with the exception of ELF (n=681) and ultrasound (n=933). Logistic regression was used to identify the strength of association between baseline scores and biomarkers, and incident cirrhosis/HCC in this cohort. Best performing existing risk scores were chosen as the base models; assessed on performance using C-statistic (to assess discrimination), the Hosmer-Lemeshow test for Logistic Regression (to assess calibration, >0.05 accepted) and Aikaike Information Criterion (AIC) (a measure of overall model performance). Correlation between FIB-4 and APRI risk scores was assessed using Pearson correlation coefficient.

The strength of association of additional baseline variables that have been previously reported as potentially associated with pathogenesis or progression of liver disease were assessed. These were demographics (sex, deprivation index (SIMD), smoking status and alcohol intake), duration of Type 2 diabetes and HbA1c, metabolic variables (BMI, waist-hip ratio, cholesterol), markers of liver function and injury (ALP, γGT, bilirubin, albumin and hyaluronic acid) and markers of inflammation (IL-6, CRP and TNF∝). Hyaluronic acid, TNFα and γGT were log-transformed (natural log) to ensure linearity of response to the logit. Biomarkers that remained significantly associated with outcome after correction for markers in the base models, age and sex, were assessed individually and in combination when added to the base models using C-statistic, AIC and Hosmer-Lemeshow test. Because of the number of cases of cirrhosis/HCC in this cohort (n=43) a maximum of 3 additional biomarkers were added.

Due to this cohort's mixed population of screen-detected and clinician-diagnosed outcomes, possibly skewing time-to-event data as those who were screen-detected were often diagnosed at a pre-symptomatic stage, the primary analysis (logistic regression) did not include a time component. A sensitivity analysis was undertaken using competing risks regression to assess whether there was a significant impact of the competing risk of non-liver death on final model performance. The Bayesian

Information Criterion (BIC) was used to assess model performance for the competing risks regression. A second sensitivity analysis was undertaken excluding any participant with definite non-NAFLD disease.

The impact of adding the biomarkers that best improved the performance of models by AIC was assessed through calculation of sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), false positive and false negative rate. To undertake this, dichotomous cut-points needed to be allocated for values of the base model and for biomarkers used. A complete-case analysis was undertaken for model development, with only those participants with all biomarker information available included (n=999, of whom 39 developed cirrhosis/HCC).

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.

4.4 Results

Participant characteristics and incident events

Baseline characteristics are detailed in Table 14. Mean age was 67.9 years and 51.3% were male. Mean duration of Type 2 diabetes was 8 years, HbA1c 7.4% (57mmol/mol) and BMI 31.4 kg/m2. Participants were predominantly of Caucasian ethnicity (98.3%) and 7 (0.01%) had cirrhosis/HCC. During follow-up, 43 participants were identified with incident cirrhosis/HCC. Of these, 39 developed cirrhosis of whom 58% developed varices, ascites or encephalopathy. There were 13 cases of HCC (9 participants developed both cirrhosis and HCC). The aetiology of incident disease was NAFLD (n=31), mixed NAFLD and alcohol (n=6), mixed NAFLD and α-1 antitrypsin deficiency (n=1), alcohol (n=2), autoimmune (n=1), or no clear diagnosis (n=3).

Identification of base risk-stratification model

The performance of 5 pre-selected risk scores in the ET2DS study population is shown in

Table 15. The risk scores that showed best association between score and outcome (cirrhosis/HCC) by logistic regression assessment were FIB-4 (C-statistic 0.86, AIC 244.5) and APRI (C-statistic 0.85, AIC 246.5) and these were chosen as base models to assess any incremental benefit of additional biomarkers. Correlation between APRI and FIB-4 scores in the individuals with and without incident cirrhosis/HCC was high (Pearson's r>0.9).

Association of individual biomarkers with incident cirrhosis/HCC

Individual baseline biomarkers, in addition to those already in the FIB-4 and APRI risk scores, were assessed for their association with incident cirrhosis/HCC by odds ratio (OR) (Table 16). SIMD, HbA1c, BMI, ALP, γGT, bilirubin, hyaluronic acid, TNF∝, IL-6 and CRP were associated ($p<0.1$) in univariable analysis. SIMD, BMI, HbA1c, γGT, hyaluronic acid, IL-6 and CRP remained associated ($p<0.05$) after adjustment for

age, sex and individual factors already in the base models (AST and platelets in both FIB-4 and APRI, plus ALT and age in FIB-4) (Table 16).

Addition of individual biomarkers to base prediction model

Individual biomarkers were added to the base models, and the association with incident cirrhosis/HCC was assessed using logistic regression (Table 17). Those that improved FIB-4 model performance most in terms of AIC were HbA1c (improvement in AIC of base model from 238.2 to 228.7), hyaluronic acid (209.4) and γGT (205.5). Hyaluronic acid and γGT addition also showed the greatest increase in C-statistic performance (from 0.85 to 0.89 and 0.93 respectively). For APRI, improvement in AIC was also seen most clearly with HbA1c (from 243.8 to 236.2), hyaluronic acid (211.2) and γGT (219.1), though with only modest C-statistic improvements. When hyaluronic acid alone was added to APRI, the Hosmer-Lemeshow test was significant, indicating poor calibration.

Hyaluronic acid, γGT and HbA1c were chosen to fit to mixed models (Table 17). Regardless of the base model used, the addition of both hyaluronic acid and γGT further improved model performance, with AIC decreasing to 184.5 (FIB-4 as base model) or 192.9 (APRI as base model). The addition of HbA1c to either hyaluronic acid, γGT or both did not improve AIC or C-statistic substantially beyond the improvement gained by hyaluronic acid and γGT alone.

A sensitivity analysis was undertaken using competing risk regression analysis (non-liver death as the competing risk), which supports the finding that the addition of hyaluronic acid and/ or γGT provides the best improvement in model performance (Table 18).

Predictive accuracy of the base models plus additional biomarkers

The models that performed best according to AIC and C-statistic (base models plus hyaluronic acid, γGT, HbA1c or combinations) were assessed for accuracy in the prediction of incident cirrhosis/HCC using sensitivity, specificity, PPV, NPV, false positive and negative rates. APRI plus hyaluronic acid alone was not assessed further due to poor calibration. Cut-points used were: for FIB-4 the ‘high risk’ of fibrosis (>2.67), ‘medium to high risk’ of fibrosis (≥1.3) and the ‘medium to high risk’ adjusted for age (>2) cut-point; for APRI the ‘medium to high risk’ of fibrosis (>0.5) cut-point;

for hyaluronic acid ≥100μg/L (appropriate for identification of fibrosis) and ≥50μg/L; for γGT the laboratory cut-point of >55U/L, and >20U/L; for HbA1c >7.5. The second lower cut-points for hyaluronic acid and γGT were chosen arbitrarily, with the aim of attempting to reduce false negative results.

Hyaluronic acid (cut-point >50μg/L) plus FIB-4 (≥1.3) was the only model with a false negative rate ≤25% (n=10/40), thus correctly identifying the majority of those truly at high risk at baseline (

Table 19). FIB-4 plus hyaluronic acid (cut-point ≥50μg/L) reduced the number of people assessed as 'high-risk' that did not develop cirrhosis/HCC during follow-up (i.e. false positive rate) by 46% (399 to 214). Results were similar using the combined fibrosis marker as part of the EASL-EASD-EASO algorithm. Using APRI as a base model, false negative rates were ≥50%.

Two sensitivity analyses were undertaken, one excluding participants with definite non-NAFLD disease and another excluding participants who developed HCC in a non-cirrhotic liver. Neither analysis materially changed the results (Table 20,

Table 21).

4.5 Discussion

Serum hyaluronic acid in conjunction with the FIB-4 risk-stratification score reduced the number of false positive results in this cohort, without substantially increasing false negative results, either in isolation or within the EASL-EASD-EASO algorithm. To this team's knowledge, this is the first study to examine the use of hyaluronic acid for risk-stratification of liver disease in a community population with Type 2 diabetes.

Addition of hyaluronic acid improved the association of the FIB-4 model with incident cirrhosis/HCC. Moreover, when hyaluronic acid (cut-point ≥50μg/L) was added to the FIB-4 risk-stratification tool, the number of people inappropriately classified as 'high-risk' was reduced by 46% (n=399 to n=214), while increasing those inappropriately classified as 'low-risk' from 18% to 25% (n=7 to n=10). APRI performed similarly to FIB-4 as a base model. Both have similar component factors and the scores were highly correlated. Therefore, the additive effect of using both markers in combination was not assessed. The addition of hyaluronic acid to APRI had poor calibration and was not assessed further in isolation. A 'high-risk' FIB-4 plus hyaluronic acid score was associated with a median time-to-diagnosis of cirrhosis/HCC of approximately 3 years, with the majority presenting within 6 years. Due to the often asymptomatic course of NAFLD, it seems likely that a significant proportion of these individuals had undiagnosed cirrhosis at the time of the baseline assessment, while the remainder had at least advanced fibrosis.

Strengths and weaknesses of the study

The ET2DS specifically studied liver outcomes in a community population of otherwise asymptomatic individuals with Type 2 diabetes, who did not necessarily have liver disease. Almost all other studies have identified outcomes in cohorts recruited from secondary care hepatology clinics, with established NAFLD diagnoses and likely advanced pathology. European guidelines recommend screening in populations like the one represented by the ET2DS cohort, making this a suitable testbed for assessing the impact of potential population screening strategies. [217] The ET2DS is a moderate sized cohort. Participants were well-characterised at baseline to allow accurate determination of any potential additional baseline risk factors and were followed up using multiple sources of information to accurately identify incident disease.

There are limitations to this study. The ET2DS is a single-centre study, undertaken in people aged 60-75 years, of predominantly Caucasian origin (98.3%) and care should be taken in extrapolating results to other populations. While the aetiology of incident disease was almost entirely NAFLD, cirrhosis/HCC of other aetiologies was included. There are known difficulties in determining the exact contributions of different aetiologies (or cofactors) in cirrhosis/HCC development, thus investigation of all liver disease seemed more relevant in a real-world setting. [268] A sensitivity analysis that excluded participants who developed definite non-NAFLD disease did not materially affect results. Medication exposure data was not analysed.

The incidence data may be an underestimate as it is possible that some asymptomatic participants who developed cirrhosis/HCC during follow-up were not identified, as screening for cirrhosis/HCC at 11-year follow-up was not repeated. Alternatively, our the incidence data may overestimate the clinical burden as a substantial proportion of diagnoses were made after hepatology referral following year 1 and year 4 screening investigations. NAFLD cirrhosis can have a silent natural history and may not manifest clinically for many years. Thus, some people who may never have developed overt cirrhosis, or may have died before their disease became clinically apparent may have been identified. However, 58% of those identified with cirrhosis developed varices, ascites and/or encephalopathy and 23% developed HCC, so it is likely that a large majority would have presented with clinical sequelae during follow-up.

Those who were diagnosed following screening in year 1 may have had undiagnosed cirrhosis/HCC at baseline. However, the study considered prevalent disease to be only that which was clinically apparent at baseline because the diagnosis of cirrhosis for some referred post-screening came many years following that referral (people were kept under active follow-up due to high-risk features for progression). Additionally, the time-to-diagnosis for those who were diagnosed following year 1 screening and those diagnosed following routine clinical referral significantly overlapped, suggesting that stage of disease in the two groups at baseline did not differ significantly. [262]

ELF was measured at the year 1 clinic (all other biomarkers at baseline), so this analysis used slightly different ‘baseline’ time points. However, this group has demonstrated previously that there is no significant difference in model performance using baseline or year 1 data; in addition, no participant was identified with incident

disease prior to the year 1 clinics. [262] Hyaluronic acid is known to be raised in the context of joint, as well as liver disease. As accurate data on joint disease prevalence for the whole cohort at baseline were not available, individuals with joint disease were not excluded. However, as hyaluronic acid was used in conjunction with other markers of liver fibrosis, isolated elevation of hyaluronic acid due to joint disease should not have had a material impact on the models.

This group has previously described the performance of current risk-stratification models in predicting cirrhosis and HCC in different cohorts. [262] In addition, risk-stratification scores perform worse in populations with diabetes than in those without. [237] Previous cohort studies have failed to consistently identify individual non-invasive biomarkers that are associated NAFLD progression. [179,180] This study demonstrates that using serum hyaluronic acid in conjunction with the FIB-4 risk-stratification score can reduce the burden of false positive results. Hyaluronic acid is a glycosaminoglycan found in connective tissue that is almost exclusively cleared by liver metabolism. Raised levels of hyaluronic acid are known to be associated with cirrhosis. [269] However, few studies have assessed it as a prognostic marker. In combination with other biomarkers as part of the ELF risk-stratification tool, hyaluronic acid is associated with fibrosis in NAFLD. [224] One study found a significant association with rising hyaluronic acid and all-cause mortality, liver mortality and liver transplant-free survival. [270] Thus, the present data, finding suggesting its utility in predicting those who are at ‘high-risk’ of developing incident cirrhosis/HCC, is consistent with published data.

The present findings derive from a single moderately-sized cohort and need validation in other independent cohorts. A change in FIB-4 plus hyaluronic acid over time was not examined. Moreover, there were too few individuals who developed cirrhosis/HCC to determine reliably if the median time-to-diagnosis was more prolonged in those with a ‘low-risk’ score compared to those with a ‘high-risk’ score. If the time-to-diagnosis was more prolonged in those with a ‘low-risk’ score, repeat assessment at intervals of several years might successfully identify additional individuals who would develop cirrhosis/HCC.

In conclusion, the prevalence of both NAFLD and Type 2 diabetes are rising in association with the rising population prevalence of obesity. Type 2 diabetes is associated with an increased risk of cirrhosis/HCC. [21,75] As a result, both European

and American guidelines advocate a high index of suspicion for liver disease in Type 2 diabetes, with European guidelines recommending routine screening. [7,217] However, current risk-stratification tools perform suboptimally, especially in diabetes. [237,262] This study shows that using a combination of FIB-4 and hyaluronic acid for risk-stratification can significantly reduce false positive rates without substantially increasing false negative rates. This makes this combination a possible candidate for community screening, as it would lead to identification of a substantial proportion of cases while reducing stress on health systems from false positive results. These findings are promising, but require further validation. Furthermore, the false positive rates for the FIB-4 and hyaluronic acid combination remain high and so it is acknowledged that better biomarkers are required for the identification of people with Type 2 diabetes at risk of developing cirrhosis/HCC.

Table 14. Baseline characteristics of the study population

Baseline Characteristic		ET2DS Population (n=1066)	
Age		67.9	(4.2)
Sex (male)		547.0	(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 T2DM (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)		154.0	(14.4)
Alcohol (excess)†		207.0	(19.9)

Values are mean (sd) or n (%)
† Defined as females >14 units/week, males >21 units/week or patient disclosed history of a current or prior alcohol problem
T2DM Type 2 diabetes, **HbA1c** glycated haemoglobin

Table 15. Odds ratios for the development of cirrhosis and HCC by fibrosis score

Fibrosis marker	Range in Population	OR adjusted for age and sex (95% CI)	p-value	AIC	C-Statistic	Hosmer-Lemeshow p-value
ELF	6.89-17.40	3.20 (2.18-4.84)	<0.001	195.2	0.83	<0.001
APRI	0.07-1.76	3.02 (2.37-3.94)	<0.001	246.5	0.85	0.10
AST:ALT	0.33-1.67	2.03 (1.61-2.57)	<0.001	318.4	0.73	0.03
NFS	-5.91-2.98	3.11 (2.21-4.46)	<0.001	297.3	0.80	0.001
FIB-4	0.41-7.82	3.42 (2.60-4.62)	<0.001	244.5	0.86	0.16

OR odds ratio, **AIC** Akaike Information Criterion.
OR calculated per increase of one standard deviation in marker.

Table 16. Association of additional predictive variables with cirrhosis or HCC

Variable		Total population (n=1059)	Population with cirrhosis/HCC (n=43)	Population without cirrhosis/ HCC (n=1016)	Univariable Analysis		Analysis adjusted for factors in existing models, age and sex	
					OR (95% CI)[§]	p	APRI (p)	FIB-4 (p)
Age		67.9 (4.2)	68.5 (4.7)	67.9 (4.2)	1.17 (0.86-1.60)	0.31	-	-
Sex (male)		544 (51.4)	18 (41.9)	526 (51.8)	0.67 (0.36-1.24)	0.21	-	-
SIMD quintile	1 (most deprived)	125 (11.8)	9 (20.9)	116 (11.4)	3.78 (1.38-10.79)	**0.01**	**0.01**	**0.01**
	2	206 (19.5)	8 (18.6)	198 (19.5)	1.97 (0.7-5.69)	0.20	**0.03**	0.06
	3	187 (17.7)	8 (18.6)	179 (17.6)	2.18 (0.77-6.3)	0.14	0.09	0.14
	4	193 (18.2)	11 (25.6)	182 (17.9)	2.94 (1.14-8.12)	**0.03**	**0.01**	**0.02**
	5 (least deprived)	348 (32.9)	7 (16.3)	341 (33.6)				
Duration T2DM (years)		8.1 (6.5)	9.1 (6.2)	8.0 (6.5)	1.17 (0.87-1.51)	0.27	-	-
HbA1c (%)		7.4 (1.1)	8.1 (1.5)	7.4 (1.1)	1.53 (1.21-1.90)	**<0.001**	**<0.001**	**<0.001**
HbA1c (mmol/mol)		57 (12)	65 (16.4)	57 (12)	-	-	-	-
BMI (kg/m^2)		31.4 (5.7)	33.7 (6.2)	31.3 (5.7)	1.44 (1.09-1.87)	**0.008**	**0.02**	**0.02**
Waist-Hip Ratio		0.97 (0.1)	0.98 (0.1)	0.96 (0.1)	1.20 (0.88-1.63)	0.25	-	-
Smoker (current)		153 (14.4)	8 (18.6)	145 (14.3)	1.37 (0.58-2.87)	0.43	-	-
Alcohol (excess)[†]		204 (19.8)	13 (30.2)	191 (18.8)	1.81 (0.9-3.46)	0.08	0.73	0.71
Cholesterol (mmol/L)		4.3 (0.9)	4.2 (0.8)	4.3 (0.9)	0.84 (0.59-1.16)	0.31	-	-
ALT (U/L)		43.2 (14.3)	53.4 (19.9)	42.8 (13.9)	1.56 (1.26-1.94)	**<0.001**	0.07	-
AST (U/L)		31.0 (10.4)	45.9 (15.4)	30.4 (9.7)	2.20 (1.78-2.74)	**<0.001**	-	-

Variable	Total population (n=1059)	Population with cirrhosis/HCC (n=43)	Population without cirrhosis/ HCC (n=1016)	Univariable Analysis		Analysis adjusted for factors in existing models, age and sex	
				OR (95% CI)§	p	APRI (p)	FIB-4 (p)
ALP (U/L)	91.7 (27.3)	106.1 (33.5)	91.1 (26.9)	1.45 (1.15-1.82)	**0.001**	0.21	0.12
γGT (U/L)†	29.4 (40.3)	96.7 (86.7)	26.7 (34.7)	3.55 (2.66-4.86)	**<0.001**	**<0.001**	**<0.001**
Bilirubin (μmol//L)	10.0 (4.7)	11.2 (4.1)	9.9 (4.7)	1.24 (0.94-1.56)	0.09	0.63	0.57
Albumin (g/L)	44.8 (3.3)	44.7 (3.8)	44.8 (3.3)	0.97 (0.71-1.33)	0.86	-	-
Platelets (10^9/L)	258.7 (69.3)	201.5 (77.4)	261.1 (68.0)	0.33 (0.22-0.49)	**<0.001**	-	-
Hyaluronic Acid (μg/L)‡	56.1 (46.6)	132.2 (85.3)	52.8 (41.3)	5.29 (3.42-8.47)	**<0.001**	**<0.001**	**<0.001**
TNF-$\propto$ (pg/mL)‡	1.4 (1.5)	1.6 (0.8)	1.3 (1.5)	1.63 (1.19-2.23)	**0.002**	0.05	0.08
IL-6 (pg/mL)	3.9 (3.5)	5.7 (3.9)	3.8 (3.5)	1.38 (1.12-1.66)	**0.001**	**0.01**	**0.02**
CRP (mg/L)	3.9 (6.0)	6.0 (8.3)	3.8 (5.9)	1.26 (1.00-1.52)	**0.03**	**0.03**	**0.03**

Values are mean(sd) or n (%)
† Defined as females >14 units/week, males >21 units/week or patient disclosed history of a current or prior alcohol problem **‡** results for the natural log of these values **§** for continuous variables, odds ratio represents change in odds for standard deviation change in variable
SIMD Scottish Index of Multiple Deprivation, **T2DM** Type 2 diabetes**, HbA1c** glycated haemoglobin**, ALT** alanine aminotransferase**, AST** aspartate aminotransferase**, ALP** alkaline phosphatase**,** γGT gamma-glutamyltransferase**, TNF-**$\propto$ tumour necrosis factor-alpha**, IL-6** Interleukin-6**, CRP** C-reactive protein

Table 17. Performance of the baseline models (FIB-4 and APRI) with the addition of complementary biomarkers

