INVESTIGATING THE PERFORMANCE OF NON-INVASIVE FIBROSIS TESTS IN ALCOHOL RELATED LIVER DISEASE

Freya Alison Rhodes

Thesis submitted to University College of London
for the degree of Doctor of Philosophy

2021

DECLARATION

I, Freya Alison Rhodes, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

ABSTRACT

Alcohol-related Liver Disease (ArLD) often presents late, when opportunities to improve prognosis are limited. In the last 18 months, the SARS-CoV-2 pandemic has exacerbated already climbing mortality rates from ArLD, with a 20% increase in ArLD-related deaths between 2019 and 2020, on the background of a 43% increase between 2001 and 2019. Non-invasive-tests for liver fibrosis are increasingly advocated for use in ArLD but are not yet established in routine practice. In this thesis I aimed to investigate current alcohol referral practice from primary-care to specialist hepatology services, and the performance of commonly used non-invasive tests to detect fibrosis severity and predict mortality in ArLD.

A systematic review and meta-analysis on four non-invasive tests in ArLD revealed a paucity of studies on alcohol compared to other liver aetiologies, but found good performance (AUROC >0.7) of all four tests (FIB4/FibroTest/ELF/FibroScan) in detecting F2/F3/F4 fibrosis-stages. A 3-year retrospective evaluation of alcohol-referrals to secondary-care found two-thirds of referrals were 'unnecessary', in that they had no evidence of advanced fibrosis or cirrhosis. Only 16% had a non-invasive fibrosis test performed prior to referral, and I applied modelling of simple fibrosis scores (FIB4 and APRI) to demonstrate that the proportion of unnecessary referrals could be reduced by 50% if simple non-invasive fibrosis tests were used in primary-care before referral.

I report the design and set-up of a pathway involving use of non-invasive tests in primary-care (specifically, the Enhanced Liver Fibrosis (ELF) test in people with alcohol-use-disorder (AUD), based in Camden and Islington practices). A prospective 1-year study in 99 inpatients diagnosed with AUD found a third of

patients had elevated ELF scores indicative of advanced fibrosis that had not been detected previously, despite multiple previous hospital attendances. ELF was not associated with recent alcohol intake or AST/ALT values, differentiating it from FibroScan. A second systematic review on prognostic-performance of non-invasive fibrosis tests found FIB4/ELF/FibroScan/FibroTest performed well (AUROCs all >0.7) in predicting mortality, and ELF/FibroTest performed equally well or better than liver histology.

Finally, in a cohort of 162 serum samples from patients with Alcoholic Hepatitis (AH) from the 'STOPAH' cohort (a published Randomised Controlled Trial of steroids and pentoxifylline conducted in over 1,000 patients with AH), I provide the first evidence that the Enhanced Liver Fibrosis (ELF) test can be used to predict outcomes in alcoholic hepatitis (AH), and discovered a new prognostic biomarker combining ELF and ABIC, which outperformed traditional prognostic biomarkers in predicting 90-day mortality in AH.

Impact statement

This thesis has investigated the current use, diagnostic, and prognostic performance of non-invasive fibrosis tests in Alcohol related Liver Disease (ArLD). In two systematic reviews, (one with meta-analysis), I have confirmed the value of non-invasive fibrosis tests for the diagnosis of liver fibrosis and cirrhosis in ArLD, and their ability to predict outcomes (including liver-related mortality, all-cause mortality and variceal bleeding).

Whilst I discovered that two-thirds of patients referred from primary to secondary care with suspected ArLD for liver specialist assessment were 'unnecessary referrals' (in that they had no evidence of advanced fibrosis and could be discharged), the use of a simple score based on routinely available blood tests in primary care (FIB4) would have reduced the proportion of unnecessary referrals by 50%. I report the design process and launch of a new primary care alcohol pathway incorporating ELF testing in people drinking at hazardous or harmful levels, with planned analysis of FIB4 as a first step in this pathway. This has the potential to benefit:

1. Patients through a reduction in unnecessary hospital appointments and investigations; improved detection of liver damage due to alcohol; better understanding of the consequences of harmful drinking.

2. The NHS through reduction in the unnecessary use of secondary-care resources; improved detection of liver disease at a point at which intervention can avoid harm; cost savings should accrue through the reduction in referrals and investigations and avoidance of harms from unnecessary investigations, and cost-utility should arise through the reduction in harms from advanced liver disease.

I report the use of opportunistic ELF testing to detect occult liver fibrosis in patients with AUD presenting to hospital.

Finally, I have shown for the first time that not only can a non-invasive fibrosis marker (ELF) be used in Alcoholic Hepatitis (AH) to predict mortality, but when combined with a traditionally used simple score based on blood tests and age (ABIC), the combined ELF-ABIC score performs superiorly in 90-day mortality prediction to simple blood test scores alone. This is the first study where the performance of a biomarker in predicting 90-day mortality in AH has reached an AUROC above 0.8. If validated, this could be readily adopted in routine clinical practice and would be of interest to clinicians to enable more accurate prognosis prediction in this patient cohort.

This thesis has the potential to benefit people living in the community with AUD engaging in primary care, primary care physicians seeking to identify appropriate patients for referral to secondary care, secondary care physicians seeking to identify chronic liver disease in people presenting to hospital with AUD, and intensive care doctors managing patients with AH. At all levels this thesis with have health economic implications for the wider society and the NHS.

To date, my work in this thesis has resulted in three peer-reviewed publications and several abstracts at national and international conferences.

ACKNOWLEDGEMENTS

I am extremely grateful to Professor William Rosenberg for giving me the opportunity to conduct this research, for teaching me how to 'think like a scientist' and for all the wisdom he has imparted to me along the way. I have gained a wide range of skills under his supervision, from the work within this thesis, and also additional work for other studies in which I participated in alongside this PhD, that have not only been integral to this work, but I strongly believe will also make me a better doctor.

This process has also taught me the importance of collaboration during research, and I have learnt a great deal from conducting and attending meetings with PPI (patients and public involvement), and being a part of a working research group to help set up a new study together with health economists, statisticians, and primary care physicians. I am also grateful for the collaboration with the STOPAH trial team at St Mary's, whose work has been highly inspirational to me, and who kindly allowed me to investigate the performance of the ELF test in patient samples from their STOPAH trial.

I would like to thank Professor Alison Rodger, to whom I am eternally grateful for her fantastic support as a secondary supervisor throughout my PhD. I have learnt so much from her ideas and suggestions which have no doubt improved the quality of my work. She has been a real positive inspiration in my career.

I am very grateful to Dr Rachel Westbrook, Dr Sudeep Tanwar and Dr Jasmina Panovska-Griffiths, for their roles as supervisors and on my thesis committee, who have each provided me with invaluable feedback on my studies that has shaped the way I have conducted and written up the research. Rachel has also been a fantastic clinical mentor in hepatology.

I would like to thank my Dad, who has always encouraged me, believed in me, and inspired me.

Finally, when more than half of my PhD time has been during a global pandemic, balancing research work with extra clinical shifts, I am especially grateful to my husband and close friends for their patience, support, and encouragement.

Supervisors

Primary supervisor

Professor William Rosenberg
Peter Scheuer Professor of Liver Diseases Institute for Liver and Digestive Health Division of Medicine
University College London
And consultant hepatologist
Royal Free London NHS Foundation trust

Secondary supervisors

Professor Alison Rodger
Professor of Infectious Diseases
Infection & Population Health
Institute for Global Health
Faculty of Population Health Sciences
University College London

Dr Rachel Westbrook
Consultant hepatologist
Royal Free London NHS Foundation trust
And Institute for Liver and Digestive Health
Division of Medicine
University College London

Dr Jasmina Panovska-Griffiths
Department of Applied Health Research
And Institute of Global Health
University College London

Thesis committee: Secondary supervisors plus Dr Sudeep Tanwar, Consultant Gastroenterologist and Hepatologist at Barts Health NHS Trust, Associate Professor at University College London Institute for Liver and Digestive Health.

RELEVANT PUBLICATIONS AND AWARDS

PUBLICATIONS

Arising directly from this thesis:

Rhodes FA, Trembling P, Panovska-Griffiths J, Tanwar S, Westbrook RH, Rodger A, et al. Systematic review: Investigating the prognostic performance of four non-invasive tests in alcohol-related liver disease. J Gastroenterol Hepatol. **2021**;36(6):1435-49.

Rhodes FA, Cococcia S, Patel P, Panovska-Griffiths J, Tanwar S, Westbrook RH, et al. Is there scope to improve the selection of patients with alcohol-related liver disease for referral to secondary care? A retrospective analysis of primary care referrals to a UK liver centre, incorporating simple blood tests. BMJ Open. 2021;11(6):e047786.

Rhodes F, Cococcia S, Panovska-Griffiths J, Tanwar S, Westbrook RH, Rodger A, et al. Uncovering unsuspected advanced liver fibrosis in patients referred to alcohol nurse specialists using the ELF test. BMC Gastroenterol. **2021**;21(1):143.

(In submission):

Rhodes FA, Trembling P, Smith, C, Tanwar S, Westbrook RH, Rodger A, Rosenberg WM. Systematic review: Investigating the diagnostic performance of four non-invasive tests in alcohol-related liver disease *(In submission **2021**).*

Published alongside the thesis:

Hussain A, Patel PJ, **Rhodes F**, Srivastava A, Patch D, Rosenberg W. Decompensated cirrhosis is the commonest presentation for NAFLD patients undergoing liver transplant assessment. Clin Med (Lond). **2020** May;20(3):313-318.

Patel PJ, Connoley D, **Rhodes F**, Srivastava A, Rosenberg W. A review of the clinical utility of the Enhanced Liver Fibrosis test in multiple aetiologies of chronic liver disease. Ann Clin Biochem. **2020**;57(1):36-43.

Connoley D, Patel PJ, Hogan B, Tanwar S, **Rhodes F,** Parkes J, et al. The Enhanced Liver Fibrosis test maintains its diagnostic and prognostic performance in alcohol-related liver disease: a cohort study. BMC Gastroenterol. **2021**;21(1):268.

Tanwar S, **Rhodes F**, Srivastava A, Trembling PM, Rosenberg WM. Inflammation and fibrosis in chronic liver diseases including non-alcoholic fatty liver disease and hepatitis C. World J Gastroenterol. **2020** Jan 14;26(2):109-133.

Abstracts presented during PhD

International

Rhodes F, Greenham O, Hunt E, Shah S, Ryan J, Patch D, et al. *Cirrhotic patients with vitamin d deficiency fail to respond to oral replacement therapy*. J Hepatol. 2019;70:E674-E. (Poster presentation EASL **2019** (selected for 'poster tour': 10 minutes oral presentation next to poster)

F Rhodes, Greenham O, Hunt E, Shah T, Ryan J, Patch D, et al. *Bone disease does not correlate with severity of liver disease in cirrhosis*. J Hepatol. **2019**;70:e674. (Poster presentation at EASL 2019)

Connoley D, Patel P, Hogan B, Tanwar S, **Rhodes F,** Parkes J, et al. *The utility of the enhanced liver fibrosis test in alcoholic liver disease*. J Hepatol. **2019**;70:e815. (Poster presentation EASL 2019)

Rhodes F, Cococcia S, Vergis N, Thursz M, Atkinson S, The-STOPAH-Trial-Group, et al. *An algorithm combining Hyaluronic acid, Procollagen III Amino-Terminal peptide, and the Glasgow Alcoholic Hepatitis score in a single number predicts 90-day mortality in alcohol-related hepatitis*. Hepatology. **2019**;70. (Poster presentation AASLD 2019)

Cococcia S, **Rhodes F**, Patel P, Connoley C, Di Sabatino A, Rosenberg W. *Are Patients with BAFLD at higher risk to have advanced liver fibrosis? A retrospective analysis of referrals.* Hepatology **2019;** 70 (S1) p1067A-1067A. (Poster presentation AASLD 2019)

Rhodes F, Cococcia S, Connoley D, Patel P, Rosenberg W. *Alcohol Use disorder and liver fibrosis - cases are missed through failure to test.* J Hepatol. **2020**;73:S176-7. (Poster presentation EASL 2020)

Rhodes F, Vergis N, Thursz M, The STOPAH Trial Management group, Westbrook R, Rodger A, Tanwar S, Rosenberg W. *The combination of ABIC+ELF as a single marker outperforms ABIC alone in predicting 90-day survival in alcoholic hepatitis*. a/w publication in J Hepatol **2021**. (Poster presentation EASL 2021, awarded bursary).

National

Chan SC, **Rhodes F**, Greenham O, Hunt E, Westbrook R. *Vitamin D deficiency does not increase the risk of post-liver transplant allograft cellular rejection (ACR).* Gut **2019**; 68 (S2) (Poster presentation BSG 2019)

Rhodes F, Cococcia S, Patel P, Connelly D, Rosenberg W. Alcohol use disorders and liver fibrosis-Can we improve the referral pathway to secondary care? (Poster presentation BSG **2019***)*

Patel P, Hussain A, **Rhodes F**, Srivastava A, Patch D, Rosenberg W. *Decompensated cirrhosis is the commonest presentation for NAFLD patients undergoing liver transplant assessment*. Gut **2019**, 68 (S2) PTU-029. (Poster presentation BSG 2019)

Rhodes F, Cococcia S, Patel P, Connoley D, Rosenberg W. *Alcohol use disorders and liver fibrosis – can we improve the referral pathway to secondary care?* Gut **2019**; 68 (S2) A128 (Poster presentation BSG 2019)

Rhodes F, Greenham O, Hunt E, Kearney O, Patch D, Westbrook RH. *Bone disease does not correlate with severity of liver disease in cirrhosis.* J Hepatol **2019**; 70 (e674) (Poster presentation ILC 2019)

Rhodes F, Cococcia S, Connoley D, Patel P, Rosenberg WM. *Enhanced-Liver-Fibrosis score was not influenced by alcohol consumption in a patient cohort with Alcohol-Use-Disorders*. Gut. **2020**;70:Issue Suppl 1. (Poster presentation BSG **2021)**

Rhodes F, Cococcia S, Connoley D, Patel P, Rosenberg WM. *Alcohol use disorders and liver fibrosis – cases are missed through failure to test.* Gut. 2020;70:A152. (Poster presentation BSG **2021)**

PRIZES FROM PhD

- **BASL poster of distinction award 2019** (£250)
- **EASL ILC Young Investigator bursary award for abstract 2021** (conference fees paid for and free EASL membership for 1 year)
- **BSG Campus Prize winner 2021** (£50 voucher)

TABLE OF CONTENTS

Declaration……2

Abstract……3

Impact statement……5

Acknowledgements……7

Supervisors……8

Relevant publications and awards……9

Table of contents……12

List of abbreviations ……17

List of tables……19

List of figures……22

Chapter 1: Aims and objectives……26

1.1 Background: Alcohol-related Liver disease……27

1.2 Background: Fibrosis tests……29

1.3 Thesis rationale……32

1.3.1 Research questions……33

1.3.2 Aims and objectives……34

1.4 Outline of thesis……39

Chapter 2: Introduction to Alcohol-related Liver disease and non-invasive liver fibrosis tests……41

2.1 Alcohol and public health: The extent of the problem……42

2.2 How much alcohol is 'too much'? ……55

2.3 Spectrum of Alcohol-related Liver Disease……58

2.4 Pathophysiology of ArLD……61

2.5 Complications of ArLD……66

2.6 Impact of late presentation of symptoms of ArLD……73

2.7 Detection of fibrosis and cirrhosis……74

2.7.1 Liver biopsy – the imperfect 'gold standard'……74

2.8 Dealing with inaccuracies of liver biopsy……84

2.9 Serum markers of fibrosis……89

2.9.3 The Enhanced Liver Fibrosis (ELF) test……94

2.9.6 Imaging methods of Fibrosis assessment……98
2.9.7 Identified gaps in the evidence base……106
Chapter 3: Investigating the diagnostic performance of four non-invasive tests in Alcohol-related Liver Disease: a systematic review with meta-analysis……109
3.1. Abstract……111
3.2.1 Introduction……112
3.2.2 Ethics and patient consent……113
3.2.3 Methods……113
3.2.4 PICO……114
3.2.5 Search strategy……115
3.3 Quality assessment……122
3.4 Data synthesis and analysis……123
3.5 Results……133
3.6 Discussion……167
3.7.1 Strengths……172
3.7.2 Limitations……173
3.8 Conclusion……174
Chapter 4: Diagnosing advanced fibrosis in Alcohol-related Liver Disease in practice: examining current referral strategies from primary to secondary care, and risk factors associated with advanced fibrosis……175
4.1 Abstract……176
4.2 Introduction……178
4.3 Methods……180
4.4 Results……186
4.5 Discussion……201
Chapter 5: Implementing a community referral pathway involving the ELF test in patients with Alcohol Use Disorder – the 'Camden and Islington alcohol pathway'
5.1 Background……211
5.2.1 Current guidelines on non-invasive fibrosis testing in ArLD……213
5.2.2 The need for validation of the BSG Guidelines……217
5.3 The Camden and Islington (C&I) Alcohol pathway……218
5.3.1 Aim and objectives……218

5.3.2 Reason for choosing ELF over FibroScan……218
5.3.3 Objectives……219
5.3.4 Methods……220
5.3.5 Setting……220
5.3.6 Designing the pathway……221
5.3.7 Patient and public involvement……222
5.3.8 Ethics……223
5.3.9 Pathway design process……225
5.4 Deciding on the ELF threshold for the pathway……228
5.5.1 Camden 'Abnormal Liver Function Test (LFT) pathway'……230
5.5.2 Approval and launch of the pathways……233
5.5.3 Communicating the new pathway launch……234
5.6.1 Evaluation plan……234
5.7.3 Challenges of the pathway set-up and data collection……247
Chapter 6: Uncovering unsuspected advanced liver fibrosis in patients referred to the alcohol nurse specialist using the ELF test……259
6.1 Abstract……260
6.2 Introduction……262
6.3.1 Methods……264
6.3.2 Study Design and ethics……264
6.4 Statistical analysis……268
6.5 Results……269
6.6 Discussion……282
Chapter 7: Investigating the prognostic performance of four non-invasive tests in Alcohol-related Liver Disease: a systematic review……289
7.1 Abstract……290
7.2 Introduction……292
7.3.1 Patients and methods……294
7.3.2 PICO……295
7.3.3 Ethics and patient consent……295
7.3.4 Search strategy……295

7.3.8 Quality assessment……303
7.3.9 Data synthesis and analysis……304
7.4 Results……309
7.7 Discussion……332
Chapter 8: Investigating the prognostic performance of ELF in alcoholic hepatitis
8.1 Abstract……336
8.2 Introduction……339
8.3 Methods……343
8.3.1 STOPAH Methodology……343
8.3.2 Ethics……344
8.3.3 Patients and data collection……344
8.3.4 Measurement of key outcomes……344
8.3.5 Selection of prognostic models……344
8.3.8 Statistical analyses……346
8.4 Results……348
8.7.1 Discussion……372
Chapter 9: Summary of main findings and Discussion……380
9.1 Outline of thesis aims and methodology……381
9.2 Summary of key points emerging from the new work presented in this thesis……381
9.3 Discussion……388
9.3.1 Trends in alcohol use since the SARS-CoV2 pandemic……388
9.4 Current practice of non-invasive fibrosis testing in ArLD- scope for change……391
9.5 Diagnostic tests for fibrosis in ArLD- which to use?……393
9.6 Impact of inflammation and alcohol on non-invasive test results..396
9.7 Which non-invasive fibrosis tests can be used to predict prognosis in ArLD? ……398
9.8.1 Limitations……401
9.8.2 Unanswered questions and areas for future research……403
9.9.1 Final remarks……407

9.9.2 Overall conclusions……408

References……411-439

List of abbreviations:

ABIC = Age-Bilirubin-INR-Creatinine score

ACM= All-cause Mortality

AFLD: Alcoholic Fatty Liver Disease

AH = Alcoholic Hepatitis

ArLD: Alcohol related Liver Disease

ALT: Alanine Aminotransferase

ALP: Alkaline Phosphatase

APRI: AST to Platelet Ratio Index

AST: Aspartate Aminotransferase

AUD: Alcohol Use Disorder

AUDIT: Alcohol Use Disorder Identification Test

AUROC: Area Under the Receiver Operating Characteristics curve

BAFLD: Both Alcohol and Fatty Liver Disease

BMI: Body Mass Index

BSG: British Society of Gastroenterology

C&I: Camden and Islington

CCG: Clinical Commissioning Group

CI: Confidence Interval

CLD: Chronic Liver Disease

CP: Child Pugh score

ELF: Enhanced Liver Fibrosis

FIB4: Fibrosis 4 score

G/day = grams per day

GAHS = Glasgow Alcoholic Hepatitis Score

GPs: General Practitioners

HCC: Hepatocellular Carcinoma

HE: Hepatic Encephalopathy

HR: Hazard Ratio

HTN: Hypertension

IQR: Inter Quartile Range

LFTs: Liver Function Tests

LRE: Liver Related Event

LRM = Liver Related Mortality

LSM = Liver Stiffness Measurement

MELD: Model for End stage Liver Disease

MeSH = Medical Subject Headings

NAFLD: Non-Alcoholic Fatty Liver Disease

NHS: National Health Service

NLR = Neutrophil to Lymphocyte Ratio

OR: Odds Ratio

PC: Primary Care

PICO = Participants, interventions, comparators, outcomes

PRISMA = Preferred Reporting Items for Systematic reviews and Meta-Analyses

QALY: Quality-Adjusted Life Year

QUIPS = Quality In Prognosis Studies

RR: Relative Risk

SC: Secondary Care

SD: Standard deviation

SNLRD = Survival or Non-liver related Death

T2DM: Type II Diabetes Mellitus

UCL: University College London

UCLH: University College London Hospital

US: Ultrasound

U/w: Units per week

WOS: Web of Science

List of tables

Table 2.1: DSM-5 definition of Alcohol Use Disorder

Table 2.2: The AUDIT questionnaire (Alcohol Use Disorder Identification Tool)

Table 2.3: Classification of fibrosis stages from four commonly used staging tools

Table 2.4: 2 x 2 contingency table for test result and disease state

Table 2.5: AUROC interpretation

Table 2.6: Overview of non-invasive methods of liver fibrosis evaluation

Table 2.7: Diagnostic performance of serum biomarkers for F3 (advanced fibrosis) in ArLD

Table 2.8: Formulas for combination serum marker panels

Table 3.1: Databases searched

Table 3.2A: DETAILS OF SEARCH STRATEGY: Ovid MEDLINE and EMBASE

Table 3.2B: DETAILS OF SEARCH STRATEGY: Cochrane database

Table 3.2C: DETAILS OF SEARCH STRATEGY: Web of Science

Table 3.3: Selection criteria

Table 3.4 QUADAS2 Questionnaire

Table 3.5 Baseline study characteristics

Table 3.6 Diagnostic performance of FIB4, FibroTest, FibroScan, ELF

Table 3.7: Influence of alcohol, steatosis, obesity, age and inflammation on FIB4/FibroTest/FibroScan/ELF

Table 3.8: QUADAS2 risk of bias results

Table 4.1: Inclusion and exclusion criteria

Table 4.2: Baseline characteristics

Table 4.3: Demographics within each re-classified aetiology group

Table 4.4: Multivariable logistic regression analysis of factors influencing odds of a diagnosis of advanced fibrosis in whole study cohort

Table 4.5: Accuracy of indirect fibrosis markers in detecting advanced fibrosis in a cohort of 231 patients referred from primary care with suspected ArLD.

Table 5.1: Alcohol use disorders identification test consumption (AUDIT C)

Table 5.2 Diagnostic accuracy of ELF for advanced fibrosis in ArLD by Thiele et al.

Table 6.1: Comparison of clinical characteristics of patients with and without advanced fibrosis (as determined by ELF score of $\geq$ 10.5).

Table 6.2: Factors associated with advanced fibrosis (ELF $\geq$ 10.5), as determined by univariable and multivariable regression analyses

Table 6.3: Summary of multiple regression analysis of factors associated with continuous ELF score.

Table 6.4a: Multiple logistic regression analysis to investigate for effect of alcohol unit quartiles on presence or absence of advanced fibrosis (as per ELF $\geq$10.5)

Table 6.4b: Multiple linear regression analysis to investigate for effect of alcohol unit quartiles on continuous ELF score.

Table 7.1: Databases searched

Table 7.2a: Details of search strategy: Cochrane database

Table 7.2b: Details of search strategy: Ovid MEDLINE and EMBASE

Table 7.2c: Details of search strategy: Web of Science

Table 7.3: Selection criteria

Table 7.4: Summary of the Bias Domains, Rating of "Risk of bias" and Prompting Items and Considerations using an adapted QUIPS Tool, based on the following references:

Table 7.5: Baseline characteristics of included studies

Table 7.6: Baseline participant characteristics

Table 7.7: Prognostic performance of non-invasive tests in each study

Table 7.8: Study–Level Quality Assessment using the Quality In Prognosis Studies (QUIPS) Tool

Table 8.1: Patient characteristics: Comparison of sub-study with STOPAH original cohort

Table 8.2: Areas Under Receiver Operating Characteristic curves for variables at 28, 90 and 120-day mortality time points.

Table 8.3: Multivariable backwards stepwise logistic regression for 28, 90, 120-day mortality end-points

Table 8.4: Comparison of AUCs (whole cohort): combined direct fibrosis marker (ELF) with indirect fibrosis marker (GAHS, MELD or ABIC) versus indirect fibrosis marker alone.

Table 8.5: Performance of ELF, GAHS, MELD, ABIC in predicting 28, 90 and 120-day mortality

List of figures

Figure 2.1: Standardised UK Mortality Rates for the commonest causes of death, from 1970-2010

Figure 2.2: Cumulative alcohol related deaths in England and Wales.

Figure 2.3: Liver mortality in England and Wales, 1980-2013

Figure 2.4: Alcohol related Liver Disease hospital admissions in England

Figure 2.5: Trends in affordability of alcohol in the UK

Figure 2.6: The spectrum of alcohol-related liver disease

Figure 2.7: Representative diagram illustrating the liver anatomy, showing the hepatocytes surrounding the central portal tract.

Figure 3.1: PRISMA flow diagram

Figure 3.2: QUADAS2 risk of bias

Figure 3.3: Forest plot of AUROCs for all studies, all tests at F2/F3/F4 fibrosis stages

Figure 3.4: A/B FibroScan for the detection of ≥ F3 fibrosis

Figure 3.5: A/B FibroScan for the detection of ≥ F4 fibrosis

Figure 4.1: Proportion of patients referred from GP with suspected ArLD or NAFLD who had a diagnosis (composite clinical judgement) of advanced fibrosis

Figure 4.2a: Boxplot of ALP by diagnosis of advanced fibrosis in patients referred with suspected ArLD

Figure 4.2b: Boxplot of units of alcohol per week by diagnosis of advanced fibrosis in patients referred with suspected ArLD

Figure 4.3: Flow chart depicting reclassification of aetiologies into ArLD, BAFLD and NAFLD

Figure 4.4: ROC analysis of the performance of indirect tests for fibrosis and simple liver blood tests in the detection of advanced fibrosis (composite clinical judgement) in patients referred with suspected ArLD.

Figure 5.1: BSG alcohol-related liver disease guideline for non-invasive testing.

Figure 5.2: Flow diagram illustrating new pathway for ELF testing, for incorporation into the Camden and Islington alcohol pathway.

Figure 5.3: Final new alcohol pathway on Camden CCG website 2020

Figure 5.4A: First page of the updated 'abnormal liver function test' guidance

Figure 5.4B: Second page of the updated 'abnormal liver function test' guidance

Figure 5.5: Pathway evaluation data sources

Figure 5.6: Impact of SARS-CoV-2 pandemic on outpatient referral numbers

Figure 5.7: Data flow diagram

Figure 5.8: New NCL CCG fatty liver pathway

Figure 5.9: New NCL CCG ArLD pathway

Figure 6.1a: Pie chart of the principal recorded reasons for presentation to hospital

Figure 6.1b: Pie chart of the discharging hospital specialty team.

Figure 6.2: Proportion of patients in study cohort with advanced fibrosis as assessed by ELF $\geq$ 10.5

Figure 6.3 A, B, C, D: Influence of alcohol and age on binary and continuous ELF scores

Figure 6.4 A, B, C, D: Influence of ALT and AST on binary and continuous ELF scores

Figure 6.5 A, B: Influence of AST:ALT ratio on binary and continuous ELF scores

Figure 7.1: PRISMA flow Diagram ('Preferred Reporting Items for Systematic Reviews and Meta-Analyses')

Figure 7.2: QUIPs tool results: Quality assessment of included studies using the quality in prognosis studies tool.

Figure 7.3: Forest plot of AUROCs + 95% CI for outcome prediction

Figure 8.1 Boxplot of ELF against mortality status at 28, 90 and 120 days

Figure 8.2 A/B: Scatter plot of A) ELF by AST and B) ELF by CRP

Figure 8.3 A/B/C: ROC curves for ABIC+ELF, ABIC alone, and ELF alone at 28/90/120-day mortality end points.

Figure 8.4 A/B/C: Kaplan Meier Survival curves for combined ABIC+ELF at 28, 90, and 120-day mortality end-points.

Figure 8.5 A/B/C: Kaplan Meier survival curves for high or low ELF with and without prednisolone at 28/90/120 days.

Figure 8.6 A/B/C: Kaplan Meier Survival curves for incongruous ABIC+ELF categories, with or without prednisolone, at 28, 90, and 120-day mortality end-points.

Figure 8.7 A/B/C: Kaplan Meier Survival curves for combinations of low ABIC/ELF, high ABIC/ELF, with or without prednisolone, at 28, 90, and 120- day mortality end-points.

Figure 8.8 A/B/C: Kaplan Meier Survival curves for combined ABIC+ELF, using high ABIC threshold of 9, at 28, 90, and 120-day mortality end-points.

Figure 8.9 A/B/C: Kaplan Meier Survival curves for combined ABIC+ELF, using high ABIC threshold of 9, at 28, 90, and 120-day mortality end-points.

Figure 8.9.1: Kaplan Meier survival curve for high or low ELF/ABIC with and without Lille score at threshold of 0.45, at 90 days.

Figure 8.9.2: Kaplan Meier survival curve for high ABIC/ELF groups, with Lille score, on vs off prednisolone at 90 days.

Chapter 1

Aims and objectives

1.0 Background

1.1.1 Public health burden of alcohol

Alcohol is a major public health problem in the UK, with a quarter of the adult population estimated to be drinking over the government recommended limits of 14 units per week (2). Hospital admissions related to alcohol are rising annually, with 350,000 alcohol-related admissions per year in 2019, (an increase of 20% in a decade) (3) and with a cost to the NHS of £3.5 billion per year (4). This is likely to be due to a shift in drinking behaviours from low-strength beer in pubs to higher-strength beer, wine and spirits sold in supermarkets for home consumption (5). In addition, the price of alcohol is now 64% more affordable than it used to be 30 years ago (3).

Whilst alcohol causes a wide array of health and social harms, the greatest morbidity and mortality are associated with alcohol-related liver disease (ArLD)(6) .

1.1.2 Alcohol-related Liver Disease (ArLD)

ArLD encompasses a spectrum of liver damage, from steatosis ('fatty liver') to steatohepatitis (including the severe clinical syndrome Acute Alcoholic Hepatitis (AH)), liver fibrosis and cirrhosis. Whilst up to 90% of people who have an Alcohol Use Disorder (AUD) will develop hepatic steatosis (7, 8), only approximately 20% will progress to advanced fibrosis or cirrhosis both of which may benefit from

secondary care management (9). Genetic susceptibility, nutritional status and BMI are amongst the factors thought to predispose to the risk of cirrhosis (10).

Cirrhosis is characterised by a disruption to the normal relationship between substructures and cells in the liver that may result in liver failure, obstruction of blood flow through the liver (portal hypertension) and liver cancer. Cirrhosis is now the third commonest cause of premature death in working age, (11) with alcohol as its leading cause. Liver damage and the severity of fibrosis rise exponentially with increasing alcohol consumption (3) and liver fibrosis has been shown to be the strongest predictor of mortality from ArLD (5).

Mortality from cirrhosis has increased by 400% since 1970, mainly due to trends in alcohol consumption (12). As such, this is an area of vital importance to research.

1.1.3 The need for earlier detection of advanced fibrosis and cirrhosis

Advanced fibrosis and cirrhosis are usually asymptomatic prior to the development of complications of end stage liver disease, and simple liver function blood tests (LFTs) and ultrasound are not sensitive or specific for their detection (1). This likely contributes to the fact that 75% of patients with ArLD first present when their liver disease is advanced (i.e. decompensated cirrhosis) and when it is often too late for interventions to have a major effect on prognosis (13).

Conversely, it is suspected that many patients with suspected chronic liver disease (CLD) are referred unnecessarily from primary-care to secondary-care when they do

not have evidence of advanced fibrosis (equivalent to ≥ METAVIR F3) or cirrhosis. This was found to be the case in a recent study on NAFLD (non-alcoholic fatty liver disease) – which found that 92% of NAFLD referrals were 'unnecessary', in that they had steatosis but no evidence of advanced fibrosis (14).

The proportion of referrals with ArLD that could be considered 'unnecessary' (no evidence of advanced fibrosis or cirrhosis) is unknown.

There is, therefore, an urgent need for earlier detection of advanced fibrosis in people with Alcohol Use Disorder (AUD) in primary-care, so that patients who are at low-risk for advanced fibrosis may remain safely in primary-care and avoid an unnecessary referral, and so that those who are at high risk can be identified and appropriately selected and referred to secondary-care, where interventions can be implemented to improve their outcomes. This can include the use of beta-blockers to reduce variceal haemorrhage (15, 16) and screening and early treatment of hepatocellular cancers (17, 18).

1.2 Fibrosis tests

The 'reference standard' diagnostic test for liver fibrosis has traditionally been considered to be a liver biopsy (19), however its drawbacks include risk to the patient due to its invasiveness, sampling error and inter-observer variation in interpretation (20). To overcome these limitations, there is increasing interest in the development and use of non-invasive approaches for assessing liver fibrosis to

improve earlier diagnosis, to initiate management plans and monitor response. Whilst there are several available non-invasive tests for liver fibrosis (both blood biomarkers and physical tests such as elastography), the optimum test of choice for the diagnostic and prognostic use in ArLD remains open to debate.

Liver fibrosis tests include both direct and indirect biochemical markers of fibrosis. Direct markers of fibrosis measure constituents of liver matrix and molecules that regulate fibrogenesis and fibrolysis. They include tests that measure collagens, glycoproteins, metalloproteinases and their inhibitors (21). Indirect tests reflect the consequences of disease processes that impair hepatic function and include the AST to Platelet Ratio Index (22), FIB-4 score (23), Forns index (23), AST:ALT ratio (24) and gamma GT: platelet ratio (25). Physical methods of assessing liver fibrosis such as transient elastography (FibroScan) have also been shown to be accurate in the assessment of liver fibrosis but these have some limitations including the requirement for dedicated equipment and skilled operators, a recognised failure-rate particularly in obese and older subjects (26), and they are operator dependent (27). In contrast, assessment using blood tests is more advantageous as these can be incorporated into routine primary-care assessment, do not require skilled clinicians to undertake them and can be standardised and automated to deliver consistency.

1.2.1 The Enhanced Liver Fibrosis (ELF) test

The ELF test is an automated blood test generating a score derived from a logarithmic algorithm combining serum measurements of three “direct” markers of hepatic

extracellular matrix metabolism: Hyaluronic acid (HA), N-terminal peptide of procollagen III (PIIINP) and Tissue Inhibitor of Metalloproteinase-1 (TIMP-1). The ELF score has a linear relationship to fibrosis severity, with higher scores reflecting more advanced fibrosis (28). Since its original conception and validation in 2004 (29) the diagnostic value of the ELF test has been validated in cohorts of patients with primary biliary cholangitis (formerly cirrhosis) (30), primary sclerosing cholangitis (31), non-alcoholic fatty liver disease (32), chronic hepatitis C (33), chronic hepatitis B (34) and alcohol-related liver disease (ArLD) (23). A unit value of 9.8 is associated with a 65% sensitivity and 90% specificity for detecting advanced fibrosis and $\geq$9.8 is associated with a 76% sensitivity and 90% specificity for detecting advanced fibrosis (35).

A recent publication on the use of non-invasive tests (NIT) in primary-care for diagnosing liver fibrosis in NAFLD found that implementing a two-step pathway using FIB4 score and then the ELF test (in those with indeterminate FIB4 scores) produced an 88% reduction in 'unnecessary referrals' to hepatology with a five-fold increase in the detection of advanced fibrosis, and significant cost-saving (14, 36). This study influenced national guidelines on non-invasive testing in NAFLD (37).

In ArLD, Thiele et al. (23) reported a large multi-centre prospective study of NIT for liver fibrosis in 289 patients with AUD conducted in primary and secondary-care comparing NIT including the ELF test to liver biopsy as a reference standard. The authors concluded that the ELF test was an excellent diagnostic test, which was effective and safe at a threshold of 10.5 for triaging patients from primary to secondary-care, with a NPV (negative predictive value) of 98% in the primary-care cohort (sensitivity 75%, specificity 97%, AUROC 0.92).

However, although fibrosis testing is now recommended in national guidelines for patients with AUD (1), with BSG recommending ELF or FibroScan in AUD (1) and NICE: FibroScan, (38) this approach is not yet in widespread use in the NHS for people with AUD, and the impact of using such strategies in primary-care in the UK is unknown. Furthermore, it is not known if ELF or FibroScan can be used effectively in people with AUD to improve the detection of advanced fibrosis, if one is 'superior' to the other in terms of diagnostic performance, and if there are other comparable non-invasive tests.

1.3 Rationale for thesis

The importance of testing for advanced fibrosis before a patient develops decompensated cirrhosis or hepatocellular carcinoma is clear. As advanced fibrosis is asymptomatic and only present in around 20% heavy drinkers, it is vital that there is a shift in clinical practice to early detection and prevention in 'at-risk' people with AUD in primary and secondary care and this could be achieved with the use of non-invasive tests. This is not yet in widespread practice in the UK for AUD, and whilst there are several available NIT for liver fibrosis as described above, the optimum choice of test for the diagnostic and prognostic use in ArLD remains open to debate.

As fibrosis is the strongest predictor of prognosis, there is also evidence that NIT can help prognosticate in ArLD (39, 40), but again there is no consensus about which is the superior prognostic score – either for ArLD or alcoholic hepatitis – an

acute severe inflammatory illness that develops after recent heavy alcohol intake in people with AUD.

Whilst the ELF test has been shown to be effective in allowing people at low-risk without advanced fibrosis (ELF score <10.5) to remain in primary-care (23), this strategy has not been evaluated in the UK in terms of efficacy in increasing the detection of advanced fibrosis and cirrhosis, reducing the proportion of 'unnecessary referrals' (those without advanced fibrosis) and whether it is practical and cost-effective.

1.3.1 This thesis aims to address the following unanswered questions:

1) Which non-invasive test performs the best in the accurate detection of fibrosis stages in ArLD? Whilst transient elastography (FibroScan) has been suggested for non-invasive testing, blood tests may be more advantageous in a community setting for practical reasons – are blood-based tests effective in ArLD? Which one performs the best?

2) What proportion of ArLD referrals could be considered 'unnecessary'? Whilst it has been shown that 92% of NAFLD referrals to secondary care are 'unnecessary', in that the referred patients had no evidence of advanced fibrosis or cirrhosis; is this the same for ArLD?

3) What methods are currently being used to aid referral decisions in ArLD to secondary care? Is there scope for optimization of current practice?

4) Can a blood-based fibrosis test be readily adopted in a community referral pathway from primary care to secondary care hepatology clinic for patients with suspected ArLD?

5) What is the proportion of patients seen in hospital with AUD who are thought to *not* have any existing liver disease, that actually have 'occult' advanced fibrosis or cirrhosis when tested? Are there missed opportunities for testing for liver fibrosis in secondary care, and can this be addressed with non-invasive tests?

6) Can non-invasive fibrosis tests be used to predict prognosis in ArLD? Which test performs the best? How do they compare to histology?

7) Latest research has shown prognostic markers in alcoholic hepatitis to be suboptimal. Can the ELF test (a blood-based test for liver fibrosis) be applied in alcoholic hepatitis to predict prognosis? If so, how does it perform, and how does it compare to more traditional prognostic markers such as MELD (Model for End-stage Liver Disease score), and GAHS (Glasgow Alcoholic Hepatitis Score)?

1.3.2 Aims and objectives

The overall aim of this thesis is to investigate the performance and current practice of non-invasive testing to risk-stratify liver fibrosis in ArLD.