Model	C-statistic	Hosmer-Lemeshow p-value	AIC
Base Model			
FIB-4	0.85	0.35	238.2
Addition of one additional variable			
FIB-4 + HbA1c	0.87	0.43	228.7
FIB-4 + γGT†	0.93	0.98	205.5
FIB-4+ HA†	0.89	0.06	209.4
FIB-4 + BMI	0.87	0.32	232.9
FIB-4 + SIMD	0.87	0.79	239.6
FIB-4 + IL-6	0.87	0.16	235.0
FIB-4 + CRP	0.88	0.76	235.9
Mixed Models			
FIB-4, Hba1c, γGT†	0.93	0.86	199.2
FIB-4, Hba1c, HA†	0.90	0.10	203.5
FIB-4, γGT†, HA†	0.93	0.23	184.5
Full Model FIB-4, HbA1c, γGT†, HA†	0.94	0.71	181.0
Base Model			
APRI	0.85	0.92	243.8
Addition of one additional variable			
APRI + HbA1c	0.86	0.93	236.2
APRI + γGT †	0.91	0.84	219.1
APRI + HA†	0.88	<0.01	211.2
APRI + BMI	0.86	0.52	238.6
APRI + SIMD	0.87	0.18	242.5
APRI + IL-6	0.88	0.01	239.5
APRI + CRP	0.87	0.28	241.3
Mixed Models			
APRI, Hba1c, γGT †	0.91	0.84	213.6
APRI, Hba1c, HA†	0.89	<0.01	206.3
APRI, γGT †, HA†	0.92	0.20	192.9
Full Model APRI, HbA1c, γGT †, HA†	0.93	0.14	189.7

† log-transformed γGT/ HA variable
HbA1c glycated haemoglobin, **γGT** gamma glutamyltransferase, **HA** Hyaluronic Acid, **SIMD** Scottish Index of Multiple Deprivation, **IL-6** Interleukin-6, **CRP** C-reactive protein, **AIC** Akaike Information Criterion

Table 18. Performance of the baseline models (FIB-4 and APRI) with the addition of complementary biomarkers, re-run using competing risk regression analysis with non-liver death as the competing risk

Model	BIC (null=535.75)
Base Model	
FIB-4	467.17
Addition of one additional variable	
FIB-4 + HbA1c	460.27
FIB-4 + γGT†	431.97
FIB-4+ HA†	445.68
FIB-4 + BMI	469.52
FIB-4 + SIMD	470.47
FIB-4 + IL-6	469.94
FIB-4+ CRP	472.90
Mixed Models	
FIB-4, Hba1c, γGT†	433.04
FIB-4, Hba1c, HA†	444.16
FIB-4, γGT†, HA†	418.36
Full Model FIB-4, HbA1c, γGT†, HA†	420.65
Base Model	
APRI	458.58
Addition of one additional variable	
APRI + HbA1c	457.17
APRI + γGT†	436.10
APRI + HA†	439.09
APRI + BMI	459.60
APRI + SIMD	462.82
APRI + IL-6	460.00
APRI + CRP	463.22
Mixed Models	
APRI, Hba1c, γGT†	439.63
APRI, Hba1c, HA†	439.80
APRI, γGT†, HA†	419.49
Full Model APRI, HbA1c, γGT†, HA†	423.29

† log-transformed γGT / HA variable, **HbA1c** glycated haemoglobin, **BMI** body mass index, **γGT** gamma glutamyltransferase, **HA** Hyaluronic Acid, **SIMD** Scottish Index of Multiple Deprivation, **IL-6** Interleukin-6, **CRP** C-reactive protein, **BIC** Bayesian Information Criterion

Table 19. Predictive ability of models by sensitivity, specificity, PPV, NPV, false positive rate and false negative rate

Model	Sens (%, 95% CI)	Spec (%, 95% CI)	PPV (%, 95% CI)	NPV (%, 95% CI)	False $^{+}$ve n (%)	False $^{-}$ve n (%)
FIB-4 >2.67	40 (25-57)	98 (97-99)	46 (29-63)	98 (96-98)	19 (2)	24 (60)
FIB-4 >2.0	62 (46-77)	92 (90-93)	23 (16-33)	98 (97-99)	82 (8)	15 (37)
FIB-4 ≥1.3	82 (67-93)	59 (56-62)	8 (5-11)	99 (98-100)	399 (41)	7 (18)
As addition of further variables will increase false negative values, only FIB-4 ≥1.3 was taken forward.						
FIB-4 ≥1.3, γGT >55	45 (29-62)	95 (94-97)	28 (18-41)	98 (97-99)	46 (5)	22 (55)
FIB-4 ≥1.3, γGT >20	72 (56-85)	82 (79-84)	14 (10-20)	99 (98-99)	176 (18)	11 (28)
FIB-4 ≥1.3, HA ≥100	62 (46-77)	95 (93-96)	32 (22-44)	98 (97-99)	53 (5)	15 (38)
FIB-4 ≥1.3, HA ≥50	75 (59-87)	78 (75-81)	12 (8-17)	99 (98-99)	214 (22)	10 (25)
FIB4 ≥1.3, HbA1c >7.5	47 (32-64)	88 (85-90)	13 (8-20)	98 (96-99)	122 (12)	21 (53)
FIB-4 ≥1.3, HA ≥50, γGT >20	65 (48-79)	90 (88-92)	22 (15-30)	98 (97-99)	94 (10)	14 (35)
FIB4 ≥1.3, HA ≥50, HbA1c >7.5	45 (29-62)	93 (91-94)	20 (12-30)	98 (96-99)	72 (7)	22 (55)
FIB4 ≥1.3, γGT >20, HbA1c >7.5	40 (25-57)	94 (92-95)	22 (13-33)	97 (96-98)	58 (6)	24 (60)
Fib4 ≥1.3, HA ≥50, GGT >20, HbA1c >7.5	38 (23-54)	97 (95-98)	31 (19-46)	97 (96-98)	33 (3)	25 (63)
APRI >0.5	53 (36-68)	94 (93-96)	27 (18-38)	98 (97-99)	57 (6)	19 (48)
APRI >0.5, γGT >55	35 (21-52)	98 (97-99)	45 (27-64)	97 (96-98)	17 (2)	26 (65)
APRI >0.5, γGT >20	50 (34-66)	96 (95-97)	36 (24-50)	98 (97-99)	35 (4)	20 (50)

Model	Sens (%, 95% CI)	Spec (%, 95% CI)	PPV (%, 95% CI)	NPV (%, 95% CI)	False $^{+}$ve n (%)	False $^{-}$ve n (%)
APRI >0.5, HbA1c >7.5	33 (19-49)	98 (97-99)	37 (21-55)	97 (96-98)	22 (2)	27 (68)
APRI >0.5, HA ≥50, γGT >20	50 (34-66)	98 (96-98)	45 (30-61)	98 (97-99)	24 (2)	20 (50)
APRI >0.5, HA ≥50, HbA1c >7.5	33 (19-49)	98 (97-99)	45 (26-64)	97 (96-98)	16 (2)	27 (68)
APRI >0.5, γGT >20, HbA1c >7.5	30 (17-47)	99 (98-99)	50 (29-71)	97 (96-98)	12 (1)	28 (70)
APRI >0.5, HA ≥50, GGT >20, HbA1c >7.5	30 (17-47)	99 (98-100)	57 (34-78)	97 (96-98)	9 (1)	28 (70)
EASL guidelines- USS steatosis + FIB-4 >=1.3 OR ALT>50 OR AST>45 OR γGT>55	86 (71-95)	60 (57-63)	8 (6-11)	99 (98-100)	346 (40)	5 (14)
EASL (USS)+ HA ≥50	81 (64-92)	81 (78-84)	15 (10-21)	99 (98-100)	163 (19)	7 (19)
EASL guidelines- FLI positive + FIB-4 ≥1.3 OR ALT>50 OR AST>45 OR γGT>55	90 (76-97)	58 (55-61)	8 (6-11)	99 (98-100)	411 (42)	4 (10)
EASL (FLI)+ HA ≥50	78 (62-89)	79 (76-81)	13 (9-18)	99 (98-99)	206 (21)	9 (23)

γGT gamma glutamyltransferase, **HA** Hyaluronic Acid, **USS** ultrasound assessed, **ALT** alanine aminotransferase, **AST** aspartate aminotransferase, **FLI** Fatty liver index, **HbA1c** glycated haemoglobin, **sens** sensitivity, **spec** specificity, **PPV** positive predictive value, **NPV** negative predictive value

Table 20. Predictive ability of models by sensitivity, specificity, PPV, NPV, false positives and false negatives- final models, participants with definite non-NAFLD disease excluded (n=3)

Model	Sens (%, 95% CI)	Spec (%, 95% CI)	PPV (%, 95% CI)	NPV (%, 95% CI)	False $^{+}$ve n (%)	False $^{-}$ve n (%)
FIB-4 ≥1.3	84 (68-94)	59 (56-62)	7 (5-10)	99 (98-100)	399 (41)	6 (16)
FIB-4 ≥1.3, HA ≥50	78 (62-90)	78 (75-81)	12 (8-17)	99 (98-100)	214 (22)	8 (22)
EASL guidelines- USS steatosis + FIB-4 >=1.3 OR ALT>50 OR AST>45 OR γGT>55	89 (75-97)	60 (57-64)	8 (6-11)	99 (98-100)	376 (44)	4 (11)
EASL (USS)+ HA ≥50	78 (62-90)	81 (79-84)	14 (10-19)	99 (98-100)	179 (21)	8 (22)
EASL guidelines- FLI positive + FIB-4 ≥1.3 OR ALT>50 OR AST>45 OR γGT>55	89 (75-97)	58 (55-61)	7 (5-10)	99 (98-100)	413 (43)	4 (10)
EASL (FLI)+ HA ≥50	78 (62-90)	79 (76-81)	12 (8-17)	99 (98-100)	208 (21)	8 (22)

γGT gamma glutamyltransferase, **HA** Hyaluronic Acid, **USS** ultrasound assessed, **ALT** alanine aminotransferase, **AST** aspartate aminotransferase, **FLI** Fatty liver index, **HbA1c** glycated haemoglobin, **sens** sensitivity, **spec** specificity, **PPV** positive predictive value, **NPV** negative predictive value

Table 21. Predictive ability of models by sensitivity, specificity, PPV, NPV, false positives and false negatives- final models, participants who developed HCC in a non-cirrhotic liver excluded (n=4)

Model	Sens (%, 95% CI)	Spec (%, 95% CI)	PPV (%, 95% CI)	NPV (%, 95% CI)	False $^{+}$ve n (%)	False $^{-}$ve n (%)
FIB-4 ≥1.3	84 (68-94)	59 (56-62)	7 (5-10)	99 (98-100)	394 (41)	6 (16)
FIB-4 ≥1.3, HA ≥50	76 (59-88)	78 (75-81)	12 (8-16)	99 (98-99)	211 (22)	9 (24)
EASL guidelines- USS steatosis + FIB-4 >=1.3 OR ALT>50 OR AST>45 OR γGT>55	92 (78-98)	60 (57-63)	8 (6-12)	99 (98-100)	371 (40)	2 (8)
EASL (USS)+ HA ≥50	92 (78-98)	67 (63-70)	10 (7-13)	100 (99-100)	315 (33)	3 (8)
EASL guidelines- FLI positive + FIB-4 ≥1.3 OR ALT>50 OR AST>45 OR γGT>55	92 (78-98)	57 (54-61)	8 (5-11)	99 (98-100)	408 (43)	3 (8)
EASL (FLI)+ HA ≥50	89 (75-97)	64 (61-67)	9 (6-12)	99 (98-100)	344 (36)	4 (11)

γGT gamma glutamyltransferase, **HA** Hyaluronic Acid, **USS** ultrasound assessed, **ALT** alanine aminotransferase, **AST** aspartate aminotransferase, **FLI** Fatty liver index, **HbA1c** glycated haemoglobin, **sens** sensitivity, **spec** specificity, **PPV** positive predictive value, **NPV** negative predictive value

4.6 Additional Information

Embedded in the discussion section of this chapter is the discussion of the biological rationale for hyaluronic acid and its use as a biomarker of liver disease. In this paragraph I wish to additionally discuss the biological rationale for the other biomarker that best performed in addition to FIB-4 in our model building; GGT. GGT is found in the cell membranes of many tissues and is thought to be involved in amino acid transport, glutathione and leukotriene metabolism and maintenance of intracellular homeostasis in the context of oxidative stress. [271] Serum levels rise in the context of liver disease associated with liver cell disruption and cholestasis. One UK based observational study assessing the predictive ability of individual liver enzymes in identifying 2 year all-cause clinical liver outcomes showed that GGT was one of the most strongly associated with incident disease. [272] However, care must be taken in the interpretation of this in the context of our study as the proportion of people with diabetes was low (1% derivation cohort, 12% validation cohort). Furthermore, many GGT readings were imputed. Interestingly, the cut-point of GGT used to provide best prognostication in this cohort was similar to our second cut point ot 20 U/L. Using a similar cut point in this study showed that >25% of our participants would have been inappropriately identified as low risk for the development of cirrhosis/ HCC over 10 years.

Chapter 5 The association of non-alcoholic fatty liver disease screening tests with incident advanced chronic liver disease and mortality in people with Type 2 diabetes

This section has been submitted and is under review in Diabetic Medicine under the same title by Sheila M Grecian (SMG), Jonathan A Fallowfield (JF), Stela McLachlan (SM), Peter C Hayes (PH), Indra Neil Guha (NG), Joanne R Morling (JM), Stephen Glancy (SG), Rachel M Williamson (RW), Rebecca M Reynolds (RR), Brian M Frier (BF), Nicola N Zammitt (NZ), Jackie F Price (JP) and Mark WJ Strachan (MS). SMG wrote the manuscript. JP was principal investigator of the ET2DS, designed the study, analysed and interpreted the data. MS was lead investigator of the ET2DS liver sub-study, designed the study, analysed and interpreted the data. RR, BF, PH, JF, RW, NG and SG contributed to study design. SMG, SM, RW and JM contributed to data collection, analysis and interpretation. All authors contributed to revision and final approval of the article.

In summary, in the previous chapters it was discussed that screening for NAFLD in the context of T2DM would be potentially beneficial and is recommended in international guidelines. [217] Whilst the presence of hepatic fibrosis is acknowledged as a key factor in predicting progression of NAFLD to cirrhosis and associated HCC, none of the non-invasive tools designed to identify fibrosis and thus additionally used as predictive tools for the identification of risk of progressive disease perform optimally in the context of T2DM. We then discussed that the addition of hyaluronic acid to the FIB-4 risk prediction tool may reduce the false positive rate. In this study we compared the predictive ability of any tests for NAFLD (ultrasound and the Fatty Liver Index as markers of hepatic steatosis, serum liver enzymes, fibrosis markers) or any combination in the identification of incident cirrhosis and HCC. Furthermore, as NAFLD is thought to be associated with an increased mortality we assessed whether any of these tests were associated with deaths during the study follow up period.

Formulae used for the calculation of predictive ability can be found in appendix 2.

Please note- all tables for this section sit at the end of the chapter text.

5.1 Abstract

Background: Type 2 diabetes is associated with an increased risk of non-alcoholic fatty liver disease (NAFLD) and of progression to cirrhosis and hepatocellular carcinoma (HCC). Screening of people with Type 2 diabetes for NAFLD is recommended, but the optimum test to use is uncertain.

Aims: To compare the ability of non-invasive tests for NAFLD to identify incident cirrhosis/HCC and mortality in a community cohort of older people with Type 2 diabetes.

Methods: Participants in the Edinburgh Type 2 Diabetes Study (n=1066, age 60-75 at baseline) were followed for over 11 years. Serum liver enzymes, fatty liver index (FLI), hepatic steatosis on ultrasound, Fibrosis-4 index (FIB-4) and hyaluronic acid were measured at baseline and year 1. Individual and composite tests were analysed for their ability to accurately identify incident cirrhosis/HCC and mortality.

Results: Incidence of cirrhosis/HCC was 4.1% and 320 deaths occurred. All tests investigated had false positive or negative rates of >20% or 35% respectively for the identification of cirrhosis/HCC. A ‘positive’ FLI was associated with significantly increased mortality (hazard ratio (95% confidence interval 1.45 (1.13-1.87), p=0.004). FLI and other tests showed high false positive or negative rates (>20% or 75% respectively) for mortality.

Conclusion: None of the tests provided a ‘good balance’ between false positive and negative rates in the identification of cirrhosis/HCC and are unlikely to be helpful in mortality assessment. Clinicians could choose tests with a low false positive rate to identify a proportion of cases of incident cirrhosis/HCC, while minimising unnecessary investigation.

5.2 Introduction

People with Type 2 diabetes are at high risk of developing non-alcoholic fatty liver disease (NAFLD). Furthermore, Type 2 diabetes is associated with increased risk of progression to cirrhosis and hepatocellular carcinoma (hazard ratio (HR) 1.8).[21,75] Concurrent diagnoses of NAFLD and diabetes have been associated with increased mortality rates; in one study of 337 people with diabetes, the HR for death with concurrent NAFLD was 2.2.[40] The European Association for the Study of the Liver, European Association for the Study of Diabetes and The European Association for the Study of Obesity (EASL-EASD-EASO) NAFLD guidelines recommend screening for NAFLD in people with Type 2 diabetes to enable targeted lifestyle modification to reduce the rate of (or reverse) disease progression and to allow prompt identification and treatment of complications.[217] Several non-invasive biomarkers and tests for NAFLD are available, including serum liver enzymes (aspartate aminotransferase (AST), alanine aminotransferase (ALT), gamma-glutamyl transferase (GGT)), serum biomarker scores associated with NAFLD and/or fibrosis and imaging tests such as liver ultrasound or magnetic resonance imaging-proton density fat fraction for steatosis and transient elastography or MR elastography for fibrosis. In addition, algorithms that combine several tests have been developed, such as in the EASL-EASD-EASO guideline (Figure 2).[217]

Although many tests for NAFLD have been developed, no consensus exists as to which are most effective for screening large numbers of individuals in a diabetes out-patient setting, with the aim of detecting those who are at greatest risk of developing cirrhosis/HCC. The ability of the chosen screening test to predict clinically significant disease is paramount, as well as having a test that is sufficiently practical and cost-effective to be used at scale. To date, existing tests have been found to be sub-optimal in predicting cirrhosis/HCC in people with diabetes. For example, in one study (n=284, follow-up over 4 years), 15% of people with diabetes with a ‘low-risk’ Fibrosis-4 (FIB-4) score developed decompensated cirrhosis, whereas in people without diabetes, no-one with a ‘low-risk’ score developed decompensated cirrhosis.[227,237] Similarly, we have previously reported that in older people with diabetes, 18% of those with a ‘low-risk’ FIB-4 score developed cirrhosis or HCC, while 41% of those who did not develop cirrhosis or HCC had a ‘high-risk’ FIB-4 score at baseline.[262]

5.3 Aims

Given the on-going uncertainty regarding the optimal NAFLD screening test to use, the aim of the present study was to compare the ability of a range of commonly used NAFLD tests to identify incident cirrhosis/HCC in a community population of people with Type 2 diabetes. We also aimed to determine whether these tests were predictive of death during 11-years follow-up.

5.4 Methods

The Edinburgh Type 2 Diabetes Study

The Edinburgh Type 2 Diabetes Study (ET2DS) is a population based prospective cohort study which recruited 1066 participants aged 60-74 with Type 2 diabetes in 2006/07. Detailed methodology has been described previously.[255] All who attended the baseline clinic were invited to re-attend an 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 (deceased n=88, unsuitable for clinical reasons n=26, unable to contact n=23, unable to attend n=98).

Data collection

Detailed biomarker assessment was undertaken at the baseline clinic. Assessments were undertaken at dedicated research clinics at the Wellcome Trust Clinical Research Facility, Western General Hospital, Edinburgh, UK, by specially trained research staff using Standard Operating Procedures.[255] Fasting venous blood samples were collected at baseline. ALT, AST, alkaline phosphatase (ALP), GGT), triglycerides and platelets were analysed using a Vitros Fusion chemistry system (Ortho Clinical Diagnostics, Bucks, UK) at the Western General Hospital. Hyaluronic acid was measured using a radiometric assay (Pharmacia, Uppsala, Sweden). Height, weight and waist circumference were measured at the baseline clinic. Liver ultrasound was undertaken at the year 1 clinic (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 graders and ^{1}H MRI spectroscopy in a subset, as previously described.[259] 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 significant overlap between grade 0-2 steatosis was identified, only those with grade 3 steatosis on ultrasound assessment were deemed to have ‘definite steatosis’.

Participants attending the year-1 research clinic who had an abnormal liver ultrasound or liver enzymes above the laboratory reference range underwent a diagnostic liver screen (including Hepatitis B and C serology, liver autoantibody titres, alpha-fetoprotein, ferritin). All participants completed standard questions about alcohol consumption (AUDIT-C questionnaire), 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 (ALP) >125 U/L), AST:ALT ratio >1, hyaluronic acid >100μg/L (in the absence of known joint disease), positive liver autoantibodies, ferritin >1000ng/mL, alpha-feto protein >6ng/mL, positive hepatitis B or C serology, spleen diameter >13cm, platelets <150x10^9/L (in the absence of known haematological cause), or suspected cirrhosis on ultrasound, was referred for specialist hepatology review.

Only one fibrosis risk-stratification tool (FIB-4 at the low to medium risk cut-point of 1.3) is presented in the present study - previous work by this group showed this to be the best-performing of the fibrosis scores in our cohort, especially when considering the combination of false positive and false negative rates.[262] We have shown previously that other published FIB-4 cut points of 2.67 (the medium to high-risk cut point (6)) or 2.0 (a suggested age-specific cut point (10)) resulted in >35% false negative rate (7), and so a cut point of 1.3 was chosen for the present analysis. A novel combination of biomarkers (FIB-4 and hyaluronic acid assessment) was included, which was shown previously in our cohort to improve the false positive level, and additional combinations of biomarkers from those suggested in the EASL-EASD-EASO algorithm.[273]

Scores were calculated and cut-point levels used as per published work:

- Fibrosis-4 index (FIB-4) was calculated as: ((age(years)xAST(U/L))/(plt(x10^9/L)xsqrt ALT(U/L))) [262]
- Fatty liver index (FLI) was calculated as: $e^y/(1+e^y)$x100 where y=0.953 x ln(triglycerides, mg/dl) + 0.139 x BMI, kg/m^2 + 0.718 x ln (GGT, U/L) + 0.053 x waist circumference, cm – 15.745) [274]
- The EASL-EASD-EASO referral decision algorithm was used [217]

For the purpose of this analysis, a ‘positive test’ was defined:

- Abnormal liver enzymes: above the upper limit of normal in the laboratory reference range of ALT, AST or GGT
- Fatty liver index level: ≥60
- Ultrasound: 'definite' steatosis
- FIB-4 level: ≥1.3
- Hyaluronic acid level: ≥50 μg/L

Participants were followed-up until death or the end of study (11 years). Mean follow-up was 9.6 years (standard deviation 2.8 years). Incident cirrhosis and HCC were identified using multiple information sources (hospital patient record review (TrakCare, InterSystems Corp., Cambridge, USA), patient and GP questionnaire at year 4 and year 10 follow-up, discharge summary coding of hospital admissions and death coding (Information Services Division, NHS Scotland)). Cases were confirmed if a clinician diagnosis was recorded in medical notes. Death records were obtained from hospital patient records, national death coding (Information Services Division, NHS Scotland) and death certificates (National Records Scotland). We have previously published detailed data on prevalent and incident liver disease.[262] In brief, 7 people had prevalent cirrhosis/HCC and 43 people developed incident cirrhosis/HCC (an incidence of 4.1%). Of those 43, 37 cases were attributed to NAFLD, NAFLD with alcohol as a co-factor or mixed aetiology NAFLD and alpha-1-antitrypsin deficiency. 320 participants died during study follow-up.

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/.). Event rates were calculated for the cohorts with 'positive' and 'negative' test results. Association between test result and survival was assessed using Cox proportional hazards regression. Association between test result and liver outcomes has been reported previously.[262] Performance was assessed through calculation of positive predictive value, negative predictive value, false positive rate (FPR) and false negative rate (FNR). We have previously reported these data for FIB-4, 'FIB-4 with hyaluronic acid' and the EASL-EASD-EASO algorithm, but results are presented here to allow direct comparison.[262]

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.

5.5 Results

Performance of NAFLD tests in predicting incident cirrhosis/HCC

Table 22 shows the ability of tests to predict incident cirrhosis or HCC. All tests either had a FPR >20% or FNR >35%. For the three tests with a FNR <10% ('positive FLI', 'raised liver enzymes OR positive FLI', 'raised liver enzymes OR [positive FIB-4 with hyaluronic acid]'), all had a FPR >40%. For those tests with a FPR <20% ('raised liver enzymes PLUS positive FLI', 'positive FLI PLUS [positive FIB-4 with hyaluronic acid]', 'raised liver enzymes PLUS [positive FIB-4 with hyaluronic acid]'), FNRs were >35%.

In terms of clinical utility, the use of 'raised liver enzymes PLUS FLI' or 'raised liver enzymes PLUS [FIB-4 with hyaluronic acid]' would result in appropriate 'negative' results for 879/981 (89.6%) or 936/1013 (92.4%) of the disease-free cohort respectively but would lead to missing 23 (56.1%) or 19 (45.2%) incident cases respectively. Conversely, the use of 'FLI OR [FIB-4 with hyaluronic acid]', 'raised liver enzymes OR FLI' or 'raised liver enzymes OR [FIB-4 with hyaluronic acid]' would limit missed incident cases to <10% (n≤3) but would result in inappropriate referral of 728 (72.9%), 713 (71.3%) or 411 (40.9%) of the cohort respectively.

Association of NAFLD tests with survival

Cause of death was predominantly cardiovascular disease (35%), with 30% of deaths due to malignancy and 2% from cirrhosis and HCC. We assessed five tests (raised liver enzymes, USS, FLI, FIB-4 and [FIB-4 with hyaluronic acid]) for their association with survival (

Figure 4). People with a negative FLI, FIB-4 or [FIB-4 with hyaluronic acid] test tended to have slightly better survival compared to those with a positive test, HR (95% confidence intervals (CI)) were respectively 1.45 (1.13-1.87; p=0.004), 1.18 (0.93-1.50; p=0.17) and 1.29 (1.00-1.66; p=0.05). Survival curves for people with negative and positive 'raised liver enzyme' and 'hepatic steatosis on ultrasound' tests showed no difference with HR (CI) 0.85 (0.65-1.11, p=0.23) and 0.92 (0.72-1.18, p=0.50), respectively.

Table 23 presents test performance statistics for all-cause mortality. All tests either had a FPR >20% or a FNR >75%. The lowest FNR (20.8%) was for the 'FLI OR [positive FIB-4 with hyaluronic acid]' test, but the corresponding FPR was 71.6%. Similarly, for the two tests that had FPRs <20% ('positive FLI PLUS [positive FIB-4 with hyaluronic acid]', 'raised liver enzymes PLUS [positive FIB-4 with hyaluronic acid]'), FNRs were above 75%.

Repeating either analysis using hepatic steatosis on ultrasound as the steatosis marker did not improve the accuracy of the tools (Table 24, Table 25).

5.6 Discussion

When considering a population screening programme, WHO criteria state that any test should be acceptable to the population, suitable for use in a screening programme and have a high level of accuracy.[274] False negative results can result in a patient ignoring important symptoms or lead to delays in diagnosis and treatment. False positive results can lead to unnecessary and expensive follow-up, tests that may cause the patient harm and the psychological distress of an inappropriate diagnosis. In the present study we have observed that in older people with Type 2 diabetes, none of the tests used had a 'good balance' of FPR and FNR with respect to the prediction of incident cirrhosis and HCC. These findings are consistent with a previous smaller study the present study was larger with longer follow-up.[237]

FLI, FIB-4 and [FIB-4 with hyaluronic acid] were associated with increased mortality. This was statistically significant only for FLI. The incidence of cirrhosis and HCC were relatively low and it is likely that the biomarkers that showed a positive association with mortality were also identifying factors associated with other pathology (most likely cardiovascular disease). For example, FLI is calculated using triglycerides, BMI and waist circumference, all known to be associated with increased cardiovascular risk. No association was observed between USS-detected steatosis and mortality. This may reflect the fact that USS is a relatively insensitive test for identifying low levels of steatosis, so that an USS finding of 'no steatosis' or 'indeterminate steatosis' did not exclude the presence of at least some steatosis, which in turn may be associated with cardiovascular risk factors.[258] Previous biopsy studies in cohorts attending secondary care services have shown an association between NAFLD related liver fibrosis and mortality, though a community cohort examining NAFLD in an unselected population (diagnosis based on USS steatosis) showed no increase in mortality.[76,77,189,192]

Although associations were identified between some of the tests and increased mortality, when applied at individual patient level their performance was poor, with unacceptably high FPR and FNR. This is presumably because the overall effect size was small.

Strengths and weaknesses of the study

The ET2DS examined outcomes of liver disease in a community population of people with Type 2 diabetes, who did not necessarily have symptoms of liver disease. Almost

all other studies have identified outcomes in cohorts with established NAFLD diagnoses and were likely to have advanced pathology. EASL-EASD-EASO guidelines recommend screening in populations like the one represented by the ET2DS cohort, making this an appropriate cohort in which to examine the potential utility of screening tools.[217] Participants were well-characterised at baseline and were followed up using multiple sources of information to accurately identify incident disease.

ET2DS is a single centre study, undertaken in people with Type 2 diabetes, aged 60-75 years at baseline, and of predominantly Caucasian origin (98.3%), so care should be taken in extrapolating these findings to other populations. All-cause cirrhosis/HCC were investigated. While aetiology was predominantly NAFLD, a small number of individuals with cirrhosis/HCC from other causes were included [262]. However, determining the precise aetiology of cirrhosis/HCC can be difficult, particularly the relative contributions of alcohol excess and obesity, and so it seemed more clinically relevant to include individuals with all causes of liver disease. We have previously undertaken sensitivity analyses to show that excluding participants who had definite non-NAFLD pathology did not materially change outcomes.[262] Participants did not undergo a liver biopsy, which is the gold standard technique for identification of liver disease. However, liver biopsy is an invasive procedure, and it would neither have been ethical nor feasible to perform this in an asymptomatic community population. Transient elastography data has not been included as this was only measured at year 4. With respect to serum fibrosis biomarkers, we utilised FIB-4, cut-point ≥1.3 (with or without hyaluronic acid), as we have demonstrated that this was overall the best performing biomarker in this study; we have previously published data on the performance of the 2.67 and 2.0 cut-offs.[262]

Our incidence data may be an underestimate as it is possible that we did not identify some participants who developed asymptomatic cirrhosis/HCC, as screening for cirrhosis/HCC at final follow-up was not repeated. Alternatively, our incidence may overestimate the clinical burden as a substantial proportion of our diagnoses were made after hepatology referral following year 1 or 4 investigations. NAFLD cirrhosis can have a silent natural history for many years. Thus, some people who may never have developed overt cirrhosis, or may have died before their disease became clinically apparent, may have been identified. However, 58% of those identified with

cirrhosis developed varices, ascites and/or encephalopathy and 23% developed HCC, so it is likely that a large majority would have presented with clinical sequelae.

Implications for practice

NAFLD in the context of Type 2 diabetes fulfills many of the WHO recommended criteria for population screening, especially the ability to detect NAFLD at a pre-cirrhotic stage, the ability to undertake timely surveillance for varices and HCC and commence specific treatments for complications. Current available tests do not have a 'good balance' between FPR and FNR. In an individual clinic setting, given the poor overall performance of the tests, clinicians may choose to use a test with a low FPR (such as 'raised liver enzymes PLUS FLI' or 'raised liver enzymes PLUS [FIB-4 with hyaluronic acid]'), which would identify a proportion of cases while minimising service pressures through large numbers of people receiving unnecessary investigation. In our cohort, the Enhanced Liver Fibrosis test (cut point ≥10.51) and FIB-4 (cut-point >2.67) also had very low FPR, though both had a FNR >50%.[262] The combinations of 'FLI OR [FIB-4 with hyaluronic acid]', 'raised liver enzymes OR FLI' or 'raised liver enzymes OR [FIB-4 with hyaluronic acid]' would identify the majority of those who will develop cirrhosis/HCC, but would result in large numbers of people who will not develop advanced liver disease undergoing additional investigation. Care would have to be taken in both situations to explain the limitations of the tests to patients and clinical staff. Furthermore, while USS may not in isolation be a good screening tool (56.8% FPR, 38.5% FNR), it will always likely form part of an investigation pathway because of its ability to identify structural liver disease.

Assessment of different combinations of existing biomarkers or development of alternative biomarkers is required. It would be interesting to investigate the utility of transient elastography in this context, acknowledging that this test may be less reliable in obese subjects and that employing imaging for population screening can impose system challenges and be more resource-intensive.[275] Consideration also needs to be given as to whether serial screening may reduce FNR, increasing the utility particularly of those tests with low FPR.

Table 22. Performance of NAFLD tests for predicting incident cirrhosis or HCC

Non-invasive test		Event rate for those with +/- test % (95% CI)	Test performance			
			PPV n (95% CI)	NPV n (95% CI)	False +ve n (%)	False -ve n (%)
Raised liver enzymes	Test +	10.2 (6.9-14.5)	10 (7-14)	99 (98-99)	273 (27.1)	10 (24.4)
	Test -	1.3 (0.6-2.5)				
Positive Fatty Liver Index (FLI)	Test +	5.3 (3.7-7.3)	5 (4-7)	99 (97-100)	662 (66.4)	4 (9.8)
	Test -	1.2 (0.3-3.0)				
Positive FIB-4	Test +	7.6 (5.2-10.7)	8 (5-10)	99 (98-100)	402 (40.6)	7 (17.5)
	Test -	1.2 (0.5-2.4)				
Positive FIB-4 with HA	Test +	12.2 (8.3-17.5)	12 (8-17)	98 (97-99)	215 (21.4)	12 (28.6)
	Test -	1.5 (0.8-2.6)				
Combinations of individual components of the algorithm						
Raised liver enzymes PLUS positive FLI	Test +	15 (8.9-23.7)	15 (9-23)	97 (96-98)	102 (10.4)	23 (56.1)
	Test -	2.5 (1.6-3.8)				
Raised liver enzymes OR positive FLI	Test +	4.1 (3.6-6.9)	5 (4-7)	99 (97-100)	713 (71.3)	3 (7.3)
	Test -	1.0 (0.2-3.0)				
Positive FLI PLUS [positive FIB-4 with HA]	Test +	15.3 (10.1-22.3)	15 (10-22)	98 (97-99)	149 (14.9)	15 (35.7)
	Test -	1.7 (1.0-2.9)				
Positive FLI OR [positive FIB-4 with HA]	Test +	5.2 (3.7-7.1)	5 (4-7)	100 (98-100)	728 (72.9)	1 (2.4)
	Test -	0.4 (0.0-2.0)				
Raised liver enzymes PLUS [positive FIB-4 with HA]	Test +	23.0 (14.6-34.5)	23 (15-32)	98 (97-99)	77 (7.6)	19 (45.2)
	Test -	2.0 (1.2-3.1)				

Non-invasive test		Event rate for those with +/- test % (95% CI)	Test performance			
			PPV n (95% CI)	NPV n (95% CI)	False +ve n (%)	False -ve n (%)
Raised liver enzymes OR [positive FIB-4 with HA]	Test +	8.5 (6.0-11.6)	8 (6-11)	99 (9-100)	411 (40.9)	3 (7.9)
	Test -	0.5 (0.1-1.5)				
Full algorithm						
EASL-EASD-EASO algorithm[+] (raised liver enzymes OR [positive FLI PLUS [positive FIB-4 with HA]])	Test +	9.0 (6.2-12.4)	9 (6-12)	99 (98-100)	356 (35.5)	6 (14.6)
	Test -	0.9 (0.3-2.0)				

Raised Liver enzymes: Alanine aminotransferase, aspartate transaminase or gamma-glutamyl transferase above the reference range, **HA** hyaluronic acid, **FLI** Fatty Liver Index, **FIB-4** Fibrosis-4 Index, **CI** confidence interval, **PPV** positive predictive value, **NPV** negative predictive value

[+] EASD-EASL-EASO Algorithm [217]

Table 23. Performance of NAFLD tests for prediction of mortality

Baseline Predictor-raised required for positive test		Event rate for those with +/- test % (95% CI)	Test performance			
			PPV n (95% CI)	NPV n (95% CI)	False +ve n (%)	False -ve n (%)
Raised liver enzymes	Test +	25.7 (20.3-32.0)	26 (21-31)	69 (66-72)	226 (30.5)	231 (74.8)
	Test -	31.0 (27.1-35.3)				
Positive Fatty Liver Index (FLI)	Test +	31.0 (27.1-35.5)	31 (28-35)	75 (70-80)	482 (65.4)	84 (27.9)
	Test -	24.8 (19.8-30.7)				
Positive FIB-4	Test +	34.9 (29.6-41.0)	35 (30-40)	74 (71-78)	283 (39.0)	152 (50.0)
	Test -	25.5 (21.6-29.9)				
Positive FIB-4 with HA	Test +	38.8 (31.4-47.4)	39 (33-45)	73 (70-76)	150 (20.4)	213 (69.2)
	Test -	26.6 (23.2-30.5)				
Combinations of individual components of the algorithm						
Raised liver enzymes PLUS positive FLI	Test +	27.0 (21.0-34.2)	73 (67-78)	30 (27-33)	184 (24.9)	238 (77.8)
	Test -	30.0 (26.3-34.0)				
Raised liver enzymes OR positive FLI	Test +	30.2 (26.4-34.4	30 (27-34)	73 (68-78)	524 (71.1)	77 (25.3)
	Test -	26.6 (21.0-33.2)				
Positive FLI PLUS [positive FIB-4 with HA]	Test +	40.9 (32.1-51.5)	41 (34-49)	73 (70-76)	104 (14.1)	234 (76.5)
	Test -	27.0 (23.6-30.7)				
Positive FLI OR [positive FIB-4 with HA]	Test +	31.3 (27.4-35.5)	31 (28-35)	7 (71-82)	528 (71.6)	63 (20.8)
	Test -	23.2 (17.8-29.6)				
Raised liver enzymes PLUS [positive FIB-4 with HA]	Test +	37.0 (26.1-51.0)	37 (28-47)	71 (68-74)	63 (8.5)	273 (88.1)
	Test -	32.7 (29.1-36.5)				

Baseline Predictor- raised required for positive test		Event rate for those with +/- test % (95% CI)	Test performance PPV n (95% CI)	NPV n (95% CI)	False +ve n (%)	False -ve n (%)
Raised liver enzymes OR [positive FIB-4 with HA]	Test +	30.3 (25.4-35.8)	30 (26-35)	71 (67-75)	313 (42.4)	171 (55.7)
	Test -	28.7 (24.6-33.3)				
Full algorithm						
EASL-EASD-EASO algorithm[+] (raised liver enzymes OR [positive FLI PLUS [positive FIB-4 with HA]])	Test +	29.4 (24.3-35.3)	29 (25-34)	71 (67-74)	276 (37.4)	191 (62.4)
	Test -	29.3 (25.3-33.8)				

Raised Liver enzymes: Alanine aminotransferase, aspartate transaminase or gamma-glutamyl transferase above the reference range, **HA** hyaluronic acid, **FLI** Fatty Liver Index, **FIB-4** Fibrosis-4 Index, **CI** confidence interval, **PPV** positive predictive value, **NPV** negative predictive value

[+] EASD-EASL-EASO Algorithm [217]

Table 24. Performance of NAFLD tests for predicting incident cirrhosis or HCC, using ultrasound for assessment of steatosis

Non-invasive test		Event rate for those with +/- test % (95% CI)	Test performance			
			PPV n (95% CI)	NPV n (95% CI)	False +ve n (%)	False -ve n (%)
Raised liver enzymes	Test +	10.2 (6.9-14.5)	10 (7-14)	99 (98-99)	273 (27.1)	10 (24.4)
	Test -	1.3 (0.6-2.5)				
Positive Ultrasound for steatosis (USS)	Test +	4.5 (2.9-6.7)	5 (3-7)	96 (94-98)	508 (56.8)	15 (38.5)
	Test -	3.7 (2.1-6.2)				
Positive FIB-4	Test +	7.6 (5.2-10.7)	8 (5-10)	99 (98-100)	402 (40.6)	7 (17.5)
	Test -	1.2 (0.5-2.4)				
Positive FIB-4 PLUS HA	Test +	12.2 (8.3-17.5)	12 (8-17)	98 (97-99)	215 (21.4)	12 (28.6)
	Test -	1.5 (0.8-2.6)				
Combinations of individual components of the algorithm						
Raised liver enzymes PLUS positive USS	Test +	9.5 (5.6-15.1)	10 (6-15)	98 (96-99)	171 (17.4)	20 (52.6)
	Test -	2.4 (1.5-3.7				
Raised liver enzymes OR positive USS	Test +	5.7 (4.0-7.9)	6 (4-8)	98 (96-99)	610 (66.2)	5 (11.9)
	Test -	1.6 (0.5-3.7)				
Positive USS PLUS [positive FIB-4 with HA]	Test +	15.0 (8.9-23.7)	15 (9-23)	97 (96-98)	102 (10.4)	23 (56.1)
	Test -	2.5 (1.6-3.8)				
Positive USS OR [positive FIB-4 with HA]	Test +	5.5 (3.8-7.6)	5 (4-8)	99 (97-100)	621 (67.9)	4 (10.0)
	Test -	1.3 (0.4-3.4)				
Raised liver enzymes PLUS [positive FIB-4 with HA]	Test +	23.0 (14.6-34.5)	3 (15-32)	98 (97-99)	77 (7.6)	19 (45.2)
	Test -	2.0 (1.2-3.1)				

Baseline Predictor- raised required for positive test		Event rate for those with +/- test % (95% CI)	Test performance PPV n (95% CI)	NPV n (95% CI)	False +ve n (%)	False -ve n (%)
Raised liver enzymes OR [positive FIB-4 with HA]	Test +	8.5 (6.0-11.6)	8 (6-11)	99 (9-100)	411 (40.9)	3 (7.9)
	Test -	0.5 (0.1-1.5)				
Full algorithm						
EASL-EASD-EASO algorithm[+] (raised liver enzymes OR [positive USS PLUS [positive FIB-4 with HA]])	Test +	9.7 (6.7-13.4)	10 (7-13)	99 (98-100)	327 (33.1)	6 (14.6)
	Test -	0.9 (0.3-2.0)				

Raised Liver enzymes: Alanine aminotransferase, aspartate aminotransferase or gamma-glutamyl transferase above the reference range, **HA** hyaluronic acid, **FLI** Fatty Liver Index, **FIB-4** Fibrosis-4 Index, **CI** confidence interval, **PPV** positive predictive value, **NPV** negative predictive value

[+] EASD-EASL-EASO Algorithm [217]

Table 25. Performance of NAFLD tests for prediction of mortality, using ultrasound for assessment of steatosis