Objectives

1. To perform a systematic review, with meta-analysis to investigate the diagnostic performance of four NIT (FIB4, FibroTest, ELF and FibroScan) in ArLD. Specifically:

    i. The diagnostic accuracy of ELF, FibroScan, FibroTest, and FIB-4 in distinguishing advanced fibrosis (equivalent to ≥ METAVIR F3) from patients without advanced fibrosis (<METAVIR F3) in all patients at risk of ArLD, compared with the reference standard liver histology as judged by Area Under Receiver Operator Characteristic curve (AUROC), sensitivity, specificity, positive and negative predictive value

    ii. The diagnostic accuracy of ELF, FibroScan, FibroTest, and FIB-4 in distinguishing cirrhosis (equivalent to ≥METAVIR F4) from patients without cirrhosis (≤ METAVIR F0-3) in all patients at risk of alcohol-related liver disease, compared with the reference standard liver histology as judged by AUROC, sensitivity, specificity, positive and negative predictive value

    iii. The diagnostic accuracy of ELF, FibroScan, FibroTest, and FIB-4 in distinguishing significant fibrosis (METAVIR F2) from patients without any significant fibrosis (METAVIR F0-1) in all patients at risk of alcohol-related liver disease, compared with the reference standard

liver histology as judged by AUROC, sensitivity, specificity, positive and negative predictive value

iv. Numbers of test failures for each non-invasive test

2. To perform a retrospective analysis of referrals from primary care to the hepatology service at the Royal Free hospital over a 3-year period to determine the reasons for referral and demography of patients with AUD referred from primary-care to secondary-care. This will include an exploration of:

i. Reasons for referral to secondary care, and current use of non-invasive fibrosis tests in the community

ii. What proportion of referred patients with AUD to secondary-care have advanced fibrosis ('necessary referrals')

iii. Demographic risk factors for a diagnosis of advanced fibrosis (including BMI, alcohol consumption, smoking status, age, sex, and deprivation score)

iv. Using modelling, to predict the impact of the use of simple blood-based fibrosis scores in predicting a diagnosis of advanced fibrosis.

3. To plan and set-up a new community pathway involving the use of the non-invasive testing for triage of patients with AUD to secondary-care services, following current national guidelines. This will be achieved following:

    i. A comprehensive review of the relevant literature on non-invasive tests to improve detection of chronic liver disease, detailed in Chapters 2 and 3

ii. Close collaboration between primary care local Clinical Commissioning Group leads and secondary care alcohol and hepatology services

iii. Focus groups with patients and the public, hepatologists and health economist, to design the pathway.

4. To plan and conduct a prospective study over 1 year at the Royal Free Hospital involving the use of the ELF test on consecutive patients presenting to hospital that are identified to have an active AUD but not recognised to have liver disease by their admitting team and who were not previously investigated for or known to have liver disease. The objective is to evaluate if there are potentially missed opportunities for the detection of advanced fibrosis in hospital in-patients with AUD. Specifically, this will include an investigation into:

   i. The proportion of patients presenting to hospital and being recognised to have an alcohol use disorder, who had previously undetected advanced fibrosis as determined by the ELF test.

   ii. Which demographic factors are associated with an elevated ELF score (indicating advanced fibrosis), including alcohol consumption, BMI, age, sex, deprivation score and smoking status.

iii. What proportion of patients with AUD in hospital had previously missed opportunities for the assessment and diagnosis of liver disease.

5. To conduct a systematic review to assess the prognostic performance of four NIT (FIB4, FibroTest, ELF, FibroScan) in ArLD. Specifically:

    i. The ability of ELF, FibroTest, FibroScan and FIB4 to predict all cause and liver-related mortality
    ii. The ability of ELF, FibroTest, FibroScan and FIB4 to predict liver-related cirrhotic decompensation events (LRE) including ascites, variceal bleeding, encephalopathy, need for liver transplantation and development of hepatocellular carcinoma (HCC)

6. To collaborate with the 'STOPAH' working group to gain access to stored sera from the 'STOPAH' study (Steroids Or Pentoxifylline in Alcoholic Hepatitis) (41). I then aim to measure and analyse ELF test scores in these trial participants, and use statistical analysis methods (described in Chapter 8) to investigate the prognostic ability of ELF and ELF components, compared to 'traditional' prognostic scores in AH. Specifically, this will include an evaluation of:

i. The performance of ELF, in comparison to and in combination with ABIC (Age-Bilirubin-INR-Creatinine), GAHS, and MELD, in the prediction of survival at 90 days.

ii. The performance of ELF, in comparison to and in combination with ABIC (Age-Bilirubin-INR-Creatinine), GAHS, and MELD, in the prediction of survival at 28 and 120 days.

iii. Any association between ELF and inflammation (assessed by AST and CRP where biopsy data are lacking)

iv. Any association between ELF and episodes of infection or gastrointestinal bleeding.

1.4 Outline of thesis

This thesis begins with a literature review, examining the extent of the problem of ArLD, current trends in alcohol use, and trends in mortality from ArLD, covering the period of time up to starting the research. I will then describe the disease spectrum of ArLD, including alcoholic hepatitis, and the current knowledge of the pathophysiology of alcohol induced injury to the liver.

In this introductory chapter I will introduce the concept of non-invasive tests for liver fibrosis, and conduct a literature review of available tests, and describe the concept behind the value of sensitivity versus specificity, and Positive Predictive Value and Negative Predictive Value of diagnostic tests, and the issues surrounding liver biopsy as a reference standard with which to compare these tests.

The thesis will then move on to include five results chapters, and one chapter describing the set-up of a new community alcohol referral pathway incorporating fibrosis assessment. The five results chapters will report the results from a retrospective evaluation of primary care alcohol referrals over three years, exploring the use of non-invasive tests in triaging patients to secondary care, and the proportion of 'necessary' versus 'unnecessary referrals'. This chapter also includes modelling of simple blood tests to try and improve the proportion of 'necessary' referrals. I will also describe findings from two systematic reviews- one on the diagnostic performance and the other on the prognostic performance of five non-invasive tests in ArLD. In addition, I will describe results from a prospective study investigating liver fibrosis in patients presenting to hospital with AUD who were not thought to have liver disease, using the ELF test. Finally, I will report on the results of the prognostic performance of ELF as a prognostic marker in a cohort of patients with alcoholic hepatitis, in comparison to traditional prognostic scores – namely MELD, ABIC and GAHS.

The thesis will end with a discussion of key findings, the generalisability of the results with implications of the study findings placed in the context of current literature, and an appraisal of areas that warrant future research.

Chapter 2:

Introduction to Alcohol-related Liver Disease and non-invasive tests for liver fibrosis

2.1.1 Alcohol and public health: The extent of the problem

Excessive alcohol consumption is a major global public health problem, with over two hundred diseases and injuries linked to alcohol by the World Health Organisation in 2018 (42), and a reported 3 million alcohol-related deaths per year, corresponding to 5.9% of all deaths (43). Alcohol is now the leading cause of death globally in people aged 15-49 (44).

Alcohol-related liver disease (ArLD) is the most well recognised of the alcohol-related diseases by the general public (45) and indeed is responsible for causing the greatest morbidity and mortality from alcohol (46). Non-liver alcohol-related health consequences must not be forgotten, however, and include hypertension, arrythmias, stroke, cardiomyopathy, pancreatitis, cancers (namely head and neck, oesophageal, breast, bowel, liver), dementia, depression and increased susceptibility to lower respiratory tract infections (47, 48).

In addition to the physical harms of alcohol, alcohol is also frequently responsible for a plethora of social harms to the individual and wider society. These include road traffic accidents, domestic violence, relationship breakdown, homelessness, time off work, and unemployment (49). There are also well-recognised impacts in children of parents who struggle with alcohol, including low self-esteem (50), depression (51), worse educational outcomes (52), drug use (53), and higher risk of alcohol use disorder themselves (54).

It is therefore not surprising that problems related to alcohol impose a substantial financial cost on society, with an estimated £3.5 billion per year spent on alcohol

related conditions in the NHS (55). Whilst it is difficult to get an accurate estimate of the wider financial costs to society, in their 2018 report on the 'Public health burden of alcohol', Public Health England estimated that alcohol cost the UK £47 billion per year, of which 72% was attributable to indirect costs such as loss of productivity, lost working years and unemployment; followed by health care costs (13%), other direct costs including health and social-care costs and costs attributable to welfare systems (12%) and criminality costs (3%) (2). These financial costs do not encompass the intangible costs of alcohol misuse, including the impact on quality of life, costs assigned to pain and suffering of affected individuals and their families, and displaced costs from money spent on alcohol within families (2).

2.1.2 Definition of Alcohol Use Disorder (AUD)

Having now moved away from the stigmatised term 'alcoholic', the newer term 'Alcohol Use Disorder' is the currently recognised umbrella term for problematic alcohol intake, and is the term I will be using throughout this thesis. It is defined by the DSM-5 (Diagnostic and Statistical Manual of Mental Disorders, 5th Edition) (Table 2.1), which no longer explicitly includes the term 'alcohol dependence' although features of this are included within these AUD criteria. As shown in Table 2.1, there are 11 criteria, of which two are required to make a diagnosis of AUD. AUD is graded is mild (2-3 criteria), moderate (4-5 criteria) and severe (6 or more criteria). The WHO (World Health Organisation) continues to use the terms 'hazardous' and 'harmful' alcohol use and 'alcohol dependence' to subcategorise

AUD, and these are incorporated into their most recent definition of alcohol related problems in the ICD-11 (International Classification of Diseases, 11th edition). 'Hazardous' is defined in brief as "a pattern of alcohol use that appreciably increases the risk of harmful physical or mental health consequences to the user or to others to an extent that warrants attention and advice from health professionals" (56).

'Harmful' is defined in brief as "a pattern of alcohol use that has caused damage to a person's physical or mental health or has resulted in behaviour leading to harm to the health of others"(56).

Broadly speaking, in either set of guidelines, the term 'AUD' can be applied when alcohol use is causing a physical or mental health problem to the individual, or impacting on wider social or occupational functioning (57).

NICE recommends the use of the 'AUDIT' score (Alcohol Use Disorder Identification Tool) to screen for AUD, which gives a diagnosis of 'hazardous' drinking with a score of 8-15 and 'harmful drinking with a score of 16-19, and 'possible dependence' with a score of 20 or more (Table 2.2). This terminology in the AUDIT score reflects the WHO classification of AUD with the use of 'hazardous' and 'harmful' drinking behaviours.

Table 2.1 DSM-5 definition of Alcohol Use Disorder (Adapted from the American Psychiatric Association. Diagnostic and Statistical Manual of mental disorders, 5th edition. Philadelphia: APA (58)).

A pattern of problematic pattern of alcohol use leading to clinically significant impairment or distress as manifested by at least two of the following criteria over a 12-month period:	
Presence of at least **2 of these 11** criteria indicates Alcohol Use Disorder (AUD), with severity defined as: **Mild**: 2 to 3 criteria **Moderate:** 4 to 5 criteria **Severe:** 6 or more criteria	Alcohol is often taken in larger amounts or over a longer period than was intended
	A persistent desire, or unsuccessful attempts at cutting down or stopping drinking
	A lot of time is spent on alcohol – acquiring it, drinking it, getting over its after-effects.
	Craving, or a strong desire or urge to drink alcohol.
	Alcohol use or after-effects of alcohol use are often interfering with obligations at work, school or home.
	Continued alcohol use despite it causing trouble with friends or family
	Important social, occupational or recreational activities are given up or reduced because of alcohol use
	Alcohol use has more than once led to situations whilst or after drinking that increased the chances of getting hurt (such as driving, swimming, using machinery, walking in a dangerous area, or having unsafe sex)
	Alcohol use is continued despite knowledge that it is causing or exacerbating a physical or psychological problem
	Alcohol tolerance – defined by either a need for markedly increased amounts of alcohol to achieve the desired effect or a noticeably diminished effect experienced with continued use of the same amount of alcohol
	Alcohol withdrawal – defined by experiencing the characteristic withdrawal symptoms, or requiring further alcohol (or a closely related substance such as benzodiazepine) to relieve or avoid alcohol withdrawal symptoms

Table 2.2. The AUDIT questionnaire (Alcohol Use Disorder Identification Tool)

(From 'NICE CKS: How should I screen for problem drinking? The AUDIT questionnaire. 'Last revised in February 2018. Available from https://cks.nice.org.uk/topics/alcohol-problem-drinking/diagnosis/how-to-screen/#the-audit-questionnaire (59)).

Questions	0 points	1 point	2 points	3 points	4 points
How often do you have a drink that contains alcohol?	Never	Monthly or less	2–4 times per month	2–3 times per week	4+ times per week
How many standard alcoholic drinks do you have on a typical day when you are drinking?	1–2	3–4	5–6	7–9	10+
How often do you have 6 or more standard drinks on one occasion?	Never	Less than monthly	Monthly	Weekly	Daily or almost daily
How often in the last year have you found you were not able to stop drinking once you had started?	Never	Less than monthly	Monthly	Weekly	Daily or almost daily
How often in the last year have you failed to do what was expected of you because of drinking?	Never	Less than monthly	Monthly	Weekly	Daily or almost daily
How often in the last year have you needed an alcoholic drink in the morning to get you going?	Never	Less than monthly	Monthly	Weekly	Daily or almost daily
How often in the last year have you had a feeling of guilt or regret after drinking?	Never	Less than monthly	Monthly	Weekly	Daily or almost daily
How often in the last year have you not been able to remember what happened when drinking the night before?	Never	Less than monthly	Monthly	Weekly	Daily or almost daily
Have you or someone else been injured as a result of your drinking?	No		Yes, but not in the last year		Yes, during the last year
Has a relative/friend/doctor/health worker been concerned about your drinking or advised you to cut down?	No		Yes, but not in the last year		Yes, during the last year

- Low-risk drinking: score of 1–7
- Hazardous drinking: score of 8–15
- Harmful drinking: score of 16–19
- Possible alcohol dependence: score of 20 or more.

Abbreviated alcohol use disorder scores

Whilst the AUDIT score is the current recommended tool for identification of AUD by the WHO, it takes time to complete, and may not always be feasible in time-pressured clinical consultations. There are other alternative abbreviated scoring systems which have been validated for use in clinical practice. The most common of which is the AUDIT-C score, which is a modified, shortened version of the full AUDIT score, consisting of just three questions: 1) How often did you have a drink containing alcohol in the last year? 2) How many drinks containing alcohol did you have on a typical day when you were drinking in the past year? 3) How often did you have six or more drinks on one occasion in the past year? This score is effective at identifying high-risk drinkers (AUROC >0.85 (20)), and it is recommended in national guidelines that if a person scores five points or more in the AUDIT-C, that they should then complete the full AUDIT score (10).

The CAGE score is another commonly used alcohol use disorder scores, consisting of four questions, largely related to features of alcohol dependence. The AUDIT-C score, however, is superior to the CAGE score, with a higher sensitivity for detecting heavy drinking, and better at identifying current harmful drinking patterns (AUDIT-C AUROC 0.87 (SE 0.0075), CAGE AUROC 0.67 (SE 0.018), $p < 0.0001$) (21).

Finally, the FAST score (22) is another alcohol harm assessment tool, which selects four questions from the full AUDIT questionnaire into an abbreviated tool, originally developed for use in Emergency departments, but could be applied to all health care settings. A FAST score of three or more is 'FAST positive', and, as per the

AUDIT-C score, it is then recommended that a full AUDIT questionnaire is completed.

Throughout this thesis, where an alcohol score is needed, the AUDIT-C score is used, as this is the locally used scoring tool in North London community and hospital settings.

2.1.3 Trends in alcohol use and alcohol-related mortalities

Whilst alcohol misuse is a global problem, there is a wide geographical variation in the distribution of alcohol-related morbidity and mortality, and this is closely correlated to the amount of alcohol consumed in each country (62). The WHO European region has the highest rates of alcohol intake in the world with average consumption of 10.9 litres of pure alcohol per person per year, compared to 6.2 litres per person per year globally (62), and correspondingly also has the highest burden of liver disease in the world (63). Geographic differences in alcohol intake correlate with levels of economic wealth, religion and cultural norms. For example, the prevalence of 'current drinkers' in 2016 was 59.9% in Europe compared to 32.2% in Africa (42).

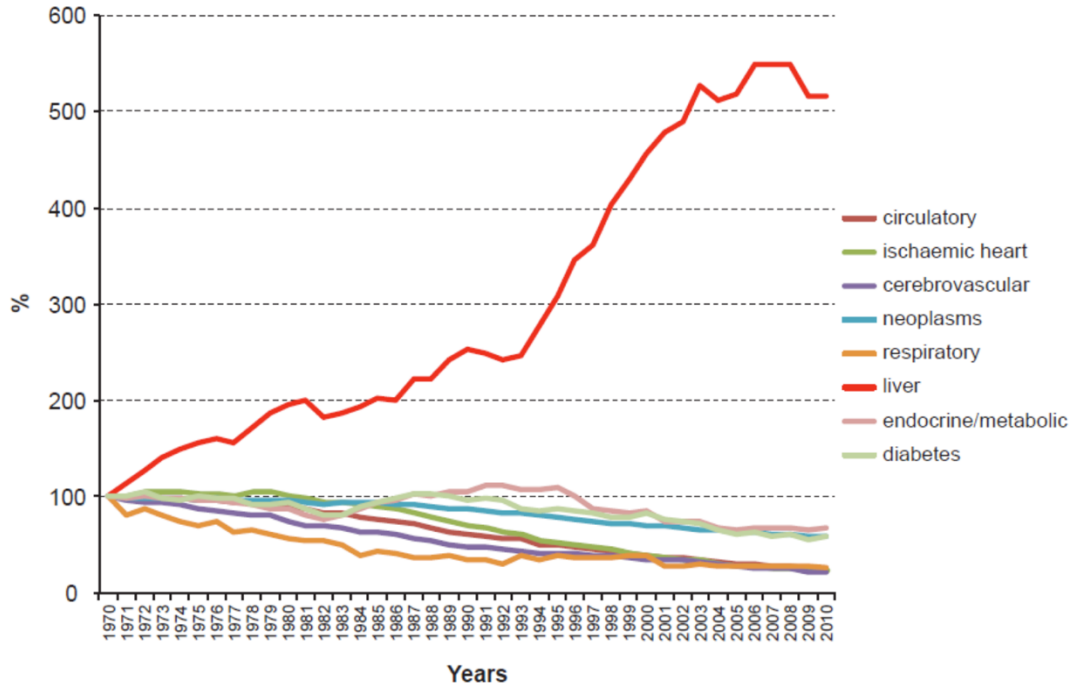
Whilst alcohol use has slightly reduced in the overall population over the past 30 years, mainly attributable to decreases in the wealthier central western European countries, it has increased significantly in eastern Europe, in the UK and in Finland, with corresponding increases in mortality rates from ArLD (64). The World Health

Organisation (WHO) have also acknowledged that the small reduction in alcohol use globally is mainly attributable to a reduction in the percentage of the population who drink by 5% (from 47.6% in 2000 to 43% in 2016), and that those who already drink are consuming higher amounts (42). Total Alcohol Per Capita (APC) consumption trends reveal that the global APC has increased from 5.5 litres of pure alcohol per person per year in 2005 to 6.4 in 2016, and, specifically in drinkers– from 11.1 litres in 2000 to 15.1 in 2016, with further predicted increases projected by the WHO over the next two decades (42). In the years 2020-2021 in particular, the SARS-CoV2 global pandemic and associated lockdowns have affected alcohol intake and I will be exploring this in Chapter 9.

2.1.4 Trends in liver related mortality in the UK

Figure 2.1: Standardised UK Mortality Rates for the commonest causes of death, from 1970-2010

(Reproduced with permission from Elsevier Ltd (Copyright © 2014): Williams R, Aspinall R, Bellis M, Camps-Walsh G, Cramp M, Dhawan A, 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. The Lancet. 2014;384(9958):1953-97 (12).)



(Data were normalised to 100% in 1970, and subsequent trends plotted using the software Statistical Package for the Social Sciences. Data are from the WHO-HFA database (65).)

Figure 2.1 depicts a graph that has been widely published amongst the scientific community, displaying a 400% increase in liver related mortality in the UK between 1970 and 2010, in comparison to other major causes of death such as ischaemic heart disease and cancers for which mortalities have improved. Alcohol is the main cause of this increase in liver mortality, as illustrated in Figure 2.2 and 2.3 (66), contributing to 77% of the increased mortality, compared to mild increase in deaths

from NAFLD (8%), and minimal change in deaths from viral hepatitis (4% increase), or autoimmune/metabolic (5%)(67).

Figure 2.2. Cumulative alcohol related deaths in England and Wales.
(Reproduced with permission from BMJ Publishing Group Ltd (Copyright 2016): Sheron N, Gilmore I. Effect of policy, economics, and the changing alcohol marketplace on alcohol related deaths in England and Wales. BMJ. 2016;353:i1860.) (66)

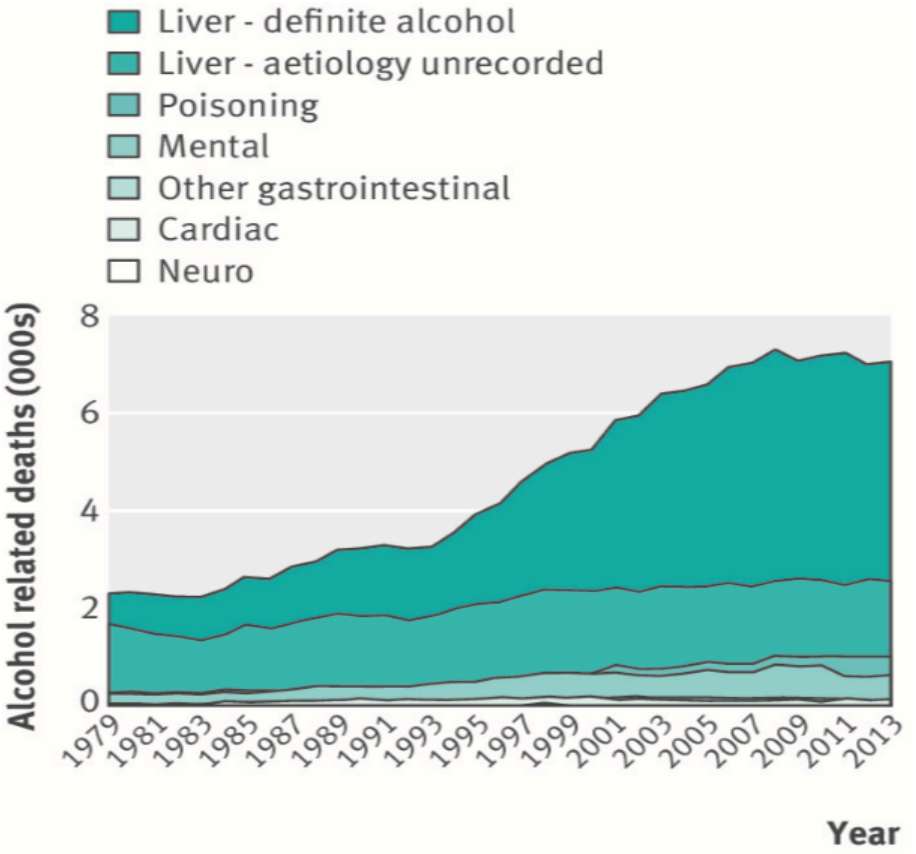
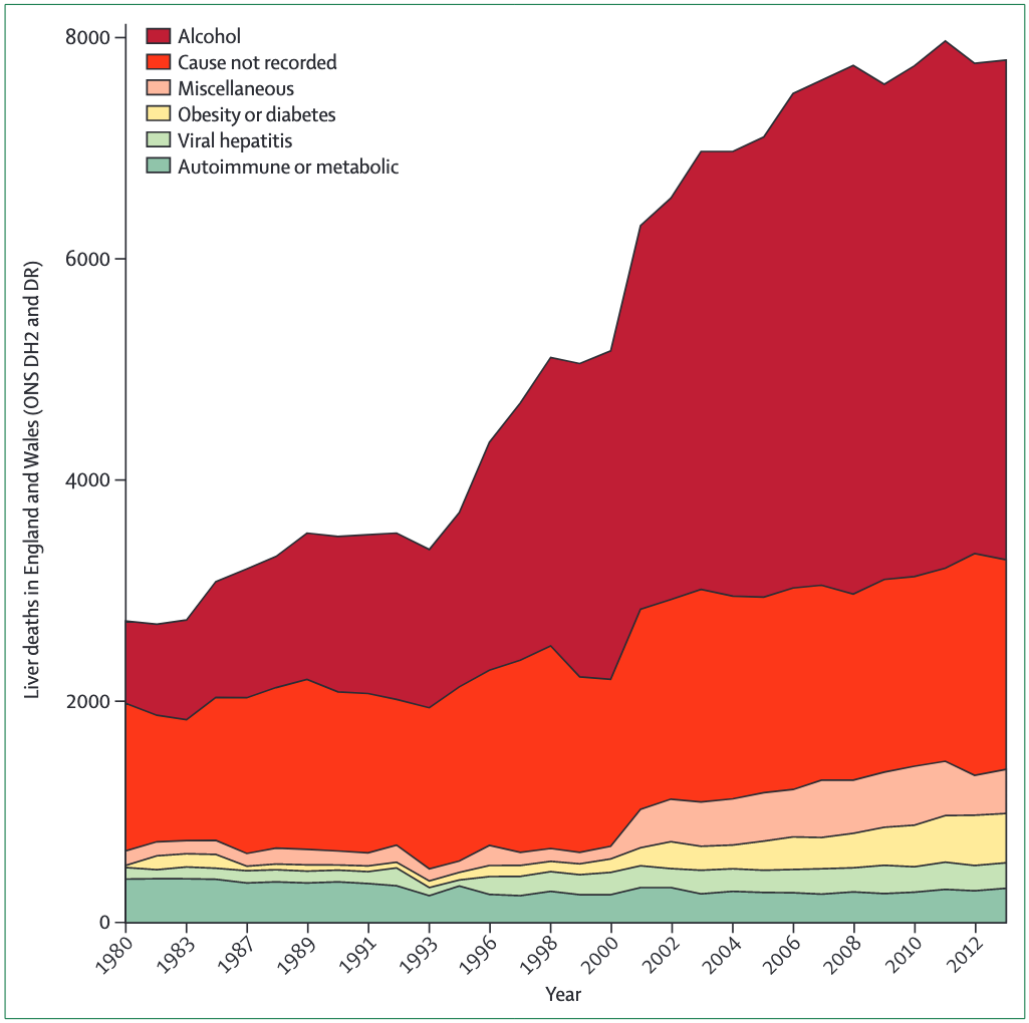


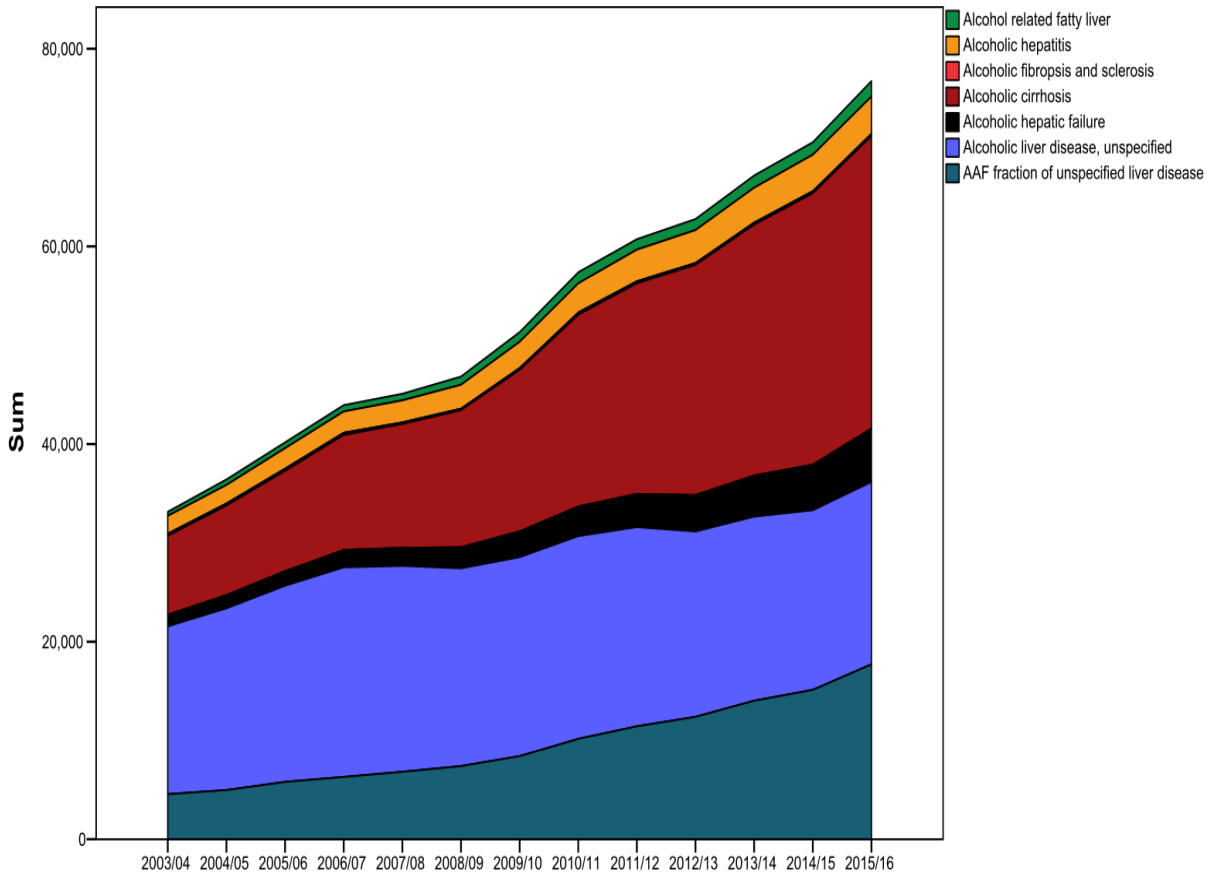
Figure 2.3. Liver mortality in England and Wales, 1980-2013



(Reproduced with permission from Elsevier Ltd. Copyright © 2015: Williams R, Ashton K, Aspinall R, Bellis M, Bosanquet J, Cramp M, et al. Implementation of the Lancet standing commission on liver disease in the UK. The Lancet 2015; 386:2098-111. (67)

More recent mortality data from 2019 shows no sign of improvement in the UK liver mortality rates (68). Liver mortality rates in England have increased by 43% between 2001 and 2019, and admissions to hospital (where alcohol was the main reason for admission) continued to rise by 19% between 2009 and 2019 (46). Figure 2.4, taken from the fourth Lancet commission report on liver disease in the UK (68), displays the increase in hospital admissions over the last two decades that is specifically due to ArLD.

Figure 2.4 Alcohol related Liver Disease hospital admissions in England
(Reproduced with permission from Williams et al, 2018, The Lancet (68), under the creative commons attribution license (CC-BY-NC-ND licence) (https://creativecommons.org/licenses/by-nc-nd/4.0/). Copyright 2017 Elsevier Ltd)



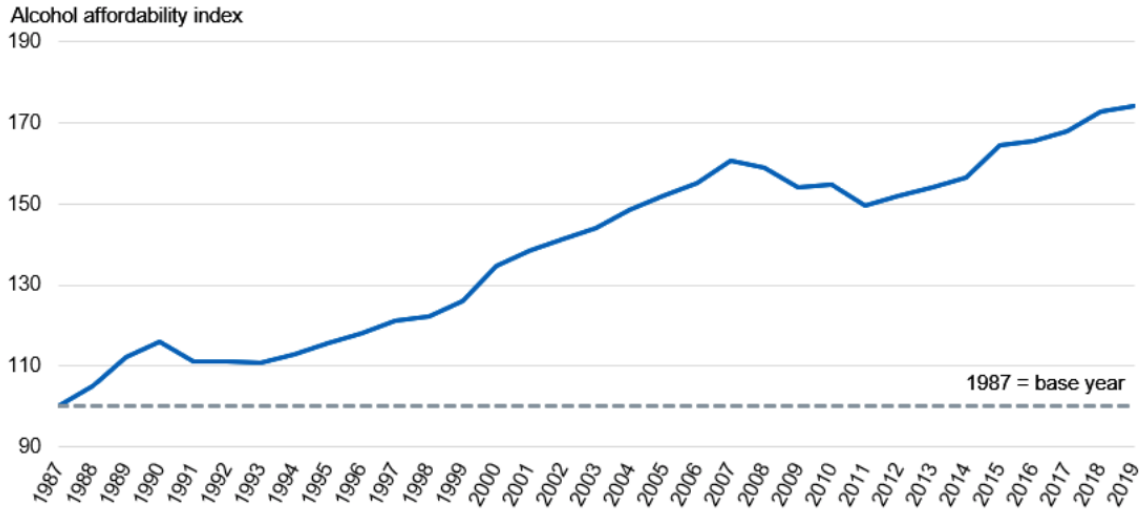
2.1.5 Why have these trends happened?

Reasons for the increase in alcohol intake, and thereby associated increase in liver related morbidity and mortality, are likely to be due to a few key changes in alcohol consumption and policy in the UK. Firstly, despite the well-documented increases in mortality and hospital admissions related to alcohol in the UK, there have been substantial government cuts to community alcohol and addiction services during this time (69). (Specifically, between 2013 and 2019, £212.2 million was disinvested from alcohol treatment services in the UK, which was a 27% decrease in funds) (70).

Secondly, there has been a shift in drinking behaviour from people drinking weak beer in pubs to people choosing stronger beer, ales, wines and spirits to drink in their own homes, with an increase in binge-drinking culture (5). Thirdly, alongside this, the affordability of alcohol has changed significantly over the past few decades, such that alcohol is now 74% more affordable than it was 35 years ago (46). This trend in affordability, as shown in Figure 2.5, corresponds with the increase in liver mortality over the same time period.

There is now good evidence from countries like Canada and Scotland that have introduced minimum unit pricing (MUP) for alcohol, that increasing the cost of alcohol reduces the amount that people drink, and can therefore be a beneficial public health approach in making improvements to the concerning alcohol statistics (71-73). Meta-analyses in 2009 and 2010 of over 50 studies found that a 10% increase in unit price of alcohol was associated with a reduction in per-capita alcohol intake of 4.4%, and a reduction in alcohol related mortality rates of 3.5% (74, 75). Efforts have since been made in Scotland (in 2018) and Wales (in 2020) to introduce MUP of alcohol, with positive preliminary findings showing this to be very successful in reducing alcohol purchases (7.7% reduction in 2020 in Scotland in 2020 when compared with Northern England) with similar preliminary findings in Wales (73)). Data are awaited on the impact of MUP on health outcomes in these countries. However, frustratingly, England remains to be the 'placebo' country with which to compare the data, because, despite the demonstrable success, MUP has not yet been approved by the government to launch in England.

Figure 2.5 Trends in affordability of alcohol in England and Wales
(Reproduced with permission from NHS Digital report 'statistics on alcohol in England, 2020') (46), licenced under the current version of the Open Government Licence).



2.2 How much alcohol is 'too much'?

The relationship between alcohol and health outcomes is complex, and it is still not fully understood (62). Whilst there is a recognised dose-response relationship between alcohol and health outcomes including mortality (63), the nature of this relationship is not entirely clear. For all-cause mortality, there is recognised to be a J-shaped curve (76), whereby moderate drinkers are thought to have a degree of cardio-protection – specifically a lower risk of coronary artery disease. However, as alcohol is known to also cause several cancers, this likely offsets any potential benefit at moderate levels of consumption (77), with the incidence of at least 4 cancers, including breast cancer, starting at a risk as low as one unit per day (78). These findings have led to a recent Lancet report (79), backed up by the Chief Medical Officer of the UK concluding that there is 'no safe level of alcohol

consumption' (80). However, current UK guidelines still advise a limit of 14 units of alcohol per week for men and women. These guidelines, unlike other countries' alcohol guidelines do not take into account the higher risk of alcohol on women's liver health when compared to men's (63). They also do not take into consideration the increasingly recognised impact of overweight and obesity on increased liver risk to alcohol, with double the risk of hepatotoxicity from alcohol in people with a BMI over 35 (81, 82)– so 14 units may actually be more like 28 units in terms of liver harm, for a person living with obesity.

For liver disease, the relationship of alcohol and risk is, perhaps surprisingly, even less well understood than that of overall mortality, with a lack of clarity on whether there is a continuous risk of alcohol on the liver (from zero units upwards), or if there is a threshold effect, above which liver harm occurs (62).

In the 2018 EASL (European Association the study of the liver) commissioned report on the burden of liver disease in Europe (63), it is stated that the relationship between alcohol and liver disease is exponential, with 'extreme drinkers' comprising a significant proportion of people with cirrhosis (63). This is supported by data showing a relative risk of 3 when 20 units per week are consumed, and a relative risk of 30 when 80 units per week are consumed (12, 83). Also, the median alcohol consumption in patients with alcohol-related cirrhosis has been noted to be 120 units per week (84).

Results from Lelbach et al. (85, 86), and Corrao et al. (meta-analysis) (83) found a continuous dose-response curve of increasing liver cirrhosis risk with increasing

alcohol intake and no evidence of a threshold. In contrast, a recent systematic review and meta-analysis in 2020 of over two and half million participants, found that an exponential effect only occurred in women, whereas in men the risk was found to increase beyond a threshold of consumption of one standard drink or more day (87). Other threshold effects have been reported, with the 'million women' population study (Lancet, 2019) finding increased risk of cirrhosis in women at a threshold of drinking 7 or more drinks per week, with incidence of cirrhosis greater in those drinking daily compared to non-daily (adjusted RR 1.61, 95% CI 1.40-1.85, p <0.0001) (88), and double the risk if alcohol was not consumed together with food (88). Other studies have shown a threshold effect at 5 drinks per day (60grams) (89), 25 grams of alcohol per day (90), one drink per day in women (12 grams) and two in men (86). BSG clinical guidelines recommend screening for liver fibrosis in men drinking 50 units or more per week and women drinking 35 units or more per week (1), although the basis for these recommendations is weak (discussed in detail in Chapter 5). Furthermore, it must be noted that the data described in this section reporting alcohol thresholds are all population-based, and do not take into consideration the possibility that there is considerable variation in risk between individuals, based on host factors (section 2.3).

Whilst threshold effects need further investigation, I agree with the Lancet report by Burton et al. (71), that it makes sense to advise zero alcohol if wanting to be sure to avoid any risk to the liver, and overall mortality risk, taking into consideration the risk of even low amounts of alcohol on the incidence of cancers. I will be further exploring alcohol thresholds in two of my studies in Chapters 4 and 5.

2.3 Spectrum of Alcohol-related Liver Disease

The spectrum of Alcohol-related Liver Disease (ArLD) includes fatty liver (steatosis), with potential for progression in some people through to steatohepatitis (including an acute, severe illness 'Acute Alcoholic Hepatitis' (AH)), various stages of fibrosis and finally cirrhosis. Progression from steatosis to fibrosis and then cirrhosis usually occurs in a relatively linear fashion in response to ongoing alcohol excess (91), but AH is a separate entity and can occur at any stage of liver disease (although cirrhosis is prevalent among people with AH). Whilst progression of liver disease was once thought to be forwards-moving only, from steatosis through to cirrhosis, it is now known that it is possible for the fibrosis to regress if an offending agent (e.g., alcohol, high Body Mass Index (BMI), viral hepatitis) can be reduced or removed (92).