Baseline Predictor-raised required for positive test		Event rate for those with +/- test % (95% CI)	Test performance			
			PPV n (95% CI)	NPV n (95% CI)	False +ve n (%)	False -ve n (%)
Raised liver enzymes	Test +	25.7 (20.3-32.0)	26 (21-31)	69 (66-72)	226 (30.5)	231 (74.8)
	Test -	31.0 (27.1-35.3)				
Positive Ultrasound for steatosis (USS)	Test +	25.8 (21.6-30.4)	26 (22-30)	71 (66-75)	395 (58.2)	117 (46.1)
	Test -	29.2 (24.1-35.0)				
Positive FIB-4	Test +	34.9 (29.6-41.0)	35 (30-40)	74 (71-78)	283 (39.0)	152 (50.0)
	Test -	25.5 (21.6-29.9)				
Positive FIB-4 with HA	Test +	38.8 (31.4-47.4)	39 (33-45)	73 (70-76)	150 (20.4)	213 (69.2)
	Test -	26.6 (23.2-30.5)				
Combinations of individual components of the algorithm						
Raised liver enzymes PLUS positive USS	Test +	23.3 (16.9-31.3)	23 (17-30)	69 (66-72)	145 (20.2)	256 (85.3)
	Test -	30.1 (27.2-34.9)				
Raised liver enzymes OR positive USS	Test +	26.4 (22.6-30.7)	26 (23-30)	71 (66-76)	476 (67.9)	92 (35.0)
	Test -	29.0 (23.4-35.6)				
Positive USS PLUS [positive FIB-4 with HA]	Test +	35.0 (25.2-47.3)	35 (27-44)	72 (69-75)	78 (10.7)	254 (85.8)
	Test -	28.1 (24.8-31.8)				
Positive USS OR [positive FIB-4 with HA]	Test +	28.9 (25.0-33.3)	29 (25-33)	74 (69-79)	467 (67.8)	76 (28.6)
	Test -	25.5 (20.1-31.9)				
Raised liver enzymes PLUS [positive FIB-4 with HA]	Test +	37.0 (26.1-51.0)	37 (28-47)	71 (68-74)	63 (8.5)	273 (88.1)
	Test -	32.7 (29.1-36.5)				

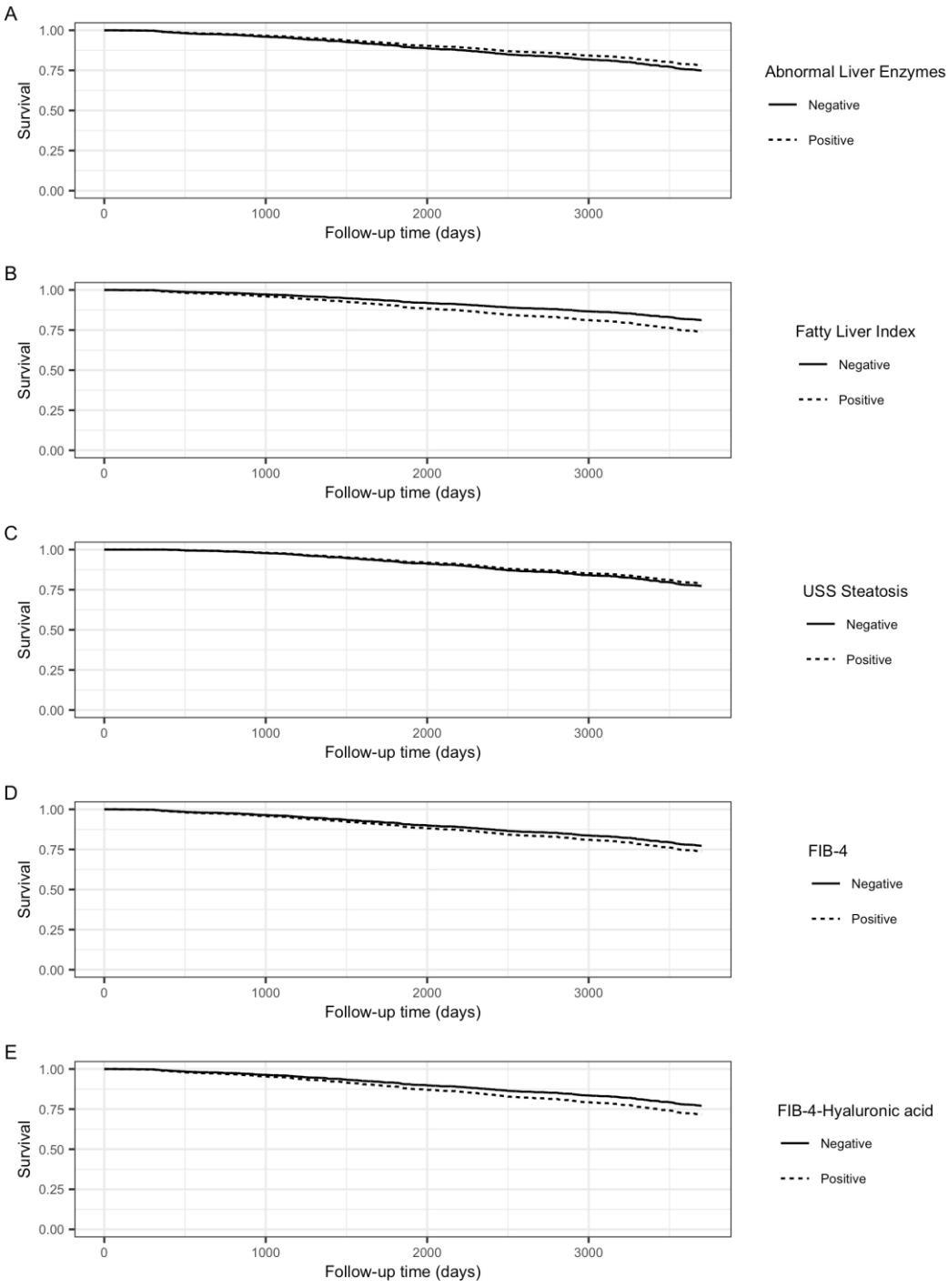
Baseline Predictor- raised required for positive test		Event rate for those with +/- test % (95% CI)	Test performance PPV n (95% CI)	NPV n (95% CI)	False +ve n (%)	False -ve n (%)
Raised liver enzymes OR [positive FIB-4 with HA]	Test +	30.3 (25.4-35.8)	30 (26-35)	71 (67-75)	313 (42.4)	171 (55.7)
	Test -	28.7 (24.6-33.3)				
Full algorithm						
EASL-EASD-EASO algorithm[+] (raised liver enzymes OR [positive USS PLUS [positive FIB-4 with HA]])	Test +	27.6 (22.5-33.6)	28 (23-33)	71 (67-74)	262 (35.7)	197 (66.3)
	Test -	29.5 (25.5-33.9)				

Raised Liver enzymes: Alanine aminotransferase, aspartate aminotransferase or gamma-glutamyl transferase above the reference range, **HA** hyaluronic acid, **FLI** Fatty Liver Index, **FIB-4** Fibrosis-4 Index, **CI** confidence interval, **PPV** positive predictive value, **NPV** negative predictive value

[+] EASD-EASL-EASO Algorithm [217]

Figure 4. Kaplan-Meier plots for association between test result and survival

Survival analysis by Cox proportional hazard regression (adjusted for age and sex)\



5.7 Additional Information

Further to the information included in this chapter there is the question of the biological plausibility of the predictive factors.

For many factors this has been discussed earlier in the thesis. In section 1.5 the pathogenesis of NAFLD is discussed including the importance of the development of steatosis and fibrosis as key stages of disease progression. Section 1.7 discusses cohort data in detail identifying which markers have been shown to be associated with disease progression; with the finding that the identification of fibrosis seems to be key. The biological derivation of existing non-invasive risk prediction tools is discussed in section 1.8.2. Hyaluronic acid and GGT are discussed in sections 4.5 and 4.6.

It is important additionally to mention that some studies have identified GGT to be an independent predictor of cardiovascular events and mortality, independent of the presence of liver disease. It has been shown that GGT can accumulate in atherosclerotic plaques although the mechanism is not fully elucidated. Furthermore, rises in GGT within the laboratory normal range is associated with cardiovascular outcome. [276] We did not see an increase in mortality associated with raised liver enzymes; although we looked only at a combined marker of AST,ALT or GGT, at a level above the laboratory range. It would be interesting in future studies to investigate, in a new diabetes cohort, if GGT was a predictor of cardiovascular morbidity and mortality.

Chapter 6 General discussion and future directions

6.1 Summary of objectives

The worldwide prevalence of NAFLD is increasing, with NAFLD related liver disease contributing an ever-greater proportion of liver-related deaths. The prevalence of NAFLD is increased in people with diabetes. Additionally, diabetes is a risk factor for progression to cirrhosis and HCC and consistently associated with worse outcomes in NAFLD. [21] It is thought that early identification of NAFLD in people with diabetes by screening as advised in European guidelines could improve outcomes through adaptations that promote disease regression, and early referral for hepatology support and surveillance programmes (for example for HCC and varices). [217] Furthermore, identifying which non-invasive markers are best associated with progressive disease could help monitor response to treatment in therapeutic trials. Identification of people at risk of progressive disease through accurate non-invasive testing could assist in the targeting of appropriate treatment to these groups.

The ET2DS cohort, a prospective cohort study of a community population of 1066 people with T2DM in Lothian, Scotland was used in this study to address four aims:

1. Define the absolute and relative cohort incidence of liver disease to date in the ET2DS cohort
2. Determine whether current non-invasive fibrosis risk prediction tools reliably identify incident cirrhosis and hepatocellular carcinoma in a community cohort of older people with T2DM
3. Determine whether the addition of other biomarkers to existing fibrosis risk prediction tools improve their performance in predicting incident cirrhosis and hepatocellular carcinoma in a community cohort of older people with T2DM
4. Identify whether potential non-invasive screening tests for NAFLD (those identifying steatosis, serum liver enzymes, markers of fibrosis) associated with incident cirrhosis, hepatocellular carcinoma and mortality in people with T2DM

6.2 Non-invasive risk scores do not reliably identify future cirrhosis or HCC in T2DM.

6.2.1 Description of incident cirrhosis and HCC in the ET2DS cohort

The understanding and interpretation of the results of the questions asked in the 3 main aims of this thesis, rests on first determining the burden of liver disease in the ET2DS cohort.

It is generally believed that the prevalence of NAFLD, and the risk of progression to cirrhosis and HCC is increased in T2DM. [6,69,70] In this community population of older people with T2DM and no known NAFLD, the incidence of cirrhosis over 11 year follow-up was 3.92 per 1000 person years and HCC was 1.28 per 1000 person years. Of those in the study with cirrhosis or who developed incident cirrhosis, 58% developed variceal disease, ascites related to their liver disease or hepatic encephalopathy. NAFLD contributed to the aetiology of incident disease in 37/43 (86%) participants.

The incidence of cirrhosis and HCC in this population was substantially higher than reported population rates (for cirrhosis age-matched UK data report 0.36-0.54 per 1000 person-years, for liver cancer age-matched Scottish data report 0.41-0.58 per 1000 person years) (www.isdscotland.org). In addition, the prevalence of NAFLD as the predominant aetiology was greater than in the general population, where <10% cirrhosis (all age groups) and 19% HCC (age >60) has been attributed to NAFLD. [277] We therefore found an increased burden of liver cirrhosis and HCC, of primary NAFLD aetiology, in a community cohort of older people with T2DM, consistent with previous observations in other populations. The findings have clinical implications for holistic care of people with T2DM and provoke consideration of whether screening or surveillance strategies for uncommon but severe liver disease should be implemented in routine care.

6.2.2 Ability of non-invasive fibrosis risk scores to identify incident cirrhosis or HCC

The presence of hepatic fibrosis is considered to be the most important indicator of disease likely to progress to cirrhosis and HCC, and to be associated with increased

liver-related and cardiovascular mortality. However, the gold standard test for fibrosis is liver biopsy which, as an invasive procedure, is not suitable for population screening. There is thus considerable interest in the development of non-invasive tools to identify hepatic fibrosis, and their use to identify people at high risk of developing clinically significant liver disease. Both European and American guidelines for NAFLD advocate consideration of the use of such tools in high risk community populations, such as people with T2DM. [7,217] Yet previous studies have suggested that the non-invasive tools recommended may be less accurate in people with T2DM than in the general population and as such there is significant uncertainty about their utility in T2DM. [237,238]

This study found that, whilst the non-invasive tools assessed (AST:ALT, APRI, ELF, FIB-4, NFS) all had a significant association with incident cirrhosis or HCC, the ability of any risk score to correctly identify people who were going to develop incident disease was poor, with scores exhibiting low PPVs (5-46%) and demonstrating either exceptionally high false positive or false negative rates. Similar results were seen when using the non-invasive fibrosis risk prediction tools in the EASL-EASD-EASO screening algorithm.

A brief discussion of the regression analysis technique chosen

Regression analysis was undertaken to look at the association of the risk prediction tool result at baseline with incident cirrhosis or HCC. There are several regression analysis techniques that could have been used. Logistic regression looks at whether, over a study, there is an association with outcome; there is no time element. Taking time to event into account (such as in regression analyses based on Cox-Hazards regression) can be helpful, especially if looking at association over a long period of time, because it allows people to die or to leave follow-up for other reasons without biasing results. However, adding in a time element if there is inaccurate time to event data can introduce error and uncertainty into the estimate, and this error is exacerbated if there is a small number of events. It can also be useful to consider the fact that someone may experience a different, ‘competing’, event (for example, death) which may prevent them developing the predicted outcome (for example, cirrhosis or HCC) had they not died of a different cause in the interim. Competing risks regression analysis takes into account both time to event, and potential competing risks that may bias results. It is a type of proportional hazards regression model, where the

exponential of the regression coefficient of the sub-distribution hazard model (Fine and Gray) can be interpreted as a relative change in the rate at a given time point (hazard) of the outcome of interest occurring in those who have not experience that event yet (including those who have succumbed to a competing event (e.g. death)). For example, one baseline factor could be associated with an x fold increase in hazard of cirrhosis, which could be interpreted as evidence that scores above the baseline factor cut-point are associated with an x fold increase in cirrhosis in participants who have not yet developed cirrhosis or experienced a mutually exclusive event (e.g. death from another cause). [278] Some caution is required in interpreting the magnitude of effect but the direction of change can be interpreted more confidently and if comparing models, bigger changes imply bigger changes in hazard. Likewise, exact values of output from logistic regression, cox-hazards regression and competing risk regression analysis cannot be directly compared whereas trends can be.

In this study two issues needed addressing. Firstly, as a large proportion of cases were diagnosed at the pre-symptomatic stage following clinical test evaluation the study team was concerned about the utility of time to event data. Cirrhosis can be asymptomatic for many years and so, compared to those cases which were clinician-identified during routine care and thus most likely as a result of symptomatic disease, it is likely that those cases which were screen-identified obtained a diagnosis relatively sooner in their clinical course, potentially by many years. So comparators of time to event for the screen-identified and clinician-identified cases seemed inappropriate for primary analysis. Secondly, a large number of participants (320/1066) died during the course of the study. Therefore, there would have been a risk of bias by not accepting the competing risk of non-liver death into the analysis. The decision was made, therefore, to undertake the primary analysis using logistic regression, using C-statistic and AIC to compare models. The entire analysis was then re-run using competing risks regression methodology, using BIC to compare models. Regression analysis was corrected for age and sex.

Comparison with existing literature and importance of this study

Several studies have shown that incident cirrhosis and HCC are associated with non-invasive scores but they did not describe predictive ability. [185,231,234] Other studies have described varying specificity and sensitivity of non-invasive scores when compared with measures of fibrosis. [221,229,232] One study has specifically compared the

performance of these tools in people with and without diabetes, but within the context of a selected hepatology clinic population; it found that the tools' ability to predict cirrhosis and HCC was less good in people with diabetes. [237] The mechanisms for this are not fully elucidated but may relate to the reliance on AST and ALT in the predictive tools, and the thought that their serum levels in people with diabetes may be altered by factors above and beyond those related to liver pathology. [242,243]

This community population study is important because it looks at the utility of these tools in a representative population without a prior diagnosis of NAFLD. This is exactly the population for which NAFLD screening is advocated in European guidelines despite there being very limited data on their performance in this setting, i.e. in a population that will naturally have a lower pre-test probability for disease than a population drawn from referrals to a secondary care hepatology service. Also, we have been able to examine the performance of these tools in a cohort of people all of whom had T2DM - a condition known to be associated with increased risk of liver disease and disease progression, but where doubt has been cast on the performance of standard risk prediction tools. This study shows that these currently recommended non-invasive risk prediction tools for NAFLD outcomes perform only modestly in an unselected group of people with T2DM. Further work to improve prediction methods in this population is necessary before routine surveillance can be advocated.

6.3 Hyaluronic acid improves the ability of the FIB-4 liver fibrosis score to predict incident cirrhosis and hepatocellular carcinoma in T2DM

Whilst this study has shown that existing risk prediction tools perform sub-optimally in people with T2DM, accurate risk prediction tools would be valuable as, though there is increased risk of cirrhosis and HCC in the context of diabetes, the absolute risk of progression is low. If those who could benefit most from heptatology review and intensified risk management could be identified reliably it would help manage resources well and not subject those at low risk to unnecessary investigation, burden of care and health anxiety.

This study identified that, in this cohort, combining a hyaluronic acid measurement (cut-point >50μg/L) with the FIB-4 risk prediction tool (cut point ≥1.3) reduced the number of people identified as 'high-risk' but who did not develop cirrhosis or HCC during follow-up by 46% whilst retaining a false negative value of ≤25%.

Discussion of the comparators used to evaluate model performance

Tools have been developed which enable comparisons between the utility of regression models in practice. Tools such as C-statistic assess the discriminatory ability of a model - the ability of a model to split individuals appropriately into those who will or will not develop the outcome. [279] The C-statistic calculates the probability that an increased probability of outcome is assigned to those who develop an outcome, comparing the odds of each individual having the outcome based on the model variables and the actual outcome as a ratio. Any rise in score above 0.5 represents an incremental improvement in model function above chance with a score of ≥0.8 considered as reasonable. It is known to be insensitive if there are a small number of people at high risk and a large number at low risk. [279] Tools such as the Hosmer-Lemeshow test assess the calibration of a model. Calibration looks at a model's ability to accurately estimate absolute risk by measuring how well predicted probabilities agree with observed risk. [279] For Hosmer-Lemeshow testing, a p>0.05 is considered acceptable calibration. Measures such as AIC (for logistic regression) or the closely related Bayesian Information Criterion (BIC) (which can be used for competing risks regression) combine an assessment of discrimination and calibration, assessing the likelihood that a fitted model would produce the data that is truly observed. [279] They have no scale, but a lower value in comparison to another model

is considered to have improved performance. For the model building in this study, it was decided to compare models using one method to assess discrimination (C-statistic), one to assess calibration (Hosmer-Lemeshow test) and one to assess overall model performance (AIC).

Discussion of model building methodology

Clinical risk prediction models combine multiple predictor variables and can be useful in identifying those individuals who may be more/ less likely to develop a condition or have a better/ worse prognosis from a condition. This can help guide clinicians in the identification of those who may benefit most from an intervention. It is important that to be useful in clinical practice factors included in models are plausible (i.e. have a plausible pathophysiological association) rather than just statistically associated and that they can be measured reliably. [280-282] Models are frequently built using regression techniques. Stepwise regression (either forwards, where increasing numbers of variables are added, or backwards where all variables are assessed and any detrimental impact of removing one at a time is assessed) are commonly used but, especially if event numbers are low, the model performance can be overestimated, and, additionally, especially using backwards regression a poorly parsimonious model can be created. [280] An alternative is to look at all sub-sets regression where the performance of regression models with all possible combinations of factors is assessed. This can be helpful, and can certainly identify factors that feature in most or all of better performing models, but, nonetheless, require care as there is a risk of overfitting the model (where the model fits the initial study data well but is not generalisable to other populations). [280] Following this reasoning, factors were restricted, from those assessed at baseline assessment, to those where thorough literature review showed they had plausible pathophysiological association. All-subsets regression was used due to the relatively low event numbers. The risk of overfitting is acknowledged and the findings will need to be validated in larger, more diverse cohorts.

Comparison with existing literature and importance of this study

To the study team's knowledge, this is the first study to examine the effect of hyaluronic acid in risk prediction in a community population of people with T2DM. Hyaluronic acid has previously been assessed as a prognostic marker of liver disease only in a few studies, but a significant association has been found between rising

hyaluronic acid and liver mortality. [270] Although hyaluronic acid is known to be raised in rheumatological as well as liver disease, it is hoped that its use in people with diabetes in conjunction with other markers of liver fibrosis would not have a material impact on the model. This study found that, in a community population with T2DM, the numbers of people inappropriately identified as ‘high risk’ for the development of cirrhosis or HCC could be reduced by adding serum hyaluronic acid to the FIB-4 assessment.

6.4 The association of NAFLD screening tests with incident advanced chronic liver disease and mortality in people with T2DM

To conclude this body of work, I investigated whether any existing potential screening tests for NAFLD would be acceptable for use for population screening in people with T2DM. Whilst there are many indications to screen for NAFLD in people with T2DM (it can be detected at a pre-cirrhotic stage with the potential for reversibility, while surveillance can prevent complications of cirrhosis and identify HCC at treatable stage), it was identified that none of the currently available screening tests that were evaluated (serum liver enzymes, FLI, hepatic steatosis on USS, FIB-4, FIB-4 with hyaluronic acid, and combinations of these) had a 'good' balance of false positive and false negative rates when considering the prediction of incident cirrhosis and HCC. This analysis was undertaken with the full cohort compared to the selected complete-cases model cohort used in the previous chapter. The FIB-4 and hyaluronic acid combination marker, whilst still reducing false positives significantly in the whole cohort, resulted in more false negatives in the whole compared to the model cohort.