Whilst 90% of people with AUD will develop hepatic steatosis (which can develop in as short a time as two weeks of heavy drinking (93)), only a minority (around 20%) will actually progress to cirrhosis (92) (Figure 2.6). Reasons for this are likely to be multifactorial, with genetics thought to be one of the main reasons (2) (implicated genes: PNPLA3, TM6SF2, MBOAT7, HSD17B13 (92)). Other factors include the pattern of drinking (increased liver risk when drinking in a fasted state, drinking every day, and binge drinking (92, 94)), smoking (95), high BMI (81), and female sex (92). In addition, the microbiome is also thought to be implicated in the progression of liver disease, with gut dysbiosis thought to be a contributing factor to the degree of fibrosis and inflammation (91). Moreover, the greater the level of dysbiosis, the

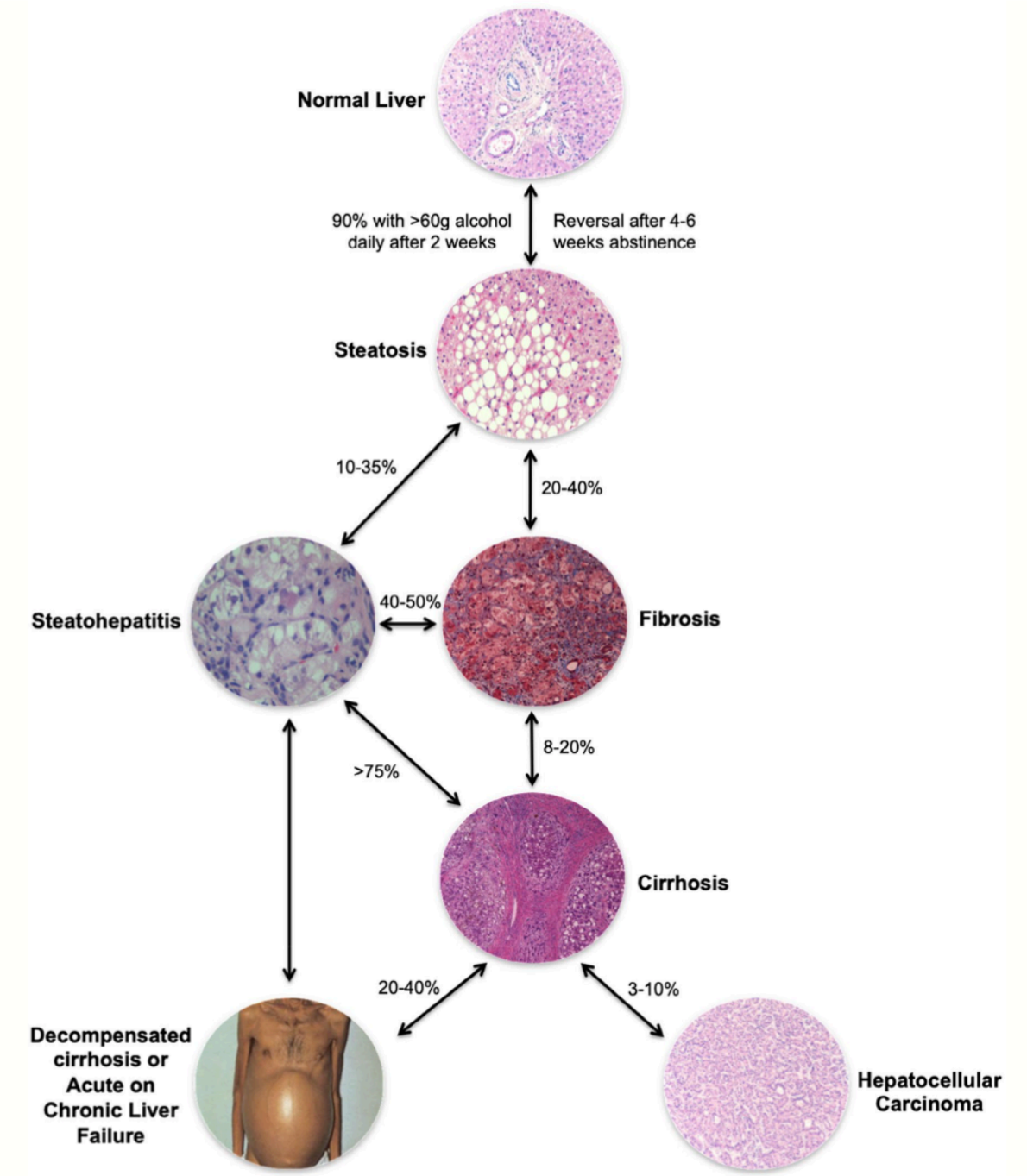
worse the severity of cirrhosis, and progression of cirrhosis in ArLD appears to be associated with reduced diversity of the microbiome (96, 97).

It is important to mention here that socio-economic status is strongly related to outcomes in ArLD (2). People of lower socioeconomic status have worse outcomes (greater alcohol-related morbidity (98) and mortality) (99) when they drink alcohol at the same levels or lower than more affluent people in the same population. In England, it has been shown that as deprivation levels increase, so does the rates of alcohol related mortality (100). This has been termed the 'alcohol harm paradox'. Reasons for this are not fully understood, and likely to be complex, but possibly include the presence of co-morbid conditions in lower socioeconomic groups that may impact on outcomes (2).

A protective factor against progression of liver disease, along with keeping a healthy weight (62), is thought to be coffee consumption, with three to four cups of coffee a day offering some protection to the liver (101-105). Further work is needed to explore whether it makes a difference if the coffee is taken black or white, if sugar is added, and if it is caffeinated or not.

Figure 2.6: The spectrum of alcohol-related liver disease

(Reproduced with permission from Wiley Publishers Ltd (Copyright 2020): Crabb DW, Im GY, Szabo G, Mellinger JL, Lucey MR. Diagnosis and Treatment of Alcohol-Associated Liver Diseases: 2019 Practice Guidance From the American Association for the Study of Liver Diseases. Hepatology. 2020;71(1):306-33(92)



2.4.1 Pathophysiology of ArLD

The initial insult to the liver from alcohol is in the form of hepatocellular steatosis. Whether or not steatosis is in itself a benign condition or not is debated (62). However, it is well accepted that the severity of the ensuing fibrosis in those susceptible (described above) is the greatest factor influencing prognosis (106).

Fibrosis is the hallmark of progressive liver injury. Before detailing the pathophysiology of fibrogenesis, I will first describe the basic functional anatomy of the liver, related to fibrosis.

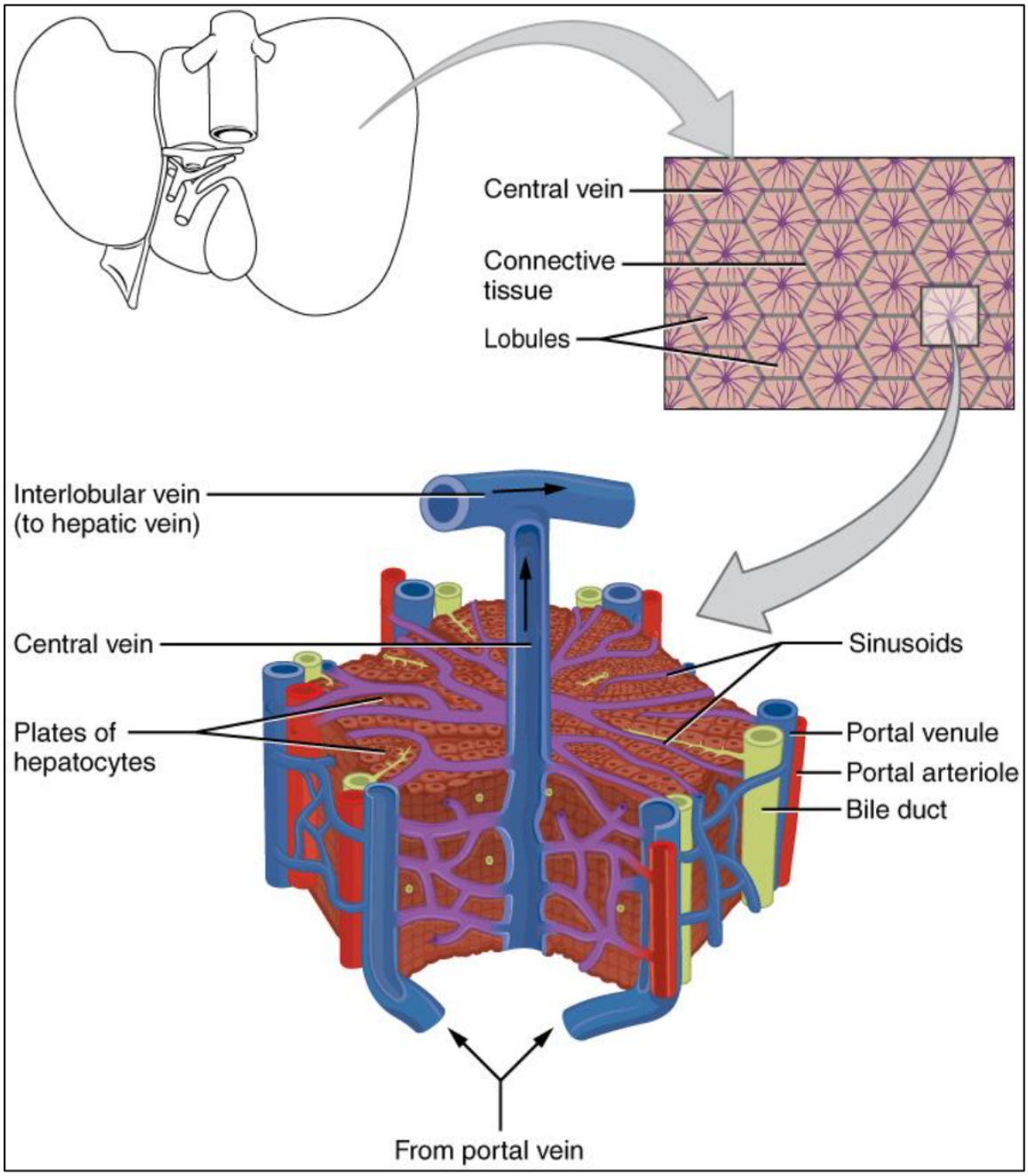
2.4.2 Anatomy and Physiology of the liver, relevant to fibrosis and cirrhosis.

The liver is made up of two lobes, containing 8 functional segments. Classically, these have been described as consisting of thousands of 1-2mm lobules, with the lobules being the functional units of the liver, comprising rows of hepatocytes that radiate out around a central vein connected to a portal tract. The tract comprises the bile duct branch, arteriole of the hepatic artery and venule of hepatic vein (Figure 2.7). A more contemporary view, is the description of acinar, rather than lobules, as the functional units in the liver. The acinus is an elliptical unit containing a central vein at each pole and a portal triad in the centre. It can be described as having three zones, with zone one at the centre around the portal triad, and zone 3 furthest from the porta triad and closest to the central vein.

The hepatocytes are close in proximity to blood-filled sinusoids, which transport blood into the central vein and out of the liver. In between the hepatocytes and

sinusoids is the Space of Disse which consists of blood plasma, and is the location of the Hepatic Stellate Cells (HSC), the major cell type involved in hepatic fibrosis. Endothelial cells and Kupffer cells are also located here, adjacent to the sinusoid. The Kupffer cells are essentially macrophages, that perform scavenger and phagocytic functions to remove small particles, senescent red blood cells, cell debris, and protein and immune complexes from the portal blood flow. They also perform a key role in immune regulatory pathways in the liver, producing pro-inflammatory cytokines including TNFα (Tumour Necrosis Factor alpha), interleukins 1 and 6 and interferon.

Figure 2.7: Representative diagram illustrating the liver anatomy, showing the hepatocytes surrounding the central portal tract.
(Reproduced with permission under the creative commons attribution license (version 4) from Anatomy and Physiology textbook, OpenStax Publisher, 2017 (107) (CC-BY-NC-ND licence, https://creativecommons.org/licenses/by-nc-nd/4.0/)



In a normal liver, the Hepatic Stellate Cells (HSC) are in a ‘quiescent state’, however,

in response to a liver insult from alcohol or other toxins, the HSC become activated

leading to a cascade of events resulting in extracellular matrix production deposition, and collagen secretion, which leads to fibrosis and cirrhosis.

2.4.3 Extracellular matrix

The hepatic extracellular matrix (ECM) is essentially a protein 'scaffolding' of the liver. All three cell types that surround the space of Disse, as described above (Kupffer cells, HSC, and endothelial cells) produce components of the ECM. The most prevalent of the proteins in the ECM are collagens, including type I, III, IV, and V collagen. Other ECM components include glycoproteins (including laminin, fibronectin, TIMP1 (Tissue Inhibitor of Metalloproteinase 1), tenascin, nidogen, and SPARC) and proteoglycans (including hyaluronic acid, heparan, dermatan, perlecan, biglycan, decorin, chondroitin sulphate) (108).

It is important to note that ECM production is not just pathological in fibrogenesis – but in the normal healthy liver, the ECM components are restricted to portal tracts, sinusoid walls and central veins, and serve as a 'scaffold', sustaining the architecture of the liver. They also serve as a signalling network, allowing signalling of hormones and cytokines, etc, between cells that maintains the cell's microenvironment (108). There is a 'homeostasis' of ECM production and degradation that maintains these functions in health and a stable amount of ECM. In liver cirrhosis and fibrogenesis, in response to an insult to the liver such as alcohol, the ECM multiplies (three to five-fold in fibrosis) leading to fibrous deposition and liver scarring. Most of the ECM (that is secreted by the activated

HSC) is made up of type I collagen and type III collagen, which is why direct biomarkers such as the ELF test (an algorithm incorporating the amino-terminal peptide of type III pro-collagen, (PIIINP) is able to detect this pathological fibrosis (109),

2.4.4 Fibrosis and Cirrhosis development in ArLD

Patterns of fibrosis seen in liver histology differ between aetiologies of liver disease. In all causes of liver fibrosis, the perpetrators of fibrogenesis are the HSC. In response to liver injury (for example from ongoing heavy alcohol use), the accumulation of lipid peroxides and apoptotic bodies in the damaged hepatocytes leads to activation of the HSC by profibrogenic cytokines, fibronectin, apoptotic bodies, immune cells and platelets (110). Following this, the HSC transform into profibrogenic myofibroblasts and proliferate rapidly (111). This leads to the accumulation of ECM resulting in fibrous tissue formation or 'scarring'. Apoptosis, tissue hypoxia and ECM signalling pathways maintain the ongoing HSC activation, eventually leading to cirrhosis in those people who are susceptible, due to underlying genetic or lifestyle risks described earlier in this chapter. The pattern of ECM deposition varies with aetiology. Alcohol and NAFLD are characterized by early pericellular fibrosis while viral hepatitis characterised by early periportal fibrosis.

The accumulation of ECM results in distortion of the liver parenchyma, with shrinking of the liver seen in cirrhosis, and increased resistance to sinusoidal blood flow, resulting in portal hypertension. On histological examination of a liver

affected by ArLD, the most typical feature in the presence of steatohepatitis is hepatocyte 'ballooning', often with the presence of Mallory-Denk bodies and neutrophilia, and along with diffuse steatosis. With subsequent scarring, 'bridging' fibrosis can be seen, and the development of cirrhotic nodules (112).

2.5 Complications of ArLD

Fibrosis and cirrhosis are generally 'silent' diseases, in that the sufferer will usually not have any symptoms whilst cirrhosis is in the 'compensated' state. In the first year after diagnosis of compensated alcohol-related cirrhosis, 37.6% will decompensate (95% CI 34.1%-41.5%) (113). This means the development of signs of portal hypertension (namely ascites, Spontaneous Bacterial Peritonitis, variceal bleeding, and encephalopathy), impaired synthetic function (jaundice, coagulopathy), or the development of HCC (Hepatocellular Carcinoma).

2.5.1 Ascites

The development of ascites in chronic liver disease (CLD) is thought to be due to a combination of factors, including increased capillary hydrostatic pressure in the splanchnic bed due to changes in the liver architecture from scarring, leading to fluid being forced in to the peritoneal space. Systemic vasodilation accompanies cirrhosis, leading to reduction in the effective circulating arterial blood volume which triggers activation of the renin-angiotensin system as a homeostatic response

to maintain blood pressure, resulting in renal vasoconstriction, causing salt and water reabsorption, promoting further transudation of fluid into the peritoneal cavity and leading to worsened renal function. Finally, there is increasing recognition of the importance of the contribution of the microbiome to the development of portal hypertension. Portal hypertension results in changes to the gut, with reduced gut motility leading to bacterial overgrowth, and a 'leaky' gut wall, resulting in bacterial translocation to the systemic circulation, triggering a pro-inflammatory cytokine cascade that results in further splanchnic vasodilatation (114).

The development of ascites is an indicator of poor prognosis, with 50% mortality at 3 years (115). It is associated with the development of Spontaneous Bacterial Peritonitis (SBP) and hepatorenal syndrome (HRS). Management is based around no-added-salt diet, diuretics, and either large-volume paracentesis, or Trans jugular Intrahepatic Portosystemic Shunt (TIPS) where indicated. The presence of ascites should prompt assessment for transplant suitability, where appropriate.

2.5.2 Variceal Haemorrhage

As a result of parenchymal changes to the liver in cirrhosis, there is increased resistance between the splanchnic vessels and the right side of the heart which leads to retrograde flow and increased portal pressure. Collaterals develop and over time, slowly enlarge and connect the portal venous system to the systemic circulation. This leads to the presence of tortuous dilated veins or 'varices' in the oesophagus, and stomach, rectum and other points of anastomosis of the portal

and systemic venous systems which can rupture causing life-threatening bleeds. Once diagnosed with cirrhosis, patients should be entered into an endoscopic variceal surveillance programme so that varices can be identified and the risk of bleeding mitigated by non-cardio-selective beta-blockers to reduce portal pressures. The use of endoscopic variceal band ligation as primary prophylaxis for oesophageal varices has been questioned by a recent meta-analysis by Roccarina et al., 2021 (116), with further research needed to determine if band ligation is better than beta-blockers for primary prophylaxis. Alternatively TIPS can be effective management in appropriately selected patients after variceal bleeding (62).

2.5.3 Encephalopathy

Hepatic Encephalopathy (HE) is brain dysfunction as a consequence of liver insufficiency and portal hypertension. It presents as a spectrum of impairment of higher mental function ranging from mild/sub-clinical 'covert' encephalopathy to coma. Overt HE can be classified using the West Haven criteria: Grade 0 (no abnormality), Grade I (trivial lack of awareness, shortened attention span, altered sleep rhythm), Grade II (lethargy or apathy, disorientation for time, personality change, inappropriate behaviour, dyspraxia and asterixis or 'flap'), Grade III (somnolence to semi stupor, confusion, gross disorientation, bizarre behaviour), Grade IV (coma) (117). HE can have significant impacts on sleep, associated impacts on employment with driving restrictions, and detrimental impacts to quality of life. Treatment is supportive, with lactulose (aiming for two to three bowel movements daily to reduce the gut nitrogenous load), and rifaximin. Patients should be assessed for transplantation if suitable.

2.5.4 Hepatocellular carcinoma (HCC)

Whilst the highest incidence of HCC is seen in patients with viral hepatitis, the risk in alcohol-related cirrhosis is still significant – estimated to be 8% at 5 years (118). Pathophysiology is likely related to the development of regenerative nodules in the liver, with progression to small cell dysplasia through to HCC. Screening for HCC is advised for patients with cirrhosis, with 6-monthly ultrasounds being the standard practice. Alpha Fetoprotein (AFP) levels may be elevated, although a low AFP should not rule out presence of HCC (119). Treatment depends on size, number and spread of tumours, with options including liver resection or transplantation, radiofrequency ablation (RFA), trans arterial embolization (TAE) or chemotherapies such as Sorafenib (an oral multi-kinase inhibitor). Treatment may be with either curative or palliative intent, depending on staging (usually staged by the Barcelona 'BCLC' classification) (119). Coffee consumption has been shown to reduce the incidence of HCC in people with chronic liver disease (119).

2.5.5 Jaundice (liver failure)

Decompensated liver disease may present with jaundice, and associated elevated serum bilirubin levels, and coagulopathy and reduced serum albumin levels related to synthetic dysfunction of the cirrhotic liver. Management is based around reversal of any ongoing liver insult such as alcohol intake, screening for other causes of jaundice such as gallstones and HCC/cholangiocarcinoma, screening for sepsis, and ensuring optimum nutritional support.

2.5.6 Acute Alcoholic Hepatitis, (AH)

Alcoholic Hepatitis (AH) is a specific cause of jaundice and must be differentiated from the occurrence of jaundice in cirrhosis. AH is a severe clinical syndrome characterised by new onset jaundice in patients with ongoing alcohol excess, commonly after a recent period of heavy drinking. Specifically, patients present with a bilirubin >50 μMol/L, an elevated AST (50-400 U/L) and AST:ALT ratio >1.5 with no other cause for hepatitis suspected (120).

Whilst cirrhosis is prevalent in AH, cirrhosis does not have to be present, and AH can develop with any degree of pre-existing liver damage. For the diagnosis, heavy alcohol use should have occurred for, at least, the last 6 months, with fewer than 60 days abstinence prior to the onset of jaundice (121). Liver histology is not mandated for the diagnosis of AH, but where obtained it typically shows steatosis, hepatocyte ballooning, with neutrophil infiltration, and sometimes the presence of Mallory-Denk bodies and megamitochondria. Advanced fibrosis or cirrhosis are also present in the majority of patients with AH (62).

The pathophysiology of AH is complex, and still not fully understood. It involves multiple processes that result in an inflammatory cascade, leading to steatohepatitis, resulting from both the intra-hepatic and extra-hepatic effects of alcohol. Intra-hepatically, the metabolites of alcohol (including acetaldehyde) directly induce hepatocyte cell death through apoptosis and necrosis, and this hepatic cell injury leads to the release of Damage-Associated Molecular Patterns (DAMPs) (122). These DAMPs bind to Toll-Like Receptors (TLR) in Kupffer cells in the

liver, leading to an inflammatory cascade (122). The extra-hepatic effects include the effects of alcohol on the intestine, which in turn worsen hepatic failure. This is thought to be through the direct toxic effect of alcohol on intestinal epithelial cells which reduces the expression of 'tight-junction proteins' leading to increased permeability of intestinal mucosa, or 'leaky gut', allowing bacterial translocation across the gut wall, leading to further 'switching on' of inflammatory cytokines and further inflammation of the liver, ultimately resulting in hepatic failure (122-124).

AH carries a high mortality, in the region of 30% at 90 days, with currently no universally effective pharmacotherapeutic option that improves medium or long-term outcomes (125). Nutritional support and abstinence of alcohol are the mainstays of management, with corticosteroids being the only pharmacotherapeutic option currently recommended (122), although this is an area of contention, with differing opinions amongst specialist hepatologists about the benefits versus risks. The STOPAH trial (steroids or pentoxifylline for alcoholic hepatitis) (41) will be described in more detail in Chapter 8, but essentially found that prednisolone demonstrated a trend towards improved mortality at 28 days compared to placebo (14% versus 18%, $p = 0.06$), but no benefits in survival were seen beyond this period. Two recent meta-analyses have also concluded that steroids can improve 28-day outcomes, but not at 90 or 180 days (126, 127). However, the most recent Cochrane meta-analysis (2019) was not able to draw any firm conclusions and recommended further prospective randomised controlled trials (128). The risk of steroids is higher in the presence of infection, bleeding and

renal failure, but in careful selection of patients, there is an argument in favour of its careful use in the absence of any other currently available effective treatment.

Trials are ongoing to investigate new therapies (122, 129, 130). These include drugs such as antibiotics (amoxicillin/clavulanic acid), Anakinra (IL-1R antagonist), and Selonsertib (ASK-1 and MAPK inhibitor) (122).

Prognosis is strongly related to presence of advanced fibrosis or cirrhosis. Liver biopsy is not routinely indicated in AH, and is not without risk, particularly in this cohort of patients with AH who often have coagulopathy and ascites which increases the risk of adverse outcome with biopsy. Therefore, it can be difficult to prognosticate without knowing if there is underlying cirrhosis. Currently used prognostic scores (MELD, ABIC, GAHS, Maddrey's DF) are based on clinical, biochemical and haematological markers, such as age, degree of jaundice, coagulopathy and presence of renal failure. These scoring systems whilst clinically useful, are suboptimal with AUROCs <0.8 in predicting 90-day mortality (131). There is therefore a need for exploration of further prognostic scores. Whilst commonly used in ArLD, non-invasive fibrosis tests such as FibroScan are not currently advocated in AH due to the recognised impact of inflammation on the readings (132, 133). However, the ELF test (Enhanced Liver Fibrosis test) has not yet been evaluated in this condition, but is known to perform well for prognosticating in CLD (134, 135). I will be exploring this further in Chapter 8.

2.6 Impact of late presentation of symptoms of ArLD

The silent nature of fibrosis and compensated cirrhosis contributes to the fact that 75% of patients with ArLD first present to health care providers when they already have 'end-stage liver disease', with complications of cirrhosis such as ascites or variceal bleeding (12). In these situations, it is often 'too late' for any interventions to have a significant impact on health outcomes, with many patients either needing a liver transplant if they meet eligibility criteria, or palliative care. Whilst stopping drinking in late disease leads to dramatic improvements with fibrosis regression on histology and improvements in portal pressure, a third of patients will die before their liver recovers (5, 136). Informing patients of an early diagnosis of liver disease leads to two-thirds of harmful or dependent drinkers to stop drinking (5, 137).

It is therefore vital that improvements are made to tackle the earlier detection of fibrosis in people with AUD, so that the trajectory can be influenced by interventions. These can include providing patients with support from alcohol services to stop drinking (potentially leading to regression of their fibrosis), screening for portal hypertension (with endoscopy) and HCC (with ultrasound) and initiating relevant treatments (beta blockers, variceal band ligation, and potentially curative HCC treatment which can impact on outcomes (15-17, 138)). The earlier detection of fibrosis needs to be proactively sought in patients at risk with AUD. As fibrosis and cirrhosis are usually asymptomatic (until complications develop), and routine liver function blood tests (LFTs) and ultrasound are not sensitive or specific for their detection (1), the use of specific fibrosis tests is needed. In fact, as many as

90% of people with early alcohol-related fibrosis, and 75% with severe fibrosis have normal LFTs (standard liver function tests) (139, 140).

As liver biopsy is not routinely indicated in ArLD, and there are now a variety of non-invasive options for fibrosis testing, it is imperative that these non-invasive tests, (including Enhanced Liver Fibrosis test (ELF), FibroScan, FibroTest) are evaluated for their use in routine practice to enable proactive investigation for fibrosis in people with liver disease risk factors like AUD.

2.7 Detection of fibrosis and cirrhosis

2.7.1 Liver biopsy – the imperfect 'gold standard'

Whilst not routinely advocated for use in ArLD, except where clarification of aetiology is required (62), liver histology remains to be the reference standard for staging liver fibrosis, with which non-invasive tests are measured against. It is known, however, to be an 'imperfect' reference standard due to its recognised limitations. A liver biopsy is an invasive procedure that can be performed either trans-abdominally (percutaneously) or via a trans-jugular approach, or less commonly via a surgical procedure, typically at the time of liver transplantation. It is not without risk, even in experienced liver centres, with reported complications including pain (in 84%, including 'mild pain' (141)), bleeding and pneumothorax (or haemothorax) in around 2% (142). Other complications include perforation of intra-abdominal organs, bile leak, infection (bacteraemia, abscess, sepsis)(143), and

rarely ventricular arrythmia with trans jugular biopsy (144). Mortality rates from liver biopsy are quoted at 1: 10,000 (143). The risk of complications is increased by the presence of ascites, coagulopathy, obesity, and influenced by operator experience, use of image guidance, number of needle passes, and ability of the patient to cooperate/stay still during the procedure (143).

Careful selection of patients for liver biopsy is therefore paramount, with due consideration for the information required from biopsy (would non-invasive assessment be an alternative?) and optimisation of risk, for example coagulopathy/platelet correction where indicated (trans-jugular is preferred route in coagulopathy or thrombocytopaenia), using image-guidance, reducing the number of needle passes if possible.

Aside from the complications of the biopsy procedure itself, the results may be inaccurate due to sampling bias, with only 1/50,000 of the liver being sampled leading to the severity of fibrosis being over- or under-estimated in as many as 20% of liver biopsies (143). In addition, it is recognised that the reporting accuracy of biopsies can be affected by inter and intra-observer variability (145). Ideally the histological specimens should be examined by pathologists who are experienced in liver disease, and in conjunction with the clinician who knows the patient to provide clinical context. In the absence of this, diagnostic errors by non-specialist pathologists at an academic centre have been observed to occur in over 25% of patients, and so liberal use of second opinion among histopathologists is encouraged (146, 147).

To maximise reliability and accuracy of liver biopsy, it is important that the length of the biopsy specimen is at least 15mm, and includes a minimum of six portal tracts (38). Quality of the biopsy specimen, in particular biopsy length, is considered one of the most important factors that could cause bias in the interpretation of histology results (148).

2.7.2 Histological staging tools for the grading of liver fibrosis severity

The histological progression of chronic liver disease evolves from no fibrosis, to fibrous portal expansion, then bridging fibrosis, ending with cirrhosis (149). Several different scoring tools exist for the classification of these stages of liver fibrosis. The METAVIR score is commonly used for all liver disease aetiologies, although it was originally designed for use in hepatitis C (150). Aside from METAVIR, other commonly used staging tools include Brunt (151), Kleiner (152), Scheuer (153), and Ishak (154). Like METAVIR, Ishak was also designed for use in viral hepatitis, whilst Brunt, Kleiner and Scheuer were created for NAFLD (Non-alcoholic Fatty Liver Disease). There is currently no 'universally accepted' staging tool for ArLD (106). In fact, there is a notable absence of any official fibrosis staging system that has been developed by consensus and validation for routine use and inter-intra-observer variability specifically for ArLD (106, 155). In 2006, a 7-tier staging system for the assessment of fibrosis severity in ArLD was proposed by Yip et al. (156), but this has not been validated or applied in routine clinical practice (157). Despite its viral hepatitis origin, METAVIR continues to be the most commonly used staging tool in ArLD (148, 158-164). As NAFLD and ArLD are more similar to each other histologically, it has been suggested by some authors that it would be preferable to

adopt Brunt, or other existing NAFLD staging tools, for use in ArLD (157, 165, 166). Lackner et al. have questioned this, however, highlighting that there are both clinical and morphological differences between NAFLD and ArLD that, in their opinion, affects the reliability of using NAFLD fibrosis histological scoring systems in ArLD to translate into prognosticating for clinical outcomes (157). Instead, the authors advocate for the validation of an alcohol-specific histological staging tool. In any regard, it makes sense that, given the choice, a NAFLD fibrosis scoring system such as Brunt would be preferable for use in ArLD rather than one of the scoring systems developed for viral hepatitis, but this is not the current practice. The continued widespread use of METAVIR in ArLD was justified by Pavlov et al.'s meta-analysis on the staging of ArLD with FibroScan, by highlighting the lack of an existing staging system for ArLD, and that METAVIR is the most widely used scoring system (167).

On direct communication with a specialist liver histopathologist at the Royal Free, Dr Jennifer Watkins, her opinion on this was that it does not matter too much which scoring system is used to detect advanced fibrosis (METAVIR F3 equivalent) or cirrhosis (METAVIR F4) as features of these are generally very similar or the same across aetiologies. It is the differentiation of the early fibrosis stages (for example, discerning mild fibrosis (METAVIR F1) from significant fibrosis (METAVIR F2), where, in her opinion, it is more important to use a staging tool relevant to the aetiology.

However, until an alcohol-specific staging system is validated, I think that it is likely that METAVIR will continue to be widely used in ArLD. Full details of the commonly used staging tools can be seen in Table 2.3.

Table 2.3: Classification of fibrosis stages from four commonly used staging tools

Brunt (151)	**METAVIR** (150)	**Kleiner** (152)	**Ishak** (154)	**Sheuer** (153)
F0: No fibrosis	F0: No fibrosis	F0: No fibrosis	F0: No fibrosis	F0: No fibrosis
F1: Zone 3 perisinusoidal/pericellular fibrosis (focal or extensive)	F1: Stellate enlargement of portal tracts without septa formation	F1: Mild/moderate zone 3 perisinusoidal fibrosis, or portal fibrosis only	F1: Fibrous expansion of some portal areas, with or without short fibrous septa	F1: Enlarged fibrotic portal tracts
F2: Zone 3 perisinusoidal/pericellular fibrosis with focal or extensive periportal fibrosis	F2: Enlargement of portal tract with rare septa formation	F2: perisinusoidal and portal/periportal fibrosis	F2: Fibrous expansion of most portal areas, with or without short septa	F2: Periportal or portal-portal septa, but intact architecture
F3: Zone 3 perisinusoidal/pericellular fibrosis and portal fibrosis with focal or extensive bridging fibrosis	F3: Numerous septa without fibrosis	F3: Bridging fibrosis	F3: Fibrous expansion of most portal areas with occasional portal to portal bridging	F3: Fibrosis with architectural distortion, but no obvious cirrhosis
F4: Cirrhosis	F4: Cirrhosis	F4: Cirrhosis	F4: Fibrous expansion of portal areas with marked bridging (portal to portal as well as portal to central)	F4: Probable, or definite cirrhosis
			F5: Marked bridging (portal to portal and/or portal to central) with occasional nodules (incomplete cirrhosis)	
			F6: Cirrhosis, probable or definite	

In application to clinical use, fibrosis stage is usually simply referred to as: F4 'cirrhosis', F3 'advanced fibrosis', F2 'moderate' or 'significant' fibrosis, F1 'minimal' or 'mild' fibrosis and F0 'no fibrosis' in keeping with the Brunt, METAVIR, Kleiner, and Scheuer scales, but not Ishak which is a 6-point scale.

It is important to note that whilst these scores are categorical, liver fibrosis progression is a continuous process, therefore F4 is not 'double the fibrosis' compared to F2. Non-invasive tests for fibrosis, which I will cover next in section 7.1, are arguably more flexible in this regard, with ability to produce continuous scores that would particularly enable more accurate information to the clinician about interval changes in the fibrosis severity.

The most clinically useful stages to detect are advanced fibrosis and cirrhosis. At the advanced fibrosis stage, there is opportunity for intervention to reduce risk of further progression, and potential for fibrosis regression with stopping alcohol in the case of ArLD, and other potentially modifiable changes such as stopping smoking, drinking coffee, and optimising nutrition and weight (77). It is also clinically useful to detect cirrhosis, so patients can be plugged into regular surveillance for varices and HCC.

To overcome the recognised limitations of liver biopsy, there is increasing interest in the development and use of non-invasive approaches for assessing liver fibrosis to improve earlier diagnosis, to initiate management plans and monitor response. Whilst there are several available non-invasive tests for liver fibrosis (both blood

biomarkers and physical tests such as elastography), the optimum test of choice for the diagnostic and prognostic use in ArLD remains in debate.

2.7.3 Non-invasive fibrosis tests: What makes a good test?

In order to answer this question, this section will cover the basic statistical principles needed to evaluate the diagnostic performance of non-invasive tests.

2.7.4 Sensitivity, specificity, Negative Predictive Value (NPV), Positive Predictive Value (PPV)

Sensitivity is the 'true positive rate', that detects the proportion of the population who have the disease in question. If a non-invasive test is sensitive, a negative result (i.e., below a particular threshold), will accurately rule out the disease. A popular way to remember this is 'SNOUT' (**S**ensitive test when **N**egative rules **OUT** the disease.

Specificity is the 'true negative rate', that detects the proportion of the population who do not have the disease in question. If a non-invasive test is specific, a positive result (above a particular threshold) will accurately rule in the disease ('**SPIN**' **SP**ecific test when positive rules **IN** the disease).

Sensitivity and specificity can be calculated by creating a 2 x 2 table (see example in Table 2.4 below) and using the formulas: Sensitivity = (TP/TP +FN), and Specificity = (TN/FP + TN).

Table 2.4: 2 x 2 contingency table for test result and disease state

		Disease state	
		Positive (disease present)	Negative (No disease)
Test result	Positive (above a defined threshold	True Positive (TP) n=	False Positive (FP) n=
	Negative (below a defined threshold)	False Negative (FN) n=	True Negative (TN) n=

It is important to note that the 'SPIN' and 'SNOUT' rules only work if disease prevalence and SE (standard error) are taken into consideration. For example, in a high-prevalence population, with high pre-test probability of disease, a sensitive test is less likely to be able to accurately rule out the disease than in a low-prevalence population, and conversely in a low-prevalence population, a specific test is less likely to accurately rule in a disease than when in a high-prevalence situation.

It is essential to bear this in mind when designing a new non-invasive test, and defining the threshold for a positive or negative test. If the disease prevalence is low, and the intention of the test is for screening, or 'ruling out' a disease, then a threshold that produces a highly sensitive test should be used. If the disease prevalence is high, or there is a high pre-test probability (for example, a patient with AUD has been referred to secondary care by their GP for testing for liver

damage), then the test should be designed with a threshold that will produce a highly specific test to 'rule in' the disease.

Two thresholds can be used for the same test – a low threshold below which rules out a disease in a low prevalence population (for example, in primary care as a screening test), and a separate high threshold for the same test, above which can be used to rule in the disease in a situation with higher pre-test population and disease prevalence. This leaves a 'grey area' in the middle, in which a different test is sometimes used. (For example, in NAFLD, a successful published pathway for detection of liver fibrosis using non-invasive tests advocates the use of one test (FIB4 blood test) in the first instance, at a low threshold to rule out advanced fibrosis, and high threshold above which prompts referral to secondary care. In the middle of these two thresholds, a second test (ELF test) is recommended (168). This method has been shown to be cost-effective (36) and an effective way to select out the patients with fatty liver who have advanced fibrosis and need referral to secondary care, and those which are low risk for advanced fibrosis and do not need referral.

Positive and Negative predictive values (PPV and NPV) are alternative ways to describe the performance of tests- PPV is the proportion of a population with a positive test result that actually has the disease. NPV is the proportion of the population with a negative test that actually does not have the disease. NPV and PPV require knowledge of the disease prevalence and test performance (sensitivity and specificity) for their calculation, with the equations:

$$\textbf{PPV} = \frac{\text{sensitivity x prevalence}}{\text{sensitivity x prevalence + (1 – specificity) x (1-prevalence)}}$$

$$\textbf{NPV} = \frac{\text{specificity x (1 - prevalence)}}{\text{(1 – sensitivity) x prevalence + specificity x (1 - prevalence)}}$$

When a non-invasive test is applied to a population with a lower disease prevalence than the one in which the test was validated, the PPV would decrease.

Conversely, NPV improves when the population prevalence is lower.

When a test is applied in a higher prevalence population, one would expect the NPV to decrease, and PPV to increase.

2.7.5 AUROC (Area under receiver operator characteristic curve)

AUROCs are commonly used to measure the diagnostic performance of a non-invasive test, using a continuous measure such as an ELF or FibroScan score, compared to the binary outcome in question – for example presence or absence of advanced fibrosis, or cirrhosis. AUROC ranges from 0 (a test with 0% specificity and sensitivity) to 1 (a test with 100% specificity and sensitivity), with clinical interpretation of the AUROC score shown in Table 2.5. ROC curves plot sensitivity against 1-specificity for all possible values of a non-invasive test as a continuous variable. AUROC plots can be used to determine the optimum threshold of a

particular test. If, as described above, two thresholds are desired - then the plot can be used to select the test value (for example ELF score or FibroScan reading) for the 'low threshold' that gives a high sensitivity (for example 90% sensitivity), and the test value for the 'high threshold' that gives a high specificity (for example 90%).

Alternatively, the 'Youden index 'can be used, which derives the point on the ROC curve, and associated test value, that maximises both sensitivity and specificity. It is computed by calculating 'sensitivity + specificity -1' for each value on the curve, then the highest value is the Youden's index 'J', which can be corresponded to the test value to determine the optimal threshold. This 'optimum threshold' is therefore a trade-off between the sensitivity and specificity.

Table 2.5: AUROC interpretation

AUROC	Interpretation of diagnostic test ability
> 0.90	**Excellent**
0.80-0.90	**Good**
0.70-0.80	**Fair**
0.60-0.70	**poor**
0.50-0.60	**Fail**
< 0.50	**Worse than chance**

2.8 Considerations when using AUROC for a test compared to liver biopsy as the reference standard

2.8.1 Inaccuracies of liver biopsy

Firstly, it must be noted that AUROC is ideally used for tests capable of achieving 100% sensitivity and specificity, but, as described earlier, liver biopsy is not a perfect reference standard – affected by sampling bias (only a tiny proportion of the liver is sampled), and inter and intra-observer variation in the histological

interpretation. Therefore, when there is discordance between the severity of fibrosis detected by a non-invasive test and that of the biopsy result, the error may be with *either* test. There is now extensive data showing that liver biopsy is only accurate for staging in 80-90% of patients, with the remainder having up to 1 stage difference in accuracy (169-174). Mathematical modelling by Mehta et al. found that, assuming the sensitivity and specificity of liver biopsy is 90% (for detection of significant liver fibrosis), and the population prevalence is 40%, a 'perfect' non-invasive test with AUROC of 0.99 could only achieve AUROC of 0.90 versus liver biopsy. Afdhal et al. report similar findings from modelling, suggesting that liver biopsy has an accuracy of 80-90%, and any non-invasive tests compared to liver biopsy could not achieve a higher AUROC than 0.90, and results more likely to lie between 0.75-0.88, with most likely value 0.85. Strikingly, Bedossa et al. compared results from 10,659 virtual liver biopsy samples from 17 surgical liver sections, finding AUC increased significantly with length of biopsy, and that even with a biopsy length of 25mm (above the usual recommended length of 15mm in clinical trials), only 75% of the biopsy specimens correctly staged the fibrosis according METAVIR. This reduced to 65% for 15mm biopsy specimens (169). Therefore, even if a non-invasive test is a 'perfect' match for histology, it may be impossible to prove this (173).