In an individual clinical setting, clinicians may wish to choose a test with a low false positive rate (for example 'raised liver enzymes PLUS FLI' or 'raised liver enzymes PLUS [FIB-4 with hyaluronic acid]') as this would identify a proportion of cases at least (as opposed to not screening at all) but would minimise service pressure from people receiving unnecessary investigation. Care would need to be taken to explain the implications of a 'negative' result to the patient in this situation.

No strong association was identified between NAFLD as determined by a 'positive' test result and increased mortality. Previous studies have found conflicting results when comparing NAFLD and mortality, although it has been previously identified as a risk factor for mortality in the context of diabetes. [40,76,77,192] In this cohort the trend was for FLI, FIB-4 and [FIB-4 with hyaluronic acid] to be associated with increased mortality - and it may be that in a larger study or with longer follow-up the trends may become statistically significant, though whether clinically significant is uncertain. Caution should also be taken in defining an increase in mortality to being a consequence of NAFLD as, for example in FLI where triglycerides, BMI and waist circumference are used in the calculation, biomarkers used in the tests are known to be associated with other risk (e.g., cardiovascular risk) apart from NAFLD.

Further evaluation needs to be undertaken in larger and more diverse cohorts. Whilst the combination of FIB-4 and hyaluronic acid holds promise by reducing false positive rates, more work needs to be undertaken to reduce false negative rates.

6.5 Strengths and weaknesses of this work

6.5.1 Strengths

The ET2DS is a study of moderate size that has reviewed long-term liver outcomes in individuals in a community population with T2DM who were asymptomatic of liver disease at baseline. This represents precisely the scenario in which existing guidelines recommend screening for liver disease providing an ideal study population. Almost all other studies have examined outcomes in people recruited from secondary care hepatology clinics, with known NAFLD and a higher likelihood of cirrhosis and HCC.

Participants were recruited through the Lothian diabetes register, which includes almost all those with T2DM, and participants were randomly selected from this register. Thus, it included people with T2DM on all treatment types (diet, oral and injectable agents) and those under both community and hospital care. It has previously been confirmed that the participants at baseline were representative of the larger group of people selected at random from the register in terms of age, HbA1c, duration of T2DM, proportion requiring insulin and total cholesterol, and thus considered to be representative of the target population (Table 4). The study was conducted with a prospective design. Participants were well characterised with extensive phenotyping at baseline providing accurate documentation of risk factors.

Completeness of data collection for incident events was maximised through the use of multiple sources of information (including patient and GP reporting at study follow-up clinics, review of electronic patient records, review of death records and records of admissions to hospital).

6.5.2 Weaknesses

There are limitations to this study. ET2DS is a single centre study, undertaken in people with T2DM aged 61-76 years at the time of USS, and of predominantly Caucasian origin (98.3%). In addition an unusually large proportion of our population were in the least deprived SIMD quintile. This was a representative sample of people with T2DM in the age-matched population sampled (Lothian, Scotland, UK). However we are aware that these results may as such not be generalisable to the general population and care should be taken in extrapolating these findings to younger or

more diverse populations, or those who may have different lifestyle choices (particularly with regards to alcohol intake, dietary composition and exercise) and those managed using newer glucose-lowering agents.

All-cause cirrhosis and HCC were investigated in this cohort. Whilst aetiology was predominantly NAFLD, individuals with cirrhosis and HCC from other causes were also included. Determining the precise aetiology of cirrhosis and HCC can be difficult, particularly the relative contributions of alcohol excess and obesity, and so including individuals with all-causes of liver disease seemed more clinically relevant. Sensitivity analyses were undertaken to exclude those where NAFLD was confirmed not to be a contributory aetiology and results did not differ significantly. Another limitation is that subjects did not undergo a liver biopsy, the gold standard technique for identification of steatosis and fibrosis. However, liver biopsy is an invasive procedure and it would not have been ethical or feasible to perform in an asymptomatic population of this size.

It is important to note that a significant proportion of 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 prolonged, it is possible that those who were diagnosed following referral from screening had cirrhosis or HCC at baseline and thus 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 and several of those who were ‘screen-detected’ were not identified with cirrhosis or HCC on initial hepatology review but were diagnosed several years later. Therefore, prevalent disease has been termed as only that which was clinically apparent at baseline. Secondly, the screening process may have led to an earlier diagnosis of cirrhosis or HCC, of whom some may have died from other causes before cirrhosis or HCC was clinically apparent and thus inflating incidence data. However, 58% of all those identified with cirrhosis developed varices, ascites and/or encephalopathy; and 23% developed HCC. Thus, it appears likely that a majority of participants would have been identified through routine clinical care during the course of the study. Additionally, sensitivity analyses were run using competing risks regression analysis with non-liver death as the competing risk - these sensitivity analyses demonstrated no material change in results and thus it is thought that this was unlikely to have significantly contributed to a results error. Cirrhosis and HCC were not screened for at the year 11 clinic follow-up so it is possible that some

participants may have developed asymptomatic incident disease during follow-up that was not identified. This is important as there is inherent selection bias in using records from routinely collected data (admissions and electronic patient records) as it will include only those who attended hospitals and only those who attended the hospitals sampled. It will also record incidental diagnoses but these would be recorded through study screening. Through collection of data from multiple sources (including the ISD data which is Scotland-wide) the study aimed to minimise this error.

While all other biomarkers were measured at baseline, ELF and liver USS were undertaken at the year 1 clinic, so analyses using these markers have examined slightly different ‘baseline’ time points. However, no participant was diagnosed with incident disease during that year and given the time course of NAFLD progression it is likely that anyone with an abnormal result at year 1 would have had an abnormal result at the baseline clinic.

Finally, this study examined a medium sized cohort, with a modest incidence of cirrhosis and HCC. The sample size was designed to be powered for cognitive outcomes, which are more frequent than liver outcomes in this population. The liver arm was introduced after the commencement of the study. Although the study was adequately powered to identify a large effect size, it is possible that a smaller effect may not have been identified. Therefore, a small absolute difference in the proportion of people developing cirrhosis or HCC between groups may have become statistically significant in a larger study. However, for example, USS examinations were performed to a ‘research standard’, with a formally validated process. In routine clinical practice, USS examinations may not be performed using such robust criteria. The present study has shown that USS failed to identify a significant proportion of individuals who develop cirrhosis/HCC (40%) and it is likely that USS may perform worse in a clinical setting. Statistical power alone would not explain the lack of an observed effect of ‘definite steatosis’ on mortality, as absolute mortality rates were higher in the group without ‘definite steatosis’.

6.6 Future Directions

The data presented in this thesis has shown that the cirrhosis and HCC risk prediction tools developed thus far perform only modestly in in a community population with T2DM. The addition of hyaluronic acid to the FIB-4 tool does reduce the number of people identified incorrectly as being at high risk of developing cirrhosis or HCC but still leaves it well short of being a clinically reliable surveillance test. It is acknowledged that the cohort was of moderate size, in an older and predominantly Caucasian population. Thus, to examine generalisability, it would be necessary to validate results in more diverse populations. Validation in a larger population is also needed to address the possibility that effects and performance of a model can be over or underestimated in small populations. [280]

Furthermore, there are aspects of risk prediction that should be considered but were unable to be assessed due to the constraints of the size and characteristics of the cohort. Firstly, the number of outcomes among people in our cohort was too small to determine whether median time to diagnosis of cirrhosis or HCC was longer in those who had lower baseline risk scores. It is possible that, especially in the context of T2DM where it is thought that progression of NAFLD to fibrosis, cirrhosis and HCC can be accelerated, markers in the risk prediction tools for those who developed cirrhosis or HCC later in follow-up were not raised at baseline assessment. It would thus be interesting to assess this in a larger cohort and examine whether serial measurements show dynamic change in the pathway to cirrhosis/HCC. If this was the case, then serial measurement may result in fewer false negative results.

Secondly, there is increasing interest in the subgroup of people who develop HCC in a non-cirrhotic liver. This is mechanistically intriguing, but also of clinical concern because cirrhosis is normally the clinical trigger for HCC surveillance monitoring. In this cohort, 29% of HCC instances were identified in people with a non-cirrhotic liver, but this represented just four individuals. This suggests that it may be valuable to investigate this subset of people in more depth in larger populations to identify factors that predispose to this rarer outcome.

Thirdly, it is thought that genetic tendency plays a role in both the development of NAFLD and the aggressiveness of disease progression. [132] The study's modestly sized population, in combination with the population prevalence of the alleles known to be consistently associated with NAFLD meant that it did not have enough power to

be able to investigate this association. However, it will likely be important to consider this in the development of future risk prediction strategies in larger cohorts.

Fourthly, there is an expanding field of interest in pharmacological agents (including GLP-1 receptor agonists, SGLT-2 inhibitors, novel antifibrotic agents including Galectin 3 inhibitors and probiotics (section 1.9.2)) that may be able to treat NAFLD. A detailed understanding of how pharmacological agents affect disease progression needs to be studied. To have contributed any meaningful analysis based on drug exposure our study would have needed details on medication use not only for the duration of the study (which in terms of NAFLD pathogenesis is relatively short) but for the years prior to the study. Unfortunately the study did not have access to that data, nor were many of the agents of current research interest (such as GLP-1 agonists) in common or indeed any use at recruitment. It would be very interesting to look prospectively at the outcomes of people now commenced on such medications, perhaps using pharmaco-epidemiological surveillance rather than a dedicated study, given the infrequency of the outcomes and length of their pathogenesis.

Finally, it is important to acknowledge that there are existing and emerging risk prediction markers (and combinations) that were not examined within the context of this study; some of these novel markers may also have improved predictive ability. For example, recent publications discuss the potential for use of serum metabolite panels or the incorporation of PRO-C3 (a marker of type III collagen formation) into risk prediction tools for the identification of advanced hepatic fibrosis. {Harrison:2020cu, [283,284] Transient elastography is also increasingly been seen as a more reliable test for liver fibrosis than serum markers, and interest is growing in its use in combination with serum markers. [218,275,285] However whether the role that transient elastography could play a role in population screening as an imaging tool is uncertain.

6.7 Concluding Statement

This study investigated the incidence of cirrhosis and HCC and factors that are associated with incident cirrhosis and HCC in a community population of older people with T2DM who were asymptomatic of liver disease at baseline. This study found an increased incidence of cirrhosis and HCC in our population compared to whole population data, confirming that people with T2DM do experience higher rates of cirrhosis and HCC than the general population. In addition, it was identified that existing fibrosis risk prediction tools performed only modestly and did not accurately identify those who developed incident disease. Further study identified that combining serum hyaluronic acid measurement with the FIB-4 risk prediction tool reduced the number of people who were identified as 'high risk' at baseline but did not develop incident cirrhosis or HCC. However, when several potential NAFLD screening tests were assessed for predictive value, inclusive of the fibrosis risk prediction tools, no 'good balance' between false positive and negative rate was found. In an individual clinic setting, if clinicians wish to undertake any liver screening, they may want to choose a test with a low false positive rate (for example 'raised liver enzymes PLUS FLI' or 'raised liver enzymes PLUS [FIB-4 with hyaluronic acid]') as this would identify a proportion of cases but would minimise unnecessary investigation. But, most importantly, further investigation of this uncommon but increasing and clinically important accompaniment of T2DM is required in larger and more diverse cohorts.

Appendix 1- Permission for reproduction of EASL-EASD-EASO NAFLD screening algorithm in thesis

ELSEVIER LICENSE
TERMS AND CONDITIONS
Aug 06, 2020

This Agreement between Dr. Sheila Grecian ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	4882981411694
License date	Aug 06, 2020
Licensed Content Publisher	Elsevier
Licensed Content Publication	Journal of Hepatology
Licensed Content Title	EASL–EASD–EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease
Licensed Content Date	Jun 1, 2016
Licensed Content Volume	64
Licensed Content Issue	6
Licensed Content Pages	15
Start Page	1388
End Page	1402
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Title	Investigating the Natural History of Liver Disease in Type 2 Diabetes
Institution name	Edinburgh University
Expected presentation date	Feb 2021
Portions	Figure 1
Requestor Location	Dr. Sheila Grecian c/o Edinburgh Type 2 Diabetes Study Usher Institute, Edinburgh University Old Medical School, Teviot Place Edinburgh, EH8 9AG United Kingdom Attn: Dr. Sheila Grecian
Publisher Tax ID	GB 494 6272 12

Appendix 2- Formulae used in calculations for incidence and model predictive ability

Formulae for Calculation of Incidence

Total incidence over time was calculated as (total number of cases/ total number of participants).

Incidence was calculated as ((number new cases) / (patient years in follow-up)) x 1000 for cases per 1000 patient years. Patient follow up days were calculated for those who developed an event during follow up as (date of event – date of baseline clinic), for those who had no event but had died during follow up as (date of death – date of baseline clinic) and for those who had no event and were alive at end of follow up as (last date of data collection – date of baseline clinic). Years of follow up was calculated as total follow up days/365.25.

Formulae for Calculation of Predictive Ability

The performance of each tool in identifying incident cirrhosis or HCC was assessed using sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), false positive rate and false negative rate. This was calculated from the standard 2x2 table:

	Test positive	**Test negative**
Outcome present	True positive (TP)	False negative (FN)
Outcome absent	False positive (FP)	True negative (TN)

- Sensitivity is the proportion of participants with the outcome who test positive (TP / (TP+FN))
- Specificity is the proportion of participants without the outcome who test negative (TN / (TN+FP))
- Positive Predictive Value is the probability that following a positive test result, an individual with truly have the outcome (TP / (TP+FP))

- Negative Predictive Value is the probability that following a negative test result, an individual will truly not have the outcome (TN / (TN+FN))
- False Positive Rate is the proportion of people who do not have the outcome who test positive (FP / (FP+TN))
- False Negative Rate is the proportion of people who have the outcome who test negative (FN / (TP+FN))

Appendix 3- Supplementary information for Chapter 3

Identification of potential biomarkers

To identify potential biomarkers that may improve predictive ability, reviews of the literature both considering what is understood about the pathogenesis of NAFLD (link) and which factors have been found to be significantly associated with disease progression in other human epidemiological studies (link) were undertaken. Following this, the study identified the following baseline biomarkers to investigate from those collected in the baseline clinic (Table 26).

Additional detail on validation of model building

The following validation checks were undertaken during model building.

1. Co-correlation (to ensure there were no concerns about co-linearity) was checked using both Pearson and Spearman methodology (due to concern that all variables may not form a perfect normal distribution).
2. The effect of extreme outliers on results for both the base models and the alternative biomarkers was assessed using Cook's distance (>0.5 deemed significant) and regression re-run excluding any extreme outliers identified to look for influence.
3. Final models were checked for co-linearity with the variance inflation factor (VIF), effects of any extreme outliers with Cook's distance, and were checked for interaction terms.

Table 26. Baseline Factors to Consider During Data Analysis

Baseline Factor		Rationale
Demographics	Age	Key population indicators. Population Deprivation Indices and Age have shown to be linked to prognosis in previous analyses (section 1.7)
	Sex	
	Standard Index Multiple depravation	
Metabolic factors	BMI	NAFLD is known to be associated with the metabolic syndrome and for metabolic syndrome to be associated with progressive NAFLD (section 1.7). BMI has been shown to be associated with prognosis in previous analyses (section 1.7).
	Waist- Hip Ratio	
	Smoking	
	Alcohol Intake	May be linked to overall outcome pathogenically (it is possible to have mixed NAFLD and ALD pathology)
	Cholesterol	Although not identified in previous cohort studies to be implicated in disease progression, it would be thought pathophysiologically that lipid profile may impact on NAFLD pathogenesis (section 1.5).
Diabetes specific factors	Duration T2DM	T2DM is known to be associated with worse outcomes in NAFLD (section 1.4). Key markers of diabetes effect will be duration of T2DM and HbA1c as a marker of exposure to hyperglycaemia.
	HbA1c	
Markers of liver integrity and synthetic function	AST	Enzymes released in liver injury. They have been associated with disease progression in cohort studies (section 1.7).
	ALT	
	ALP	
	GGT	
	Bilirubin	Known markers of liver synthetic failure.
	Albumin	
	Platelets	
	Hyaluronic acid (HA)	A glycosaminoglycan found in connective tissue that is almost exclusively cleared by liver metabolism. Raised levels have been found to be associated with cirrhosis and liver mortality [269,270]
Markers of inflammation	CRP	Implicated in the progression of NAFLD pathogenesis. Mixed results in previous cohort study analyses (section 1.7).
	IL-6	
	TNFα	

References

1. Khoudari G, Singh A, Noureddin M, et al. Characterization of patients with both alcoholic and nonalcoholic fatty liver disease in a large United States cohort. *World J Hepatol*. 2019;11(10):710-718. doi:10.4254/wjh.v11.i10.710.

2. Williams R, Alexander G, Aspinall R, et al. Gathering momentum for the way ahead: fifth report of the Lancet Standing Commission on Liver Disease in the UK. In: Vol 392. 2018:2398-2412. doi:10.1016/S0140-6736(18)32561-3.

3. Williams R, Aspinall R, Bellis M, et al. Addressing liver disease in the UK: a blueprint for attaining excellence in health care and reducing premature mortality from lifestyle issues of excess consumption of alcohol, obesity, and viral hepatitis. *Lancet*. 2014;384(9958):1953-1997. doi:10.1016/S0140-6736(14)61838-9.

4. Calzadilla Bertot L, Adams LA. The Natural Course of Non-Alcoholic Fatty Liver Disease. *Int J Mol Sci*. 2016;17(5):774. doi:10.3390/ijms17050774.

5. Bril F, Cusi K. Management of Nonalcoholic Fatty Liver Disease in Patients With Type 2 Diabetes: A Call to Action. *Diabetes Care*. 2017;40(3):419-430. doi:10.2337/dc16-1787.

6. Bellentani S. The epidemiology of non-alcoholic fatty liver disease. *Liver Int*. 2017;37 Suppl 1:81-84. doi:10.1111/liv.13299.

7. 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. doi:10.1002/hep.29367.

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

9. 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*. 2011;34(5):1139-1144. doi:10.2337/dc10-2229.

10. 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*. 2009;29(1):113-119. doi:10.1111/j.1478-3231.2008.01718.x.

11. 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. doi:10.1111/dme.13174.

12. Fan N, Zhang L, Xia Z, Peng L, Wang Y, Peng Y. Sex-Specific Association between Serum Uric Acid and Nonalcoholic Fatty Liver Disease in Type 2 Diabetic Patients. *J Diabetes Res*. 2016;2016(3):1-6. doi:10.1155/2016/3805372.

13. Wilman HR, Kelly M, Garratt S, et al. Characterisation of liver fat in the UK Biobank cohort. Lu S-N, ed. *PLoS ONE*. 2017;12(2):e0172921. doi:10.1371/journal.pone.0172921.

14. Garcia-Monzón C, Martín-Pérez E, Iacono OL, et al. Characterization of pathogenic and prognostic factors of nonalcoholic steatohepatitis associated with obesity. *J Hepatol*. 2000;33(5):716-724. doi:10.1016/S0168-8278(00)80301-3.

15. Dixon JB, Bhathal PS, O'Brien PE. Nonalcoholic fatty liver disease: predictors of nonalcoholic steatohepatitis and liver fibrosis in the severely obese. *Gastroenterology*. 2001;121(1):91-100.

16. Beymer C. Prevalence and Predictors of Asymptomatic Liver Disease in Patients Undergoing Gastric Bypass Surgery. *Archives of Surgery*. 2003;138(11):1240-1244. doi:10.1001/archsurg.138.11.1240.

17. Vernon G, Baranova A, Younossi ZM. Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults. *Aliment Pharmacol Ther*. 2011;34(3):274-285. doi:10.1111/j.1365-2036.2011.04724.x.

18. Chang Y, Jung H-S, Cho J, et al. Metabolically Healthy Obesity and the Development of Nonalcoholic Fatty Liver Disease. *Am J Gastroenterol*. 2016;111(8):1133-1140. doi:10.1038/ajg.2016.178.

19. Wong VW-S, Wong GL-H, Yeung DK-W, et al. Incidence of non-alcoholic fatty liver disease in Hong Kong: a population study with paired proton-magnetic resonance spectroscopy. *J Hepatol*. 2015;62(1):182-189. doi:10.1016/j.jhep.2014.08.041.

20. El-serag HB, Tran T, Everhart JE. Diabetes increases the risk of chronic liver disease and hepatocellular carcinoma. *Gastroenterology*. 2004;126(2):460-468. doi:10.1053/j.gastro.2003.10.065.

21. Pang Y, Kartsonaki C, Turnbull I, et al. Diabetes, Plasma Glucose, and Incidence of Fatty Liver, Cirrhosis, and Liver Cancer: A Prospective Study of 0.5 Million People. *Hepatology*. 2018;68(4):1308-1318. doi:10.1002/hep.30083.

22. Ong JP, Elariny H, Collantes R, et al. Predictors of nonalcoholic steatohepatitis and advanced fibrosis in morbidly obese patients. *Obes Surg*. 2005;15(3):310-315. doi:10.1381/0960892053576820.

23. 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. doi:10.1038/ajg.2013.405.