2.8.2 Dealing with spectrum bias

A 'spectrum effect' can occur when fibrosis stages differ between the sample and reference populations. For example, if one population consists of patients with predominantly F0 or F4 fibrosis, the non-invasive test AUROC for the accuracy in

differentiating between mild (<F2) and advanced (F3-F4) would be greater than the AUROC in a population where most patients had F1 and F2 fibrosis, or F2 and F3. There is therefore the risk of a type I or type II error occurring when comparing non-invasive test performance between two different patient populations (175).

In addition, AUROC analysis necessitates dividing a population into a binary outcome – for example advanced fibrosis (F3 or above) versus no advanced fibrosis (<F3), but is being compared to liver histological staging which is usually five stages (F0-4 METAVIR/Brunt, etc).

To overcome these potential biases, two statistical methods have been proposed when comparing non-invasive test performance between two different populations. The first is called 'DANA' –with a formula proposed by Poynard et al. (176) to standardise the AUROC based on a regression equation linking observed AUROC with the DANA (difference between mean fibrosis stage of advanced fibrosis (F2, F3, F4) minus the mean stage of non-advanced fibrosis (F0, F1, F2). The equation is: [adjusted AUROC = observed AUROC + 0.1056 (2.5 – observed DANA) (177). A drawback of this is that it has not been validated outside the use of FibroTest or viral hepatitis.

The second method is called the 'Obuchowski' method, which calculates a weighted AUROC ('ordROC') to more appropriately compare non-invasive tests to the ordinal histological staging (e.g., METAVIR F0-F4) to account for the spectrum effect. It is essentially a 'multinomial' version of AUROC, with pair-wise comparisons being performed between the fibrosis stages, which are weighted to account for distance

between the stages, and a penalty function for misclassifying (134). The final result of the ordROC is analogous to the AUROC in that it ranges from 0 to 1, with 1 being the best test.

2.8.3 Introduction to non-invasive tests in ArLD

To overcome the limitations with biopsy, there has been a wealth of interest in the development and use of non-invasive approaches over the last decade for the assessment of liver fibrosis to improve earlier diagnosis, initiate management, and monitor response. Whilst there are several available non-invasive tests for liver fibrosis (both blood biomarkers and physical tests such as elastography) (Table 2.6), the optimum test of choice for the diagnostic and prognostic use in ArLD remains in debate. In Chapters 3 and 7 I will investigate this in detail through two systematic reviews. This next section will describe the breadth of currently available tests.

Table 2.6: Overview of non-invasive methods of liver fibrosis evaluation

Indirect serum markers/panels	Direct serum markers/panels	Patented serum panels	Imaging modalities
AST:ALT ratio	Hyaluronic acid	FibroTest	Transient elastography (FibroScan)
APRI	PIIINP	ELF	MR-elastography
FIB4	TIMP1	Hepascore	Acoustic radiation Force impulses (ARFI)
Forns index	MP3	Fibrospect	2D-SWE (Sheer wave elastography)
Gamma-GT: Platelet ratio	Microfibril-associated glycoprotein4 (MFAP-4)	Fibroindex	
Age-platelet index	laminin	Fibrometers	
Lok index	Metalloproteinases (MMP)- 1 and MMP-2		
	Transforming growth factor-β1 (TGF-β1)		
	PGAA index		

(AST: aspartate transaminase, ALT: alanine transaminase, APRI: aspartate transaminase to platelet ratio index, MR: magnetic resonance, CT: computed tomography, PIIINP: -terminal propeptide of type III procollagen, TIMP1: tissue inhibitor of matrix metalloproteinase 1), PGA: a2alpha-2-macroglobulin, prothrombin time, gamma-glutamyl transpeptidase, apolipoprotein A1)

Individual equations for the marker panels are shown in Table 2.8.

2.9 Serum markers of fibrosis

Liver fibrosis tests include both direct and indirect ('simple') biochemical markers of fibrosis. Table 2.7 summarises the performance of key biomarkers for the detection of advanced fibrosis.

Table 2.7: Diagnostic performance of serum biomarkers for F3 (advanced fibrosis) in ArLD

Serum biomarkers	Cut off	AUROC (95% CI)	Sensitivity % (95% CI)	Specificity % (95% CI)	Reference
Advanced fibrosis (F3)					
FIB4	≥ 3.25	0.85 (0.80-0.90)	58 (45-70)	91 (86-93)	(23)
APRI	≥ 1.0	0.80 (0.74-0.86)	38 (26-51)	90 (85-93)	(23)
AST: ALT	≥ 1.0	0.76 (0.69-0.82)	85 (74-93)	46 (39-52)	(23)
Forns' index	≥ 6.8	0.86 (0.81-0.91)	71 (59-82)	89 (84-93)	(23)
GGT-to-platelet ratio	≥ 0.32	0.80 (0.75-0.85)	88 (78-95)	64 (57-70)	(23)
Age-platelet index	≥ 6.0	0.81 (0.75-0.88)	65 (52-77)	85 (79-89)	(23)
Fibrometer	n/r	0.88 (0.80-0.95)	-	-	(178)
Hepascore	n/r	0.83 (0.74-0.93)	-	-	(178)
FibroTest	≥ 0.58	0.88 (0.84-0.92)	67 (54-78)	87 (82-91)	(23)
ELF	≥ 9.8	0.92 (0.89-0.96)	89 (79-96)	91 (86-94)	(23)
ELF	≥ 10.5	0.92 (0.89-0.96)	79 (67-88)	91 (86-94)	(23)
HA	55.5 mcg/l	0.76	83	69	(179)
TIMP1	n/r	0.68	-	-	(180)
PIIINP	16 ng/ml	n/r	71	50	(181)

Table 2.8 Formulas for combination serum marker panels

Biomarker panel	Formula	Ref
Age-platelet index	1 point for each age-decade: 30-40, 40-50, 50-60, 60-70, ≥70years. 1 point for each platelet count interval: 225-200, 200-175, 175-150, 150-125, <125	(23)
ALT/AST ratio	ALT (IU/L)/AST (IU/L)	(23)
APRI score	AST (IU/L)/ULN X 100/platelet count (10^9/L)	(23)
FIB4	(Age $_{years}$ x AST (IU/L) / (Platelets 10^9/L x (√ (ALT IU/L))	(23)
Forns index	= 7.811 −3.131 × ln platelet count G/) + 0.781 × ln γGT IU/L + 3.467 × ln age years −0.014 × cholesterol g/l	(23)
Gamma-GT: Platelet ratio	(GGT/ULN)/platelet count (10^9/L)	(23)
Lok index	-5.56-0.0089*platelet count + 1.26 * AST/ALT + 5.27 * INR	(182)
FibroTest	Patented formula of α2macroglobulin, γGT, apolipoprotein A1, haptoglobin, bilirubin, age and gender	(183)
ELF	2.278 + 0.851 ln (HA ng/mL) + 0.751 ln (PIIINP ng/mL) + 0.394 ln (TIMP1 ng/mL)	(29)
Hepascore	(−4.185818 − (0.0249 × age) + (0.7464 × 1 if male, 0 if female gender) + (1.0039 × α2 macroglobulin) + (0.0302 × hyaluronate ng/mL) + (0.0691 × bilirubin) − (0.0012 × γGT))	(184)
Fibrospect	Exp [(-4.3633 + (0.0108 * HA ng/ml) + (0.0015 * TIMP1 ng/ml) + (0.53357 * A2M mg/ml)] / 1 + / Exp [(-4.3633 + (0.0108 * HA ng/ml) + (0.0015 * TIMP1 ng/ml) + (0.53357 * A2M mg/ml)]	(185)
Fibroindex	1.738 – 0.064* platelet count + 0.005 * AST + 0.463 * gamma globulin	(186)
Fibrometer	−0.007 × platelets G/l −0.049 × prothrombin time % + 0.012 × AST IU/L + 0.005 × α2 macroglobulin mg/dl + 0.021 × hyaluronate mg/l −0.270 × urea mmol/l + 0.027 × age years + 3.718	(187)

2.9.1 Indirect markers

Indirect markers reflect the consequences of hepatic function and disease rather than having a direct involvement in liver fibrosis. They include the AST (Aminotransferase) to Platelet Ratio Index (APRI) (22), FIB-4 score (23), Forns' index (comprised of platelet count, age, gamma GT, total cholesterol) (23), AST:ALT ratio (24), gamma GT: platelet ratio, and age-platelet index (25). Indirect markers are less frequently used and studied in ArLD, compared to other aetiologies such as viral hepatitis and NAFLD, but I will be exploring some of these markers further in Chapters 3, 4, 6, and 7.

An advantage of indirect markers is that they are easily calculable from blood test results that are usually performed routinely (apart from possibly cholesterol, in the Forns index). As the blood tests are usually readily available, there is minimal, if any, additional cost for these indirect markers. Indirect markers have been shown to perform well in accurately detecting presence or absence of advanced fibrosis or cirrhosis in ArLD, but less well in classifying intermediate stages from each other, nor early fibrosis changes (188).

FIB4, APRI and Forns' index are the most commonly referenced indirect markers.

FIB4 is comprised of a formula incorporating age, AST, platelet count and ALT.

It has been validated for use in NAFLD as an 'initial test' for fibrosis screening in when investigating fatty liver, to guide further investigation for liver fibrosis in those with intermediate or high FIB4 scores (14, 38). This has proven to be a cost-

effective method at screening for risk of advanced fibrosis in a community population of people with fatty liver disease (189).

Although indirect markers have not yet been incorporated into clinical guidelines for ArLD, they have shown good performance in advanced fibrosis detection in this condition: Thiele et al. found AUROCs of 0.85 for FIB4, 0.80 for APRI, 0.76 for AST:ALT ratio and 0.86 for Forns' index for the detection of advanced fibrosis in ArLD (23).

Indirect markers can be a very useful 'crude tool' for the screening of advanced fibrosis or cirrhosis. However, the components of these indirect marker panels can be influenced by extra-hepatic factors, for example AST/ALT can be influenced by inflammation or muscle injury, and platelet count can be influenced by the direct effect of alcohol on the bone marrow.

Alcohol can also directly affect AST levels. This is thought to be in part due to the toxic effect of alcohol on mitochondria, which then release the mitochondrial isoenzyme of AST, and also due to the depletion of vitamin B6 (pyridoxine) in people with chronic alcohol use (190). ALT and AST both use pyridoxine as a coenzyme, but the synthesis of ALT is more profoundly inhibited by pyridoxine deficiency than that of AST (191).

Whilst an elevated serum AST-to-ALT ratio has been traditionally proposed as a sign that the patient has been consuming excess alcohol (192), many patients with alcohol use disorders do not display an elevated AST-to-ALT ratio. This led Nyblom et al. in 2004 to study 313 patients with alcohol dependence, comparing AST-to-ALT

ratios in groups with differing levels of cirrhosis and alcohol intakes. The authors of this study concluded that most patients with high alcohol consumption do not have an elevated AST-to-ALT ratio (above 1), and a high AST-to-ALT ratio is suggestive of advanced liver disease (192).

Nevertheless, caution should be used when interpreting tests that include AST in the context of alcohol excess, and tests such as FIB4, APRI and AST-to-ALT ratio should not be relied upon in isolation for distinguishing between fibrosis stages or confirming the definite presence of advanced fibrosis.

2.9.2 Direct markers

Direct markers of fibrosis measure constituents of liver matrix and molecules that regulate fibrogenesis and fibrinolysis. They include tests that measure collagens, glycoproteins and metalloproteinases (21). The single direct marker that has been the most extensively evaluated is Hyaluronic acid (HA) and this is a component of several of the combined marker panels, included ELF (Enhanced Liver Fibrosis test), Fibrometer, and Hepascore.

Table 2.6 demonstrates the range of indirect and direct markers and marker panels available.

Of these, the Enhanced Liver Fibrosis test (ELF) and FibroTest (FT) are the most commonly used direct marker panels for evaluation of liver fibrosis (193).

2.9.3 The Enhanced Liver Fibrosis (ELF) test

The ELF test is an automated blood test generating a score derived from a logarithmic algorithm combining serum measurements of three "direct" markers of hepatic extracellular matrix metabolism: Hyaluronic acid (HA), N-terminal peptide of procollagen III (PIIINP) and Tissue Inhibitor of Metalloproteinase-1 (TIMP-1).

The ELF equation is:

ELF score = 2.278 + 0.851 ln (CHA) + 0.751 ln (CPIIINP) + 0.394 ln (CTIMP1) for use on an ADVIA centaur XP analyser.

ELF is a CE (Conformité Européene) marked test, manufactured by Siemens Healthineers Inc., Tarrytown, NY, USA. It requires 5ml of blood, with serum separated by centrifugation at 1500 x g for 10 minutes at room temperature. The minimum serum volume for analysis is 250μl. The manufacturer's thresholds for ELF test score interpretation is <7.7 (none to mild fibrosis), 7.7-9.8 (moderate fibrosis), ≥9.8 (advanced fibrosis).

2.9.4 How does ELF work?

The three components of ELF: HA (hyaluronic acid), PIIINP (N-terminal peptide of procollagen III), and TIMP-1 (tissue inhibitor of metalloproteinase-1), are all markers of hepatic extracellular matrix metabolism.

2.9.4.1 HA

HA is a polysaccharide, synthesised by hyaluronic synthases, and made up of two alternating subunits of β-glucuronic acid and N-acetyl-D-glucosamine (194). It functions as a component of the extra-cellular matrix of the liver, a key factor in fibrogenesis as described earlier in this chapter. It has a short half-life, and is synthesised and removed by the liver, with the coordination of this via hyaluronate receptors in the hepatic sinusoid epithelial cells (195).

2.9.4.2 PIIINP

The amino terminal fragment of procollagen III (PIIINP) is cleaved from type III procollagen during collagen synthesis (196). When pathological fibrogenesis is occurring, increased collagen deposition in the liver leads to raised circulating levels of the PIIINP. Therefore, the detection of PIIINP in a blood test can be used to represent the underlying fibrotic transformation of the liver (195).

2.9.4.3 TIMP1

TIMP1 is produced by activated hepatic stellate cells in response to liver injury. Its function is inhibition of matrix metalloproteinases that cleave collagens and so its expression favours the accumulation and deposition of fibrous tissue (195, 197, 198).

2.9.5 The development of ELF

The European Group on Liver Fibrosis (EUROGOLF) was a prospective cohort study which investigated biomarkers of liver fibrosis in serum samples obtained from 1,021 patients with mixed liver aetiologies undergoing liver biopsy. The study measured markers of liver fibrosis thought to be directly involved in fibrogenesis and fibrolysis as well as a number of indirect biochemical and haematological markers of liver disease and the findings were reported in 2004 (29). This was a prospective study of patients that were under investigation for chronic liver disease who had abnormal liver biochemistry for at least 6 months, with most patients (n=496) having chronic hepatitis C, 61 with NAFLD, 81 with ArLD, and the rest with PBC (primary biliary cholangitis), PSC (primary sclerosing cholangitis), autoimmune hepatitis, or recurrent liver disease post-transplantation. All had serum blood samples taken on the same day as liver biopsy, and 9 different 'direct' serum markers and over 30 'indirect' markers were evaluated from all patients. The liver biopsies from 921 patients met pre-specified criteria for inclusion (>15 mm in length and more than 9 portal tracts) and were included in the analysis. Logistic regression analysis produced an optimal algorithm combining HA, PIIINP and TIMP1, in addition to age. It was assessed against liver histology using both the Ishak and Scheuer staging, with prevalence in each Scheuer stage as follows: F0 = 24.4%, F1 = 35.5%. F2 = 13.4%, F3: 14.9%, F4:11.8%. This algorithm produced AUROC of 0.804 for the detection of significant fibrosis, and 0.887 for advanced fibrosis/cirrhosis (29). Aetiology-specific AUROCs were 0.773 for chronic hepatitis C, 0.870 for NAFLD, and 0.994 for ArLD (29).

Future validation studies on ELF found that it discriminated between fibrosis stages just as well when age was not included in the algorithm (32) therefore age was removed to simplify the algorithm.

2.9.5.1 Validation of ELF

Since its conception, ELF has been validated in several patient populations, including NAFLD (AUROC for advanced fibrosis: 0.93 (95%CI 0.88–0.98)(32)), ArLD (AUROC for advanced fibrosis: 0.92 (0.89–0.96)(23)), PBC (AUROC 0.75 for significant fibrosis) (30), PSC (AUROC for advanced fibrosis: 0.81 (95% CI: 0.73–0.87) (31), chronic hepatitis C (33), and chronic hepatitis B (34).

It has been shown to have a linear relationship to fibrosis severity, with higher scores reflecting more advanced fibrosis (28), and can be used to monitor fibrosis progression (195, 199).

In addition, ELF has been shown to perform well at prognosticating clinical outcomes, and I will further explore this through my systematic review in Chapter 7.

2.9.5.2 FibroTest

FibroTest is another patented panel of serum markers, originally developed for use in hepatitis C (200). It is an algorithm consisting of 5 blood-based biomarkers (α2macroglobulin, gamma-GT, apolipoprotein A1, haptoglobin, bilirubin), along with age and gender. In a systematic review of 9 studies (1679 patients, predominantly

viral hepatitis), FibroTest was found to perform excellently in detecting cirrhosis (AUROC 0.9), but less well in detecting significant (≥F2) fibrosis (AUROC 0.81). It has since been compared by Thiele et al. in a cohort of 289 patients with ArLD, with performance equal to that of ELF and FibroScan (intention to diagnose protocol). For the detection of advanced fibrosis (F3), FibroTest AUROC = 0.88 (95% CI 0.84-0.92), ELF = 0.92 (95% CI 0.89-0.96), FibroScan = 0.89 (95% CI 0.83-0.96). A potential disadvantage of FibroTest is that it includes bilirubin and haptoglobin, which are affected by haemolysis and gamma-GT that may be increased by liver inflammation.

Combinatorial fibrosis marker panels such as ELF/FibroTest (Table 2.6), perform better than individual direct markers alone such as HA (175).

As yet, however, there has been no conclusive evidence that combining biomarker panels provides any benefit in ArLD – either with combinations of serum non-invasive tests, or combinations of serum with physical measures such as elastography, show any benefit when compared to individual non-invasive tests (139).

2.9.6 Imaging methods of Fibrosis assessment

A range of imaging modalities have been developed to measure liver stiffness, as a marker of fibrosis. Compared with serum biomarkers which may not be entirely liver-specific (as described above), a benefit of using imaging methods is that these are specific to the liver.

The most widely used and researched of these imaging methods is FibroScan, or 'Transient Elastography'. Other methods include magnetic resonance elastography (MRE), acoustic radiation force impulse imaging (ARFI), and sheer-wave elastography (2D-SWE).

From a practical perspective, imaging methods of fibrosis assessment require a greater time commitment from the operator and patient compared to automated measurement of serum markers in a blood sample. Whilst the procedure itself may only take a few minutes, it can take longer to position the patient, and to get optimum views, particularly in FibroScan when obesity and anatomy of the ribcage can obscure views (201), and all of these methods depend upon the skill of trained operators and maintenance of the equipment. Serum markers, in contrast, depend upon blood test that can be automated and performed at the same time as other routine blood tests and so are less time-consuming.

2.9.6.1 FibroScan (Transient Elastography)

FibroScan (or 'transient elastography') for the quantification of liver stiffness as a surrogate for liver fibrosis was first described in 2003 in France.

This technique involves the use of a transducer on the end of an ultrasound probe, which, when placed on the patient's right upper quadrant, transmits 50 MHz of pressure waves through the liver tissue. The velocity of the wave through a 1.5cm cube of liver is measured by the probe using ultrasound, and this is converted to a measurement of liver stiffness of the liver that is taken to represent the amount of liver fibrosis.

Criteria for successful FibroScan measurement are: 1) At least 10 attempts at liver stiffness measurement taken, 2) success rate is at least 60% (percentage of successful liver stiffness measurements out of attempts made to obtain a reading), 3) the interquartile range of the successful measurements of liver stiffness is less than 30% of the median liver stiffness score (202).

Since its inception, FibroScan has been evaluated in many liver disease aetiologies (37, 38). A large meta-analysis of studies of mixed aetiology liver diseases (including hepatitis B, C, NAFLD and ArLD), found FibroScan to have a pooled sensitivity of 83% and specificity of 89% for the diagnosis of cirrhosis (F4), with a threshold of 15kPa (203).

It performs well in ArLD (23, 204) with a meta-analysis of 14 studies (834 patients) on ArLD finding FibroScan to have summary sensitivity of 0.92 (95% CI 0.89-0.96) and summary specificity 0.70 (95% CI 0.61-0.79) for advanced fibrosis. However, thresholds have not yet been validated for the use of FibroScan in ArLD, and furthermore, it has been shown to be affected by alcohol intake (205), alcohol withdrawal (206), and inflammation (132, 207). It should therefore be interpreted with care in these settings.

A potential significant benefit of FibroScan over serum markers, however, is that it allows feedback of the result to patients in real-time, rather than requiring the patient to wait for a further follow-up visit to discuss the results, as is usually the case when using serum markers or other radiological investigations such as MRI or CT. The

method of communicating liver fibrosis results with patients may also be important in influencing behaviour change. This will be discussed further in section 2.9.6.5.

A further notable benefit of the FibroScan includes includes its wide availability and low cost comparable to that of ELF and less costly than liver biopsy (208). NICE 2020 guidance on FibroScan use in primary care reports a cost per unit of between £30,000 and £70,000 for FibroScan (excluding VAT) (208). However, it further details that with a manufacturer-reported shelf-life of 7 years, the estimated cost per use is between £50-£400 depending on the centre, which is cheaper than a liver biopsy (which costs approximately £500 per use).

A recent 2019 study by Srivastava et al. directly compared the cost effectiveness of FibroScan versus ELF in a cohort of NAFLD patients in primary care, reporting similar unit costs per patient of £42 for ELF, and £43 for FibroScan (36). The cost-per-detection of ≥F3 fibrosis was also comparable between the two tests (£9487 for ELF, £10,351 for FibroScan), and significantly less costly than the detection of ≥F3 fibrosis by standard care (£25,543) (predominantly due to the reduced need for secondary care referrals) (36). The 'standard care' was defined in this study as a decision process involving three primary care consultations, including full history, examination and three routine blood tests including liver function and full non-invasive liver screen, and a liver ultrasound scan. This was then classified into a binary outcome of either 'high risk of advanced fibrosis necessitating referral to a specialist or low risk, appropriate for primary care. The study concluded that the use of FibroScan or ELF for the detection of advanced fibrosis would enable substantial cost savings to the NHS (36).

I outline the performance of FibroScan compared to other non-invasive tests in ArLD in Chapter 3.

2.9.6.2 Magnetic Resonance Elastography (MRE)

Magnetic Resonance Elastography (MRE) combines elastography of the liver using a pressure impulse device with magnetic resonance imaging. It has not yet been validated for use in ArLD, but shows promise in NAFLD, with one study suggesting that it may be more accurate than FibroScan for the detection of significant fibrosis, with MRE AUROC 0.91 (95% CI 0.86-0.96), compared to FibroScan AUROC 0.82 (95% CI 0.74-0.89), p = 0.001 (209). Further research is needed on the use of MRE, particularly in ArLD, but its use is likely to be restricted by resources- cost and availability.

2.9.6.3 Acoustic Radiation Force Impulse Imaging (ARFI)

ARFI is a type of ultrasound elastography, that employs ultrasound to propagate the shear wave as well as to measure its velocity through the liver. It can be performed using standard ultrasound equipment. It is a newer test than FibroScan, and there are consequently fewer studies reporting its efficacy. The area of the liver examined with ARFI is smaller than that examined with FibroScan (10mm x 6mm with ARFI, 10mm x 40mm with FibroScan), but the area can be targeted. ARFI combines the use of ultrasound for abdominal imaging with the liver stiffness measurement, and therefore has some potential advantage over FibroScan as it can target areas of the liver, avoiding overlying bowel gas or blood vessels that can obstruct the view. It has not been validated for use in ArLD, but two small trials have evaluated ARFI in ArLD,

which have shown good diagnostic accuracy with AUROC of 0.86 (210) and 0.87 (211) for advanced fibrosis.

2.9.6.4 Two Dimension shear wave elastography (2D-SWE)

2D-SWE is a relatively recent tool for the non-invasive assessment of liver fibrosis. It uses a standard 2D ultrasound probe, but with a focused acoustic beam to generate 'shear wave' within the liver, with the speed of the wave correlating with the degree of liver fibrosis. A meta-analysis using data from 13 centres on mixed aetiology liver disease (predominantly viral hepatitis and NAFLD) reported AUROC values of 0.85 in NAFLD, 0.90 in hepatitis B, and 0.86 in hepatitis C (212) for the detection of advanced fibrosis.

Thiele et al. also evaluated 2D-SWE in a trial of 289 patients with ArLD, producing AUROC of 0.93 (0.89-0.98) for advanced fibrosis, not significantly different to FibroScan, FibroTest or ELF in the intention to diagnose protocol (23). Further validation is needed of 2D-SWE in ArLD.

2.9.6.5 Brief intervention and biofeedback of results impacting on behaviour change

With three-quarters of patients having decompensated cirrhosis at first presentation, and in the knowledge that fibrosis stage predicts mortality, it would seem logical that by opportunistically testing people with risk factors for fibrosis (for example, AUD), it would enable earlier detection of fibrosis and therefore allow time for interventions to positively influence clinical outcomes.

Studies have shown that an effective 'intervention' leading to behaviour change can be as simple as a consultation with the patient incorporating advice about reducing alcohol. A 'Brief Intervention' is an umbrella term for these types of interventions delivered by healthcare practitioners that provide advice or counselling with the aim to help patients understand risks of their alcohol consumption and explore ways they could cut down. Brief interventions typically follow a 'FRAMES' structure, encompassing **F**eedback about existing consumption, **R**esponsibility for change, **A**dvice about practical strategies to reduce drinking, a **M**enu of options for behaviour change, **E**mpathic delivery, and **S**elf-efficacy building (4).

A Cochrane meta-analysis in 2019 of 34 trials (15,197 participants) found that brief interventions in general practice or emergency care settings reduced hazardous or harmful alcohol consumption after one year (in both men and women), with an average reduction of alcohol consumption by 20g per week after brief advice compared to control groups (95% CI -28 to -12, I^2=73%) (4). Furthermore, short, advice-based interventions were as effective as extended counselling.

An addition to brief advice, there has been an increase in focus in the inclusion of information to the patient about the biological impact that their alcohol consumption has caused, with 'biofeedback' of liver tests leading to positive outcomes in reducing alcohol intake, enhancing the effect of brief interventions (5).

A systematic review by Subhani et al. found that biofeedback of results, for example non-invasive liver fibrosis tests or biomarkers of liver injury, resulted in reduction in harmful alcohol consumption, GGT levels, and alcohol-related mortality (5).

Sheron et al. used a Southampton Traffic Light (STL) score (comprised of HA, PIIINP, and platelet count) when assessing for liver fibrosis, and discovered that biofeedback of the STL score resulted in a significant reduction in AUDIT test scores across all risk groups (6). The Malmo study also found that feeding back gamma-glutamyl transferase (GGT) results to patients with high alcohol intake improved their outpatient attendance where a brief intervention could be delivered (7).

In addition, interventions delivered by technology, for example via computers and smart phone apps, have been shown to reduce alcohol consumption in hazardous and harmful drinkers (8).

From a population perspective, UK modelling suggests that delivery of a brief intervention to every patient registering with a new general practitioner would result in a reduction of 125,000 hospital admissions and 2,500 fewer alcohol-related deaths over 20 years, with associated cost savings of £282 million (9).

It must be noted, however, that most of the relevant studies described above included participants that were predominantly middle-aged men. Therefore further work is needed to assess the impact of biofeedback and brief intervention on women and younger patients (10).

Furthermore, most of the studies on brief intervention and biofeedback only followed up the participants for short periods of time, usually around 12 months. Therefore, there is currently a lack of data to support any long-term alcohol reduction or abstinence following brief intervention or biofeedback of test results.

The Nottingham research group led by Ryder et al. have recently published a study protocol for a clinical trial investigating the impact of biofeedback of FibroScan results in the community on alcohol intake, which should allow longer term follow-up (11).

2.9.7 Identified gaps in the evidence base

Through this literature review, I have discovered several unanswered questions in the field of non-invasive testing in ArLD that require further research.

First, whilst I have highlighted several different methods of non-invasive fibrosis testing, it is unknown which test performs the best in ArLD, and if blood-based biomarkers are as effective as FibroScan for fibrosis detection at all fibrosis stages, how the performance of 'direct' fibrosis markers compares with 'indirect' markers such as FIB4, what factors affect their performance, and which thresholds should be

used for ArLD. I will address these knowledge gaps by performing a systematic review and meta-analysis in Chapter 3 comparing four of the most commonly used non-invasive tests.

I have reported in this chapter that 75% patients with ArLD present to healthcare providers late, when they already have decompensated cirrhosis that significantly impacts on their prognosis. Are there missed opportunities for testing patients in hospital with AUD? Is fibrosis testing already embedded in standard practice? I will be addressing these questions in a prospective study on people with AUD presenting to hospital over a 13-month period in Chapter 6.

Whilst I have commented on the national guidelines for non-invasive testing in ArLD, do primary care physicians have access to these tests? What methods are they currently using to aid referral decisions to secondary care? Is there scope for optimisation of current practice? I also revealed the uncertainties about whether an alcohol 'threshold effect' exists, above which liver disease develops, and if so – there is no consensus on what this threshold is. I will be exploring this in Chapters 4 and 6.

Whilst fibrosis is the strongest predictor of prognosis in chronic liver disease, it is unclear if non-invasive fibrosis tests can reliably predict prognosis in ArLD, and if so, which one performs the best. I will be addressing this in Chapter 7.

Finally, currently used scoring tools for mortality prediction in AH perform sub optimally, with AUROCs <0.80 for 90-day mortality. It is not known whether non-invasive fibrosis tests can be used in this clinical situation to predict prognosis, and

how they would compare to the traditional scoring tools such as MELD, GAHS and ABIC. I will be answering these questions in the final study chapter, Chapter 8.

In the next chapter, I will investigate in more detail the diagnostic performance of the four most commonly described non-invasive tests in ArLD – FibroTest, ELF, FibroScan and FIB4, with a systematic review and meta-analysis.

Chapter 3

Investigating the diagnostic performance of four non-invasive tests in Alcohol-related Liver Disease: a systematic review with meta-analysis.

3.1 Abstract:

Background/Aims:

Fibrosis stage is the main prognostic factor in alcohol-related liver disease (ArLD). Non-invasive tests (NIT) are increasingly used to detect fibrosis, but performance varies depending on aetiology. It is not clear which NIT performs the best in ArLD. I aimed to describe the diagnostic performance of four widely used NITs (Fibrosis 4 test [FIB4], Enhanced Liver Fibrosis [ELF] test, FibroScan, and FibroTest) in ArLD.

Methods:

I applied systematic review methodology to search four databases from inception to Feb 2021, including a combination of Medical Subject Heading terms and keywords, applying pre-defined inclusion/exclusion criteria. The results were screened by myself and a second reviewer independently, along with independent data extraction and risk of bias assessment using Quality Assessment of Diagnostic Accuracy Studies (QUADAS2) tool.

Results:

Searches produced 11,000 articles. After initial screening, 782 articles were independently reviewed by myself and the second reviewer, leaving 16 articles remaining for analysis (total n=2,280): 9: FibroScan, 1: FIB4, 1: ELF, 1: FibroTest, 1: FIB4/FibroTest/ELF/FibroScan, 1: FIB4/FibroTest, 1: FIB4/FibroTest/FibroScan, 1: FibroTest/FibroScan. Studies scored low-moderate for all risk of bias domains. Results were heterogeneous for outcomes and reporting, making meta-analysis only possible for FibroScan, with pooled Area-Under-receiver-Operator-Characteristic-curve (AUROC) of 0.91 (95% CI 0.89-0.94) for F3, for F4: pooled

sensitivity 88% (95% CI 0.84-0.92), pooled specificity 84% (95% CI 0.81-0.87). AUROCs for F3 were for ELF: 0.82-0.92, for FibroTest 0.80-0.90, for FIB4 0.70-0.85.

Conclusions:

This systematic review returned 16 papers (3 of which were abstracts). Heterogeneity precluded pooling of results for FIB4/FibroTest/ELF, but all tests had good diagnostic accuracy at F2/F3/F4 (AUROC $\geq$ 0.7). FibroScan had pooled AUROC of 0.91 for F3, but was influenced by alcohol, inflammation and bilirubin/AST levels and had a failure rate of 1-22%. Whilst all four tests could be used in clinical practice, optimal thresholds for use in ArLD are yet to be determined, and this requires further prospective validation.

3.2.1 Introduction

Non-invasive fibrosis tests are increasingly used in the assessment of chronic liver disease, as there is recognition that liver biopsy is an imperfect test, due to its invasive nature and risk of sampling error (171, 218). With the flurry of non-invasive tests developed over recent decades, there is now a plethora of available tests, with varying performance across different liver disease aetiologies (203, 219-222) as described in Chapter 1.

Detection of advanced fibrosis (METAVIR F3 equivalent) by non-invasive testing is essential in people with alcohol use disorders (AUD), as this is the level of fibrosis that requires secondary care management, and also allows those with <F3 fibrosis to be re-screened at interval periods depending on ongoing risk factors (1). Prognosis in ArLD is also strongly related to the fibrosis stage (223).

Despite alcohol-related liver disease (ArLD) being the leading cause of chronic liver disease in the UK, the study of non-invasive tests in this condition is significantly less than in other aetiologies such as NAFLD and viral hepatitis (224).

Whilst UK guidelines currently recommend the use of FibroScan or ELF for the detection of advanced fibrosis in ArLD (1, 38), the optimum non-invasive test and thresholds for use in ArLD are not defined.

Similar test performance was reported for FibroTest, ELF and FibroScan in the detection of advanced fibrosis and cirrhosis in a biopsy-paired Danish study in ArLD

(23). FIB4 also performed well in prognosticating ArLD in a systematic review, with AUROCs >0.7 (224), but its diagnostic performance in ArLD is not known.

I aimed to perform a systematic review to investigate the diagnostic, performance of four non-invasive tests (FibroTest, FibroScan, ELF and FIB4) in the detection of advanced fibrosis (F3) and cirrhosis (F4) in ArLD.

3.2.2 Ethics and patient consent

Ethical approval was not required for this systematic review, since it used data from previous studies which had their own ethics and patient consent.

3.2.3 Methods

I conducted this systematic review in accordance with the Cochrane Handbook of Diagnostic test accuracy. My aim was to investigate the diagnostic performance of four non-invasive fibrosis tests in ArLD- FibroTest, FibroScan, the ELF test and FIB4. I observed the PRISMA guidelines (Preferred Reporting Items for Systematic reviews and Meta-Analyses) in the conduct of this study. Details of the databases searched are shown in Table 3.1.

Table 3.1: Databases searched

I updated the search on 18/02/2021 using the same search strategy and methodologies:

Databases	Date initial search performed	Date repeat search performed
MEDLINE (Ovid) (1946 to date of search)	23/05/2020	18/02/2021
EMBASE (Ovid) (1974 to date of search)	23/05/2020	18/02/2021
Web of Science (1900 to date of search)	23/05/2020	18/02/2021
Cochrane Database of Systematic Reviews	23/05/2020	18/02/2021

3.2.4 PICO

Participants: All adult humans (Age 18+) with Alcohol Related Liver Disease

Intervention: Studies that included ELF, FibroTest, FibroScan or FIB4 for the staging of liver fibrosis in ArLD were included

Comparisons: Liver biopsy was used as the reference standard

Outcomes:

1) The diagnostic accuracy of non-invasive tests (ELF/FibroScan/FibroTest/ FIB-4) in distinguishing advanced fibrosis (equivalent to ≥METAVIR F3) from patients without advanced fibrosis (<METAVIR F3) in all patients at risk of alcohol-related liver disease, compared with the reference standard liver histology as judged by AUROC, sensitivity, specificity, positive and negative predictive value
2) The diagnostic accuracy of non-invasive tests (ELF/FibroScan/FibroTest/ FIB-4) in distinguishing cirrhosis (equivalent to ≥METAVIR F4) from patients without cirrhosis (<METAVIR F0-3) in all patients at risk of alcohol-related

liver disease, compared with the reference standard liver histology as judged by AUROC, sensitivity, specificity, positive and negative predictive value

3) The diagnostic accuracy of non-invasive tests (ELF/FibroScan/FibroTest/ FIB-4) in distinguishing significant fibrosis (METAVIR F2) from patients without any significant fibrosis (METAVIR F0-1) in all patients at risk of alcohol-related liver disease, compared with the reference standard liver histology as judged by AUROC, sensitivity, specificity, positive and negative predictive value
4) Numbers of test failures for each NIT

3.2.5 Search strategy

Search strategies can be found in Tables 3.2A, 3.2B, and 3.2C. I searched four databases systematically from inception to 18th February 2021– EMBASE, Web of Science, Ovid Medline and Cochrane library. I first conducted pilot searches to refine the search. Search themes related to the PICO, and incorporated a combination of MeSH terms and keywords. After I had conducted the initial searches, I then hand-searched references of key studies and reviews to check for any further potentially relevant studies for inclusion. Furthermore, where information from abstracts or full text articles was sufficient to include the study, I contacted the relevant authors by email to request the data.

References were imported into EndnoteTM web basic reference manager, and subsequently once duplicates removed, into Rayyan systematic review manager (225) which allowed blinded reviewing and sorting of articles between myself and the second reviewer (Paul Trembling (PT)).

I used a combination of MeSH terms and free text words to make the search as comprehensive as possible. I used different combinations of similar words to maximise the results, for example, for ELF I used 'ELF' or 'elf adj score' or 'hyaluronic acid' or 'hyalauronate' or 'hyaluronan' or 'procollagen' or 'piiinp' or 'p3np' or 'ppcp' or 'tissue adj inhibitor adj metalloproteinase$' or 'timp$'. For a more general search of non-invasive liver fibrosis tests, I included 'biological marker$' or 'biomarker$' or 'algorithm$' or 'non adj invasive adj test' or 'non adj invasive', and these terms were included using 'AND' along with terms for alcohol-related liver disease including MeSH terms. Cirrhosis was not limited to alcohol-related in these searches, so as to be able to explore studies on mixed-aetiology cirrhosis in case they included sub-analyses on alcohol-related liver disease.

Finally, I combined these terms with 'AND' for the diagnostic terms, e.g., sensitivity, specificity, ((pre-test or pretest) adj probability), post-test probability, predictive value$, and likelihood ratio$.

Table 3.2A: DETAILS OF SEARCH STRATEGY: Ovid MEDLINE and EMBASE

No.	Searches	Search type	Total number of results
1.	(enhanced adj liver adj fibrosis).tw.	Advanced	
2.	(elf adj test$).tw.	Advanced	
3.	(elf and diagnos$).tw.		
4.	elf.tw.	Advanced	
5.	(elf adj score).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	Advanced	
6.	FibroTest.tw.	Advanced	
7.	fibroscan.tw.	Advanced	
8.	(transient adj elastograph$).tw.	Advanced	
9.	(elastograph$ and liver).tw.	Advanced	
10.	(hyaluronic adj acid).mp. or (hyalauronate or hyaluronan).tw. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	Advanced	
11.	(procollagen or piiinp or p3np or ppcp).tw.	Advanced	
12.	((tissue and inhibitor and metalloproteinase$) or timp*).tw.	Advanced	
13.	FIB 4.tw.	Advanced	
14.	FIB4.tw.	Advanced	
15.	biological markers/	Advanced	
16.	biomarker$.tw.	Advanced	
17.	algorithm$.tw.	Advanced	
18.	(non adj invasive adj test).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	Advanced	
19.	(non adj invasive).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	Advanced	
20.	exp liver cirrhosis/ or exp liver diseases, alcoholic/	Advanced	
21.	(fibros* or cirrhos*).tw.	Advanced	
22.	Exp “sensitivity and specificity”/	Advanced	
23.	Sensitivity.tw.	Advanced	
24.	Specificity.tw.	Advanced	
25.	((pre-test or pretest) adj probability). Tw.	Advanced	
26.	Post-test probability.tw	Advanced	
27.	Predictive value$.tw.	Advanced	
28.	Likelihood ratio$.tw	Advanced	
29.	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19	Advanced	
30.	20 or 21	Advanced	
31.	22 or 23 or 24 or 25 or 26 or 27 or 28	Advanced	
32.	29 and 30 and 31	Advanced	
33.	limit 32 to human	Advanced	6,447

Table 3.2B: DETAILS OF SEARCH STRATEGY: Cochrane database

No.	Searches	Search type	Total number of results
1	(enhanced adj liver fibrosis) OR (elf adj test) OR (elf) OR (elf adj score) OR (FibroTest) OR (fibroscan) OR (transient adj elastography$) OR (elastography$ and liver) OR (hyaluronic adj acid) OR (hyalauronate) OR (hyaluronan) OR (procollagen) OR (piiinp) OR (p3np) OR (ppcp) OR (tissue adj inhibitor adj metalloproteinase$) OR (timp$) OR (FIB4) OR (FIB 4) OR (biological marker$) OR (biomarker$) OR (algorithm$) OR (non adj invasive adj test) OR (non adj invasive)	Advanced	26122
2	MeSH descriptor: [Fatty Liver, Alcoholic] explode all trees	Advanced	20
3	MeSH descriptor: [Liver Diseases, Alcoholic] explode all trees	Advanced	481
4	MeSH descriptor: [Liver Cirrhosis} explode all trees	Advanced	2861
5	MeSH descriptor: [Sensitivity and Specificity] explode all trees	Advanced	15150
6	(sensitivity OR specificity OR pre-test adj probability OR pretest adj probability OR post-test probability OR predictive value$ OR likelihood ratio$ OR diagnos$)	Advanced	82495
7	#2 OR #3 OR #4	Advanced	3101
8	#5 OR #6	Advanced	83007
9	#1 AND #7 AND #8	Advanced	22

Table 3.2C: DETAILS OF SEARCH STRATEGY: Web of Science

No.	Searches	Search type	Total number of results
1.	TS = enhanced adj liver adj fibrosis	Advanced	
2.	TS = elf adj test$	Advanced	
3.	TS = elf	Advanced	
4.	TS = elf adj score	Advanced	
5.	TS = FibroTest	Advanced	
6.	TS = fibroscan	Advanced	
7.	TS = transient adj elastograph$	Advanced	
8.	TS = (elastograph$ and liver)	Advanced	
9.	TS = (hyaluronic adj acid OR hyalauronate OR hyaluronan)	Advanced	
10.	TS = (procollagen OR piiinp or p3np or ppcp)	Advanced	
11.	TS = (tissue adj inhibitor adj1 metalloproteinase$) OR TS = timp$	Advanced	
12.	TS = FIB4	Advanced	
13.	TS = FIB 4	Advanced	
14.	TS = biological adj marker$	Advanced	
15.	TS = biomarker$	Advanced	
16.	TS = algorithm$	Advanced	
17.	TS = non adj invasive	Advanced	
18.	TS = non adj invasive adj test$	Advanced	
19.	TS = cirrhosis	Advanced	
20.	TS = liver adj fibrosis	Advanced	
21.	TS = (sensitivity OR specificity)	Advanced	
22.	TS = ((pre-test OR pretest) adj probability)	Advanced	
23.	TS = post-test probability	Advanced	
24.	TS = predictive value$	Advanced	
25.	TS = likelihood ratio$	Advanced	
26.	#18 OR #17 OR #16 OR #15 OR #14 OR #13 OR #12 OR #11 OR #10 OR #9 OR #8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1	Advanced	
27.	#20 OR #19	Advanced	
28.	#25 OR #24 OR #23 OR #22 OR #21	Advanced	
29.	#28 AND #27 AND #26	Advanced	2,859

3.2.6 Process for reviewing articles

The first sift of articles by title and abstract was performed by myself (FR), and then the remaining 782 articles were independently screened by both me and PT. I predefined the exclusion and inclusion criteria that I used to decide on articles for inclusion (see Table 3.3). I documented reasons for exclusion for each study, and any discrepancies of decisions between myself and second reviewer were resolved by discussion between ourselves, or if consensus not agreed then by input from a third reviewer (William Rosenberg). The resulting articles were reviewed again in full text independently by myself and PT, resulting in a final list of included papers (see Figure 3.1 for PRISMA flow).

3.2.7 Selection criteria

The full list of inclusion and exclusion criteria can be found in Table 3.3. I included all levels of evidence except descriptive review articles and opinion pieces. Pre-clinical and non-human studies were excluded. Grey literature (conference abstracts and unpublished manuscripts) was not excluded, in keeping with Cochrane guidance (226).

Studies that investigated mixed aetiology chronic liver disease were included as long as they incorporated at least 30 participants with ArLD, and these data were extractable separately. Studies needed to have used liver histology as the reference standard to compare diagnostic performance, with results displayed as sensitivity

and specificity or AUROC, or False Negative/True Negative/False Positive/True Positive, or if this information was calculable from the provided data.

Table 3.3: Selection criteria

Inclusion criteria	**Exclusion Criteria:**
All adult humans (age 16+)	Review articles
Participants have ArLD	Opinion pieces
≥ 30 participants (as per Parkes et al systematic review - as smaller studies would be underpowered to give accurate estimates of test performance, more likely to produce zero denominator effects in a 2 x 2 table, and give wide CIs which may result in unreliable results.) (227)	Non-human studies
Study relates to at least one of the four non-invasive tests of interest (FIB4, ELF, FibroTest, FibroScan)	Pre-clinical and biological studies
Study uses liver biopsy as reference standard	Aetiology of liver disease other than alcohol
Data are presented as sensitivity or specificity or AUROC or TP/TN/FP/FN, or if this information is able to be calculated from the provided information in the study.	Data not extractable by fibrosis stage
Article written in English	Mixed aetiology studies where alcohol data not able to be extracted separately.

CI: Confidence Interval; AUROC: Area Under Receiver Operating Characteristic curve, TP; True Positive. TN: True negative; FP: False positive; FN: False negative

3.2.8 DATA EXTRACTION STRATEGY

I created pre-defined data extraction forms prior to the independent extraction of the data by both myself and PT. Any disagreements were resolved through discussion.

I attempted to extract the following data: Study design, year of publication, participant's epidemiological and laboratory characteristics, type of non-invasive test investigated, biopsy length (mm) and number of portal tracts, definition of alcohol-related liver disease as defined by study authors, number of participants included, average stage of liver fibrosis, any failures of the test, thresholds used for non-invasive tests and diagnostic accuracy data (sensitivity, specificity, NPV, PPV, AUROC with 95% CI), LR+. LR-, or FP/TP/FN/TN) (See Tables 3.5 and 3.6). Where available, I also extracted data on if the non-invasive test of interest was influenced by inflammation (either histological or AST/ALT), BMI, histological steatosis, age or alcohol intake or withdrawal (Table 3.7). In the cases where data were unclear, or where data were reported for mixed aetiology liver disease patients but not specifically for those with alcohol-related disease, I contacted the authors by email for clarification or to request data. This also included instances where the same author had published several different articles and I wanted to clarify if any or all of the articles included overlapping sample cohorts, so that I could avoid duplication of results. Where authors failed to respond, the studies were excluded.

3.3 Quality assessment

The quality of the included diagnostic studies was assessed independently by myself and PT, using the QUADAS2 tool (Quality Assessment of Diagnostic Accuracy Studies) (228). This allowed grading of each publication for risk of bias as 'low-risk', 'unclear-risk' or 'high-risk', on the basis of answers to signalling questions relating to four key domains: patient selection, index test, reference standard, and flow and timing. Each of these is assessed in terms of risk of bias, and the first three are additionally

assessed in terms of concerns regarding applicability. I adapted definitions of some scoring criteria from the systematic review by Pavlov et al. (229). I also added in criteria for minimum biopsy length of 15mm and minimum of six portal tracts, as per best evidence (38). The QUADAS2 tool can be found in Table 3.4. Where there were differences in judgements between myself and PT these were resolved by discussion.

3.4 Data synthesis and analysis

I present results of the collected data in full tabulation for each of the included studies. Where available, the data for the diagnostic performance of each test were reported as sensitivity, specificity, Positive Predictive Value (PPV), Negative Predictive Value (NPV), likelihood ratios (LR) and Areas Under Receiver Operating Characteristic curves (AUROCs) with 95% Confidence intervals (CI). Severity of fibrosis was defined by authors from histology results. METAVIR was the most commonly used histological grading method, but where alternative scoring systems were used, I converted these to the METAVIR equivalent for consistency in my results (as per the conversion grid in Pavlov et al.'s systematic review (219). As such, fibrosis was defined as F0: No fibrosis; F1: portal fibrous expansion; F2: thin fibrous septa emanating from portal triads; F3: fibrous septa bridging portal triads and central veins; F4: cirrhosis (219). On this METAVIR scale from F0 to F4, clinically significant fibrosis is generally defined as F2 or above, advanced fibrosis as F3 or above, and cirrhosis F4.

I report AUROCs and 95% CI where available for each non-invasive test in forest plots. Risk of bias results are plotted in an individual study graph, and in a

summary-graph, which I produced using 'RevMan' software (230). Sensitivities/specificities/likelihood ratios and confidence intervals, where needed to be calculated, I did so using MedCalc statistical software 2020. I used SPSS (version 26, Armonk, NY: IBM Corp) to obtain AUROCs for one of the included articles where raw data on FibroScan and biopsy results were provided to me from the first author after I emailed to request data for alcohol patients (231).

I calculated pooled estimates for those tests where there were more than 3 qualifying studies in addition to there being an acceptable level of heterogeneity (defined as I^2 statistic (inconsistency index) of less than 50%) (232), and hence combining results was reasonable. Where results were able to be pooled, I applied a random effects meta-analysis model to account for any remaining heterogeneity. I performed these tests using STATA IC (version 16.1 *StataCorp* LLC, College Station, TX).

I calculated Cohen's kappa for risk of bias agreement between myself and the second reviewer, PT, using SPSS. I produced Risk of bias summary graphs from Reference Manager (RevMan) version 5.4 (230), and forest plots from GraphPad and STATA.

TABLE 3.4

QUADAS2 Questionnaire (adapted from Pavlov et al., (229))

DOMAIN	PARTICIPANT SELECTION	INDEX TEST	REFERENCE STANDARD	FLOW AND TIMING
Description	**Describe methods of participant selection: describe included participants (prior testing, presentation, intended use of index test and setting):**	**Describe the index test and how it was conducted and interpreted**	**Describe the reference standard and how it was conducted and interpreted**	**Describe any people who did not receive the index test(s) or reference standard (or both) or who were excluded from the 2 x 2 table: describe the time interval and any interventions between index test(s) and reference standard:**
	Studies fulfilling inclusion criteria should include adults above 16yrs old, any sex and ethnic origin, outpatients or inpatients, with a diagnosis of ArLD (not including acute alcoholic hepatitis). The diagnosis of ArLD should have been established by history of excessive alcohol intake, plus clinical, biochemical or imaging-based evidence of liver disease. To confirm the diagnosis the patients should all have had a liver biopsy plus one of the 4 non-invasive tests in question in this study (FIB4/ELF/FibroTest/ FibroScan)	For FibroScan: Transient elastography used for grading liver fibrosis either before or after liver biopsy. Recommended FibroScan parameters are at least 10 validated stiffness measurements at the same measurement point, and IQR of no more than 30%, and the ratio of number of successful	Liver biopsy is the reference standard in this study. It is used to stage liver fibrosis, with differing stage definitions depending on scoring system used (e.g., METAVIR, Knodell, Ishak, Kleiner, Brunt, Scheuer). Supplementary table 1 shows comparison of fibrosis stage definitions between scoring systems.	As fibrosis may develop rapidly, the liver biopsy and non-invasive test under investigation should be performed within 6 months of each other.

		measurements to total investigation number should be no less than 60% (www.echosens.com/pdf/FS402_WEB.pdf) For ELF, FibroTest, FIB4: As they are blood tests, no specific recommendations for conduct of test.		
Signalling questions: yes/no/ unclear	**Was a consecutive or random sample of participants enrolled?**	**Were the index test results interpreted without knowledge of the results of the reference standard?**	**Is the reference standard likely to classify the target condition correctly?**	**Was there an appropriate interval between index test(s) and reference standard?**
	Yes: All patients with ArLD that were included in this study were either a random sample or consecutive participants. No: Selected patients were not included. Unclear: insufficient information	Yes: FIB4/FibroTest/FibroScan/ELF results were interpreted without knowledge of the results of the liver biopsy	Yes: if all participants had undergone a liver biopsy and the morphological results were reported correctly, and if average biopsy length ≥15mm and with ≥ 6 portal tracts	Yes: The interval between the liver biopsy and the ELF test/FibroScan/FibroTest /FIB4 was ≤ 6 months No: The interval between the liver biopsy and the ELF test/FibroScan/FibroTest /FIB4 was > 6 months.

	reported to allow a judgement.	No: FIB4/FibroTest/FibroScan/ELF results were interpreted with knowledge of the liver biopsy result. Unclear: Unclear: Insufficient information reported to allow a judgement.	No: If all participants had not undergone liver biopsy or results were not reported correctly or if average biopsy length <15mm or < 6 portal tracts. Unclear: insufficient information reported to allow a judgement.	Unclear: Insufficient information reported to allow a judgement.
	Was a case-control design avoided?	**If a threshold was used, was it pre-defined?**	**Were the reference standard results interpreted without knowledge of the results of the index test?**	**Did all participants receive the reference standard?**
	Yes: Case-control design was avoided No: Case-control design was not avoided Unclear: Insufficient information reported to allow a judgement.	Yes: The threshold for a positive test was pre-defined. No: The threshold for a positive test was not pre-defined. Unclear: Insufficient information reported to	Yes: Liver biopsy results were interpreted without knowledge of the results of the ELF/FIB4/FibroTest/FibroScan. No: Liver biopsy results were interpreted with knowledge of the results of the ELF/FIB4/FibroTest/FibroScan.	Yes: All participants underwent a liver biopsy. No: Not all participants underwent a liver biopsy Unclear: Insufficient information reported to allow a judgement.

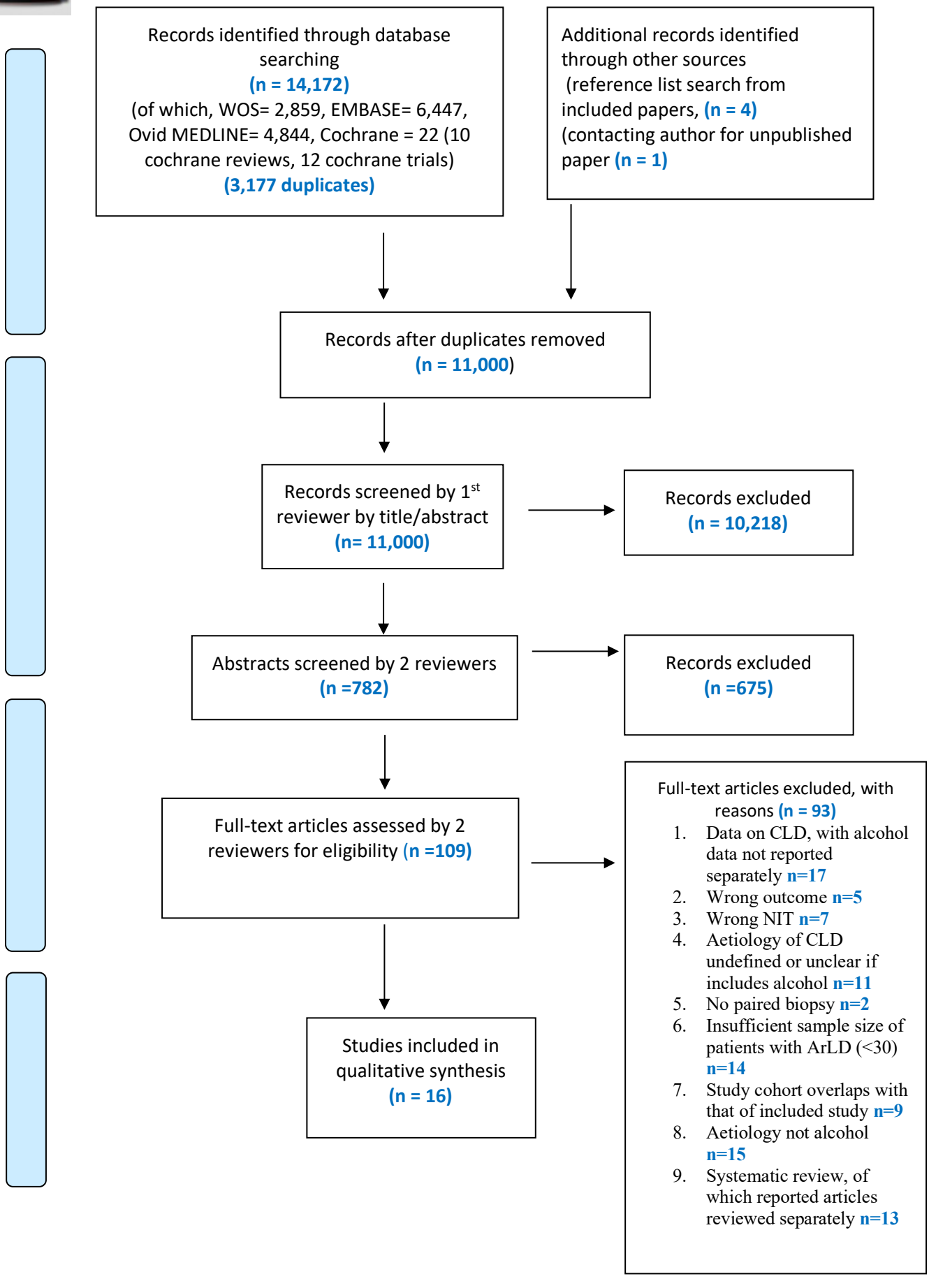
		allow a judgement.	Unclear: Insufficient information reported to allow a judgement.	
	Did the study avoid inappropriate exclusions?	**Did all participants receive the same reference standard?**	**Were all participants included in the analysis?**	
	Yes: The study avoided inappropriate exclusions (i.e. difficult to diagnose participants) No: The study excluded patients inappropriately Unclear: Insufficient information reported to allow a judgement.	Yes: All participants received the same reference standard, i.e., a liver biopsy. No: Not all participants received the same reference standard, i.e., a liver biopsy Unclear: Insufficient information reported to allow a judgement.	Yes: All participants meeting the selection criteria (Selected participants) were included in the analysis, or data on all the selected participants were available so that AUROCS, sensitivity/specificity/NPV/PPV could be calculated No: Not all participants meeting the selection criteria (Selected participants) were included in the analysis, or data on all the selected participants were not available so that AUROCS, sensitivity/specificity/NPV/PPV	

			could not be calculated. Unclear: Insufficient information reported to allow a judgement.	
Risk of bias: High/ Low/ Unclear	**Could the selection of participants have introduced bias?**	**Could the conduct or interpretation of the index test have introduced bias?**	**Could the reference standard, its conduct, or its interpretation have introduced bias?**	**Could the participant flow have introduced bias?**
	High risk of bias: Yes, the selection of participants introduced bias Low risk of bias: No, the selection of participants did not introduce bias. Unclear risk of bias: Insufficient information reported to allow a judgement.	High risk of bias: If the answer to the signalling questions on the conduct or interpretation of the index test was 'no'. Low risk of bias: If the answer to the 2 signalling questions on the conduct or interpretation of the index test was either 'unclear' or any combination of 'unclear'	High risk of bias: If the answer to the signalling questions on the reference standard, its conduct, or its interpretation was 'no'. Low risk of bias: If the answer to the signalling questions on the reference standard, its conduct or its interpretation was 'yes'. Unclear risk of bias: If the answers to the 3 signalling questions on the reference standard, its conduct, or its interpretation was either	High risk of bias; If the answer to the signalling questions on flow and timing was 'no'. Low risk of bias: If the answer to the signalling questions on flow and timing was 'yes'. Unclear risk of bias: If the answers to the 4 signalling questions on flow and timing was either 'unclear' or any combination of 'unclear' with 'yes' or 'no'.

		with 'yes' or 'no'. Unclear risk of bias: Insufficient information to allow a judgement	'unclear' or any combination of 'unclear' with 'yes' or 'no'.	
Concerns regarding applicability: High/Low/ Unclear	**Were there concerns that the included participants did not match the review question?**	**Were there concerns that the index test, its conduct, or interpretation differed from the review question?**	**Were there concerns that the target condition as defined by the reference standard did not match the review question?**	
	High concern: There was high concern that the included participants do not match the review question. Low concern: There was low concern that the included participants did not match the review question. Unclear: If it was unclear.	High concern: There was high concern that the conduct or interpretation of the FIB4/ELF/FibroTest/FibroScan results differs from the way they are likely to be used in clinical practice. Low concern: There was low concern that the conduct or	High concern: All participants did not undergo liver biopsy for grading liver fibrosis Low concern: All participants underwent liver biopsy for grading liver fibrosis. Unclear concern: If it was unclear.	

		interpretation of the FIB4/ELF/FibroTest/FibroScan results differs from the way they are likely to be used in clinical practice. Unclear concern: If it was unclear.		

PRISMA FLOW DIAGRAM (date of searches 18/02/2021) (Figure 3.1)



3.5 Results

3.5.1 Study selection

A total of 14,172 articles were returned by searching the four databases, of which 3,177 were duplicates (detected and removed using Endnote). A further four results were found by hand-searching reference lists of included papers and relevant review articles. Three of the articles (two conference abstracts (233, 234) and one full paper (235) reported data from the same patient cohort. The first author for the full paper (235) was co-author on both abstracts, and I consulted with them about which one to include. This resulted in permission to use unpublished data (under review) (236) from the most recent abstract (234). As this was the most recent and comprehensive of the articles (including 81 patients with alcohol-related liver disease, as opposed to 64 in the other full text paper), this was included despite not yet being published at the time of conducting this systematic review. Subsequently it has now been published (134). The other full paper (235) and abstracts reporting the same cohort (233, 234) were excluded.

I found several articles investigating the diagnostic performance of non-invasive tests in mixed-aetiology liver disease. Where studies included at least 30 patients with ArLD, I made efforts to contact authors by email to request any available data specifically for these patients, and these were included if data were received.

I found thirteen systematic reviews (227, 229, 237-247). However, only 4 out of these 13 focussed on ArLD (219, 223, 243, 248) with the rest investigating mixed-aetiology liver disease. I excluded the 13 systematic reviews as they either did not

include patients with ArLD, did not report alcohol data separately, or where they did, I had either already included the articles within these systematic reviews, or had excluded them if they did not meet my eligibility criteria. Reasons for exclusion, aside from alcohol data not being reported separately, commonly included the studies having sample sizes for alcohol cohorts less than 30, and reporting different outcomes (for example clinically significant portal hypertension or presence of varices), instead of fibrosis staging. I also excluded an article by Papatheodoridi *et al.* on "*refining the Baveno VI elastography criteria for the definition of compensated advanced chronic liver disease*" (249), as this was an amalgamation of individual study results that I had already included in this systematic review.

This resulted in 11,000 articles that I then screened by title or abstract, resulting in 10,218 being excluded and leaving 782 articles for review of abstract again by myself and PT independently. Subsequently, the full texts of 109 articles were assessed for eligibility by myself and PT (blinded to each other's decisions), and then after resolving any discrepancies between us, 16 articles remained for inclusion in the data analysis. This comprised 3 conference abstracts and 13 full-text published papers (Figure 3.1).

3.5.2 Study characteristics

Of the sixteen included studies, some investigated a single non-invasive test, and others included more than one. The single-test studies comprised of one on ELF (full paper), one on FIB4 (full paper), one on FibroTest (full paper) and 9 on

FibroScan (6 full papers, 3 conference abstracts). Of those that examined more than one test, one evaluated FIB4, FibroTest, ELF, and FibroScan (full paper), one evaluated FIB4 and FibroTest (full paper), another evaluated FIB4, FibroTest and FibroScan (full paper) and the last one evaluated both FibroTest and FibroScan (full paper).

The total number of participants with ArLD included in the analyses of these studies was 2,280 (median participants 118, range 45-289). Ten were prospective, three retrospective, one retro-prospective, and in two it was not entirely clear (both conference abstracts).

Full characteristics of each study and participants, with references, are displayed in Tables 3.5 and 3.6. Studies were conducted between 2008 and 2021. The participants' median age was 53 (range 48 to 57.1) and 76.7% were male (range 66 to 89%).

There was significant heterogeneity between studies. Firstly, the prevalence of advanced fibrosis and cirrhosis was variable, ranging from 36-80% of participants with ≥F3 (advanced) fibrosis and 15-66% with cirrhosis (≥F4) across the studies. Six out of the included eighteen studies did not investigate performance of the non-invasive-test in predicting F3 (advanced) fibrosis, instead examining F2 or F4. Different scoring systems were used for histological grading between studies: METAVIR (n=9), Ishak (n=1), Brunt (n=2), Kleiner (n=3), and Batts-Ludwig (n-1). Acceptable biopsy lengths and number of portal tracts also differed between

studies, with some accepting a minimum of 10mm, and others 15mm, and similarly >5 or >10 portal tracts.

The amount of alcohol consumption was also heterogenous between studies – with one study excluding patients who had consumed >50g/day over the preceding two months (250), and others recruiting only patients admitted for detox (251) or those consuming differing amounts of alcohol, for example >50g/day or >80g/day.

Where thresholds for detecting fibrosis were used, these also varied across the studies. For example, in the 10 out of 11 studies examining FibroScan performance for detecting F3 fibrosis that applied thresholds, each study reported different optimum thresholds for detection of advanced fibrosis, ranging from 8kPa to 17kPa. In the FibroTest studies, reported thresholds for F4 included 0.30, 0.58, 0.70 and 0.71 (only one study reported a threshold for F3). The two ELF studies did both report the same two thresholds (9.8 and 10.5) but one study reported these for advanced fibrosis (F3) (23), and the other for moderate fibrosis (236). Only one of the three studies reporting on FIB4 applied a threshold. The significant heterogeneity between studies precluded pooling of results in meta-analysis for ELF, FibroTest, and FIB4 where there were only few studies on each test. Thus, meta-analysis was only possible for FibroScan, which had the greatest number of included studies and lower heterogeneity ($I^2 < 50\%$). I summarised all included studies for all four non-invasive tests in a forest plot of AUROCs and associated 95% Confidence intervals where available (Figure 3.3).

Table 3.5 Baseline study characteristics

Study author, year, location (reference)	Publication type Retrospective, prospective, or retro-to-prospective	Aetiology	Alcohol consumption	Total no pts in study	Total no pts with ArLD included	NIT of interest investigated	Age	BMI	%male	ALT	Biopsy scoring system used length of biopsy(mm) No. of portal tracts. Time b/w index NIT and biopsy	≥ F3 Prevalence (%)	≥ F4 Prevalence (%)
Kim Moon Young, 2011, Korea (159)	Abstract Prospective	Alcohol only	'patients with alcoholic liver disease'	230	230	FibroScan	-	-	-	-	METAVIR - -	-	-
Voican, 2017, France (252)	Full paper Prospective	Alcohol only	≥80g/day over 5 yrs	193	180	FibroScan	48±0.7	23 ± 0.2	80	90.3 ± 5.8	BRUNT ≥10mm, ≥10 portal tracts <15 days apart	40	15
Thiele, 2018, Denmark (23)	Full paper Prospective	Alcohol only	>24g/day for women and >36g per day men for≥ 1yr	289	289	FibroScan, ELF, FibroTest, FIB4	Med 53 (IQR 13)	Med 26 (IQR 7)	74 PC, 75 SC	27±23 PC, 35±27 SC	KLEINER ≥10mm, ≥5 portal tracts Same day	40	15
Salavrakos, 2019, Belgium (160)	Full paper Prospective	Alcohol only	'alcohol dependent patients, ≥70g/day	118	118	FibroScan	52 ±10	25±5	66	69±50	METAVIR ≥15mm, ≥6 portal tracts <3 days apart	47	27.1

Study author, year, location (reference)	Publication type Retrospective, prospective, or retro-to-prospective	Aetiology	Alcohol consumption	Total no pts in study	Total no pts with ArLD included	NIT of interest investigated	Age	BMI	%male	ALT	Biopsy scoring system used length of biopsy(mm) No. of portal tracts. Time b/w index NIT and biopsy	≥ F3 Prevalence (%)	≥ F4 Prevalence (%)
Nguyen-Khac 2008, France (161)	Full paper Prospective	Alcohol only	>50g/day for 5 years	103	103	FibroScan, FibroTest	52.6 ±9.6	27.7 ±5.9	74	61.7 ±59.3	METAVIR Average 7.8mm±2.7 portal tracts Same day	51.4	32
Naveau, 2009, France (148)	Full paper Retrospective	Alcohol only	≥50g/day for 1 year	218	218	FibroTest FIB4	47 ±0.7	-	78	65 ±5	METAVIR - <1 month apart	41	31
Naveau, 2014, France (253)	full paper Prospective	Alcohol only	≥50g/day for 1 year	200	200	FibroTest	51± 0.7	-	79.5	89±5	METAVIR Average 12 ±0.4 <1 week	36	27
Nahon, 2008, France (254)	Full paper Prospective	Alcohol only	>80g/day >10yrs	147	147	FibroScan	54.4 ±8.9	25.6 ±4.4	76.1	56.9 ±40.8	BRUNT ≥10mm unless cirrhosis. Same day	74.9	53.7
Kim, 2009, Korea (255)	Abstract ^ Unclear	Alcohol only	'with alcoholic liver disease'	45	45	FibroScan	-	-	-	-	Batts-Ludwig - -	80	64.4
Hien, 2018, Vietnam (163)	Abstract Unclear	Alcohol only	'patients with ArLD'	93	93	FibroScan	-	-	-	-	METAVIR - 'Concomitant'	60.2	26.9

Study author, year, location (reference)	Publication type Retrospective, prospective, or retro-to-prospective	Aetiology	Alcohol consumption	Total no pts in study	Total no pts with ArLD included	NIT of interest investigated	Age	BMI	%male	ALT	Biopsy scoring system used length of biopsy(mm) No. of portal tracts. Time b/w index NIT and biopsy	≥ F3 Prevalence (%)	≥ F4 Prevalence (%)
Fernandez 2015, Brussels (256)	Full paper Retrospective	Alcohol only	>50g/day for >5yrs	135	135	FibroScan, FibroTest, FIB4	56± 0.9	26.1 ±0.5	70	43 (29-70)	METAVIR Average 19.9 ±0.8 -	65	41
Cho, 2020, Korea (257)	Full paper Prospective Prospective	Alcohol only	>60g/day for males, >40g/day females	251	251	FIB4 (sheerwave elastography)	55.8 ±10.8	23.3 ±4.3	89	46.3 ±59.5	KLEINER - Within 72HRS ≥15mm, ≥8portal tracts Within 1wk	69.7	57.4
Mueller, 2010, Germany (258)	Full paper Prospective Retrospective	Alcohol only	Mean 146.8g/d (SD100.8)	101	101	FibroScan	53.2 ±10.6	25.4 ±4.2	72	90.2 (SD 133.7)	Kleiner >15mm Same day mean 22mm ±10 <6 months	44.5	25.7
Janssens 2010, Belgium (251)	Full paper Prospective	Alcohol only	Pts admitted for alcohol detox, >70g/day	255	49	FibroScan, APRI, Forns	53 (29-73)	25 (17-38)	69.4	62±36.6	METAVIR >15mm+ 6 portal tracts <3 weeks	66	41

Study author, year, location (reference)	Publication type Retrospective, prospective, or retro-to-prospective	Aetiology	Alcohol consumption	Total no pts in study	Total no pts with ArLD included	NIT of interest investigated	Age	BMI	% male	ALT	Biopsy scoring system used length of biopsy(mm) No. of portal tracts. Time b/w index NIT and biopsy	≥ F3 Prevalence (%)	≥ F4 Prevalence (%)
Connoley 2021, UK (134)	Full paper Retro-Prospective	Mixed	Pts with ArLD	786	81	ELF, APRI, AST: ALT,	50 (41.5-57.5)	-	67.9	36 (23-66)	Ishak >15mm +9 portal tracts ELF up to 14 days prior to biopsy	72.8	66.6
Reiberger, 2012, Austria (250)*	Full paper Prospective	mixed	(Exclusion included alcohol >50g/d within previous 2 months)	695	227 (40 with biopsy)	FibroScan	57.1 (±11.4)		85.0% (M:34, W:6)	-	METAVIR ≥10mm, 10 portal tracts Within 3 days	77.5% (31/40)	65.0% (26/40)

g/d = grams per day, ± = standard deviation, M = men, W = women, ELF = Enhanced Liver Fibrosis test, ALT = Alanine Aminotransferase, AST = Aspartate Aminotransferase, APRI = AST to Platelet ratio index, Pts = patients, ArLD = Alcohol-related Liver disease, SD = standard deviation, FIB4 = Fibrosis-4 score, HRS = hours, PC = primary care, SC = secondary care, NIT = non-invasive test, BMI = Body Mass Index

*****Data provided from direct communication with first author of the study.

Table 3.6 Diagnostic performance of FIB4, FibroTest, FibroScan, ELF

Degree of fibrosis tested	Study	No.	AUROC (95%CI)	Threshold used	Sens (%)	Spec (%)	PPV	NPV	LR+ (95%CI)	LR- (95%CI)	Details of any NIT failures or adverse events
FIB4											
Cirrhosis (≥F4 vs F0123)	Thiele	289	0.89 (0.86-0.93)	≥3.25	-	-	-	-	-	-	-
	Naveau, 2009, France	218	0.80 (0.72-0.86)	Cont.	-	-	-	-	-	-	-
	Fernandez 2015, Brussels	123	0.73 (0.63-0.82)								
	Cho, 2020, Korea	251	0.75 (0.69-0.82)	Cont.	-	-	-	-	-	-	-
Advanced fibrosis (≥F3 vs F012)	Thiele 2018, Denmark	289	0.85 (0.8-0.9)	≥3.25	58	91	64	88	6.09	0.47	FT: 4 failures TE: 6 unreliable, 7 failures, 7 cases where equipment not available
	Fernandez 2015, Brussels	123	0.70 (0.60-0.80)	Cont.							12 TE failures
	Cho, 2020, Korea	251	0.83 (0.77-0.89)	Cont.	-	-	-	-	-	-	-
Significant fibrosis (≥F2 vs F01)	Thiele 2018, Denmark	289	0.77 (0.71-0.83)	≥3.25	-	-	-	-	-	-	-
	Naveau, 2009, France	218	0.70 (0.62-0.76)	Cont.	-	-	-	-	-	-	-
	Cho, 2020, Korea	251	0.88 (0.83-0.97)	Cont.	-	-	-	-	-	-	-

Degree of fibrosis tested	Study	No.	AUROC (95%CI)	Threshold used	Sens (%)	Spec (%)	PPV	NPV	LR+ (95%CI)	LR- (95%CI)	Details of any NIT failures or adverse events
FibroTest											
Cirrhosis (≥F4 vs F0123)	Thiele, 2018, Denmark	289	0.88 (0.83-0.92)	0.58 ITD	-	-	-	-	-	-	-
	Thiele, 2018, Denmark	289	0.89 (0.85-0.93)	0.58 PP	-	-	-	-	-	-	-
	Nguyen-Khac 2008, France	103	0.84 (0.72-0.97)	Cont.	-	-	-	-	-	-	-
	Naveau, 2009, France	218	0.94 (0.90-0.96)	0.30	100	50.3	47.2	100	-	-	-
	Naveau, 2009, France	218	0.94 (0.90-0.96)	0.70	86.6	86	73.4	93.5	-	-	-
	Naveau, 2014, France	200	0.86 (0.78-0.91)	0.71	78	79	58	91	-	-	-
	Fernandez 2015, Brussels	123	0.88 (0.81-0.94)	Cont.							12 TE failures
Advanced fibrosis (≥F3 vs F012)	Thiele, 2018, Denmark	289	0.88 (0.84-0.92)	0.58 ITD	67	87	60	90	5.13 (3.51-7.5)	0.38 (0.27-0.53)	
	Thiele, 2018, Denmark	289	0.90 (0.86-0.94)	0.58 PP	67	89	64	90	5.84 (3.89-8.77)	0.38 (0.27-0.53)	
	Nguyen-Khac 2008, France	103	0.80 (0.70-0.91)	Cont.	-	-	-	-	-	-	-
	Fernandez 2015, Brussels	123	(0.81 (0.73-0.89)	Cont.							12 TE failures

Degree of fibrosis tested	Study	No.	AUROC (95%CI)	Threshold used	Sens (%)	Spec (%)	PPV	NPV	LR+ (95%CI)	LR- (95%CI)	Details of any NIT failures or adverse events
Significant fibrosis (≥F2 vs F01)	**Thiele, 2018, Denmark**	289	0.85 (0.81-0.90)	0.58 ITD	-	-	-	-	-	-	-
	Thiele, 2018, Denmark	289	0.86 (0.81-0.90)	0.58 PP	-	-	-	-	-	-	-
	Nguyen-Khac 2008, France	103	0.79 (0.69-0.90)	Cont.	-	-	-	-	-	-	-
	Naveau, 2009, France	218	0.83 (0.77-0.88)	0.30	87.7	52	75.6	71.6	-	-	-
	Naveau, 2009, France	218	0.83 (0.77-0.88)	0.70	42.9	97	96.3	50.1	-	-	-
	Naveau, 2014, France	200	0.80 (0.73-0.85)	0.71	59	91	89	65			
FibroScan											
Cirrhosis (≥F4 vs F0123)	**Kim, Moon Young, 2011, Korea**	230	0.729	Cont.	-	-	-	-		-	-
	Voican, 2017, France	193	0.93 (0.88-0.97)	15kpa	93.1	85.4	52.9	98.6	-	-	2 patients excluded with unreliable TE (22 excluded due to poor biopsy quality)
	Thiele, 2018, Denmark	289	0.87 (0.79-0.95)	15kpa ITD	-	-	-	-	-	-	-
	Thiele, 2018, Denmark	289	0.97 (0.95-0.99)	15kpa PP	-	-	-	-	-	-	-
	Salavrakos, 2019, Belgium	118	0.907	21.2kpa	81	85	-	-	-	-	-
	Salavrakos, 2019, Belgium	118	0.907	19.5kpa	84	79	60	93			-
	Nguyen-Khac 2008, France	103	0.92 (0.87-0.98)	19.5kpa	85.7	84.2	68.6	87.9	-	-	2 x TE failures

Degree of fibrosis tested	Study	No.	AUROC (95%CI)	Threshold used	Sens (%)	Spec (%)	PPV	NPV	LR+ (95%CI)	LR- (95%CI)	Details of any NIT failures or adverse events
	Nahon, 2008, France	147	0.87 (0.81-0.93)	22.7kpa	0.84	0.83	0.85	0.82	-	-	15x inadequate TE, 12x inadequate biopsy
	Kim, 2009, Korea	45	0.97 (0.93-1.01)	25.8kpa	90	87	-	-	-	-	-
	Fernandez 2015, Brussels	135	0.93 (0.90-0.97)	18.0kpa	90	86	82	93			12 TE failures
	Mueller, 2010, Germany	101	0.921 (0.03)	11.5kpa	1	0.77	-	-			-
		101	0.921 (0.03)	12.5 kpa	0.96	0.8					
		86	0.944 (0.02) (Excluding GOT >100u/l)	11.5kpa	1	0.84	-	-			
		66	0.945 (0.03) (Excluding GOT >50u/l)	10.4	1	0.87	-	-			
	Janssen, 2010, Belgium	49	0.864	21.1kpa	75	80	-	-			TE failure due to obesity or ascites in 11
	Reiberger, 2012, Austria	40	0.937 (0.862-1)	12.1kpa	80.8	100	100	73.7	-	0.19 (0.09 to 0.42)	-
Advanced fibrosis (≥F3 vs F012)	**Kim, Moon Young, 2011, Korea**	230	0.884	Cont.	-	-	-	-	-	-	-
	Voican, 2017, France	193	0.90 (0.83-0.93)	12kpa	75.6	92.2	86.8	84.8	-	-	2 patients excluded with unreliable TE (22 excluded due to poor biopsy quality)
	Thiele. 2018, Denmark	289	0.89 (0.83-0.96)	15kpa ITD	86	94	80	96	13.38 (7.98-22.43)	0.15 (0.8-0.28)	2 had major bleeding from biopsy, 4 TE failures