24. Kanwal F, Kramer JR, Mapakshi S, et al. Risk of Hepatocellular Cancer in Patients With Non-Alcoholic Fatty Liver Disease. *Gastroenterology*. 2018;155(6):1828–1837.e2. doi:10.1053/j.gastro.2018.08.024.

25. Ortiz-Lopez C, Lomonaco R, Orsak B, et al. Prevalence of Prediabetes and Diabetes and Metabolic Profile of Patients With Nonalcoholic Fatty Liver Disease (NAFLD). *Diabetes Care*. 2012;35(4):873-878. doi:10.2337/dc11-1849.

26. Lomonaco R, Ortiz-Lopez C, Orsak B, et al. Effect of adipose tissue insulin resistance on metabolic parameters and liver histology in obese patients with nonalcoholic fatty liver disease. *Hepatology*. 2012;55(5):1389-1397. doi:10.1002/hep.25539.

27. Bae JC, Cho YK, Lee WY, et al. Impact of Nonalcoholic Fatty Liver Disease on Insulin Resistance in Relation to HbA1c Levels in Nondiabetic Subjects. *Am J Gastroenterol*. 2010;105(11):2389-2395. doi:10.1038/ajg.2010.275.

28. Silverman JF, O'Brien KF, Long S, et al. Liver pathology in morbidly obese patients with and without diabetes. *Am J Gastroenterol*. 1990;85(10):1349-1355.

29. Mantovani A, Byrne CD, Bonora E, Targher G. Nonalcoholic Fatty Liver Disease and Risk of Incident Type 2 Diabetes: A Meta-analysis. *Diabetes Care*. 2018;41(2):372-382. doi:10.2337/dc17-1902.

30. Targher G, Marchesini G, Byrne CD. Risk of type 2 diabetes in patients with non-alcoholic fatty liver disease: Causal association or epiphenomenon? *Diabetes Metab*. 2016;42(3):142-156. doi:10.1016/j.diabet.2016.04.002.

31. Kim C-H, Park J-Y, Lee K-U, Kim J-H, Kim H-K. Fatty liver is an independent risk factor for the development of Type 2 diabetes in Korean adults. *Diabet Med*. 2008;25(4):476-481. doi:10.1111/j.1464-5491.2008.02410.x.

32. Park SK, Seo MH, Shin HC, Ryoo J-H. Clinical availability of nonalcoholic fatty liver disease as an early predictor of type 2 diabetes mellitus in korean men: 5-year prospective cohort study. *Hepatology*. 2013;57(4):1378-1383. doi:10.1002/hep.26183.

33. Ekstedt M, Franzén LE, Mathiesen UL, et al. Long-term follow-up of patients with NAFLD and elevated liver enzymes. *Hepatology*. 2006;44(4):865-873. doi:10.1002/hep.21327.

34. Sung K-C, Wild SH, Byrne CD. Resolution of fatty liver and risk of incident diabetes. *The Journal of Clinical Endocrinology & Metabolism*. 2013;98(9):3637-3643. doi:10.1210/jc.2013-1519.

35. Yamazaki H, Tsuboya T, Tsuji K, Dohke M, Maguchi H. Independent Association Between Improvement of Nonalcoholic Fatty Liver Disease and Reduced Incidence of Type 2 Diabetes. *Diabetes Care*. 2015;38(9):1673-1679. doi:10.2337/dc15-0140.

36. Targher G, Lonardo A, Byrne CD. Nonalcoholic fatty liver disease and chronic vascular complications of diabetes mellitus. *Nat Rev Endocrinol*. 2018;14(2):99-114. doi:10.1038/nrendo.2017.173.

37. Targher G, Bertolini L, Padovani R, et al. Prevalence of nonalcoholic fatty liver disease and its association with cardiovascular disease among type 2 diabetic patients. *Diabetes Care*. 2007;30(5):1212-1218. doi:10.2337/dc06-2247.

38. Ryysy L, Häkkinen AM, Goto T, et al. Hepatic fat content and insulin action on free fatty acids and glucose metabolism rather than insulin absorption are associated with insulin requirements during insulin therapy in type 2 diabetic patients. *Diabetes*. 2000;49(5):749-758.

39. Schmid V, Wagner R, Sailer C, et al. Non-alcoholic fatty liver disease and impaired proinsulin conversion as newly identified predictors of the long-term non-response to a lifestyle intervention for diabetes prevention: results from the TULIP study. *Diabetologia*. 2017;60(12):2341-2351. doi:10.1007/s00125-017-4407-z.

40. Adams LA, Harmsen S, St Sauver JL, et al. Nonalcoholic fatty liver disease increases risk of death among patients with diabetes: a community-based cohort study. *Am J Gastroenterol*. 2010;105(7):1567-1573. doi:10.1038/ajg.2010.18.

41. Okamoto M, Takeda Y, Yoda Y, Kobayashi K, Fujino MA, Yamagata Z. The association of fatty liver and diabetes risk. *J Epidemiol*. 2003;13(1):15-21.

42. Fan J-G, Li F, Cai X-B, Peng Y-D, Ao Q-H, Gao Y. Effects of nonalcoholic fatty liver disease on the development of metabolic disorders. *J Gastroenterol Hepatol*. 2007;22(7):1086-1091. doi:10.1111/j.1440-1746.2006.04781.x.

43. Shibata M, Kihara Y, Taguchi M, Tashiro M, Otsuki M. Nonalcoholic Fatty Liver Disease Is a Risk Factor for Type 2 Diabetes in Middle-Aged Japanese Men. *Diabetes Care*. 2007;30(11):2940-2944. doi:10.2337/dc07-0792.

44. Yamada T, Fukatsu M, Suzuki S, Wada T, Yoshida T, Joh T. Fatty liver predicts impaired fasting glucose and type 2 diabetes mellitus in Japanese undergoing a health checkup. *J Gastroenterol Hepatol*. 2010;25(2):352-356. doi:10.1111/j.1440-1746.2009.05998.x.

45. Bae JC, Rhee EJ, Lee WY, et al. Combined effect of nonalcoholic fatty liver disease and impaired fasting glucose on the development of type 2 diabetes: a 4-year retrospective longitudinal study. *Diabetes Care*. 2011;34(3):727-729. doi:10.2337/dc10-1991.

46. Sung K-C, Jeong W-S, Wild SH, Byrne CD. Combined influence of insulin resistance, overweight/obesity, and fatty liver as risk factors for type 2 diabetes. *Diabetes Care*. 2012;35(4):717-722. doi:10.2337/dc11-1853.

47. Kasturiratne A, Weerasinghe S, Dassanayake AS, et al. Influence of non-alcoholic fatty liver disease on the development of diabetes mellitus. *J Gastroenterol Hepatol*. 2013;28(1):142-147. doi:10.1111/j.1440-1746.2012.07264.x.

48. Chang Y, Jung H-S, Yun KE, Cho J, Cho YK, Ryu S. Cohort study of non-alcoholic fatty liver disease, NAFLD fibrosis score, and the risk of incident diabetes in a Korean population. *Am J Gastroenterol*. 2013;108(12):1861-1868. doi:10.1038/ajg.2013.349.

49. Choi JH, Rhee EJ, Bae JC, et al. Increased risk of type 2 diabetes in subjects with both elevated liver enzymes and ultrasonographically diagnosed nonalcoholic fatty liver disease: a 4-year longitudinal study. *Arch Med Res*. 2013;44(2):115-120. doi:10.1016/j.arcmed.2013.01.007.

50. Ming J, Xu S, Gao B, et al. Non-alcoholic fatty liver disease predicts type 2 diabetes mellitus, but not prediabetes, in Xi'an, China: a five-year cohort study. *Liver Int*. 2015;35(11):2401-2407. doi:10.1111/liv.12851.

51. Li W-D, Fu K-F, Li G-M, et al. Comparison of effects of obesity and non-alcoholic fatty liver disease on incidence of type 2 diabetes mellitus. *World J Gastroenterol*. 2015;21(32):9607-9613. doi:10.3748/wjg.v21.i32.9607.

52. Shah RV, Allison MA, Lima JAC, et al. Liver fat, statin use, and incident diabetes: The Multi-Ethnic Study of Atherosclerosis. *Atherosclerosis*. 2015;242(1):211-217. doi:10.1016/j.atherosclerosis.2015.07.018.

53. Fukuda T, Hamaguchi M, Kojima T, et al. The impact of non-alcoholic fatty liver disease on incident type 2 diabetes mellitus in non-overweight individuals. *Liver Int*. 2016;36(2):275-283. doi:10.1111/liv.12912.

54. Kim JK, Gavrilova O, Chen Y, Reitman ML, Shulman GI. Mechanism of insulin resistance in A-ZIP/F-1 fatless mice. *Journal of Biological Chemistry*. 2000;275(12):8456-8460. doi:10.1074/jbc.275.12.8456.

55. Samuel VT, Shulman GI. The pathogenesis of insulin resistance: integrating signaling pathways and substrate flux. *J Clin Invest*. 2016;126(1):12-22. doi:10.1172/JCI77812.

56. Samuel VT, Liu Z-X, Qu X, et al. Mechanism of Hepatic Insulin Resistance in Non-alcoholic Fatty Liver Disease. *Journal of Biological Chemistry*. 2004;279(31):32345-32353. doi:10.1074/jbc.M313478200.

57. Samuel VT, Liu Z-X, Wang A, et al. Inhibition of protein kinase Cepsilon prevents hepatic insulin resistance in nonalcoholic fatty liver disease. *J Clin Invest*. 2007;117(3):739-745. doi:10.1172/JCI30400.

58. Perry RJ, Kim T, Zhang X-M, et al. Reversal of Hypertriglyceridemia, Fatty Liver Disease, and Insulin Resistance by a Liver-Targeted Mitochondrial Uncoupler. *Cell Metab*. 2013;18(5):740-748. doi:10.1016/j.cmet.2013.10.004.

59. Monetti M, Levin MC, Watt MJ, et al. Dissociation of hepatic steatosis and insulin resistance in mice overexpressing DGAT in the liver. *Cell Metab*. 2007;6(1):69-78. doi:10.1016/j.cmet.2007.05.005.

60. Kumashiro N, Erion DM, Zhang D, et al. Cellular mechanism of insulin resistance in nonalcoholic fatty liver disease. *Proc Natl Acad Sci USA*. 2011;108(39):16381-16385. doi:10.1073/pnas.1113359108.

61. Magkos F, Su X, Bradley D, et al. Intrahepatic diacylglycerol content is associated with hepatic insulin resistance in obese subjects. *Gastroenterology*. 2012;142(7):1444–6.e2. doi:10.1053/j.gastro.2012.03.003.

62. Arkan MC, Hevener AL, Greten FR, et al. IKK-β links inflammation to obesity-induced insulin resistance. *Nat Med*. 2005;11(2):191-198. doi:10.1038/nm1185.

63. Cai D, Yuan M, Frantz DF, et al. Local and systemic insulin resistance resulting from hepatic activation of IKK-beta and NF-kappaB. *Nat Med*. 2005;11(2):183-190. doi:10.1038/nm1166.

64. Tilg H, Moschen AR. Insulin resistance, inflammation, and non-alcoholic fatty liver disease. *Trends in Endocrinology & Metabolism*. 2008;19(10):371-379. doi:10.1016/j.tem.2008.08.005.

65. Meex RC, Hoy AJ, Morris A, et al. Fetuin B Is a Secreted Hepatocyte Factor Linking Steatosis to Impaired Glucose Metabolism. *Cell Metab*. 2015;22(6):1078-1089. doi:10.1016/j.cmet.2015.09.023.

66. Holland WL, Brozinick JT, Wang L-P, et al. Inhibition of Ceramide Synthesis Ameliorates Glucocorticoid-, Saturated-Fat-, and Obesity-Induced Insulin Resistance. *Cell Metab*. 2007;5(3):167-179. doi:10.1016/j.cmet.2007.01.002.

67. Ussher JR, Koves TR, Cadete VJJ, et al. Inhibition of de novo ceramide synthesis reverses diet-induced insulin resistance and enhances whole-body oxygen consumption. *Diabetes*. 2010;59(10):2453-2464. doi:10.2337/db09-1293.

68. Arab JP, Barrera F, Gallego C, et al. High prevalence of undiagnosed liver cirrhosis and advanced fibrosis in type 2 diabetic patients. *Ann Hepatol*. 2016;15(5):721-728. doi:10.5604/16652681.1212434.

69. Younossi ZM, Gramlich T, Matteoni CA, Boparai N, McCullough AJ. Nonalcoholic fatty liver disease in patients with type 2 diabetes. *Clin Gastroenterol Hepatol*. 2004;2(3):262-265.

70. Porepa L, Ray JG, Sanchez-Romeu P, Booth GL. Newly diagnosed diabetes mellitus as a risk factor for serious liver disease. *CMAJ*. 2010;182(11):E526-E531. doi:10.1503/cmaj.092144.

71. Alexander M, Loomis AK, van der Lei J, et al. Risks and clinical predictors of cirrhosis and hepatocellular carcinoma diagnoses in adults with

diagnosed NAFLD: real-world study of 18 million patients in four European cohorts. *BMC Med*. 2019;17(1):95. doi:10.1186/s12916-019-1321-x.

72. Wang P, Kang D, Cao W, Wang Y, Liu Z. Diabetes mellitus and risk of hepatocellular carcinoma: a systematic review and meta-analysis. *Diabetes/Metabolism Research and Reviews*. 2012;28(2):109-122. doi:10.1002/dmrr.1291.

73. Simon TG, King LY, Chong DQ, et al. Diabetes, metabolic comorbidities, and risk of hepatocellular carcinoma: Results from two prospective cohort studies. *Hepatology*. 2018;67(5):1797-1806. doi:10.1002/hep.29660.

74. Vilar-Gomez E, Calzadilla Bertot L, Wai-Sun Wong V, et al. Type 2 Diabetes and Metformin Use Associate With Outcomes of Patients With Non-alcoholic Steatohepatitis-related, Child-Pugh A Cirrhosis. *Clin Gastroenterol Hepatol*. May 2020. doi:10.1016/j.cgh.2020.04.083.

75. 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. doi:10.2337/dc12-0002.

76. Rafiq N, Bai C, Fang Y, et al. Long-term follow-up of patients with nonalcoholic fatty liver. *Clin Gastroenterol Hepatol*. 2009;7(2):234-238. doi:10.1016/j.cgh.2008.11.005.

77. Angulo P, Kleiner DE, Dam-Larsen S, et al. Liver Fibrosis, but No Other Histologic Features, Is Associated With Long-term Outcomes of Patients With Nonalcoholic Fatty Liver Disease. *Gastroenterology*. 2015;149(2):389–97.e10. doi:10.1053/j.gastro.2015.04.043.

78. Zoppini G, Fedeli U, Gennaro N, Saugo M, Targher G, Bonora E. Mortality from chronic liver diseases in diabetes. *Am J Gastroenterol*. 2014;109(7):1020-1025. doi:10.1038/ajg.2014.132.

79. de Marco R, Locatelli F, Zoppini G, Verlato G, Bonora E, Muggeo M. Cause-specific mortality in type 2 diabetes. The Verona Diabetes Study. *Diabetes Care*. 1999;22(5):756-761.

80. Wild SH, Morling JR, McAllister DA, et al. Type 2 diabetes and risk of hospital admission or death for chronic liver diseases. *J Hepatol*. 2016;64(6):1358-1364. doi:10.1016/j.jhep.2016.01.014.

81. Lo L, McLennan SV, Williams PF, et al. Diabetes is a progression factor for hepatic fibrosis in a high fat fed mouse obesity model of non-alcoholic steatohepatitis. *J Hepatol*. 2011;55(2):435-444. doi:10.1016/j.jhep.2010.10.039.

82. Guimarães ELM, Empsen C, Geerts A, van Grunsven LA. Advanced glycation end products induce production of reactive oxygen species via the activation of NADPH oxidase in murine hepatic stellate cells. *J Hepatol*. 2010;52(3):389-397. doi:10.1016/j.jhep.2009.12.007.

83. Lin J, Chen A. Curcumin diminishes the impacts of hyperglycemia on the activation of hepatic stellate cells by suppressing membrane translocation and gene expression of glucose transporter-2. *Mol Cell Endocrinol*. 2011;333(2):160-171. doi:10.1016/j.mce.2010.12.028.

84. Alberti KGMM, Eckel RH, Grundy SM, et al. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. In: Vol 120. American Heart Association, Inc.; 2009:1640-1645. doi:10.1161/CIRCULATIONAHA.109.192644.

85. Caldwell SH, Oelsner DH, Iezzoni JC, Hespenheide EE, Battle EH, Driscoll CJ. Cryptogenic cirrhosis: clinical characterization and risk factors for underlying disease. *Hepatology*. 1999;29(3):664-669. doi:10.1002/hep.510290347.

86. Day CP, James OF. Steatohepatitis: a tale of two "hits"? *Gastroenterology*. 1998;114(4):842-845.

87. Cusi K. Role of obesity and lipotoxicity in the development of nonalcoholic steatohepatitis: pathophysiology and clinical implications. *Gastroenterology*. 2012;142(4):711-725.e716. doi:10.1053/j.gastro.2012.02.003.

88. Neuschwander-Tetri BA. Hepatic lipotoxicity and the pathogenesis of nonalcoholic steatohepatitis: the central role of nontriglyceride fatty acid metabolites. *Hepatology*. 2010;52(2):774-788. doi:10.1002/hep.23719.

89. Yamaguchi K, Yang L, McCall S, et al. Inhibiting triglyceride synthesis improves hepatic steatosis but exacerbates liver damage and fibrosis in obese mice with nonalcoholic steatohepatitis. *Hepatology*. 2007;45(6):1366-1374. doi:10.1002/hep.21655.

90. Choi CS, Savage DB, Kulkarni A, et al. Suppression of diacylglycerol acyltransferase-2 (DGAT2), but not DGAT1, with antisense oligonucleotides reverses diet-induced hepatic steatosis and insulin resistance. *Journal of Biological Chemistry*. 2007;282(31):22678-22688. doi:10.1074/jbc.M704213200.

91. Day CP. Pathogenesis of steatohepatitis. *Best Practice & Research Clinical Gastroenterology*. 2002;16(5):663-678. doi:10.1053/bega.2002.0333.

92. Tilg H, Moschen AR. Evolution of inflammation in nonalcoholic fatty liver disease: the multiple parallel hits hypothesis. *Hepatology*. 2010;52(5):1836-1846. doi:10.1002/hep.24001.

93. Liu W, Baker RD, Bhatia T, Zhu L, Baker SS. Pathogenesis of nonalcoholic steatohepatitis. *Cell Mol Life Sci*. 2016;73(10):1969-1987. doi:10.1007/s00018-016-2161-x.

94. Donnelly KL, Smith CI, Schwarzenberg SJ, Jessurun J, Boldt MD, Parks EJ. Sources of fatty acids stored in liver and secreted via lipoproteins in patients

with nonalcoholic fatty liver disease. *J Clin Invest*. 2005;115(5):1343-1351. doi:10.1172/JCI23621.

95. Fujita K, Nozaki Y, Wada K, et al. Dysfunctional very-low-density lipoprotein synthesis and release is a key factor in nonalcoholic steatohepatitis pathogenesis. *Hepatology*. 2009;50(3):772-780. doi:10.1002/hep.23094.

96. Sanyal AJ, Campbell-Sargent C, Mirshahi F, et al. Nonalcoholic steatohepatitis: association of insulin resistance and mitochondrial abnormalities. *Gastroenterology*. 2001;120(5):1183-1192. doi:10.1053/gast.2001.23256.

97. Pagano G. Nonalcoholic steatohepatitis, insulin resistance, and metabolic syndrome: Further evidence for an etiologic association. *Hepatology*. 2002;35(2):367-372. doi:10.1053/jhep.2002.30690.

98. Korenblat KM, Fabbrini E, Mohammed BS, Klein S. Liver, muscle, and adipose tissue insulin action is directly related to intrahepatic triglyceride content in obese subjects. *Gastroenterology*. 2008;134(5):1369-1375. doi:10.1053/j.gastro.2008.01.075.

99. Seppälä-Lindroos A, Vehkavaara S, Häkkinen A-M, et al. Fat accumulation in the liver is associated with defects in insulin suppression of glucose production and serum free fatty acids independent of obesity in normal men. *The Journal of Clinical Endocrinology & Metabolism*. 2002;87(7):3023-3028. doi:10.1210/jcem.87.7.8638.

100. Li S, Brown MS, Goldstein JL. Bifurcation of insulin signaling pathway in rat liver: mTORC1 required for stimulation of lipogenesis, but not inhibition of gluconeogenesis. *Proc Natl Acad Sci USA*. 2010;107(8):3441-3446. doi:10.1073/pnas.0914798107.

101. Yoon H-J, Cha BS. Pathogenesis and therapeutic approaches for non-alcoholic fatty liver disease. *World J Hepatol*. 2014;6(11):800-811. doi:10.4254/wjh.v6.i11.800.

102. Shimomura I, Bashmakov Y, Ikemoto S, Horton JD, Brown MS, Goldstein JL. Insulin selectively increases SREBP-1c mRNA in the livers of rats with streptozotocin-induced diabetes. *Proc Natl Acad Sci USA*. 1999;96(24):13656-13661. doi:10.1073/pnas.96.24.13656.

103. Dowman JK, Tomlinson JW, Newsome PN. Pathogenesis of non-alcoholic fatty liver disease. *QJM*. 2010;103(2):71-83. doi:10.1093/qjmed/hcp158.