Degree of fibrosis tested	Study	No.	AUROC (95%CI)	Threshold used	Sens (%)	Spec (%)	PPV	NPV	LR+ (95%CI)	LR- (95%CI)	Details of any NIT failures or adverse events
	Thiele. 2018, Denmark	289	0.97 (0.95-0.99)	15kpa PP	91	95	84	98	19.28 (10.43-35.32)	0.09 (0.04-0.21)	2 had major bleeding from biopsy,14 TE failures
	Salavrakos, 2019, Belgium	118	0.886	15.2kpa	78	83	-	-	-	-	Biopsy: 30 refused, 1 sample lost, 6 technically impossible, 14 poor quality TE: 8 refused, 19 unsuccessful
		118	0.886	11kpa	95	57	66	92			
	Nguyen-Khac 2008, France	103	0.90 (0.82-0.97)	11kpa	86.7	80.5	81.8	84.3	-	-	2x TE failures
	Nahon, 2008, France	147	(0.94 (0.90-0.97)	11.6kpa	0.87	0.89	0.96	0.70	-	-	15x inadequate TE, 12x inadequate biopsy
	Hien, 2018, Vietnam	93	0.91	11.3kpa	86.8	81	82	85	-	-	-
	Fernandez 2015, Brussels	135	0.89 (0.83-0.95)	10.3kpa	91	67	76	87	-	-	12 TE failures
	Mueller, 2010, Germany	101	0.914 (0.03)	8kpa	0.91	0.75	-	-			5 TE failures
		80	0.922 (0.03) (excluding GOT >100u/l)	8kpa	0.87	0.87	-	-			
		67	0.946 (0.03) (excluding GOT >50u/l)	8kpa	1	0.84	-	-			
	Janssen, 2010, Belgium	49	0.766	17kpa	72	76.5	-	-			TE failure due to obesity or ascites in 11
	Reiberger, 2012, Austria	40	0.905 (0.81-0.999)	9.6kpa	80.7	77.8	92.6	53.9	3.63 (1.06-12.47)	0.25 (0.11-0.55)	TE failure in 67/794 of total study cohort)

Degree of fibrosis tested	Study	No.	AUROC (95%CI)	Threshold used	Sens (%)	Spec (%)	PPV	NPV	LR+ (95%CI)	LR- (95%CI)	Details of any NIT failures or adverse events
Significant fibrosis (≥F2 vs F01)	**Thiele 2018., Denmark**	289	0.85 (0.81-0.90)	-	-	-	-	-	-	-	-
	Thiele 2018, Denmark	289	0.88 (0.84-0.92)	-	-	-	-	-	-	-	-
	Hien, 2018, Vietnam	93	0.86	7.9kpa	80	90	91	72	-	-	-
	Nguyen-Khac	103	0.91 (0.85-0.98)	7.8kpa	80	90.5	93	70	-	-	-
	Reiberger, 2012, Austria	40	0.974 (0.923-1)	7.2kpa	89.5	100	100	33.33	-	0.11 (0.04-0.27)	
ELF											
Cirrhosis (≥F4 vs F0123)	**Thiele 2018, Denmark**	289	0.94 (0.91-0.97)	-	-	-	-	-	-	-	-
	Connoley 2021, UK	81	0.895 (0.823-0.968)	9.8	91	63	83	77	2.45 (1.49-4.04)	0.15 (0.06-0.36)	-
				10.5	85	89	94	75	7.67 (2.62-22.41)	0.17 (0.09-0.32)	-
				11.3	67	93	95	58	9 (2.34-34.61)	0.36 (0.24-0.53)	-
Advanced fibrosis (≥F3 vs F012)	**Thiele 2018, Denmark**	289	0.92 (0.89-0.96)	≥10.5	79	91	71	94	8.37 (5.46-12.81)	0.23 (0.15-0.37)	-
	Thiele 2018, Denmark	289	0.92 (0.89-0.96)	≥9.8	89	78	54	96	3.99 (3.08-5.16)	0.14 (0.07-0.28)	-

Degree of fibrosis tested	Study	No.	AUROC (95%CI)	Threshold used	Sens (%)	Spec (%)	PPV	NPV	LR+ (95%CI)	LR- (95%CI)	Details of any NIT failures or adverse events
	Connoley 2021, UK	81	0.824 (0.787-0.861)	-	-	-	-	-	-	-	-
Significant fibrosis (≥F2 vs F01)	**Thiele 2018, Denmark**	289	0.84 (0.80-0.89)	-	-	-	-	-	-	-	-
	Connoley 2021, UK	81	0.923 (0.866-0.981)	8.3	97	28	82	71	1.34 (1-1.79)	0.11 (0.02-0.54)	-
				9.8	88	83	95	68	5.33 (1.89-15.04)	0.13 (0.06-0.28)	-

Sens = sensitivity, Spec = specificity, NPV = Negative Predictive Value, PPV = Positive Predictive Value, LR+ = Positive likelihood ratio, LR- = negative likelihood ratio, kPa = kilopascals, 95% CI = 95% confidence interval, TE = transient elastography, FT = FibroTest, NIT = non-invasive test, AUROC = area under receiver operator characteristic curve

Table 3.7: Influence of alcohol, steatosis, obesity, age and inflammation on FIB4/FibroTest/FibroScan/ELF

NIT	Study	No. of participants	Was alcohol withdrawal or recent alcohol consumption reported to influence the NIT result?	Was degree of inflammation (E.g., CRP/AST/ALT) reported to influence the NIT result?	Was participant age reported to influence NIT result?	Did degree of steatosis influence NIT result?	Was the presence of comorbid obesity in participants found to influence NIT result?
FIB4	Naveau, 2009, France	218	NR	NR	NR	NR	NR
FIB4	Fernandez 2015, Brussels	123	NR	NR	NR	NR	NR
FIB4	Cho, 2020, Korea	251	NR	NR	NR	NR	NR
FIB4	Thiele, 2018, Denmark	289	NR	NR	NR	NR	NR
FibroTest	Thiele, 2018, Denmark	289	Drinking pattern did not influence FibroTest value	Histological inflammation and AST predicted FibroTest value	Increased age predicted increased FibroTest value	NR	BMI predicted FibroTest independent of histology
FibroTest	Naveau, 2009, France	218	NR	No – correlation with biopsy persisted after adjustment for alcoholic hepatitis	NR	NR	NR
FibroTest	Naveau, 2014, France	200	NR	Ns- but pts with F2-F4 had higher levels of alcoholic hepatitis, and ALT higher in F2 and above compared to F0-1	NR	NR	NR

NIT	Study	No. of participants	Was alcohol withdrawal or recent alcohol consumption reported to influence the NIT result?	Was degree of inflammation (E.g., CRP/AST/ALT) reported to influence the NIT result?	Was participant age reported to influence NIT result?	Did degree of steatosis influence NIT result?	Was the presence of comorbid obesity in participants found to influence NIT result?
FibroTest	Nguyen-Khac	103	NR	NR	NR	NR	NR
FibroTest	Fernandez 2015, Brussels	123	NR	NR	NR	NR	NR
FibroScan	Kim, Moon Young, 2011, Korea	230	NR	NR	NR	NR	NR
FibroScan	Voican, 2017, France	193	Yes, TE values significantly decreased when repeated 1/12 after alcohol withdrawal	Yes, presence of AH on biopsy increased false positive rate of TE	NR	Yes – significant correlation between steatosis & FS	NR
FibroScan	Thiele, 2018, Denmark	289	Drinking pattern did not influence TE value	Histological inflammation predicted TE value	NR	NR	NR
FibroScan	Salavrakos, 2019, Belgium	118	Yes, 57/118 had repeat TE 14d after abstinence, with mean reduction of TE value by 2.7kpa and improvement in correct histology classification from 40% to 61%	19/118 had histological AH. 42% (8/19) were misclassified by 1 stage, but no observed association b/w misclassification and having AH, so concluded mild-mod AH does not contribute to misclassification by TE. ALT/AST did not correlate with FS.	NR	Mod-severe steatosis did not influence TE results	NR

NIT	Study	No. of participants	Was alcohol withdrawal or recent alcohol consumption reported to influence the NIT result?	Was degree of inflammation (E.g., CRP/AST/ALT) reported to influence the NIT result?	Was participant age reported to influence NIT result?	Did degree of steatosis influence NIT result?	Was the presence of comorbid obesity in participants found to influence NIT result?
FibroScan	Nguyen-Khac 2008, France	103	NR	No correlation b/w liver stiffness and ALT/AST	NR	No correlation observed b/w liver stiffness and steatosis (r = 0.064, p =0.52).	NR
FibroScan	Nahon, 2008, France	147	NR	Yes – alcoholic hepatitis increased FS in univariate analysis, but not in multivariate analysis	NR	No – steatosis did not influence FS	NR
FibroScan	Kim, 2009, Korea	45	NR	NR	NR	NR	NR
FibroScan	Hien, 2018, Vietnam	93	NR	NR	NR	NR	NR
FibroScan	Fernandez 2015, Brussels	135	NR	LS values significantly higher in pts with AST >50	NR	NR	NR

NIT	Study	No. of participants	Was alcohol withdrawal or recent alcohol consumption reported to influence the NIT result?	Was degree of inflammation (E.g., CRP/AST/ALT) reported to influence the NIT result?	Was participant age reported to influence NIT result?	Did degree of steatosis influence NIT result?	Was the presence of comorbid obesity in participants found to influence NIT result?
FibroScan	Mueller 2010, Germany	101	Yes, Interval LS decreased significantly on withdrawal of alcohol over mean period of 5.3d. (mean decrease in LS 3.5kpa, Max decrease 26.3kpa.	Yes, excluding patients with high AST increased ROC and sens and spec of FS in diagnosing cirrhosis. FS correlated with GOT levels, LS remained stable once GOT levels <100u/l on withdrawal of alcohol, and FS accuracy improved when excluding pts with GOT >100u/l	NR	No – excluding those with steatohepatitis did not improve diagnostic accuracy	NR
FibroScan	Janssen 2010, Belgium	49	NR	No, 6 patients had histological ASH, of which FibroScan correctly classified 3. Inflammation was not thought to significantly affect TE in this study.	NR	Yes, steatosis influenced TE- Of 11 patients with severe steatosis, FibroScan overestimated fibrosis in 7 patients.	NR

NIT	Study	No. of participants	Was alcohol withdrawal or recent alcohol consumption reported to influence the NIT result?	Was degree of inflammation (E.g., CRP/AST/ALT) reported to influence the NIT result?	Was participant age reported to influence NIT result?	Did degree of steatosis influence NIT result?	Was the presence of comorbid obesity in participants found to influence NIT result?
FibroScan	Reiberger, 2012, Austria	40	NR	NR	NR	NR	NR
ELF	Thiele, 2018, Denmark	289	Drinking pattern did not influence ELF	Histological inflammation predicted ELF value. AST/ALT didn't influence ELF	Yes, increased false positives in >60s and increased false negatives in <30s	NR	BMI predicted ELF independent of histology
ELF	Connoley 2021, UK	81	NR	Adding ALT to regression model containing ELF did not alter performance of model.	NR	NR	NR

NIT = non-invasive test, NR = Not reported, ALT = alanine aminotransferase, AST = aspartate aminotransferase, ELF = Enhanced Liver Fibrosis test, BMI = Body Mass Index, TE = transient elastography, ASH = alcoholic steatohepatitis, FS = FibroScan, AH = Alcoholic Hepatitis, GOT = glutamic oxaloacetic transaminase

3.5.3 Risk of bias within studies

Whilst there were no studies that were scored as high risk, only 4 studies were at low-risk in all 7 domains.

The majority of 'unclear risk' scores were in the reference standard 'risk of bias' domain, where 10/16 were given this grade. The main reasons for this were if not all participants meeting selection criteria were included in the analysis, or if average biopsy length was <15mm or <6 portal tracts.

There were three studies that were scored 'unclear' risk for the majority of the domains (159, 163, 255) and this may be because these three studies were conference abstracts, and so may have omitted the necessary information for us to be able to score them 'low risk' due to word-count restrictions in the abstract.

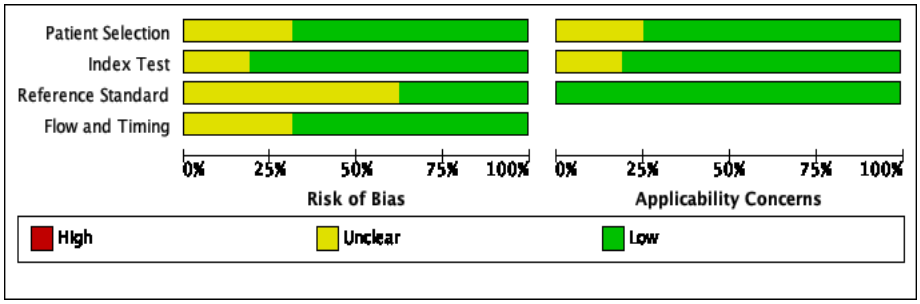
Overall, 73.2% of all the domains within the included studies were rated 'low-risk', and 26.8% 'unclear' risk (see Figure 3.2 and Table 3.8). Cohen's kappa (κ) was measured to assess for agreement between the first and second reviewers' decisions on rating the study domains as 'low risk', 'unclear risk' and 'high risk'. This showed 'good' agreement (259) with just over 10% of decisions differing between the two reviewers (myself and PT), and resulting in a Cohen's kappa of 0.761 (95% CI 0.641 to 0.881), $p < 0.001$.

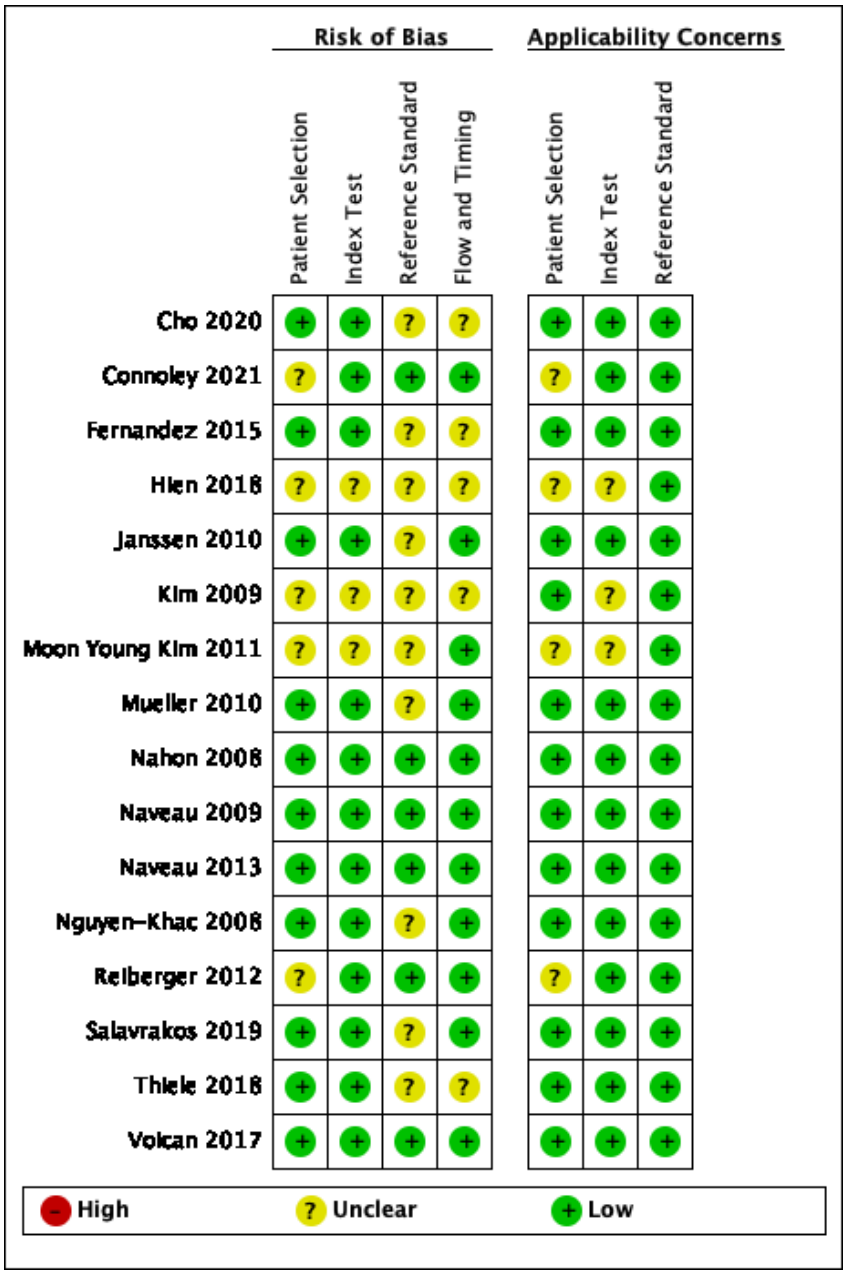
TABLE 3.8. QUADAS2 RISK OF BIAS RESULTS:

Study	Risk of bias				Applicability concerns		
	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Thiele, 2018	L	L	U	U	L	L	L
Naveau, 2009	L	L	L	L	L	L	L
Fernandez 2015	L	L	U	U	L	L	L
Cho, 2020	L	L	U	U	L	L	L
Nguyen-Khac 2008	L	L	U	L	L	L	L
Naveau, 2013	L	L	L	L	L	L	L
Kim, Moon Young, 2011	U	U	U	L	U	U	L
Voican, 2017	L	L	L	L	L	L	L
Salavrakos, 2019	L	L	U	L	L	L	L
Nahon, 2008	L	L	L	L	L	L	L
Kim, 2009	U	U	U	U	L	U	U
Hien 2018	U	U	U	U	U	U	L
Mueller 2010	L	L	U	L	L	L	L
Janssen, 2010	L	L	U	L	L	L	L
Reiberger, 2012	U	L	L	L	U	L	L
Connoley 2021	U	L	L	L	U	L	L

H = high risk, L = Low risk, U = Unclear

FIGURE 3.2
QUADAS2 RISK OF BIAS
Risk of bias graph for the 16 included articles





3.5.4 Diagnostic performance of each non-invasive test

3.5.5 FIB4

Four studies (all full papers) evaluated the diagnostic performance of FIB4 in ArLD. All four evaluated F4 (cirrhosis), with three also evaluating F2 or F3 fibrosis. All studies reported AUROCs (with 95% CIs) for continuous FIB4 data, but only one study, by Thiele et al., applied a FIB4 threshold (3.25) with associated sensitivity and specificity results for F3 fibrosis (Table 3.6). FIB4 performed similarly at each fibrosis stage, with AUROCs $\geq$ 0.70 in all four studies. For F2, AUROCs were 0.77 (95%CI 0.71-0.83), 0.70 (95% CI 0.62-0.76), 0.88 (95% CI 0.83-0.97), for F3: 0.85 (95%CI 0.80-0.90), 0.70 (95%CI 0.60-0.80), 0.83 (95%CI 0.71-0.83), and for F4: 0.89 (95%CI 0.86-0.93), 0.80 (95%CI 0.72-0.86), 0.73 (95%CI 0.63-0.82), and 0.75 (95%CI 0.69-0.82). There were no reported test failures. Heterogeneity was significant across the four studies (I^2 78% for F3 fibrosis). The prevalence of each reported fibrosis stage varied across the four studies, with prevalence of cirrhosis (F4) ranging from 15 to 57%, and varied biopsy length and number of portal tracts (Table 3.5).

In all three studies where FIB4 was directly compared with another fibrosis test, FIB4 performed inferiorly to the comparator non-invasive test. In the study by Thiele et al. (23), FIB4 performed well (AUROC 0.85) but significantly inferiorly to ELF (AUROC 0.92), for F3 (p = 0.003), and the same pattern for F2/F4 (23). In the study by Fernandez et al. (256), FIB4 performed inferiorly to FibroTest and FibroScan for F3 and F4 fibrosis detection, with an AUROC of 0.70 for F3, compared with 0.89 (FibroScan) and 0.81 (FibroTest). Finally, Naveau et al. found that FIB4

performed inferiorly to FibroTest at F2 and F4 (FIB4 AUROC of 0.70 for F3 compared with 0.83 FibroTest, p = 0.0007) (148).

None of the four studies examined any impact of inflammation/BMI/alcohol consumption/steatosis on FIB4 scores.

3.5.6 ELF

Two full papers reported on ELF in ArLD in 2018 and 2021 (23, 236). Both presented results with AUROCs and sensitivity and specificity data for the same ELF thresholds (9.8 and 10.5). However, I found considerable heterogeneity between the two studies (I^2 94%). Prevalence of F4 (cirrhosis) was 67% in one study, and 15% in the other, with the latter study including community patients from a rehab centre, and the former study only including secondary-care patients. Both studies had biopsy length ≥10mm, with ≥ 5 portal tracts. Both reported excellent ELF performance for moderate fibrosis (F2), advanced fibrosis (F3) and cirrhosis (F4), but only Thiele *et al.* reported sensitivity/specificity results for advanced fibrosis (F3), with Connoley *et al.* focussing on moderate fibrosis and cirrhosis. Thiele *et al.* evaluated 289 patients with ArLD in a biopsy-paired study, reporting an (intention-to-diagnose) ELF AUROC of 0.92 (95% CI 0.89-0.96) for F3, 0.94 (95% CI 0.91-0.97) for F4, and 0.84 (95% CI 0.80-0.89) for F2. At a threshold of 10.5, ELF had a NPV of 94% for F3 (advanced) fibrosis (sensitivity 79%, specificity 91%), and at a threshold of 9.8, NPV was 96% (sensitivity 89%, specificity 78%), with no difference in its performance between primary and secondary care patients (23).

The authors concluded that in ArLD, “ELF diagnosed advanced fibrosis with excellent discriminatory accuracy", and could be used safely and effectively at a threshold of 10.5 to evaluate for advanced fibrosis and allow triage of patients from primary to secondary care.

Connoley et al. discovered similar findings in a biopsy-paired cohort of 81 patients with ArLD, with AUROC of 0.90 (0.82-0.97) for F4, 0.82 (0.79-0.86) for F3 and 0.84 (0.80-0.89) for F2. ELF thresholds were evaluated at 8.3, 9.8 and 10.5, with authors concluding that a 10.5 threshold was highly specific to diagnose cirrhosis (specificity 94%, sensitivity 37%, NPV 87%, PPV 59%), and an 8.3 threshold was sensitive to rule out moderate fibrosis (sensitivity 78%, specificity 50%, NPV 80%, PPV 59%). The use of a 9.8 or 10.5 threshold for advanced fibrosis was not explored in this study.

No failure rate for ELF was reported. In terms of factors influencing ELF score, Thiele et al. did not find any association between alcohol drinking pattern and ELF, and whilst ALT/AST did not influence ELF score, they did find an association between histological inflammation and ELF (23). Connoley et al. did not evaluate histological inflammation against ELF, but found that adding ALT to a regression model containing ELF did not influence the performance of ELF (Table 3.7). ELF scores were more likely to be falsely positive in people aged over 60, and falsely negative in people aged under 30 in the Thiele et al. study, whilst impact of age on ELF was not reported in the Connoley study. Neither study evaluated impact of steatosis on ELF score.

3.5.7 FibroTest

Five studies reported on FibroTest – all were full-text papers from 2008 to 2018. There was significant heterogeneity between the studies (I^2 56% for F4, 48% for F3 fibrosis), and absence of sensitivity and specificity data for the majority of the studies (Table 2). Prevalence of cirrhosis (F4) varied from 15% to 41% across the studies, and biopsy length ranged from 7.8mm to 20mm. Performance of FibroTest was slightly better at F4 compared to F2 fibrosis stage, with AUROCs ranging from 0.79 to 0.85 for F2, 0.80 to 0.90 for F3, and 0.84 to 0.94 for F4 fibrosis. FibroTest was directly compared with another fibrosis test in four of the five studies. Thiele et al. found that FibroTest performed similarly to ELF and FibroScan (intention-to-diagnose), and superiorly to FIB4, with AUROC of 0.88 for F3 (advanced fibrosis) (95% CI 0.84-0.92) (sensitivity 67%, specificity 87%). Nguyen-Khac et al. compared FibroTest with FibroScan, finding numerically-higher AUROCs for FibroScan than FibroTest at F1/F2/F3/F4 fibrosis stages, but this was only significantly higher for F2 (FibroScan AUROC 0.91 (95% CI 0.85-0.98), FibroTest AUROC 0.79 (95% CI 0.69-0.90), p = 0.04 (161).

Fernandez et al. also compared FibroTest against FibroScan, finding that FibroScan outperformed FibroTest and FIB4 at F3 and F4 fibrosis stages, although the authors did not report any significance testing for AUROC comparisons. (FibroScan AUROC was 0.89 (95% CI 0.83-0.95) for F3, and 0.93 (95% CI 0.90-0.97) for F4, compared with FibroTest AUROCs of 0.81 (95% CI 0.73-0.89) for F3 and 0.88 (95% CI 0.81-0.94) for F4). FIB4 AUROCs were even lower at 0.70 (95% CI 0.60-0.80) for F3, and 0.73 (95% CI 0.63-0.82) for F4 (256). In both of these studies latterly described, (by

Fernandez et al., and Nguyen-Khac et al.) combining FibroTest with FibroScan did not improve the performance compared to FibroScan alone.

Finally, Naveau et al. compared FibroTest with FIB4, finding FibroTest performed significantly better than FIB4 at F2 and F4 fibrosis stages (p = 0.0007), with AUROC of 0.83 (95% CI 0.77-0.88) for F2 versus 0.70 for FIB4 (95% CI 0.62-0.76), and 0.94 (95% CI 0.90-0.96) for F4 versus 0.80 for FIB4 (95% CI 0.72-0.86) (148).

No failures of FibroTest were reported.

Two studies (23, 148) investigated for an association between inflammation and FibroTest score, with Naveau et al. finding that the positive correlation between FibroTest and fibrosis stage persisted after adjustment for the presence of acute alcoholic hepatitis (data not reported). Thiele et al., however, found that FibroTest score was significantly influenced by AST levels and histological inflammation. Only one study (Thiele et al.) investigated effect of alcohol on FibroTest, finding that alcohol intake did not affect FibroTest result (23) (Table 3.7).

3.5.8 FibroScan

FibroScan was the most widely studied non-invasive test out of the four included tests in this systematic review, with 12 studies reporting on FibroScan (of which 3 were conference abstracts, and the other 9 full papers). Again, there was significant heterogeneity between studies, with different thresholds used for every study. Three out of four studies that reported F2 end points applied a FibroScan threshold, ranging from 7.2 to 7.9 kPa. Ten out of the eleven studies reporting on F3 fibrosis

applied a threshold, ranging from 8 to 17 kPa, and ten out of eleven studies reporting on F4 fibrosis applied a threshold, ranging from 12.1 to 25.8. Prevalence of cirrhosis (F4) varied from 15 to 65% across studies, and advanced fibrosis (F3) prevalence from 21 to 78%. Biopsy length also varied from 10-19.9mm, with 5-10 portal tracts across the studies.

F2

For the detection of F2 fibrosis, four studies reported on this, with AUROCs ranging from 0.86-0.97 (only 3 out of 4 studies reported associated 95% Confidence Intervals). Heterogeneity between these 3 studies was significant (I^2 84%), so I did not perform meta-analysis for F2.

F3

For the detection of F3 fibrosis, 11 studies had data for this, with AUROCs ranging from 0.77-0.94. FibroScan thresholds differed for each study and ranged from 8-17 kPa. Sensitivities at the individual study thresholds ranged from 72-91%, and specificities from 67-94% for the 10/11 studies that included these data (total n=1,268). Heterogeneity testing revealed acceptable heterogeneity levels for sensitivity (I^2 26.9%), with pooled sensitivity of 0.85 (95% CI 0.81-0.89) (Figure 3.4A). However, heterogeneity was very high for specificity (I^2 78%), so results were not pooled for specificity. Because of this, I then performed a meta-analysis on AUROC for the 6/11 studies that included AUROC results with corresponding 95% Confidence Intervals. This found an acceptable level of heterogeneity between the six studies (I^2 13.4%), and pooled AUROC was 0.91 (95% CI 0.89-0.94) (Figure 3.4B).

F4

For F4 fibrosis, heterogeneity was low for sensitivity (I^2 2.2) and specificity (I^2 0.02), with meta-analysis finding a pooled sensitivity of 0.88 (95% CI 0.84-0.92) and pooled specificity of 0.84 (95% CI 0.81-0.87) across 8 studies, total n=891 (Figure 3.5A/B), although noting that again, FibroScan thresholds for each study differed.

Three studies directly compared the performance of FibroScan with the other included non-invasive tests. Thiele et al.'s prospective biopsy-paired study of n=289 (23) found FibroScan to perform better than ELF and FibroTest only when applying a per-protocol analysis, (AUROC 0.97, p 0.004 for advanced fibrosis, in comparison with ELF). However, they noted a 5% FibroScan failure-rate, and when they performed an intention-to-diagnose analysis, found that FibroScan performed equally to ELF and FibroTest (FibroScan AUROC 0.90 for advanced fibrosis), with a concluding recommendation that ELF or FibroTest could be used first-line for advanced fibrosis (using a 10.5 ELF threshold, and 0.58 FibroTest threshold) (23).

Nguyen-Khac et al. (161) found FibroScan outperformed FibroTest with numerically higher AUROCs at F3 and F4 fibrosis stages, and significantly higher AUROC at F2 (0.91, 95% CI 0.85-0.98, compared to 0.79, 95% CI 0.69-0.90). Finally, Fernandez et al. conducted a retrospective study of n=135, finding that FibroScan outperformed FibroTest and FIB4 for the diagnosis of advanced fibrosis, with AUROC of 0.89 (95% CI 0.83-0.95), compared to 0.81 (95% CI 0.73-0.89) and 0.70 (95% CI 0.60-0.80).

Failure rates for FibroScan were noted in all studies that reported failure rates (2 conference abstracts did not mention a failure rate), ranging from 1-22 % test failures.

FibroScan was noted to be influenced by recent alcohol consumption or withdrawal in 3 out of 4 studies that looked into this (increased false positives in those patients with recent alcohol intake, with reduction in liver stiffness scores observed 5, 14 and 28 days after stopping drinking (Table 3.7). Effect of age on FibroScan was not examined in any of the studies. With regards to inflammation (by histology or AST/ALT), five studies noted liver stiffness values were affected by inflammation, and three studies noted no association. The other 4 FibroScan studies did not investigate for this. Steatosis was observed to overestimate liver stiffness in 2 out of five studies examining for this, whereas no such association was noted in the other three.

Figure 3.3: Forest plot of AUROCs for all studies, all tests at F2/F3/F4 fibrosis stages

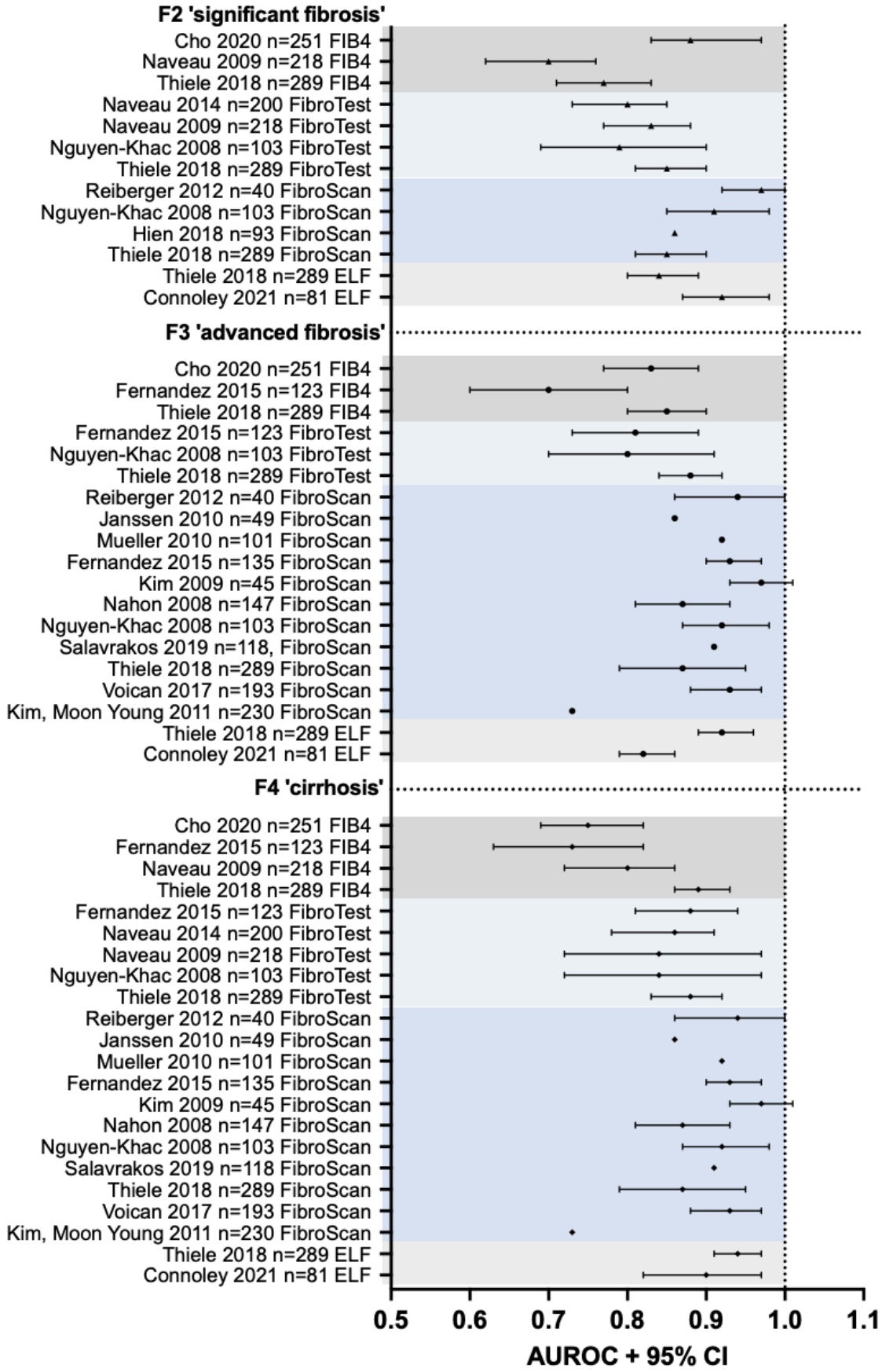
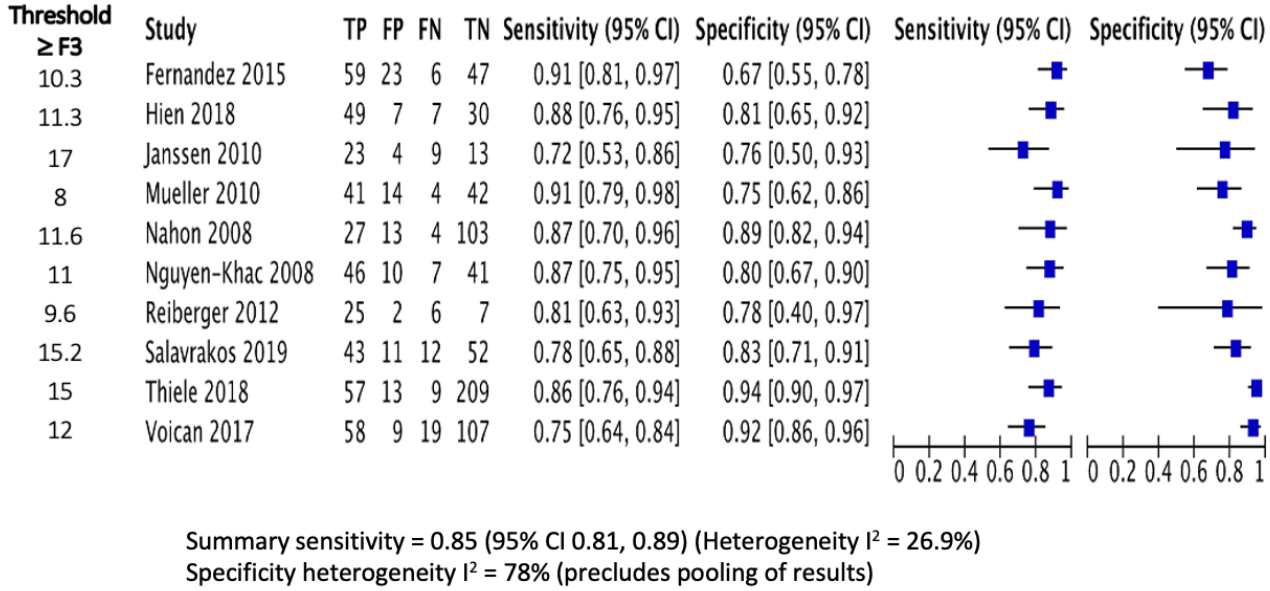


Figure 3.4 A/B FibroScan for the detection of ≥ F3 fibrosis

A)



Summary sensitivity = 0.85 (95% CI 0.81, 0.89) (Heterogeneity I^2 = 26.9%)
Specificity heterogeneity I^2 = 78% (precludes pooling of results)

B)

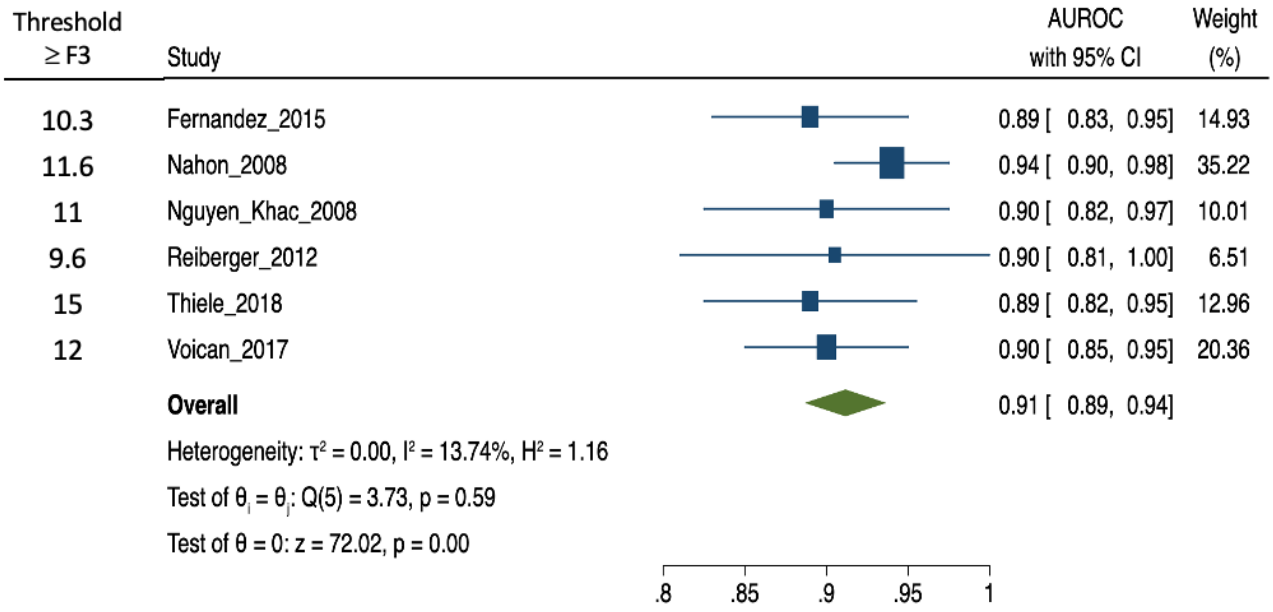
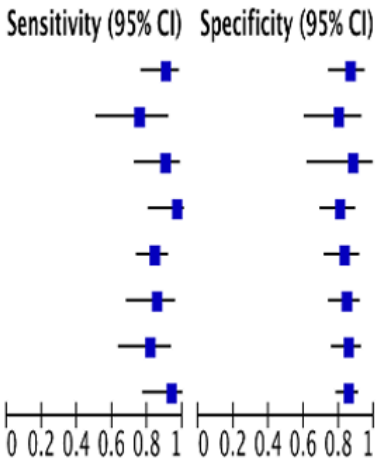


Figure 3.4 A/B:
Meta-analysis results of FibroScan for F3 (advanced fibrosis) with:
A) Forest plot of sensitivity and specificity, with pooled sensitivity and 95% confidence intervals, and between-study heterogeneity (I^2) and
B) Forest plot of AUROC (for the 6 available studies that had AUROC with associated 95% Confidence Intervals), with the green diamond representing pooled AUROC, with 95% Confidence intervals, and between-study heterogeneity (I^2) displayed.

Figure 3.5 A/B FibroScan for the detection of ≥ F4 fibrosis

A)

Threshold ≥ F4	Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
18	Fernandez 2015	36	8	4	49	0.90 [0.76, 0.97]	0.86 [0.74, 0.94]
21.1	Janssen 2010	15	6	5	23	0.75 [0.51, 0.91]	0.79 [0.60, 0.92]
25.8	Kim 2009	26	2	3	14	0.90 [0.73, 0.98]	0.88 [0.62, 0.98]
12.5	Mueller 2010	25	15	1	60	0.96 [0.80, 1.00]	0.80 [0.69, 0.88]
22.7	Nahon 2008	67	12	13	57	0.84 [0.74, 0.91]	0.83 [0.72, 0.91]
19.5	Nguyen-Khac 2008	28	13	5	67	0.85 [0.68, 0.95]	0.84 [0.74, 0.91]
21.2	Salavrakos 2019	26	13	6	73	0.81 [0.64, 0.93]	0.85 [0.76, 0.92]
15	Voican 2017	27	24	2	135	0.93 [0.77, 0.99]	0.85 [0.78, 0.90]



Summary sensitivity = 0.88 (95% CI 0.84, 0.92), (Heterogeneity I^2 = 2.19%)
Summary specificity = 0.84 (95% CI 0.81, 0.87), (Heterogeneity I^2 = 0.02%)

B)

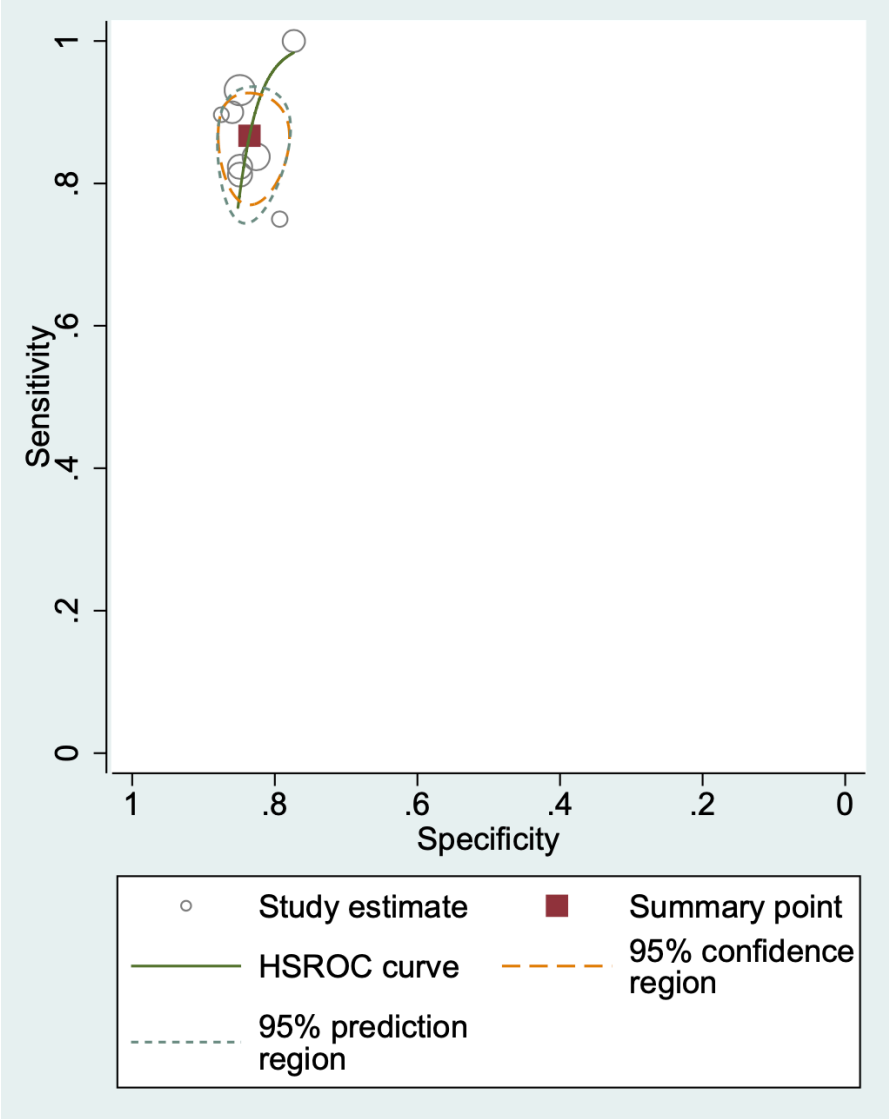


Figure 3.5 A/B: Meta-analysis results of FibroScan for F4 (cirrhosis) with:
A) Forest plot of sensitivity and specificity, with pooled sensitivity, specificity and 95% confidence intervals, and between-study heterogeneity (I2) and
B) SROC plot with circles representing individual studies and pooled sensitivity and specificity represented by the red square.

3.6 Discussion

3.6.1 Main findings

This is the first systematic review reporting on the diagnostic performance of four non-invasive tests (FIB4, ELF, FibroTest and FibroScan) in the detection of advanced fibrosis (F3) and cirrhosis (F4) for ArLD. Whilst a previous systematic review from 2012 investigated ELF, FibroTest and FibroScan in ArLD (223), it did not return any studies that were investigating ELF for ArLD at this time, and the total sample size for this review was n=989, compared to 2,280 in my current study.

Despite this significant increase in sample size in my study, this systematic review has highlighted that whilst there is now sound evidence for the diagnostic performance of non-invasive tests in mixed-aetiology liver disease (195, 260-265), there continues to be a relative paucity of studies looking at their performance in ArLD specifically. This is perhaps surprising, given ArLD is the most prevalent aetiology of liver disease, accounting for 60% of all liver disease cases (3). It is crucial that moving forward, studies on diagnostic accuracy of non-invasive tests are aetiology-specific, as their performance is known to vary according to the underlying disease aetiology, and test thresholds also differ between aetiologies (266).

All four non-invasive tests that I studied showed good diagnostic performance in detecting F2, F3 and F4 fibrosis, with AUROCs $\geq$ 0.7 for all studies and all tests at each fibrosis stage. FibroScan was the most studied of the four tests. Whilst performance of FibroScan and FIB4 appeared fairly consistent between fibrosis stages, FibroTest and

ELF both had generally higher AUROCs for F4 than F2 (FibroTest F4 AUROCs 0.84-0.94 versus 0.79-0.85 for F2, ELF F4 AUROCs 0.90-0.94 versus 0.84-0.84 for F2).

Whilst I found different test thresholds for each study, I note a previously published meta-analysis of FibroScan in ArLD by Pavlov et al. (229) that pooled results of 6 studies that used the same FibroScan threshold of 9.5Kpa for F3 and 12.5Kpa for F4. I had excluded 5 out of 6 of these studies from my analysis as they did not meet my inclusion criteria of a sample size of at least 30 patients with ArLD per study in four (n=8 (267), n=15 (268), n=16 (269), n=20 (270)), and alcohol data were not reported separately for the fifth (220) which was a mixed aetiology cohort.

3.6.2 Comparison of tests

In studies where direct comparisons were made between tests, FIB4 performed inferiorly to ELF for F2/F3/F4 (23), to FibroTest for F3/F4 (148, 256) and F2 (148), and to FibroScan for F3/F4 (256). This is in keeping with findings that indirect tests perform less well in ArLD compared to other aetiological cohorts (195, 271, 272). Despite performing inferiorly to the 'direct fibrosis markers', FIB4 still performed acceptably with AUROCs of 0.7-0.89 for F2/F3/F4. Whilst Thiele et al. found no added diagnostic benefit of combining a simple marker such as FIB4 to ELF or FibroScan (23), another study by Lannerstedt et al. discovered that FIB4 improves the diagnostic performance of FibroScan (269). As FIB4 is such a cheap and accessible test, based on a simple score from routine blood tests that are usually already available for the patient – it would be worth further exploring FIB4 in combination with FibroTest/FibroScan/ELF in future prospective studies on ArLD.

Where other non-invasive tests were directly compared, ELF, FibroScan and FibroTest performed equally well to each other (intention-to-diagnose protocol) in one study (23), whilst FibroScan outperformed FibroTest for F2 in a study by Nguyen-Khac et al. (161), and for F3 in a study by Fernandez et al. (256). However, it must be noted that out of all four non-invasive tests investigated in my study, FibroScan was the only one with a reported failure rate, of between 1-22% across studies, which corresponds to average documented failure rate of 5-15% across aetiologies, depending on criteria used for reliable results (273).

When exploring factors affecting test result score, FibroScan readings were influenced by more variables than the other three tests. Of the 8/11 FibroScan studies reporting on inflammation, 5 out of 8 found inflammation (either histological or AST) to lead to false positive FibroScan results (23, 252, 254, 256, 258). Alcohol intake or withdrawal affected FibroScan results (increased false positives) in 3 out of 4 FibroScan studies that reported on this (160, 252, 258). Steatosis influenced FibroScan values in 2 out of 6 studies examining this (251, 252). Nguyen-Khac et al.'s meta-analysis on 10 studies using individual patient data also discovered that FibroScan is affected by Bilirubin levels as well as AST, and that these two values should be taken into consideration when interpreting liver stiffness results (243).

ELF did not correlate with alcohol intake/drinking pattern in the 1 out 2 included studies reporting on this (23), but did correlate with histological inflammation in one study, but not with ALT or AST in either study. This is in keeping with findings from a further published prospective study on ELF in ArLD which did not find any correlation between ELF and alcohol or ELF and AST or ALT (274). Of the six studies on FibroTest,

2/6 examined inflammation on test result, with 1 finding that FibroTest was affected by histological inflammation (23), and the other finding no association (148). Only one study explored the impact of alcohol on FibroTest, finding no association (23).

3.6.3 Thresholds:

FibroScan:

For the use of FibroScan in ArLD, Papatheodoridi et al. in 2021 have published guidance on new FibroScan thresholds, suggesting a threshold of 8kPa to rule out advanced fibrosis ('advanced compensated liver disease'), and 12kPa to rule in. The authors acknowledge that this was based on studies where information about alcohol consumption prior to the FibroScan was unknown. In addition, only 17% of patients in this study had ArLD, with the majority of the included patients having viral hepatitis or NAFLD. The data for this study were collected retrospectively using previously published results from individual studies (included in my systematic review). In addition, I note the published response by Genesca et al. to this article (275), where they have questioned the validity of the results, given lack of availability of the XL probe for obese patients for studies included in this summary study. Approximately one third of the total cohort were obese, and as it is known that liver stiffness measurements can be inaccurate in obesity if the XL probe is not used (276), then the optimum derived thresholds may not be valid. The Baveno study authors themselves highlighted the need for further prospective validation of FibroScan thresholds, and I agree that this is a necessity particularly for ArLD, to identify an optimum threshold.

Whilst FibroScan thresholds remain in debate for ArLD, I would suggest that authors of future studies evaluate a uniform threshold in order for easier validation – as I found a different reported optimum threshold for all 11 different included FibroScan studies, so could not accurately pool results. Perhaps the best threshold to adopt for this until further validation would be the one recommended by Nguyen-Khac et al.'s meta-analysis on FibroScan in ArLD (243), which had access to individual study data of n=1026 patients, and recommended bilirubin and AST be factored in to the interpretation. Their suggested thresholds were 9kPa for F2, 12.1kPa for F3, 18.6 kPa for F4 (243).

ELF

Only one study (Thiele et al.) (23) provided a threshold for advanced fibrosis, recommending <10.5 for ruling out in primary care (NPV of 98, sensitivity 75%, specificity 97%). Thiele et al. also included data on a 9.8 threshold, which also had NPV of 98% and sensitivity of 75%, but lower specificity of 89%. As the manufacturer-recommended optimum ELF threshold for NAFLD is 9.8 (277), and there is a known overlap between NAFLD and ArLD with many patients with AUD also living with obesity, and those with NAFLD drinking excess alcohol (278) it may be beneficial to adopt a uniform threshold of 9.8kPa, particularly in a community setting to rule out advanced fibrosis, where differentiation between NAFLD and ArLD may not yet be clear, to avoid missing diagnoses. Further prospective validation of this threshold in ArLD is required.

FIB4

Only one study, by Thiele et al. (23), applied a FIB4 threshold (3.25) for advanced fibrosis in ArLD, with associated specificity of 91%, sensitivity of 58%, NPV 88%, PPV 64%. This threshold therefore requires further prospective validation in ArLD.

FibroTest

Only one study, again by Thiele et al., applied a FibroTest threshold (0.58) for advanced fibrosis in ArLD: specificity 87%, sensitivity 67%, NPV 90%, PPV 60%. This threshold therefore also requires further prospective validation for use in ArLD.

3.7.1 Strengths

This is the largest systematic review in terms of sample size of non-invasive tests in ArLD. I have included a meta-analysis for FibroScan in ArLD, with all included studies having more than 30 patients each, and a combined sample of 1,268 for advanced fibrosis – larger than any previous meta-analysis on FibroScan in ArLD. My literature searches were comprehensive, with rigorous screening of texts independently by myself and the second reviewer to minimise reporting bias. I made efforts to contact study authors by email in the event data were not able to be extracted. In addition, I hand-searched reference lists of relevant studies to maximise the number of articles included. I also made efforts to ensure there was no overlap of patients between studies, particularly where the same author had published more than one paper, by contacting authors where it was unclear, to clarify. Where there was duplication, I only

included the most recent and comprehensive studies. All included studies had available paired-histology data.

3.7.2 Limitations

I recognise there are limitations to this systematic review. Apart from one study where I had access to raw data, the rest of my data were sourced directly from published articles. This meant I could not explore the performance of a uniform threshold for each non-invasive test between studies, and that my meta-analysis on FibroScan is not as useful as it would have been if all studies had adopted the same threshold.

As there was significant heterogeneity between studies, I was not able to pool specificity results for FibroScan at F3 fibrosis stage. Furthermore, I could only include those studies which reported sensitivity/specificity data with associated 95% confidence intervals which limited the meta-analysis to 8 out of 11 studies for F4, and for 10 out of 12 studies for sensitivity for F3 (and 6/12 for AUROC for F3). It is possible therefore that I could have introduced bias by excluding the studies from analysis that did not have the necessary data.

This systematic review also found significant variation between studies in the prevalence of advanced fibrosis and cirrhosis, and the biopsy length – highlighted by Naveau et al. (148) to be the two most important variability factors that could cause bias in the direct comparison of sensitivity/specificity results.

A future meta-analysis with access to raw study data would be beneficial, to factor in biopsy length and individual study prevalence of fibrosis stages, in addition to explore

uniform test thresholds across studies, and perform sub-analyses of influence factors such as AST, bilirubin, age, steatosis, and histological inflammation across all four non-invasive tests.

3.8 Conclusion

All four non-invasive tests perform well for the detection of F2, F3 and F4 fibrosis, with AUROCs ≥0.7 at all stages. FIB4 appears to perform the least well of the four, but with AUROCs ≥0.7, as it is so readily available and easy to calculate, it would be worth exploring in combination with direct fibrosis tests in future studies. FibroScan, whilst the most studied of the four tests, is associated with histological inflammation, AST/ALT, bilirubin levels, and alcohol intake or withdrawal. It is also the only test with an associated failure rate. Whilst it can be very effective in ArLD, it must be interpreted by factoring in AST and bilirubin levels, and knowledge of any active inflammation or recent alcohol intake, and suggest interval scanning 2-4 weeks after alcohol withdrawal. In addition, as highlighted by Thiele et al., blood-based tests such as ELF or FibroTest may be more practical for patients and clinicians in preference to FibroScan for use in the community (23).

Further validation studies are needed to clarify optimum thresholds for all four tests, as described above. In the meantime, I would urge researchers to include the thresholds highlighted above in any future analysis of non-invasive tests in ArLD, so as to enable progress in validating the best thresholds. Further exploration of the potential impact of recent alcohol consumption and inflammation are also warranted for ELF, FibroTest and FIB4.

Chapter 4

Diagnosing advanced fibrosis in Alcohol-related Liver Disease in practice: examining current referral strategies from primary to secondary care, and risk factors associated with advanced fibrosis

4.1 Abstract:

Background

Twenty-percent of people with alcohol-use-disorders develop advanced fibrosis and warrant referral to secondary-care. Improving outcomes in Alcohol-related-Liver-Disease (ArLD) relies on its earlier detection in primary-care with non-invasive-tests (NIT). I aimed to determine the proportion of alcohol-related referrals who were diagnosed with advanced fibrosis in secondary-care, the prevalence of 'BAFLD' (Both Alcohol and Fatty Liver Disease), and the potential impact of NIT on referral-stratification.

Methods

I performed a retrospective analysis of all GP-referrals with suspected ArLD/NAFLD to a UK hepatology-centre between Jan2015-Jan2018.

Of 2,944 new referrals, 762 (mean age 55.5±13.53 years) met inclusion-criteria: 531 NAFLD and 231 ArLD, of which 147 (64%) could be reclassified as 'BAFLD'.

Primary outcome-measure: The proportion of referrals with suspected ArLD/NAFLD with advanced fibrosis as assessed by tertiary-centre hepatologists using combinations of FibroScan, imaging, examination and blood tests, and liver histology where indicated.

Secondary outcome-measures: The impact of BMI/alcohol consumption on the odds of a diagnosis of advanced fibrosis, and performance of NIT in predicting advanced fibrosis in planned post-hoc analysis of referrals.

Results:

Amongst ArLD referrals 147/229 (64.2%) had no evidence of advanced fibrosis and were judged 'unnecessary'. Advanced fibrosis was observed in men currently drinking ≥50U/w (OR 2.74, 95% CI 1.51-to-5.00, p = 0.001), and in women drinking ≥35U/w (OR

5.11, 95% CI 1.31-to-20.03, p = 0.019). Overall, drinking > 14 U/w doubled the likelihood of advanced fibrosis in overweight/obesity (OR 2.11; CI 1.44-to-3.09; p<0.001). Use of FIB4 could halve unnecessary referrals (OR 0.50; CI 0.32-to-0.79, p = 0.003) with false-negative rate of 22%, but was rarely used.

Conclusions:

I discovered that the majority of referrals with suspected ArLD were deemed unnecessary. NIT could improve identification of liver damage in ArLD, BAFLD and NAFLD in primary-care. I validated previously anecdotal thresholds for harmful-drinking (35U/w in women and 50U/w in men). The impact of alcohol on NAFLD highlights the importance of multi-causality in CLD.

Notice of publication

I have published a version of this chapter as a manuscript in 'BMJ Open' in 2021:

'Rhodes FA, Cococcia S, Patel P, Panovska-Griffiths J, Tanwar S, Westbrook RH, et al. Is there scope to improve the selection of patients with alcohol-related liver disease for referral to secondary care? A retrospective analysis of primary care referrals to a UK liver centre, incorporating simple blood tests. BMJ Open. 2021;11(6):e047786.'*(279)*

4.2 Introduction

Approximately 90% of all chronic liver disease (CLD) is preventable, with the commonest causes of cirrhosis attributed to ArLD and NAFLD (3). Mortality rates from cirrhosis have increased 400% since 1970, predominantly due to alcohol, although the rising prevalence of NAFLD is contributary (12). Hepatic steatosis develops in up to 90% of people with Alcohol Use Disorder (AUD) or obesity (7, 8), but advanced fibrosis or cirrhosis will affect only approximately 20% of people with AUD (9) and 5% with NAFLD (280). I explored the reasons for this in Chapters 2 and 3. Both AUD and obesity can be managed effectively in primary-care but advanced fibrosis and cirrhosis warrant management by liver specialists in secondary care. Detecting the minority of patients requiring specialist care is challenging because advanced fibrosis and most cases of cirrhosis are asymptomatic and simple liver blood tests (LFTs) and ultrasound imaging are neither sensitive nor specific in detecting advanced fibrosis or cirrhosis (1). As a consequence, it has been reported that three-quarters of people with CLD first present to healthcare late, when they already have established advanced liver disease and behaviour change or therapeutic interventions have only modest impacts on prognosis by this point (3, 140, 281).

Conversely, as many as 92% of people referred to secondary-care with suspected CLD do not have advanced fibrosis or cirrhosis requiring specialist care and could have remained in primary-care for ongoing management (14). Pathways of care employing the use of NITs for liver fibrosis (FIB-4 and the ELF test) in primary-care have been shown to be effective in the management of NAFLD, yielding an 88% reduction in 'unnecessary referrals' (no presence of advanced fibrosis) to liver specialists with a

five-fold increase in the detection of advanced fibrosis and cirrhosis, and significant cost-savings (14, 36). This NAFLD pathway is in place locally in Camden, and has influenced national guidelines (1, 38). However, the proportion of referrals from primary to secondary care with AUD who do not have advanced ArLD, that could be considered 'unnecessary' is unknown.

The ELF test has also been used successfully to triage patients from primary to secondary-care with AUD in Denmark (23). While current UK national guidelines recommend consideration of NIT in people with AUD in primary care (1), alcohol pathways employing NIT are not widely established in the UK and none have been evaluated to my knowledge.

Although NAFLD and ArLD are described as distinct entities for research purposes, the risk factors for both conditions co-exist in many patients. Moreover, it is increasingly questioned in the literature if alcohol and fat may interact to cause liver damage, with some studies finding that obese people have an increased risk of liver fibrosis for any given alcohol intake (1, 282-285).

In this study, I have used the term 'BAFLD' (Both Alcohol and Fatty Liver Disease – originally coined by the Parkes Group in Southampton, (286) to describe the combination of fat and alcohol as risk factors for CLD.

My aims for this study, as covered in Chapter 1, were to determine the proportion of patients referred for investigation of ArLD from primary-care to secondary-care hepatology clinics that had evidence of advanced fibrosis; and the prevalence of both alcohol and fat as co-contributing factors to CLD, termed 'BAFLD' to describe the

combination of Both Alcohol and Fatty Liver Disease (286). In addition, I aimed to determine the performance of simple NITs in the identification of cases of advanced fibrosis.

4.3 Methods

4.3.1 Study design

I performed a retrospective cross-sectional analysis of consecutive patients aged ≥18 years newly referred from primary-care to a hospital-based hepatology service at the Royal Free London NHS Foundation Trust (RFL), with a suspected diagnosis of ArLD or NAFLD between January 2015 and January 2018. Patients were excluded if they had any other hepatological diagnosis made prior to referral (Table 4.1).

Table 4.1: Inclusion and exclusion criteria

Inclusion Criteria	Exclusion criteria
○ Age 18 or above ○ Presence of new referral letter from GP to hepatology clinic at Royal Free during evaluation period Jan 2015 to Jan 2018 ○ Primary reason for referral from GP to hepatologist is suspected diagnosis of ArLD$ or suspected diagnosis of NAFLD^	○ Presence of pre-existing hepatological diagnosis (Including but not limited to: auto-immune hepatitis, viral hepatitis, PBC, PSC, HCC), ○ Patients are already under the care of a hepatologist/gastroenterologist for investigation or management of a liver condition.
$ 'Suspected ArLD' referrals were defined as those in which the GP referral letter requested an assessment by a liver specialist specifying concerns about suspected ArLD or expressing concerns about a patient's alcohol intake. ^ 'Suspected NAFLD' referrals were defined as those in which the GP referral letter either specified that they were referring the patient to hepatology 'with suspected NAFLD' or 'on the local NAFLD referral pathway', OR, in the absence of any other cause of liver dysfunction, where the GP specified that the patient had steatosis or chronic liver disease on ultrasound in combination with mentioning metabolic risk factors (BMI ≥25, diabetes, high waist circumference, high cholesterol or hypertension).	

4.3.2 Outcome measures

The primary outcome was the proportion of new patients referred from GP to hepatology clinic with suspected ArLD that had advanced fibrosis and could be deemed 'necessary' referrals.

Secondary outcomes included the prevalence of 'BAFLD' amongst patients referred with suspected ArLD or NAFLD, analysis of demographic data as potential risk factors for a

diagnosis of advanced fibrosis (including BMI, alcohol consumption, smoking status, age, sex, and deprivation score), and a post-hoc analysis of the performance of FIB4 and APRI in predicting a diagnosis of advanced fibrosis.

4.3.4 Study population

All electronic GP referrals for suspected ArLD or NAFLD during this period were reviewed in order to identify cases referred for NAFLD who were subsequently found to be drinking hazardous amounts of alcohol (>14 units per week). As these conditions were not always reliably coded and triaged from the outset, I reviewed every new referral from GP to hepatology clinic during a 3-year time period in order to select out the NAFLD and ArLD referrals to ensure cases were not missed. Sample size was based upon the 3-years' worth of referrals.

'Suspected ArLD' referrals were defined as those in which the GP referral letter requested an assessment by a liver specialist specifying concerns about suspected ArLD or expressing concerns about a patient's alcohol intake.

'Suspected NAFLD' referrals were defined as those in which the GP referral letter either specified that they were referring the patient to hepatology 'with suspected NAFLD' or 'on the local NAFLD referral pathway', OR, in the absence of any other cause of liver dysfunction, where the GP specified that the patient had steatosis or chronic liver disease on ultrasound in combination with mentioning metabolic risk factors (BMI $\geq$25, diabetes, high waist circumference, high cholesterol or hypertension).

4.3.5 Data Collection

I reviewed 2,944 referral letters, in order to select out those patients that met my inclusion/exclusion criteria described above. I had help with this and the extraction of data for this chapter, by another member of the research group (a research fellow, 'SC'). Anonymised data were extracted from the patients' electronic records. These included demographics, reason for referral, deprivation score, weight, height, waist circumference, alcohol intake, comorbidities, and any fibrosis assessment before and after referral. Where weight and height were unavailable, but clinical records reported that the patient was overweight or obese, they were categorised accordingly to BMI >25 (overweight) or BMI >30 (obese). FIB4 and APRI scores were calculated using the blood tests from the first attendance to clinic after referral.

The diagnosis of advanced fibrosis (equivalent to a histological stage of ≥ F3/4) or cirrhosis (≥ F4) was established by expert clinical judgement by Royal Free hepatologists based on a composite of FibroScan, imaging, blood tests, clinical examination and liver histology where available, and this information was extracted from the electronic medical records. In the minority of cases where a diagnosis of advanced fibrosis was not clearly documented, decisions were reviewed between myself and my colleague 'SC' who assisted with the data collection, and consensus achieved. FibroScan was considered diagnostic for advanced fibrosis if the elasticity of a valid scan was ≥11kpa in ArLD (38, 178) and ≥10kpa in NAFLD patients (287). For variables where any data were missing, the denominator used in the analysis was adjusted for available data.

'Unnecessary referrals' were defined as those patients that, subsequent to an assessment by a liver specialist, were deemed not to have advanced fibrosis and could be discharged back to ongoing care in the community.

In light of the frequent overlap between the two conditions, I subsequently recoded patients as having Both Alcohol and Fatty Liver Disease (BAFLD) if ArLD and NAFLD risk factors were both present. More specifically, BAFLD was applied to patients referred for suspected NAFLD who were subsequently found to be drinking more than 14 units of alcohol per week; and to patients who were referred for suspected ArLD, who also had either a BMI >25, or features of the metabolic syndrome. The metabolic syndrome was defined according to the International Diabetes Federation (IDF) and American Heart Association (AHA) as the presence of at least three of the following criteria: enlarged waist circumference (≥94cm in European men, ≥90cm South Asian men, ≥80cm women), hypercholesterolaemia, hypertension and type 2 diabetes (288).

4.3.6 Statistical analysis

Descriptive statistical analyses included calculations of the frequencies and percentages for categorical variables, while for continuous data: I used means and standard deviation (SD) for normally distributed data, or medians and interquartile range (IQR) for skewed data. For the comparison of categorical variables, I used Chi-Squared or Fischer's exact test (the latter when n = <5), and for continuous data Mann Whitney-U or Student's-t test depending on the data distribution.

For data with more than three variables to compare, I used ANOVA or Kruskall Wallis ANOVA, depending on the distribution of the data.

I then categorised alcohol consumption into groups of units consumed per week according to the perceived risk of liver damage established in the literature (1) (0-35, 36-50,51-100, >100 units per week) and into quartiles of the population distribution of alcohol consumption for the ArLD cohort in which few patients were drinking <50 units per week. I used multiple binary logistic regression analysis to determine the association between key variables and the presence of advanced fibrosis. The key variables were those risk factors for fibrosis that were of established importance in the literature, and those associated with p values <0.25 in the univariate analyses. All p values were 2-sided and significance set at <0.05. I analysed all data using SPSS software (Version 25.0. Armonk, NY: IBM Corp), except for the odds ratios (ORs) for differences in outcomes for modelling of data with FIB4 compared with current practice, together with 95% confidence intervals and chi-square for statistical significance which I performed using MedCalc statistical software 2020.

4.3.7 Ethics

This study uses secondary anonymised patient data. The project was registered with the Integrated Research Application System (IRAS 272448) and judged to not require ethical approval or informed consent according to Health Research Authority guidance as it comprises data that were collected routinely as part of a registered service evaluation at the Royal Free London NHS Foundation Trust.

4.4 RESULTS

4.4.1 Patient demographics:

Between January 2015 and January 2018, a total of 2,944 patients were referred to the RFL hepatology service from primary care and of these, 762 (mean age 55.5±13.53 years) met the inclusion criteria for this study; 231 patients were referred with suspected ArLD (mean age 54.68±12.37 years), and 531 with suspected NAFLD (mean age 55.88±14 years). One patient was deemed to have active hepatitis C virus infection as a comorbidity and three were found to have inactive chronic hepatitis B after referral. I have summarised the demographic characteristics of eligible patients in Table 4.2. There was a higher proportion of male patients in the ArLD group (76.2%) than amongst the NAFLD group (54.2%, $p<0.001$). Active or previous smoking was significantly more common among those referred for ArLD compared to the NAFLD group (47.1% vs 11.3%; $p<0.001$). The average BMI was significantly higher in the NAFLD group than the ArLD group (31.9 and 27.9 kg/m^2 respectively, $p<0.001$), while median alcohol consumption was significantly higher in the ArLD group at 70 units/week (42-135), compared to 0 units/week (0-7) in the NAFLD group. The majority of the study population lay within the lowest 4 deciles of deprivation, and no significant difference in levels of deprivation was seen when ArLD and NAFLD referrals were compared ($p=0.326$).

Table 4.2: Baseline characteristics

Patient characteristics	Overall (n=762)	Suspected ArLD referrals* (n =231)	Suspected NAFLD referrals** (n=531)	
Age (mean; sd)	55.52 ±13.53	54.68±12.37	55.88±14	*p = 0.262*
Male n (%)	464 (60.9%)	176 (76.2%)	288 (54.2%)	*p < 0.001*
BMI (mean; sd)	30.85 ± 6.23	27.9 ± 5.46 (n=174)	31.9 ±6.15	*p < 0.001*
> 25 n (%)	608/732 (83.1)	149/211 (70.6)	459/521 (88.1)	*p < 0.001*
> 30 n (%)	350/675 (51.9)	56/185 (30.3)	294/490 (60)	*p < 0.001*
Alcohol intake U/w (median, IQR)	5, (0-42.75)	70 (42-134.8)	0 (0-7)	*p < 0.001*
N =	738	226	512	
Years of harmful drinking				
Median (IQR)	0 (0-3)	20 (6-30)	0 (0-0)	*p < 0.001*
Total n =	598	143	455	
Diabetes n (%)	235/760 (30.9)	38/231 (16.5)	197/529 (37.2)	*p < 0.001*
Hypertension n (%)	397/761 (52.2)	113/231 (48.9)	284/530 (53.6)	*p = 0.236*
Hypercholesterolaemia n (%)	352/759 (46.4)	81/231 (35.1)	271/528 (51.3)	*p < 0.001*
Smoking status: Non- smoker n (%)	369/681 (54.2)	65/204 (31.9)	304/477 (63.7)	*p < 0.001*
Smoker n (%)	150/681 (22)	96/204 (47.1)	54/477 (11.3)	
Ex- smoker n (%)	162/681 (23.8)	43/204 (21.1)	119/477 (24.9)	
ALT median (IQR)	45 (30-67)	47 (30-68)	45 (30-67)	*p = 0.360*
N =	761	231	530	
Deprivation score rank Median (IQR)	11314 (6451-17642)	10648 (6100-17464)	11637 (6578-17761)	*p = 0.326*
Deprivation score decile: 1	51 (6.7%)	12 (5.2%)	39 (7.3%)	*p = 0.264*
2	146 (25.9%)	53 (28.1%)	93 (24.9%)	
3	134 (43.4%)	42 (46.3%)	92 (42.2%)	
4	107 (57.5%)	30 (59.3%)	77 (56.7%)	
5	101 (70.7%)	33 (73.6%)	68 (69.5%)	
6	82 (81.5%)	26 (84.8%)	56 (80%)	
7	64 (89.9%)	17 (92.2%)	47 (88.9%)	
8	44 (95.7%)	8 (95.7%)	36 (95.7%)	
9	22 (98.6%)	6 (98.3%)	16 (98.7%)	
10	11 (100%)	4 (100%)	7 (100%)	
Had Biopsy n (%)	122/762 (16%)	10/231 (4.3%)	112/531 (21.1.%)	*p < 0.001*
Had FibroScan n (%)	575/762 (75.5%)	158/231(68.4%)	417/531 (78.5%)	*p = 0.003*
Valid FibroScan reading***	524/575 (91%)	140/158 (89%)	389/417 (93%)	
FibroScan median kPa (IQR)	5.5 (4.5-7.7)	6 (4.7-8.5)	5.4 (4.4-7.5)	*p = 0.03*

* Where primary reason for referral from GP was for suspected alcohol-related liver disease

**Where primary reason for referral from GP was for suspected NAFLD

*** FibroScan results were considered invalid if: IQR/M >30%, success rate <60%, <10 valid readings, or if this information was not recorded in the FibroScan report (missing information about IQR/M ratio/success rate made up n=22/575 FibroScan results).

SD = standard deviation, IQR = interquartile range, ALT = alanine aminotransferase, BMI = body mass index,

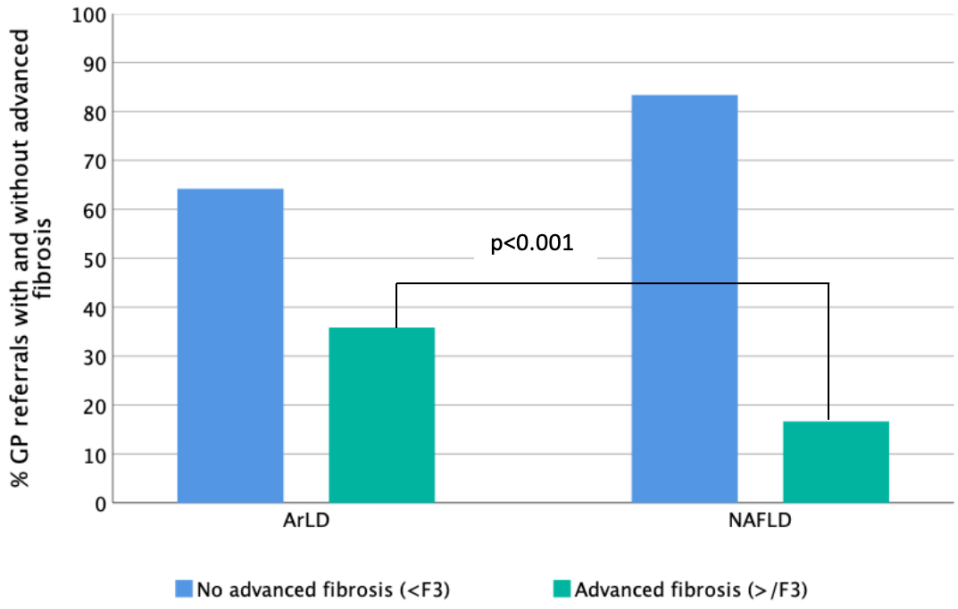
4.4.2 Reasons for referral from primary care

The presence of hepatic steatosis on an ultrasound scan and abnormal LFTs were the commonest reasons for referral to hepatology clinic regardless of the aetiology. These were followed by elevated ELF and FIB4 in the NAFLD cohort (38.2 % and 16.9% respectively). Only 38/231 (16.4%) of patients with suspected ArLD had a NIT in primary-care prior to referral (25 ELF scores, 13 FIB4) and of these, 25/38 (66%) patients had comorbid features of the metabolic syndrome and so were subsequently recoded as BAFLD. Amongst the NAFLD referrals 293/531 (55.2%) had a NIT prior to referral in accordance with the local NAFLD pathway. Of these patients 203/293 (69%) were referred on the basis of an elevated ELF test and 90/293 (31%) based on their FIB4 score.

4.4.3 Prevalence of advanced fibrosis in patients referred with suspected ArLD or NAFLD.

Data on fibrosis stage were available for 758/762 (99.5%) patients following hepatology review, with four not attending for assessment. Of patients with suspected ArLD, 64.2% (147/229) had no evidence of advanced fibrosis and could be discharged back to primary-care (Figure 4.1). This figure was even higher in the NAFLD cohort with 83.4% not having advanced fibrosis.

Figure 4.1: Proportion of patients referred from GP with suspected ArLD or NAFLD who had a diagnosis (composite clinical judgement) of advanced fibrosis



Of the patients referred with suspected ArLD who had advanced fibrosis (82/229, 36%), the frequency with which fibrosis tests were used were: liver biopsy in 10% (8/82), FibroScan in 41% (34/82) and radiology in 62% (51/82).

Of the patients referred with suspected NAFLD who had advanced fibrosis (88/529, 17%), the frequency with which fibrosis tests were used were: liver biopsy in 47% (41/88), FibroScan in 64% (56/88) and radiology in 33% (29/88).

4.4.4 Risk of advanced fibrosis (≥F3) in patients referred with suspected ArLD.

Univariate analysis of the 231 patients referred with ArLD revealed that advanced fibrosis was associated with higher alcohol consumption (alcohol data available for 224/231) (OR 1.006, 95% CI 1.002 to 1.010, p=0.006), low platelets (OR 0.99, 95% CI 0.987-0.994, p < 0.001), and raised ALP (OR 1.012, 95% CI 1.006 to 1.018 p <0.001) (Figure 4.2a/b). When categorised into alcohol unit groups of: <35 U/w, 36-50 U/w, 51-100 U/w, >101 U/w; patients drinking >50 U/w had a higher risk of advanced fibrosis in this cohort (OR 2.899, 95% CI 1.068 to 7.869, p= 0.037). The multivariable logistic regression model found that the odds of advanced fibrosis in suspected ArLD was independently associated with increased units of alcohol consumed, (OR 1.007, 95%CI 1.002-1.012, p=0.007), ALP (OR 1.009, 95% CI 1.002-1.016, p=0.01), and reduced platelets (OR 0.992, 95%CI 0.988-0.996, p<0.001). There was a trend towards higher odds of advanced fibrosis with increased age, but this did not reach significance (p=0.059).