104. Levene AP, Goldin RD. The epidemiology, pathogenesis and histopathology of fatty liver disease. *Histopathology*. 2012;61(2):141-152. doi:10.1111/j.1365-2559.2011.04145.x.

105. Riemens SC, Sluiter WJ, Dullaart RP. Enhanced escape of non-esterified fatty acids from tissue uptake: its role in impaired insulin-induced lowering of total rate of appearance in obesity and Type II diabetes mellitus. *Diabetologia*. 2000;43(4):416-426. doi:10.1007/s001250051324.

106. Frayn KN, Humphreys SM, Coppack SW. Net carbon flux across subcutaneous adipose tissue after a standard meal in normal-weight and insulin-resistant obese subjects. *Int J Obes Relat Metab Disord*. 1996;20(9):795-800.

107. van der Poorten D, Milner K-L, Hui J, et al. Visceral fat: A key mediator of steatohepatitis in metabolic liver disease. *Hepatology*. 2008;48(2):449-457. doi:10.1002/hep.22350.

108. Fernandez-Real JM. Circulating Interleukin 6 Levels, Blood Pressure, and Insulin Sensitivity in Apparently Healthy Men and Women. *The Journal of Clinical Endocrinology & Metabolism*. 2001;86(3):1154-1159. doi:10.1210/jc.86.3.1154.

109. Hotamisligil G, Shargill N, Spiegelman B. Adipose expression of tumor necrosis factor-alpha: direct role in obesity-linked insulin resistance. *Science*. 1993;259(5091):87-91. doi:10.1126/science.7678183.

110. Sabio G, Das M, Mora A, et al. A stress signaling pathway in adipose tissue regulates hepatic insulin resistance. *Science*. 2008;322(5907):1539-1543. doi:10.1126/science.1160794.

111. Singh R, Kaushik S, Wang Y, et al. Autophagy regulates lipid metabolism. *Nature*. 2009;458(7242):1131-1135. doi:10.1038/nature07976.

112. Schattenberg JM, Singh R, Wang Y, et al. JNK1 but not JNK2 promotes the development of steatohepatitis in mice. *Hepatology*. 2006;43(1):163-172. doi:10.1002/hep.20999.

113. Musso G, Gambino R, Durazzo M, et al. Adipokines in NASH: postprandial lipid metabolism as a link between adiponectin and liver disease. *Hepatology*. 2005;42(5):1175-1183. doi:10.1002/hep.20896.

114. Bugianesi E, Pagotto U, Manini R, et al. Plasma Adiponectin in Nonalcoholic Fatty Liver Is Related to Hepatic Insulin Resistance and Hepatic Fat Content, Not to Liver Disease Severity. *The Journal of Clinical Endocrinology & Metabolism*. 2005;90(6):3498-3504. doi:10.1210/jc.2004-2240.

115. Xu A, Wang Y, Keshaw H, Xu LY, Lam KSL, Cooper GJS. The fat-derived hormone adiponectin alleviates alcoholic and nonalcoholic fatty liver diseases in mice. *J Clin Invest*. 2003;112(1):91-100. doi:10.1172/JCI17797.

116. Oral EA, Simha V, Ruiz E, et al. Leptin-replacement therapy for lipodystrophy. *N Engl J Med*. 2002;346(8):570-578. doi:10.1056/NEJMoa012437.

117. Jensen T, Abdelmalek MF, Sullivan S, et al. Fructose and sugar: A major mediator of non-alcoholic fatty liver disease. *J Hepatol*. 2018;68(5):1063-1075. doi:10.1016/j.jhep.2018.01.019.

118. Lanaspa MA, Ishimoto T, Li N, et al. Endogenous fructose production and metabolism in the liver contributes to the development of metabolic syndrome. *Nat Commun*. 2013;4(1):2434-2439. doi:10.1038/ncomms3434.

119. Sanchez-Lozada LG, Mu W, Roncal C, et al. Comparison of free fructose and glucose to sucrose in the ability to cause fatty liver. *Eur J Nutr*. 2010;49(1):1-9. doi:10.1007/s00394-009-0042-x.

120. Mouzaki M, Comelli EM, Arendt BM, et al. Intestinal microbiota in patients with nonalcoholic fatty liver disease. *Hepatology*. 2013;58(1):120-127. doi:10.1002/hep.26319.

121. Miele L, Valenza V, La Torre G, et al. Increased intestinal permeability and tight junction alterations in nonalcoholic fatty liver disease. *Hepatology*. 2009;49(6):1877-1887. doi:10.1002/hep.22848.

122. Wigg AJ, Roberts-Thomson IC, Dymock RB, McCarthy PJ, Grose RH, Cummins AG. The role of small intestinal bacterial overgrowth, intestinal permeability, endotoxaemia, and tumour necrosis factor alpha in the pathogenesis of non-alcoholic steatohepatitis. *Gut*. 2001;48(2):206-211. doi:10.1136/gut.48.2.206.

123. Bäckhed F, Ding H, Wang T, et al. The gut microbiota as an environmental factor that regulates fat storage. *Proc Natl Acad Sci USA*. 2004;101(44):15718-15723. doi:10.1073/pnas.0407076101.

124. Jiang C, Xie C, Li F, et al. Intestinal farnesoid X receptor signaling promotes nonalcoholic fatty liver disease. *J Clin Invest*. 2014;125(1):386-402. doi:10.1172/JCI76738.

125. Zhu L, Baker SS, Gill C, et al. Characterization of gut microbiomes in nonalcoholic steatohepatitis (NASH) patients: a connection between endogenous alcohol and NASH. *Hepatology*. 2013;57(2):601-609. doi:10.1002/hep.26093.

126. Baker SS, Baker RD, Liu W, Nowak NJ, Zhu L. Role of alcohol metabolism in non-alcoholic steatohepatitis. Gluud C, ed. *PLoS ONE*. 2010;5(3):e9570. doi:10.1371/journal.pone.0009570.

127. Cohen JC, Horton JD, Hobbs HH. Human fatty liver disease: old questions and new insights. *Science*. 2011;332(6037):1519-1523. doi:10.1126/science.1204265.

128. Loomba R, Schork N, Chen C-H, et al. Heritability of Hepatic Fibrosis and Steatosis Based on a Prospective Twin Study. *Gastroenterology*. 2015;149(7):1784-1793. doi:10.1053/j.gastro.2015.08.011.

129. Caussy C, Soni M, Cui J, et al. Nonalcoholic fatty liver disease with cirrhosis increases familial risk for advanced fibrosis. *J Clin Invest*. 2017;127(7):2697-2704. doi:10.1172/JCI93465.

130. Cui J, Chen C-H, Lo M-T, et al. Shared genetic effects between hepatic steatosis and fibrosis: A prospective twin study. *Hepatology*. 2016;64(5):1547-1558. doi:10.1002/hep.28674.

131. Speliotes EK, Yerges-Armstrong LM, Wu J, et al. Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. McCarthy MI, ed. *PLoS Genet*. 2011;7(3):e1001324. doi:10.1371/journal.pgen.1001324.

132. Anstee QM, Day CP. The Genetics of Nonalcoholic Fatty Liver Disease: Spotlight on PNPLA3 and TM6SF2. *Semin Liver Dis*. 2015;35(3):270-290. doi:10.1055/s-0035-1562947.

133. Romeo S, Kozlitina J, Xing C, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. *Nature Genetics*. 2008;40(12):1461-1465. doi:10.1038/ng.257.

134. Kawaguchi T, Sumida Y, Umemura A, et al. Genetic polymorphisms of the human PNPLA3 gene are strongly associated with severity of non-alcoholic fatty liver disease in Japanese. Okanoue T, ed. *PLoS ONE*. 2012;7(6):e38322. doi:10.1371/journal.pone.0038322.

135. Koo BK, Joo SK, Kim D, et al. Additive effects of PNPLA3 and TM6SF2 on the histological severity of non-alcoholic fatty liver disease. *J Gastroenterol Hepatol*. 2017;149:389. doi:10.1111/jgh.14056.

136. Rotman Y, Koh C, Zmuda JM, Kleiner DE, Liang TJ, NASH CRN. The association of genetic variability in patatin-like phospholipase domain-containing protein 3 (PNPLA3) with histological severity of nonalcoholic fatty liver disease. *Hepatology*. 2010;52(3):894-903. doi:10.1002/hep.23759.

137. Anstee QM, Day CP. The genetics of NAFLD. *Nat Rev Gastroenterol Hepatol*. 2013;10(11):645-655. doi:10.1038/nrgastro.2013.182.

138. Singal AG, Manjunath H, Yopp AC, et al. The Effect of PNPLA3 on Fibrosis Progression and Development of Hepatocellular Carcinoma: A Meta-analysis. *Am J Gastroenterol*. 2014;109(3):325-334. doi:10.1038/ajg.2013.476.

139. Falleti E, Fabris C, Cmet S, et al. PNPLA3 rs738409C/G polymorphism in cirrhosis: relationship with the aetiology of liver disease and hepatocellular carcinoma occurrence. *Liver International*. 2011;31(8):1137-1143. doi:10.1111/j.1478-3231.2011.02534.x.

140. Liu Y-L, Reeves HL, Burt AD, et al. TM6SF2 rs58542926 influences hepatic fibrosis progression in patients with non-alcoholic fatty liver disease. *Nat Commun*. 2014;5:4309. doi:10.1038/ncomms5309.

141. Kozlitina J, Smagris E, Stender S, et al. Exome-wide association study identifies a TM6SF2 variant that confers susceptibility to nonalcoholic fatty liver disease. *Nature Genetics*. 2014;46(4):352-356. doi:10.1038/ng.2901.

142. Dongiovanni P, Petta S, Maglio C, et al. Transmembrane 6 superfamily member 2 gene variant disentangles nonalcoholic steatohepatitis from cardiovascular disease. *Hepatology*. 2015;61(2):506-514. doi:10.1002/hep.27490.

143. Qiao A, Liang J, Ke Y, et al. Mouse patatin-like phospholipase domain-containing 3 influences systemic lipid and glucose homeostasis. *Hepatology*. 2011;54(2):509-521. doi:10.1002/hep.24402.

144. Basantani MK, Sitnick MT, Cai L, et al. Pnpla3/Adiponutrin deficiency in mice does not contribute to fatty liver disease or metabolic syndrome. *J Lipid Res*. 2011;52(2):318-329. doi:10.1194/jlr.M011205.

145. Abul-Husn NS, Cheng X, Li AH, et al. A Protein-Truncating HSD17B13 Variant and Protection from Chronic Liver Disease. *N Engl J Med*. 2018;378(12):1096-1106. doi:10.1056/NEJMoa1712191.

146. Kashima J, Shintani-Ishida K, Nakajima M, et al. Immunohistochemical study of the autophagy marker microtubule-associated protein 1 light chain 3 in normal and steatotic human livers. *Hepatol Res*. 2014;44(7):779-787. doi:10.1111/hepr.12183.

147. Perazzo H, Dufour J-F. The therapeutic landscape of non-alcoholic steatohepatitis. *Liver Int*. 2017;37(5):634-647. doi:10.1111/liv.13270.

148. Schuppan D, Schattenberg JM. Non-alcoholic steatohepatitis: pathogenesis and novel therapeutic approaches. *J Gastroenterol Hepatol*. 2013;28 Suppl 1(Suppl. 1):68-76. doi:10.1111/jgh.12212.

149. Seki S, Kitada T, Yamada T, Sakaguchi H, Nakatani K, Wakasa K. In situ detection of lipid peroxidation and oxidative DNA damage in non-alcoholic fatty liver diseases. *J Hepatol*. 2002;37(1):56-62.

150. Pirola CJ, Gianotti TF, Burgueño AL, et al. Epigenetic modification of liver mitochondrial DNA is associated with histological severity of nonalcoholic fatty liver disease. *Gut*. 2013;62(9):1356-1363. doi:10.1136/gutjnl-2012-302962.

151. Day CP. From Fat to Inflammation. *Gastroenterology*. 2006;130(1):207-210. doi:10.1053/j.gastro.2005.11.017.

152. Puri P, Mirshahi F, Cheung O, et al. Activation and dysregulation of the unfolded protein response in nonalcoholic fatty liver disease. *Gastroenterology*. 2008;134(2):568-576. doi:10.1053/j.gastro.2007.10.039.

153. Sahai A, Malladi P, Melin-Aldana H, Green RM, Whitington PF. Upregulation of osteopontin expression is involved in the development of nonalcoholic steatohepatitis in a dietary murine model. *AJP: Gastrointestinal and Liver Physiology*. 2004;287(1):G264-G273. doi:10.1152/ajpgi.00002.2004.

154. Esposito E, Iacono A, Bianco G, et al. Probiotics reduce the inflammatory response induced by a high-fat diet in the liver of young rats. *J Nutr*. 2009;139(5):905-911. doi:10.3945/jn.108.101808.

155. Velayudham A, Dolganiuc A, Ellis M, et al. VSL#3 probiotic treatment attenuates fibrosis without changes in steatohepatitis in a diet-induced nonalcoholic steatohepatitis model in mice. *Hepatology*. 2008;49(3):989-997. doi:10.1002/hep.22711.

156. Crespo J, Cayón A, Fernández-Gil P, et al. Gene expression of tumor necrosis factor alpha and TNF-receptors, p55 and p75, in nonalcoholic steatohepatitis patients. *Hepatology*. 2001;34(6):1158-1163. doi:10.1053/jhep.2001.29628.

157. Hui JM, Hodge A, Farrell GC, Kench JG, Kriketos A, George J. Beyond insulin resistance in NASH: TNF-? or adiponectin? *Hepatology*. 2004;40(1):46-54. doi:10.1002/hep.20280.

158. Haukeland JW, Damås JK, Konopski Z, et al. Systemic inflammation in nonalcoholic fatty liver disease is characterized by elevated levels of CCL2. *J Hepatol*. 2006;44(6):1167-1174. doi:10.1016/j.jhep.2006.02.011.

159. Ueki K, Kondo T, Tseng YH, Kahn CR. Central role of suppressors of cytokine signaling proteins in hepatic steatosis, insulin resistance, and the metabolic syndrome in the mouse. *Proc Natl Acad Sci USA*. 2004;101(28):10422-10427. doi:10.1073/pnas.0402511101.

160. Nobili V, Carpino G, Alisi A, et al. Hepatic progenitor cells activation, fibrosis, and adipokines production in pediatric nonalcoholic fatty liver disease. *Hepatology*. 2012;56(6):2142-2153. doi:10.1002/hep.25742.

161. Roskams T, Yang SQ, Koteish A, et al. Oxidative stress and oval cell accumulation in mice and humans with alcoholic and nonalcoholic fatty liver disease. *Am J Pathol*. 2003;163(4):1301-1311. doi:10.1016/S0002-9440(10)63489-X.

162. Cortez-Pinto H, Baptista A, Camilo ME, de Moura MC. Hepatic stellate cell activation occurs in nonalcoholic steatohepatitis. *Hepatogastroenterology*. 2001;48(37):87-90.

163. Nakamura T, Sakata R, Ueno T, Sata M, Ueno H. Inhibition of transforming growth factor beta prevents progression of liver fibrosis and enhances hepatocyte regeneration in dimethylnitrosamine-treated rats. *Hepatology*. 2000;32(2):247-255. doi:10.1053/jhep.2000.9109.

164. Kanzler S, Lohse AW, Keil A, et al. TGF-beta1 in liver fibrosis: an inducible transgenic mouse model to study liver fibrogenesis. *Am J Physiol*. 1999;276(4 Pt 1):G1059-G1068.

165. Villanueva A. Hepatocellular Carcinoma. Longo DL, ed. *N Engl J Med*. 2019;380(15):1450-1462. doi:10.1056/NEJMra1713263.

166. Evans CDJ, Oien KA, MacSween RNM, Mills PR. Non-alcoholic steatohepatitis: a common cause of progressive chronic liver injury? *J Clin Pathol*. 2002;55(9):689-692.

167. Teli MR, James OF, Burt AD, Bennett MK, Day CP. The natural history of nonalcoholic fatty liver: a follow-up study. *Hepatology*. 1995;22(6):1714-1719.

168. Lee RG. Nonalcoholic steatohepatitis: a study of 49 patients. *Hum Pathol*. 1989;20(6):594-598.

169. Harrison SA, Torgerson S, Hayashi PH. The natural history of nonalcoholic fatty liver disease: a clinical histopathological study. *Am J Gastroenterol*. 2003;98(9):2042-2047. doi:10.1111/j.1572-0241.2003.07659.x.

170. Fassio E, Álvarez E, Domínguez N, Landeira G, Longo C. Natural history of nonalcoholic steathepatitis: A longitudinal study of repeat liver biopsies. *Hepatology*. 2004;40(4):820-826. doi:10.1002/hep.1840400411.

171. Adams LA, Sanderson S, Lindor KD, Angulo P. The histological course of nonalcoholic fatty liver disease: a longitudinal study of 103 patients with sequential liver biopsies. *J Hepatol*. 2005;42(1):132-138. doi:10.1016/j.jhep.2004.09.012.

172. Hui AY, Wong VW-S, Chan HL-Y, et al. Histological progression of non-alcoholic fatty liver disease in Chinese patients. *Aliment Pharmacol Ther*. 2005;21(4):407-413. doi:10.1111/j.1365-2036.2005.02334.x.

173. Sorrentino P, Terracciano L, D'Angelo S, Ferbo U, Bracigliano A, Vecchione R. Predicting fibrosis worsening in obese patients with NASH through parenchymal fibronectin, HOMA-IR, and hypertension. *Am J Gastroenterol*. 2010;105(2):336-344. doi:10.1038/ajg.2009.587.

174. Wong VW-S, Wong GL-H, Choi PC-L, et al. Disease progression of non-alcoholic fatty liver disease: a prospective study with paired liver biopsies at 3 years. *Gut*. 2010;59(7):969-974. doi:10.1136/gut.2009.205088.

175. Pais R, Charlotte F, Fedchuk L, et al. A systematic review of follow-up biopsies reveals disease progression in patients with non-alcoholic fatty liver. *J Hepatol*. 2013;59(3):550-556. doi:10.1016/j.jhep.2013.04.027.

176. Perazzo H, Munteanu M, Ngo Y, et al. Prognostic value of liver fibrosis and steatosis biomarkers in type-2 diabetes and dyslipidaemia. *Aliment Pharmacol Ther*. 2014;40(9):1081-1093. doi:10.1111/apt.12946.

177. McPherson S, Hardy T, Henderson E, Burt AD, Day CP, Anstee QM. Evidence of NAFLD progression from steatosis to fibrosing-steatohepatitis using paired biopsies: implications for prognosis and clinical management. *J Hepatol*. 2015;62(5):1148-1155. doi:10.1016/j.jhep.2014.11.034.

178. Pelusi S, Petta S, Rosso C, et al. Renin-Angiotensin System Inhibitors, Type 2 Diabetes and Fibrosis Progression: An Observational Study in

Patients with Nonalcoholic Fatty Liver Disease. Wong V, ed. *PLoS ONE*. 2016;11(9):e0163069. doi:10.1371/journal.pone.0163069.

179. 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. doi:10.1016/j.jhep.2009.03.019.

180. 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. doi:10.1016/j.cgh.2014.04.014.

181. Dam-Larsen S, Becker U, Franzmann M-B, Larsen K, Christoffersen P, Bendtsen F. Final results of a long-term, clinical follow-up in fatty liver patients. *Scand J Gastroenterol*. 2009;44(10):1236-1243. doi:10.1080/00365520903171284.

182. 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. Lin H-C, ed. *PLoS ONE*. 2015;10(6):e0128774. doi:10.1371/journal.pone.0128774.

183. Hashimoto E, Yatsuji S, Tobari M, et al. Hepatocellular carcinoma in patients with nonalcoholic steatohepatitis. *J Gastroenterol*. 2009;44 Suppl 19(S19):89-95. doi:10.1007/s00535-008-2262-x.

184. Ascha MS, Hanouneh IA, Lopez R, Tamimi TA-R, Feldstein AF, Zein NN. The incidence and risk factors of hepatocellular carcinoma in patients with nonalcoholic steatohepatitis. *Hepatology*. 2010;51(6):1972-1978. doi:10.1002/hep.23527.

185. 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. Lin H-C, ed. *PLoS ONE*. 2015;10(6):e0128774. doi:10.1371/journal.pone.0128774.

186. Piscaglia F, Svegliati Baroni G, Barchetti A, et al. Clinical patterns of hepatocellular carcinoma in nonalcoholic fatty liver disease: A multicenter prospective study. *Hepatology*. 2016;63(3):827-838. doi:10.1002/hep.28368.

187. Younossi ZM, Otgonsuren M, Henry L, et al. Association of nonalcoholic fatty liver disease (NAFLD) with hepatocellular carcinoma (HCC) in the United States from 2004 to 2009. *Hepatology*. 2015;62(6):1723-1730. doi:10.1002/hep.28123.

188. Wild SH, Walker JJ, Morling JR, et al. Cardiovascular Disease, Cancer, and Mortality Among People With Type 2 Diabetes and Alcoholic or Nonalcoholic Fatty Liver Disease Hospital Admission. *Diabetes Care*. 2017;41(2):341-347. doi:10.2337/dc17-1590.

189. 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. doi:10.1002/hep.26156.

190. Stepanova M, Rafiq N, Makhlouf H, et al. Predictors of all-cause mortality and liver-related mortality in patients with non-alcoholic fatty liver disease (NAFLD). *Dig Dis Sci*. 2013;58(10):3017-3023. doi:10.1007/s10620-013-2743-5.

191. Xun Y-H, Guo J-C, Lou G-Q, et al. Non-alcoholic fatty liver disease (NAFLD) fibrosis score predicts 6.6-year overall mortality of Chinese patients with NAFLD. *Clin Exp Pharmacol Physiol*. 2014;41(9):643-649. doi:10.1111/1440-1681.12260.

192. Ekstedt M, Hagström H, Nasr P, et al. Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. *Hepatology*. 2015;61(5):1547-1554. doi:10.1002/hep.27368.

193. Chang Y, Ryu S, Sung K-C, et al. Alcoholic and non-alcoholic fatty liver disease and associations with coronary artery calcification: evidence from the Kangbuk Samsung Health Study. *Gut*. November 2018:gutjnl–2018–317666. doi:10.1136/gutjnl-2018-317666.

194. Matteoni CA, Younossi ZM, Gramlich T, Boparai N, Liu YC, McCullough AJ. Nonalcoholic fatty liver disease: a spectrum of clinical and pathological severity. *Gastroenterology*. 1999;116(6):1413-1419.

195. Xun Y-H, Guo J-C, Lou G-Q, et al. Non-alcoholic fatty liver disease (NAFLD) fibrosis score predicts 6.6-year overall mortality of Chinese patients with NAFLD. *Clin Exp Pharmacol Physiol*. 2014;41(9):643-649. doi:10.1111/1440-1681.12260.

196. Caruso MG, Veronese N, Notarnicola M, et al. Fatty liver and mortality: a cohort population study in South Italy. *BMJ Open*. 2019;9(6):e027379. doi:10.1136/bmjopen-2018-027379.

197. Kim Y, Chang Y, Cho YK, Ahn J, Shin H, Ryu S. Obesity and Weight Gain Are Associated With Progression of Fibrosis in Patients With Nonalcoholic Fatty Liver Disease. *Clin Gastroenterol Hepatol*. 2019;17(3):543–550.e2. doi:10.1016/j.cgh.2018.07.006.

198. Lee T-Y, Wu J-C, Yu S-H, Lin J-T, Wu M-S, Wu C-Y. The occurrence of hepatocellular carcinoma in different risk stratifications of clinically noncirrhotic nonalcoholic fatty liver disease. *Int J Cancer*. 2017;141(7):1307-1314. doi:10.1002/ijc.30784.

199. Akuta N, Kawamura Y, Arase Y, et al. Hepatocellular carcinoma is the most common liver-related complication in patients with histopathologically-confirmed NAFLD in Japan. *BMC Gastroenterol*. 2018;18(1):165–10. doi:10.1186/s12876-018-0900-1.

200. Vilar-Gomez E, Chalasani N. Non-invasive assessment of non-alcoholic fatty liver disease: Clinical prediction rules and blood-based biomarkers. *J Hepatol*. 2018;68(2):305-315. doi:10.1016/j.jhep.2017.11.013.

201. Önnerhag K, Hartman H, Nilsson PM, Lindgren S. Non-invasive fibrosis scoring systems can predict future metabolic complications and overall mortality in non-alcoholic fatty liver disease (NAFLD). *Scand J Gastroenterol*. 2019;54(3):328-334. doi:10.1080/00365521.2019.1583366.

202. 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. doi:10.1093/qjmed/hcv191.

203. Angulo P, Keach JC, Batts KP, Lindor KD. Independent predictors of liver fibrosis in patients with nonalcoholic steatohepatitis. *Hepatology*. 1999;30(6):1356-1362. doi:10.1002/hep.510300604.

204. Miyaaki H, Ichikawa T, Nakao K, et al. Clinicopathological study of nonalcoholic fatty liver disease in Japan: the risk factors for fibrosis. *Liver Int*. 2008;28(4):519-524. doi:10.1111/j.1478-3231.2007.01614.x.

205. Singh DK, Sakhuja P, Malhotra V, Gondal R, Sarin SK. Independent predictors of steatohepatitis and fibrosis in Asian Indian patients with non-alcoholic steatohepatitis. *Dig Dis Sci*. 2008;53(7):1967-1976. doi:10.1007/s10620-007-0074-0.

206. Hossain N, Afendy A, Stepanova M, et al. Independent predictors of fibrosis in patients with nonalcoholic fatty liver disease. *Clin Gastroenterol Hepatol*. 2009;7(11):1224–9–1229.e1–2. doi:10.1016/j.cgh.2009.06.007.

207. Williams CD, Stengel J, Asike MI, et al. Prevalence of Nonalcoholic Fatty Liver Disease and Nonalcoholic Steatohepatitis Among a Largely Middle-Aged Population Utilizing Ultrasound and Liver Biopsy: A Prospective Study. *Gastroenterology*. 2011;140(1):124-131. doi:10.1053/j.gastro.2010.09.038.

208. Tomita K, Teratani T, Yokoyama H, et al. Serum immunoglobulin a concentration is an independent predictor of liver fibrosis in nonalcoholic steatohepatitis before the cirrhotic stage. *Dig Dis Sci*. 2011;56(12):3648-3654. doi:10.1007/s10620-011-1771-2.

209. Petta S, Amato MC, Di Marco V, et al. Visceral adiposity index is associated with significant fibrosis in patients with non-alcoholic fatty liver disease. *Aliment Pharmacol Ther*. 2012;35(2):238-247. doi:10.1111/j.1365-2036.2011.04929.x.

210. Ong JP, Pitts A, Younossi ZM. Increased overall mortality and liver-related mortality in non-alcoholic fatty liver disease. *J Hepatol*. 2008;49(4):608-612. doi:10.1016/j.jhep.2008.06.018.

211. Labenz C, Huber Y, Kalliga E, et al. Predictors of advanced fibrosis in non-cirrhotic non-alcoholic fatty liver disease in Germany. *Aliment Pharmacol Ther*. 2018;48(10):1109-1116. doi:10.1111/apt.14976.

212. Petit JM, Guiu B, Masson D, et al. PNPLA3 polymorphism influences liver fibrosis in unselected patients with type 2 diabetes. *Liver Int*. 2011;31(9):1332-1336. doi:10.1111/j.1478-3231.2011.02566.x.

213. Chan W-K, Nik Mustapha NR, Mahadeva S. A novel 2-step approach combining the NAFLD fibrosis score and liver stiffness measurement for predicting advanced fibrosis. *Hepatol Int*. 2015;9(4):594-602. doi:10.1007/s12072-014-9596-7.

214. Hagström H, Nasr P, Bottai M, et al. Elevated serum ferritin is associated with increased mortality in non-alcoholic fatty liver disease after 16 years of follow-up. *Liver Int*. 2016;36(11):1688-1695. doi:10.1111/liv.13144.

215. Castro PCS de, Alberton HCP, Pedroso MLA, Morsoletto DBG, Pissaia Junior A, Ivantes CAP. EVALUATION OF PROGRESSION OF HEPATIC FIBROSIS IN A GROUP OF PATIENTS WITH NON-ALCOHOLIC FATTY LIVER DISEASE ACCOMPANIED FOR 10 YEARS. *Arq Gastroenterol*. 2019;56(3):256-260. doi:10.1590/S0004-2803.201900000-48.

216. Kleiner DE, Brunt EM, Wilson LA, et al. Association of Histologic Disease Activity With Progression of Nonalcoholic Fatty Liver Disease. *JAMA Netw Open*. 2019;2(10):e1912565-e1912565. doi:10.1001/jamanetworkopen.2019.12565.

217. 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(6):1388-1402. doi:10.1016/j.jhep.2015.11.004.

218. Anstee QM, Lawitz EJ, Alkhouri N, et al. Noninvasive Tests Accurately Identify Advanced Fibrosis due to NASH: Baseline Data From the STELLAR Trials. *Hepatology*. 2019;70(5):1521-1530. doi:10.1002/hep.30842.

219. Giannini EG, Testa R, Savarino V. Liver enzyme alteration: a guide for clinicians. *CMAJ*. 2005;172(3):367-379. doi:10.1503/cmaj.1040752.

220. 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*. 2003;38(2):518-526. doi:10.1053/jhep.2003.50346.

221. 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*. 2010;59(9):1265-1269. doi:10.1136/gut.2010.216077.

222. Sorbi D, Boynton J, Lindor KD. The ratio of aspartate aminotransferase to alanine aminotransferase: potential value in differentiating nonalcoholic steatohepatitis from alcoholic liver disease. *Am J Gastroenterol*. 1999;94(4):1018-1022. doi:10.1111/j.1572-0241.1999.01006.x.

223. Rosenberg WMC, Voelker M, Thiel R, et al. Serum markers detect the presence of liver fibrosis: a cohort study. *Gastroenterology*. 2004;127(6):1704-1713. doi:10.1053/j.gastro.2004.08.052.

224. 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. doi:10.1002/hep.21984.

225. 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*. 2016;36(3):370-377. doi:10.1111/liv.12896.

226. Non-alcoholic fatty liver disease (NAFLD): assessment and management. August 2018:1-16.

227. 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. doi:10.1002/hep.21178.

228. Vallet-Pichard A, Mallet V, Nalpas B, et al. FIB-4: an inexpensive and accurate marker of fibrosis in HCV infection. comparison with liver biopsy and fibrotest. *Hepatology*. 2007;46(1):32-36. doi:10.1002/hep.21669.

229. 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. doi:10.1016/j.cgh.2009.05.033.

230. 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. doi:10.1002/hep.21496.

231. Treeprasertsuk S. NAFLD fibrosis score: A prognostic predictor for mortality and liver complications among NAFLD patients. *World J Gastroenterol*. 2013;19(8):1219-1229. doi:10.3748/wjg.v19.i8.1219.

232. 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. doi:10.1016/j.jhep.2008.07.035.

233. Staufer K, Halilbasic E, Spindelboeck W, et al. Evaluation and comparison of six noninvasive tests for prediction of significant or advanced fibrosis in nonalcoholic fatty liver disease. *United European Gastroenterol J*. 2019;7(8):1113-1123. doi:10.1177/2050640619865133.

234. 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):782-789.e784. doi:10.1053/j.gastro.2013.06.057.

235. Siddiqui MS, Yamada G, Vuppalanchi R, et al. Diagnostic Accuracy of Noninvasive Fibrosis Models to Detect Change in Fibrosis Stage. *Clin*

Gastroenterol Hepatol. 2019;17(9):1877–1885.e5. doi:10.1016/j.cgh.2018.12.031.

236. Stefan N, Häring H-U, Cusi K. Non-alcoholic fatty liver disease: causes, diagnosis, cardiometabolic consequences, and treatment strategies. *Lancet Diabetes Endocrinol*. August 2018. doi:10.1016/S2213-8587(18)30154-2.

237. 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;59:2188. doi:10.1111/liv.13739.

238. Davyduke T, Tandon P, Al-Karaghouli M, Abraldes JG, Ma MM. Impact of Implementing a “FIB-4 First” Strategy on a Pathway for Patients With NAFLD Referred From Primary Care. *Hepatol Commun*. 2019;3(10):1322-1333. doi:10.1002/hep4.1411.

239. Singh A, Le P, Peerzada MM, Lopez R, Alkhouri N. The Utility of Noninvasive Scores in Assessing the Prevalence of Nonalcoholic Fatty Liver Disease and Advanced Fibrosis in Type 2 Diabetic Patients. *J Clin Gastroenterol*. 2018;52(3):268-272. doi:10.1097/MCG.0000000000000905.

240. Morling JR, Fallowfield JA, Williamson RM, et al. Non-invasive hepatic biomarkers (ELF and CK18) in people with type 2 diabetes: the Edinburgh type 2 diabetes study. *Liver Int*. 2014;34(8):1267-1277. doi:10.1111/liv.12385.

241. Harrison SA, Oliver D, Arnold HL, Gogia S, Neuschwander-Tetri BA. Development and validation of a simple NAFLD clinical scoring system for identifying patients without advanced disease. *Gut*. 2008;57(10):1441-1447. doi:10.1136/gut.2007.146019.

242. Sheng X, Che H, Ji Q, et al. The Relationship Between Liver Enzymes and Insulin Resistance in Type 2 Diabetes Patients with Nonalcoholic Fatty Liver Disease. *Horm Metab Res*. 2018;50(5):397-402. doi:10.1055/a-0603-7899.

243. Tanaka K, Nanbara S, Tanaka T, Koide H, Hayashi T. Aminotransferase activity in the liver of diabetic mice. *Diabetes Res Clin Pract*. 1988;5(1):71-75. doi:10.1016/s0168-8227(88)80081-0.

244. Friedman SL, Neuschwander-Tetri BA, Rinella M, Sanyal AJ. Mechanisms of NAFLD development and therapeutic strategies. *Nat Med*. 2018;24(7):908-922. doi:10.1038/s41591-018-0104-9.

245. Zelber-Sagi S, Godos J, Salomone F. Lifestyle changes for the treatment of nonalcoholic fatty liver disease: a review of observational studies and intervention trials. *Therap Adv Gastroenterol*. 2016;9(3):392-407. doi:10.1177/1756283X16638830.

246. Vilar-Gomez E, Martinez-Perez Y, Calzadilla Bertot L, et al. Weight Loss Through Lifestyle Modification Significantly Reduces Features of Nonalcoholic Steatohepatitis. *Gastroenterology*. 2015;149(2):367–78.e5–quize14–5. doi:10.1053/j.gastro.2015.04.005.

247. Lazo M, Hernaez R, Bonekamp S, et al. Non-alcoholic fatty liver disease and mortality among US adults: prospective cohort study. *BMJ*. 2011;343(nov18 2):d6891-d6891. doi:10.1136/bmj.d6891.

248. Lean MEJ, Leslie WS, Barnes AC, et al. Durability of a primary care-led weight-management intervention for remission of type 2 diabetes: 2-year results of the DiRECT open-label, cluster-randomised trial. *Lancet Diabetes Endocrinol*. 2019;7(5):344-355. doi:10.1016/S2213-8587(19)30068-3.

249. Katan MB. Weight-loss diets for the prevention and treatment of obesity. *N Engl J Med*. 2009;360(9):923-925. doi:10.1056/NEJMe0810291.

250. Lassailly G, Caiazzo R, Buob D, et al. Bariatric Surgery Reduces Features of Nonalcoholic Steatohepatitis in Morbidly Obese Patients. *Gastroenterology*. 2015;149(2):379–88–quize15–6. doi:10.1053/j.gastro.2015.04.014.

251. Sanyal AJ, Chalasani N, Kowdley KV, et al. Pioglitazone, vitamin E, or placebo for nonalcoholic steatohepatitis. *N Engl J Med*. 2010;362(18):1675-1685. doi:10.1056/NEJMoa0907929.

252. Eriksson JW, Lundkvist P, Jansson P-A, et al. Effects of dapagliflozin and n-3 carboxylic acids on non-alcoholic fatty liver disease in people with type 2 diabetes: a double-blind randomised placebo-controlled study. *Diabetologia*. 2018;61(9):1923-1934. doi:10.1007/s00125-018-4675-2.

253. Newsome PN, Buchholtz K, Cusi K, et al. A Placebo-Controlled Trial of Subcutaneous Semaglutide in Nonalcoholic Steatohepatitis. *N Engl J Med*. November 2020:NEJMoa2028395. doi:10.1056/NEJMoa2028395.

254. 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 diseas. *Diabet Med*. 2018;35(3):368-375. doi:10.1111/dme.13565.

255. Price JF, Reynolds RM, Mitchell RJ, et al. The Edinburgh Type 2 Diabetes Study: study protocol. *BMC Endocr Disord*. 2008;8(1):18. doi:10.1186/1472-6823-8-18.

256. Marioni RE, Strachan MWJ, Reynolds RM, et al. Association between raised inflammatory markers and cognitive decline in elderly people with type 2 diabetes: the Edinburgh Type 2 Diabetes Study. *Diabetes*. 2010;59(3):710-713. doi:10.2337/db09-1163.

257. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. *Arch Intern Med*. 1998;158(16):1789-1795. doi:10.1001/archinte.158.16.1789.

258. 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. doi:10.1016/j.crad.2010.09.021.

259. 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 Gastroenterol*. 2006;6(1):33. doi:10.1186/1471-230X-6-33.

260. Raghunathan TE. What do we do with missing data? Some options for analysis of incomplete data. *Annu Rev Public Health*. 2004;25(1):99-117. doi:10.1146/annurev.publhealth.25.102802.124410.

261. Sterne JAC, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*. 2009;338(jun29 1):b2393-b2393. doi:10.1136/bmj.b2393.

262. 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;55(1):2005. doi:10.1111/liv.14590.

263. 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*. 2007;51(3):444-450. doi:10.1007/s00125-007-0897-4.

264. 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. doi:10.1016/j.jhep.2013.10.017.

265. 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;36(Supplement 1):19.

266. Ciardullo S, Sala I, Perseghin G. Screening strategies for nonalcoholic fatty liver disease in type 2 diabetes: Insights from NHANES 2005-2016. *Diabetes Res Clin Pract*. 2020;167:108358. doi:10.1016/j.diabres.2020.108358.

267. McPherson S, Hardy T, Dufour J-F, et al. Age as a Confounding Factor for the Accurate Non-Invasive Diagnosis of Advanced NAFLD Fibrosis. *Am J Gastroenterol*. 2017;112(5):740-751. doi:10.1038/ajg.2016.453.

268. Ntandja Wandji LC, Gnemmi V, Mathurin P, Louvet A. Combined alcoholic and non-alcoholic steatohepatitis. *JHEP Rep*. 2020;2(3):100101. doi:10.1016/j.jhepr.2020.100101.

269. Plevris JN, Haydon GH, Simpson KJ, et al. Serum hyaluronan--a non-invasive test for diagnosing liver cirrhosis. *European Journal of Gastroenterology & Hepatology*. 2000;12(10):1121-1127.

270. Plevris N, Sinha R, Hay AW, McDonald N, Plevris JN, Hayes PC. Index serum hyaluronic acid independently and accurately predicts mortality in patients with liver disease. *Aliment Pharmacol Ther*. 2018;48(4):423-430. doi:10.1111/apt.14897.

271. Dominici S, Paolicchi A, Corti A, Maellaro E, Pompella A. Prooxidant reactions promoted by soluble and cell-bound gamma-glutamyltransferase activity. *Methods Enzymol*. 2005;401:484-501. doi:10.1016/S0076-6879(05)01029-3.

272. McLernon DJ, Donnan PT, Sullivan FM, et al. Prediction of liver disease in patients whose liver function tests have been checked in primary care: model development and validation using population-based observational cohorts. *BMJ Open*. 2014;4(6):e004837. doi:10.1136/bmjopen-2014-004837.

273. Grecian SM, McLachlan S, Fallowfield JA, et al. Clinical care and other categories posters: Lipids and fatty liver. *Diabet Med*. 2020;37(S1):138-138. doi:10.1111/dme.39_14245.

274. Wilson JMG, Jungner G. PRINCIPLES AND PRACTICE OF SCREENING FOR DISEASE. *Public Health Papers*. 1968;(34):1-168.

275. Eddowes PJ, Sasso M, Allison M, et al. Accuracy of FibroScan Controlled Attenuation Parameter and Liver Stiffness Measurement in Assessing Steatosis and Fibrosis in Patients With Nonalcoholic Fatty Liver Disease. *Gastroenterology*. 2019;156(6):1717-1730. doi:10.1053/j.gastro.2019.01.042.

276. Emdin M, Pompella A, Paolicchi A. Gamma-glutamyltransferase, atherosclerosis, and cardiovascular disease: triggering oxidative stress within the plaque. *Circulation*. 2005;112(14):2078-2080. doi:10.1161/CIRCULATIONAHA.105.571919.

277. Yang JD, Harmsen WS, Slettedahl SW, et al. Factors that affect risk for hepatocellular carcinoma and effects of surveillance. *Clin Gastroenterol Hepatol*. 2011;9(7):617–23.e1. doi:10.1016/j.cgh.2011.03.027.

278. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *Journal of the American Statistical Association*. 1999;94(446):496-509.

279. Alba AC, Agoritsas T, Walsh M, et al. Discrimination and Calibration of Clinical Prediction Models: Users' Guides to the Medical Literature. *JAMA*. 2017;318(14):1377-1384. doi:10.1001/jama.2017.12126.

280. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J*. 2014;35(29):1925-1931. doi:10.1093/eurheartj/ehu207.

281. Royston P, Moons KGM, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. *BMJ*. 2009;338(mar31 1):b604-b604. doi:10.1136/bmj.b604.

282. Moons KGM, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? *BMJ*. 2009;338(feb23 1):b375-b375. doi:10.1136/bmj.b375.

283. Caussy C, Ajmera VH, Puri P, et al. Serum metabolites detect the presence of advanced fibrosis in derivation and validation cohorts of patients with non-alcoholic fatty liver disease. *Gut*. 2019;68(10):1884-1892. doi:10.1136/gutjnl-2018-317584.

284. Daniels SJ, Leeming DJ, Eslam M, et al. ADAPT: An Algorithm Incorporating PRO-C3 Accurately Identifies Patients With NAFLD and Advanced Fibrosis. *Hepatology*. 2019;69(3):1075-1086. doi:10.1002/hep.30163.

285. Newsome PN, Sasso M, Deeks JJ, et al. FibroScan-AST (FAST) score for the non-invasive identification of patients with non-alcoholic steatohepatitis with significant activity and fibrosis: a prospective derivation and global validation study. *Lancet Gastroenterol Hepatol*. 2020;5(4):362-373. doi:10.1016/S2468-1253(19)30383-8.