Figure 4.2a: Boxplot of ALP by diagnosis of advanced fibrosis in patients referred with suspected ArLD

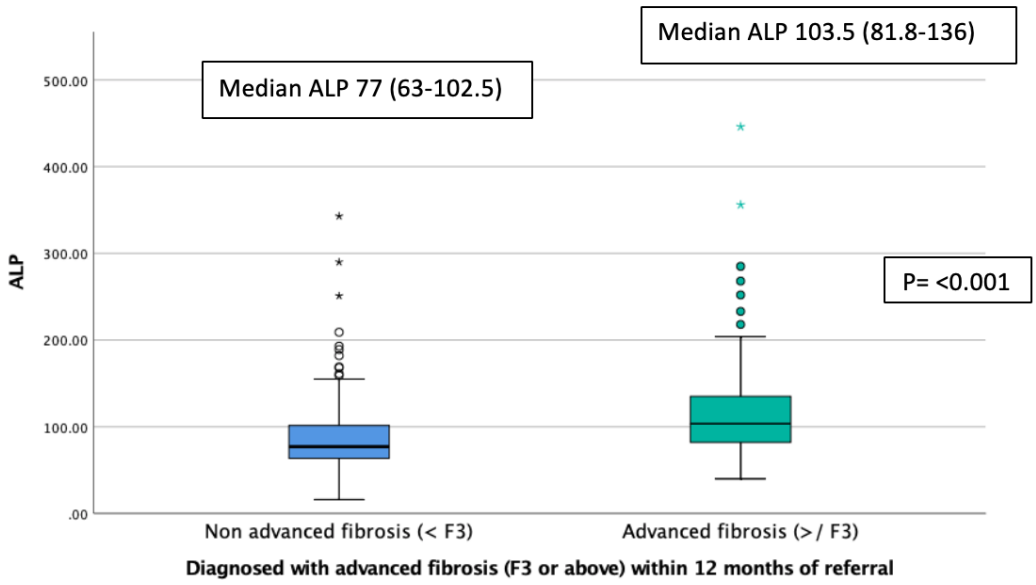
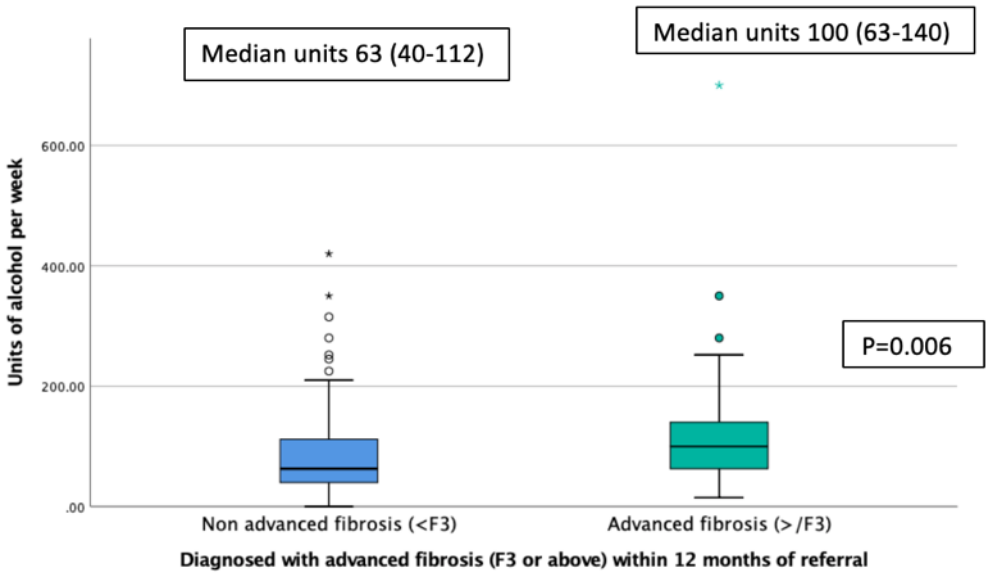


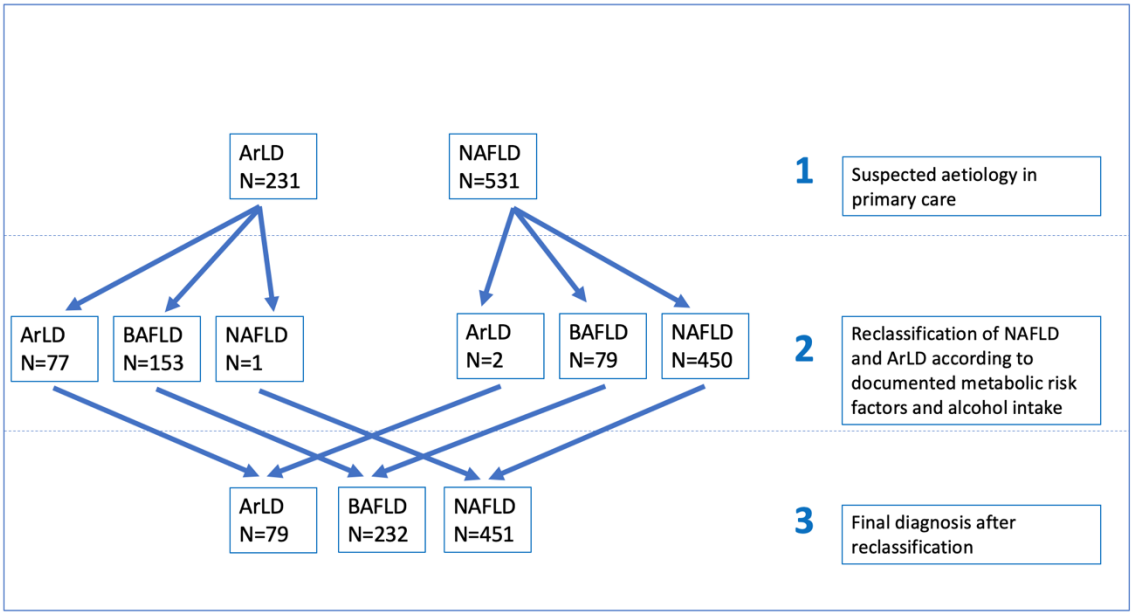
Figure 4.2b: Boxplot of units of alcohol per week by diagnosis of advanced fibrosis in patients referred with suspected ArLD



4.4.5 Patients with risk factors for both ArLD and NAFLD: 'BAFLD'.

Patients with risk factors for both ArLD and NAFLD were classified as BAFLD (as defined by: patients referred for suspected NAFLD who were subsequently found to be drinking more than 14 units of alcohol per week; or patients who were referred for suspected ArLD, who also had either a BMI >25, or features of the metabolic syndrome, as defined earlier) and the whole cohort was re-classified into three categories: ArLD, NAFLD and BAFLD, in order to evaluate further risk factors for advanced fibrosis (Figure 4.3).

Figure 4.3: Flow chart depicting reclassification of aetiologies into ArLD, BAFLD and NAFLD



From the GP referral letters, 147 (63.6%) patients out of the 231 patients referred to the hepatology clinic with suspected ArLD were overweight (BMI>25), or met the diagnostic criteria of the metabolic syndrome (defined on page 10), and were therefore reclassified as BAFLD. Of the 531 patients referred to hepatology as suspected NAFLD, 80 of them (15.1%) also regularly consumed an average of more

than 14 units per week and were reclassified as BAFLD. Overall, 83.1% of the whole cohort were overweight and 50% obese. As expected, the proportion of patients who were overweight and obese was significantly higher in the NAFLD cohort compared to ArLD cohort ($p<0.001$). The main characteristics of the three cohorts can be found in Table 4.3.

Table 4.3: Demographics within each re-classified aetiology group

Overall characteristics (n=762)	**ArLD (n =79)**	**NAFLD (n=451)**	**BAFLD (n=232)**	
Non advanced fibrosis (<F3) n (%)	46/78 (60)	377/450 (83.8)	165/230 (71.7)	*p < 0.001*
Advanced fibrosis (≥F3) n (%)	32/78 (40)	73/450 (16.2)	65/230 (28.3)	
Age (mean; sd)	51.85 ± 13.1	55.3 ± 14.07	57.2 ± 12.3	*p = 0.009*
BMI (mean; sd)	21.9 ± 2.32	32.1± 6.17	30.6 ± 5.03	*p < 0.001*
> 25 n (%)	0/59 (0)	393/443 (88.7)	215/230 (93.3)	*p < 0.001*
> 30 n (%)	0/57 (0)	252/416 (60.6)	98/202 (48.5)	*p < 0.001*
Alcohol intake median U/w (IQR)	79.9 (49.3-140)	0 (0-4)	49.5 (30-88.5)	*p <0.001*
N=	76	434	228	
Years of harmful drinking Median (IQR)	13 (5-20)	0 (0-0)	20 (8-30)	*p < 0.001*
N=	47	427	124	
ALT median, (IQR)	43 (28-68)	45 (31-68.25)	47 (30-67)	*p = 0.752*
N=	79	450	232	
Community ELF score (mean, sd)	9.96 ± 0.42	10.33 ± 0.74	10.5 ± 0.84	*p = 0.215*
N=	7	169	54	
Community FIB4 median (IQR)	2.75 (1.22-5.19)	1.56 (1.38 -2.2)	2.2 (1.5-3.25) 24	*p = 0.043*
N=	4	75		

ArLD = Alcohol-related Liver Disease, NAFLD = Non-Alcoholic-Fatty-Liver-Disease, BAFLD = Both Alcohol and Fatty Liver Disease, sd = standard deviation, BMI = Body Mass Index, IQR = interquartile range, ALT = alanine aminotransferase, ELF = Enhanced Liver Fibrosis score

Patients with BAFLD had almost double the prevalence of advanced fibrosis when compared to NAFLD (29% and 16.2% respectively, (OR 2.11, 95% CI 1.441 to 3.094, p <0.001)), suggesting that hazardous drinking potentially doubled the risk of fibrosis in people who are overweight or obese in this study population (Table 4.3).

Patients in the ArLD cohort had the highest prevalence of advanced fibrosis (38%), and their weekly alcohol intake was almost double that of the BAFLD patients, precluding the opportunity to compare the impact of overweight/obesity on heavy alcohol consumption in this cohort.

4.4.6 Influence of alcohol on fibrosis risk

As the number of ArLD patients drinking <50 units per week (U/w) was small, the entire cohort (n=762) was examined in an attempt to identify a potential threshold for the effect of alcohol on fibrosis risk. Other factors influencing fibrosis risk including age and BMI were also studied. Alcohol data were available for 733/762 (96%) patients.

Increased alcohol U/w predicted advanced fibrosis (OR 1.009, 95%CI 1.006 to 1.012, p = <0.001) on univariate analysis.

Alcohol units were categorised into quartiles of the reported distribution of consumption (0-42 U/w, 43-70 U/w, 71-135 U/w, >136 U/w). Binary logistic regression revealed that patients consuming ≥43 U/w were at greater risk of advanced fibrosis than those drinking less than 43 U/w. (OR 1.814, 95%CI 1.038 to 3.172, p = 0.037), and those drinking ≥70 U/w were at more than four times the risk of having advanced

fibrosis compared with those drinking less than 43 U/w (OR 4.25, 95% CI 2.334 to 7.740, p = <0.001).

Alcohol consumption was then evaluated at literature-based unit thresholds of interest (0-35 U/w (n=521), 36-50 U/w (n=49), 51-100 U/w (n=82), >101 U/w (n=81)) revealing that drinking more than 35 U/w was associated with double the odds of developing advanced fibrosis compared with those drinking <35 U/w (OR 2.173, 95% CI 1.119 to 4.219, p = 0.022) and the odds increased to over five-fold in those drinking more than 100 units per week (OR 5.044, 95% CI 3.071 to 8.284, p <0.001).

I discovered a different threshold effect when these data were analysed separately for men (61%) and women. In the overall cohort of 762 patients, the risk of having advanced fibrosis was higher in those men drinking >50 U/w (OR 2.743, 95% CI 1.506 to 4.998, p = 0.001), while in women the risk of having advanced fibrosis increased significantly at only >35 U/w (OR 5.115, 95% CI 1.306 to 20.030, p = 0.019), compared to <35 U/w).

In the overall cohort of 762 patients with ArLD/NAFLD/BAFLD (of which complete data for this model were available for 625/762), multivariable regression analysis revealed that increased units of alcohol, age, ALP, BMI and decreased platelet count were significantly associated with increased odds of a diagnosis of advanced fibrosis (Table 4.4).

Table 4.4: Multivariable logistic regression analysis of factors influencing odds of a diagnosis of advanced fibrosis in whole study cohort

Variables (N= 625/762)	OR	95% CI		P value
		Lower	Upper	
Alcohol (U/w)	1.008	1.005	1.012	.000
AGE	1.017	1.000	1.034	.049
Bilirubin	1.023	.999	1.048	.062
ALP	1.005	1.001	1.009	.026
ALT	1.003	.999	1.007	.118
Platelet	.993	.990	.996	.000
BMI	1.053	1.019	1.089	.002

Modelling the impact of indirect fibrosis tests on the detection of advanced fibrosis in patients referred from primary care with suspected ArLD.

I used blood test results from the first attendance at the secondary care clinic to calculate FIB4 and APRI scores for 225/231 (97%) patients referred with suspected ArLD (6 patients did not have an AST value available).

The equation for FIB4 is: Age (years) × AST (U/L)/ [Platelet count (10^9/L) × ALT (U/L)].

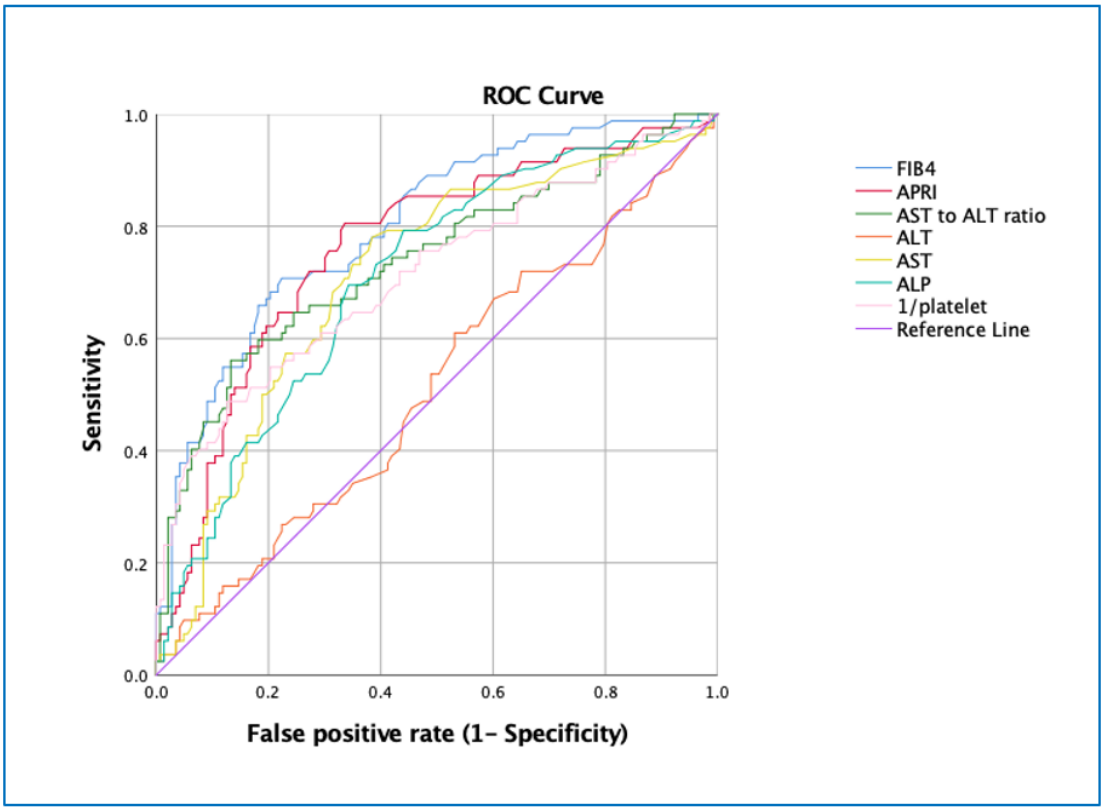
The equation for APRI is: [(AST/ULN AST) x 100]/Platelet count

Median FIB4 and APRI were 1.58 (IQR 0.97-3.29) and 0.68 (IQR 0.36-1.53) respectively. Both scores independently predicted the clinical diagnosis of advanced fibrosis in secondary-care in multivariable regression analysis (for FIB4, OR=1.658, 95% CI 1.397 to 1.967, p <0.001; for APRI, OR=1.485, 95% CI 1.204 to 1.832, p <0.001).

When ROC analysis was used to examine the ability of NIT based on routine blood tests to predict a diagnosis of advanced fibrosis, FIB4 performed the best (AUROC 0.801), compared with APRI, AST, ALT, ALP and platelet count (all p <0.005 using DeLong

comparison) and numerically but not significantly better than APRI (p = 0.06) (Figure 4.4).

Figure 4.4: ROC analysis of the performance of indirect tests for fibrosis and simple liver blood tests in the detection of advanced fibrosis (composite clinical judgement) in patients referred with suspected ArLD. (N=231)



AUROCs with 95% CI in brackets: FIB4: 0.801 (0.742 to 0.860); APRI: 0.763 (0.697 to 0.829); AST:ALT ratio: 0.739 (0.668 to 0.809); ALT: 0.512 (0.433 to 0.591); AST: 0.711 (0.640 to 0.782); ALP: 0.708 (0.638 to 0.777); 1/platelet: 0.714 (0.641 to 0.787). (All p values <0.001 apart from ALT which was non-significant at p = 0.758)

Amongst the cohort of patients with ArLD referred to secondary-care, 35.81% were judged to have advanced fibrosis and thus 64.2% could be considered 'unnecessary' referrals. Use of a FIB4 threshold of ≥3.25 (289) could have improved the detection of patients with advanced fibrosis nearly five-fold (OR=*4.82; 95% CI 2.56 to 9.09, p <0.0001)*, leading to a 79.3% reduction in unnecessary referrals to secondary care (64.2% to 27.1%) (OR = 0.21; 95% CI 0.11 to 0.39, p <0.001) However, this would be

associated with the exclusion of 39 patients judged to have advanced fibrosis (false negative rate of 47.6%) (Table 4.5).

When modelling the referrals using a FIB4 threshold of ≥1.45,(289) the detection of advanced fibrosis improved two-fold compared with standard-care (OR=1.98; 95% CI 1.27 to 3.09, p = 0.0027) and reduced the number of unnecessary referrals from 64.2% to 47.5% (OR=0.5; CI 0.32 to 0.79, p = 0.003), with 103 patients (45.7%) having a FIB4 score below 1.45 that could have remained in primary care. The false negative rate was lower using FIB4 ≥1.45 compared to threshold ≥3.25 (18/103, 22% compared to 39/103, 47.5%; X^2=10.60; p=0.001).

Table 4.5: Accuracy of indirect fibrosis markers in detecting advanced fibrosis in a cohort of 231 patients referred from primary care with suspected ArLD. (N= 225/231.)

Indirect fibrosis test (n=225/231)	Correctly classifies	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	LR+ (95% CI)	LR- (95% CI)	TP FP	FN TN	False negative rate (%)	False Positive rate (%)
APRI	165	64.6%	78.3%	63.1%	80%	3.02	0.44	53	29	35.4	21.7
≥ 1	(73.3%)	(54-75)	(70-85)	(52-73)	(72-86)	(2.13-4.28)	(0.33-0.6)	31	112		
FIB4	170	52.4%	88.8%	72.9%	76.5%	4.69	0.54	43	39	47.6	11.2
≥ 3.25	(75.6%)	(41.2-63.5)	(82.2-93.3)	(59.5-83.3)	(69.1-82.6)	(2.83-7.77)	(0.43-0.67)	16	127		
FIB4 ≥1.45	149 (66%)	78%	59.4%	52.4%	82.5%	1.92	0.37	64	18	22	40.6
		(67.3-86.1)	(50.9-67.4)	(43.3-61.5)	(73.5-89)	(1.53-2.42)	(0.24-0.56)	58	85		

4.5 Discussion

In this chapter, through analysis of 231 patients referred to secondary care for suspected ArLD, I found that two-thirds had no evidence of advanced fibrosis, representing unnecessary referrals. This can be explained in part because the commonest reasons for referral were abnormal LFTs and ultrasound scans, neither of which are sensitive or specific tests for advanced fibrosis (1). While some of these patients may have benefited from a hepatologist's advice about the wider consequences of their drinking, many primary-care physicians consider that they are actually better placed to deliver brief advice about hazardous or harmful drinking and that referral to liver specialists should be restricted to patients with ArLD who have suspected advanced fibrosis or cirrhosis. In fact, not only is there no evidence that patients benefit from getting this counselling from a hepatologist, but previous research has actually identified nursing staff as best placed to deliver alcohol brief interventions (290), and another study by Eggleston et al. found that patients regard advice delivered by nurses as equal to that delivered by doctors (291), therefore community alcohol advice does not appear to be inferior to, and might even be superior to, secondary care alcohol advice.

In having lots of 'unnecessary referrals' to secondary-care (as per NICE and BSG guidelines, patients without advanced fibrosis can remain in primary care (1, 38)), it is also not only an inefficient use of NHS resources, but is likely to be putting patients through the stress of unnecessary investigation through the hospital system when they could be safely managed in primary care. This is supported by the Cochrane study on brief interventions in AUD (213), which found that brief advice in primary

care was effective in reducing alcohol intake in people drinking at harmful levels, and moreover that extended interventions were no more effective than brief interventions. The authors opined that this seems counter-intuitive compared to other areas of health care, where more intervention can amount to a larger impact, but they noted that brief intervention usually opportunistically targets people who are not seeking help with their alcohol consumption, and so an extended intervention may be perceived as excessive, and alienate those who have not accepted that they have problematic drinking (213).

There is also an argument that GPs or practice nurses may actually be best placed to deliver the alcohol intervention, as they are likely to know the patient better than a secondary care doctor, and may have built rapport with them and their family over a long period of time.

I also found that only 38/231 (16%) patients with suspected ArLD had any kind of fibrosis assessment prior to referral to secondary-care, the majority of whom had features of metabolic syndrome or were overweight and received FIB4 and ELF tests suggesting that their GPs had followed the local NAFLD pathway that incorporates these investigations. The lack of fibrosis testing in 84% of this cohort is perhaps surprising, given that current national guidelines recommend the use of non-invasive tests for people with AUD (1), and reflects the need for further education in primary and secondary care, the need for clear, easy-to-follow guidelines and pathways for primary care physicians to undertake the investigations and refer appropriately into secondary care, and also for further prospective

research validating such community pathways involving the use of non-invasive tests in people with AUD.

The majority (64%) of patients referred with suspected ArLD were overweight, obese or had other features of metabolic syndrome and were reclassified as BAFLD. These patients with 'BAFLD' had double the odds of advanced fibrosis when compared to the NAFLD cohort suggesting that hazardous drinking in this cohort was associated with a doubling of the risk of liver fibrosis in people who are overweight or obese but who did not have hazardous drinking habits. This both highlights the increased risk of liver disease in patients with dual pathology and the importance of considering multimorbidity in chronic liver disease. That is, the need to consider the patient holistically and investigate for and treat each co-existing condition with referral as appropriate to relevant members of the multi-disciplinary team such as dieticians, alcohol specialist nurses, diabetes clinics, etc. This is in keeping with recent findings from a recent systematic review and meta-analysis by Glyn-Owen et al in 2020 (278), who found that compared to people of normal BMI (<25) who drink alcohol within UK recommended limits (<14 units/week), the relative risk of chronic liver disease in overweight people drinking $\geq$ 14 units/week was 3.32 (95% CI 2.88-3.83), and relative risk in obese people drinking $\geq$ 14 units/week was 5.39 (95% CI 4.62-6.29). This meta-analysis included results of a study by Trembling et al. which incorporated over 95,000 participants (292).

I believe there is considerable benefit in raising the profile of the added adverse risk of alcohol and obesity together, and that the term 'BAFLD' captures this concept well. Whilst a new term 'MAFLD' (Metabolic Associated Fatty Liver Disease) (293) has been recently suggested as a new term for NAFLD (in part to remove the stigma

associated with the term ‘alcoholic’ in ‘Non-Alcoholic Fatty Liver Disease’, and to remove association of the word ‘alcohol’ with NAFLD); there isn’t yet any term in use for the combined condition of fatty liver (NAFLD) with Alcohol-related Liver Disease. This is despite my study highlighting how common this combination is, and the importance in terms of risk in identifying people with dual risk factors. Whilst I recognise that this terminology is not yet widely accepted, I think that the term ‘BAFLD’ (originally coined by Parkes et al.) (286) is easy to apply in clinical practice, is less ambiguous, and will help raise awareness of the added risks of combined pathology and hopefully improve knowledge mobilisation within public health, policy makers and the general public.

With regards to alcohol unit thresholds, although national guidelines state that the risk of advanced fibrosis develops at a lower alcohol unit threshold for women than men (<35 U/w for women, <50 U/w for men), (1) these thresholds do not appear to be based on published data. Few studies have investigated the association between levels of alcohol consumption and the risk of advanced fibrosis, and those that did have reported a wide range of thresholds (83, 89, 90, 294-296). Furthermore, the levels of drinking that cause harm in the context of people who are overweight or obese are not known. Through this analyses of 762 patients that included a high prevalence of overweight and obese people, I derived thresholds identical to those in national guidance, i.e. 35 U/w in women and 50 U/w. It should be noted that these thresholds focus purely on the risk of advanced liver fibrosis and cannot be generalized to other health measures such as cardiovascular risk, mental health.

National guidelines state that there is an increased risk to general health above 14 U/w.

The performance of “indirect” serum fibrosis tests is well reported in NAFLD, but less so in ArLD. In this study cohort of 231 ArLD patients, FIB4 and APRI outperformed simple liver blood tests (ALP, ALT, AST and platelet count) in predicting a diagnosis of advanced fibrosis on AUROC analysis, with FIB4 having the highest AUROC of 0.801. This FIB4 AUROC is in keeping with findings from my systematic review in Chapter 3, with reported AUROCs of >0.8 . Furthermore, when examining FIB4 at literature-derived binary thresholds of 3.25 and 1.45, (23, 289) I found that it did not perform as well in detecting clinically defined advanced fibrosis as has been reported in a recent study in which all participants were required to undergo liver biopsy (23). In this study, by Thiele et al. in 2018, 289 patients (primary or secondary care) with a history of excessive alcohol for at least a year (>24g/day for women and >36g/day for men) all underwent a liver biopsy plus non-invasive fibrosis tests on the same day. Thiele et al. found that FIB4 at a threshold of 3.25 detected advanced fibrosis with specificity of 91% (95% CI 86-94), sensitivity 58% (95% CI 45-70), PPV 64% (95% CI 51-76), NPV 88% (95% CI 83-92), and false negative rate of 42%.

In my study, stratifying patients in primary care using a FIB4 threshold of 3.25 could have reduced unnecessary referrals by 79.3%, with PPV and NPV for the detection of advanced fibrosis of 72.9% and 76.5% respectively (NPV lower than in the Thiele et al. study). However, the associated false negative rate was 47.5% (similar to Thiele et al. 42%) suggesting that nearly half the cases of advanced fibrosis would be not

referred to secondary care, making it unsuitable for case stratification. A FIB4 threshold of 1.45 produced a lesser false negative rate of 22%, and although it reduced the proportion of unnecessary referrals by 50%, the PPV was 52.4% and overall, this threshold correctly classified only 66% of patients into presence or absence of advanced fibrosis. Thiele et al. found that the Forns Index (an algorithm based on age, platelet count, GGT, and cholesterol level) at a threshold of 6.8, performed slightly better than FIB4, with improved sensitivity of 71% (95% CI 59-82), similar specificity of 89% (95% CI 84-93), similar NPV and PPV to FIB4 (91% and 66% respectively), and improved false negative rate of 29% (23). The authors of this study went on to report a post-hoc analysis of a two-step approach using Forns index first as an exemplar 'indirect fibrosis test' to use in the first instance in primary care as a cheap and easy test, followed by ELF in those with high Forns index. When adopting a Forns threshold of 4.1, 58/128 (45%) had Forns below 4.1 and were deemed low risk of advanced fibrosis and did not require further investigation. Those with Forns ≥4.1 went on to have an ELF test, of which only 10/70 (14%) had high ELF (≥ 10.5). This strategy resulted in the prevention of referral to secondary care in 92% (118/128) of primary care patients. Unfortunately, whilst Thiele et al. included in their published study the diagnostic accuracy data (NPV, PPV, sensitivity, specificity, etc.) for Forns index at a threshold of 6.8, they did not report the corresponding data for threshold at 4.1, despite advocating the use of this threshold.

I also note that Forns index necessitates the use of a cholesterol level which actually makes it a less practical 'first step' test for use in primary care where this test is not performed routinely, unlike FIB4 which would be calculable from the patients routine

Liver function blood tests that would be already available, and would not require the patient to be called back in for a repeat blood test.

Overall, my results suggest that an effective ArLD pathway would require the use of either a NIT with better diagnostic performance or the use of two or more NIT in series, as employed in the Camden and Islington NAFLD pathway (14) . As FIB4 is a readily available and easy to adopt test, and has been shown to perform well (AUROC >0.8) in my current study, potentially reducing unnecessary referrals by 50%, and also having been found to perform well in my systematic review in chapter 3– I think it would be worth exploring this further. This could perhaps be prospectively examined at a lower threshold of 1.45 to minimize false negative rates, and evaluating it as a 'first step' primary care test to see if it could reduce the need for further fibrosis tests and referral in those with low FIB4 scores.

This retrospective study has limitations. Firstly, it lacked access to liver biopsy as a reference-standard to stage fibrosis severity. Self-reported alcohol intake at the point of referral to secondary-care was also used to record drinking behavior and this may not be reliable. However, this clinic-based sample of 'real-world' cases reflects current practice in the UK and many other countries and highlights the opportunity to stratify patients with ArLD community settings to ensure that only those with a high likelihood of advanced fibrosis are referred for liver specialist care.

Having so many 'unnecessary referrals' to secondary-care is not only an inefficient use of resources, but also exposes patients to unnecessary investigation and the associated time, risk and anxiety. These patients could be managed more appropriately in community settings with an appropriate focus on the wider harms associated with their drinking. This could include allowing a focus on mental health

problems which often co-exist with AUD, providing advice and support on cardiovascular risk (hypertension and obesity frequently associated with AUD), as well as providing community support in alcohol reduction. Conversely, emphasis on finding those with advanced fibrosis might improve the early detection of those people living with AUD who are likely to progress to cirrhosis and suffer life-limiting effects of their drinking.

Based on the performance of APRI and FIB4 in this cohort, I would not recommend routine use on their own to risk stratify patients with AUD. Instead, further evaluation of pathways incorporating non-invasive tests such as ELF or FibroScan (1, 38) (23) would be preferable, with evaluation of FIB4 or APRI as a potential cost-saving 'first step' as per the NAFLD guidance (14).

In this study I have also highlighted the multi-causality and multi-morbidity endured by patients with ArLD and NAFLD. Although the dual-existence of excess alcohol and obesity is recognized, the low threshold of alcohol consumption at which the risk of advanced fibrosis nearly doubled in this cohort highlights the importance of communicating this risk to patients with fatty liver disease in clinics and primary care, and through public health messaging. There is a need for greater awareness about this increased liver risk with alcohol consumption in those people that are living with overweight and obesity amongst healthcare professionals, policy makers and the public and a need for a multi-disciplinary approach to address the lifestyle risk factors that are likely to influence the morbidity and mortality of those with BAFLD.

In summary, I found that the current referral strategy for patients with alcohol use disorders at risk of liver disease from a primary care borough is inefficient and ineffective. It is essential that there is increased awareness of the need to proactively

search for fibrosis in primary care in those patients at risk, using appropriate strategies incorporating non-invasive testing, and education of the guidelines for fibrosis testing in both AUD and NAFLD. In addition, there is a need for improved collaboration between primary and secondary-care services to develop referral pathways employing NIT, with planned prospective evaluation to further refine thresholds for referral and education to improve awareness and the advice provided to patient about the impact of overweight/obesity and alcohol on liver health.

These results led me to set-up a new referral pathway from primary to secondary care, with a planned prospective evaluation of the use of non-invasive tests in primary care to stratify patients with AUD into secondary care. This process is outlined in Chapter 5.

CHAPTER 5

Implementing a community referral pathway involving the ELF test in patients with Alcohol Use Disorder – the 'Camden and Islington alcohol pathway'

5.1 Background

As I discovered in Chapter 2, 75% of patients with ArLD present 'late' to healthcare, being unaware of their liver condition until they have developed complications of cirrhosis when it can be too late for interventions at this point to alter outcomes. Whilst stopping drinking in late disease can lead to dramatic improvements with fibrosis regression on histology and improvements in portal pressure, a third of patients will die before their liver recovers (5, 136). Informing patients of an early diagnosis of liver disease leads two-thirds of harmful or dependent drinkers to stop drinking (5, 137). As cirrhosis and fibrosis are 'silent' conditions, not detectable by routine liver function blood tests, it is essential for clinicians to proactively test for fibrosis in people with risk factors.

Conversely, as I discovered in the previous chapter, two-thirds (64%) of patients with AUD who were referred from primary care for evaluation by a secondary care hepatologist were considered 'unnecessary referrals', as they did not have a diagnosis of advanced fibrosis, and were discharged back to primary care. Had they had a non-invasive fibrosis test in primary care before a decision was made to refer, my modelling predicted that this percentage would be lessened.

A recent study by Srivastava et al., (168) investigated if a pathway incorporating blood-based fibrosis tests could improve the proportion of patients referred to secondary care who had advanced fibrosis. This was a two-step pathway, applying the FIB4 score in the first instance in people found to have fatty liver on ultrasound, and no history of excess alcohol (< 14 units per week in females, <21 units per week

in males), and raised transaminases. Data for 3,011 patients were collected over a 26-month period from 2014-2016, with 48% uptake of the pathway by GP practices. On this pathway a low FIB4 score (< 1.30) was used to exclude advanced fibrosis (≥F3), a high FIB4 score (>3.25) prompted referral to a secondary care hepatology clinic for further assessment for chronic liver disease, and an intermediate score (1.3-3.25) prompted the use of an ELF test, with a threshold of 9.5 for advanced fibrosis, triggering referral to a hepatologist. The primary outcome was the reduction in 'unnecessary' NAFLD referrals in patients referred after FIB4 or ELF using the pathway, compared to patients referred without using the pathway. The unnecessary referrals were those that were deemed not to have advanced fibrosis on clinical evaluation in secondary care, and this was based on composite clinical assessment by the hepatologists (including use of imaging, FibroScan, and biopsy where deemed indicated by the hepatologist). Results from the pathway evaluation were striking, with an 88% reduction in the proportion of 'unnecessary referrals' (no advanced fibrosis) by use of the pathway (79/83 unnecessary pre-pathway (95.2%), compared to 107/152 (70.4%) unnecessary when the pathway was followed (OR 1.12, 95% CI 0.042-0.349, p<0.0001). Use of the pathway also led to an increase in the detection of advanced fibrosis by five-fold (OR 5.18, 95% CI 2.97-9.04, p<0.0001) (168). This study validated the British Society of Gastroenterology guidelines that recommended this use of FIB4 and ELF in NAFLD for two-stage stratification. The strategy has since been proven to be cost-effective (297).

In ArLD, Thiele et al. (23) described a large multi-centre prospective study of non-invasive tests for liver fibrosis in 289 patients with AUD conducted in primary and secondary-care in Denmark. This compared several non-invasive tests including the ELF test, and FibroScan, to liver biopsy as the reference standard, with all tests performed on the same day. The authors concluded that the ELF test was an excellent diagnostic test in ArLD, which was accurate and safe using a threshold of 10.5 for advanced fibrosis to triage patients from primary to secondary-care, with a NPV (negative predictive value) of 98% in the primary-care cohort (sensitivity 75%, specificity 97%, AUROC 0.92). It performed equally to FibroScan in the intention-to-diagnose protocol, but FibroScan was noted to have a failure rate of 5%.

However, the impact of using a similar strategy of non-invasive testing of patients with AUD in primary-care in the UK is undetermined.

5.2.1 Current guidelines on non-invasive fibrosis testing in ArLD

Non-invasive fibrosis testing is now recommended in national guidance for patients with AUD. National Institute of Clinical Excellence (NICE) recommend FibroScan in men drinking > 50 units of alcohol per week, and women drinking > 35 units of alcohol per week ("and who have done for several months") (38). The British Society of Gastroenterology (BSG) guidelines (Figure 5.1) recommend use of FibroScan or ELF test, also for people drinking at or above 35/50 units per week for women/men (these unit thresholds were based on NICE guidelines (1)). However, the basis for the recommendations of 35 and 50 alcohol unit thresholds is unclear, and does not seem to be based on any clear evidence for these thresholds. The full NICE guideline

explains that they found 6 studies (94, 101, 298-301) reporting on the association of alcohol consumption and either risk of diagnosis of alcohol-related cirrhosis, or risk of hospitalisation or death from cirrhosis (38). The studies were all rated as either 'low quality' or 'very low quality' evidence, and were not able to be pooled in meta-analysis as they each reported different categorisations for the level of alcohol consumption, and also because of heterogeneity in the confounding factors adjusted for in the analyses. NICE noted there was a general increase in the hazard or odds ratios related to cirrhosis or cirrhosis-related deaths associated with alcohol intake above 7 drinks per week, or 5 grams per day (just over 4 units a week) (1 unit = 8 grams). However, due to the differences between studies, the NICE GDG (Guideline Development Group) found it impossible to come to a conclusion on alcohol thresholds from these studies, above which to recommend a diagnostic assessment for cirrhosis. Instead, the committee came to a general consensus of opinion that it was not necessary to test everyone drinking above the government recommended 'safe limits' (14 units per week) for cirrhosis, and that people who fall within the "NHS definition of higher risk drinking" (> 50 units men, > 35 units women) would be a reasonable group to target for cirrhosis testing. They noted, however, that there is likely to be a case for testing people below these unit thresholds. These unit thresholds of 35 and 50 units are arbitrary rather than having been based on definitive evidence.

As discussed in Chapter 2, there is now convincing evidence that there is 'no safe lower limit' of alcohol for *overall* mortality, when taking into consideration the risk of cancers at even a very low alcohol intake (80). For the risk of alcohol on the liver,

however, consensus does not seem to have been achieved among the scientific community, with a variety of numbers of units reported as the threshold for harm. My study in Chapter 4 found a threshold effect of 35 units per week in women and 50 units per week in men, above which the patients had greater odds of a diagnosis of advanced fibrosis or cirrhosis (OR advanced fibrosis for men ≥50 units = 2.74, 95% CI 1.51-to-5.00, p = 0.001, OR advanced fibrosis for women ≥35 units = 5.11, 95% CI 1.31-to-20.03, p = 0.019). These thresholds are in keeping with the BSG (Figure 5.1) and NICE guidelines on unit thresholds above which to test for liver fibrosis.

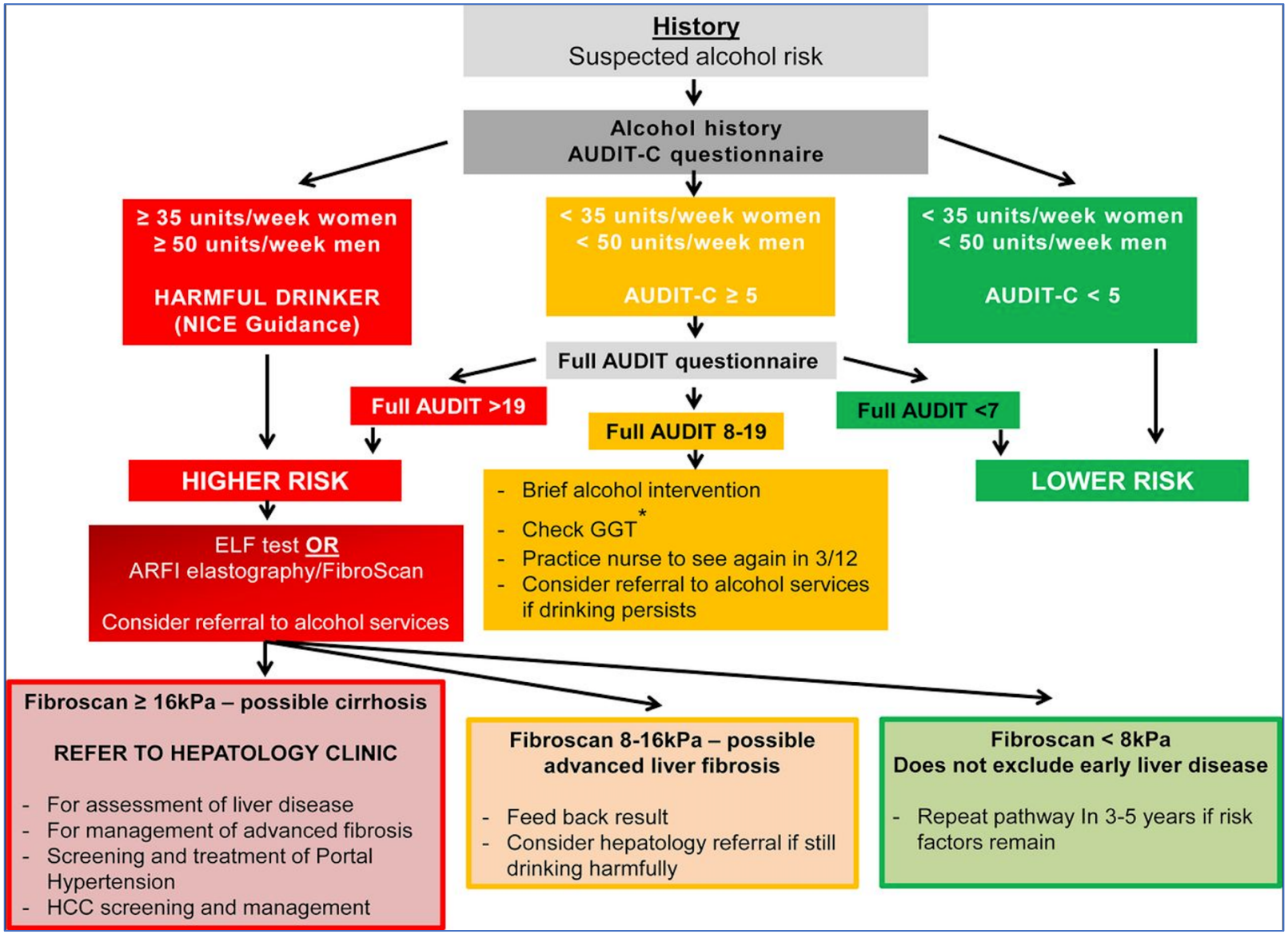


Fig 5.1
BSG alcohol-related liver disease guideline for non-invasive testing.

(Figure reproduced from Newsome et al., BMJ, 2018(1) with permission under the creative commons license, CC BY 4.0, https://creativecommons.org/licenses/by-nc-nd/4.0)

5.2.2 The need for validation of the BSG Guidelines

As I outlined in Chapter 4, although recommended in national guidelines, the use of non-invasive fibrosis assessments is not yet in widespread use in the UK, with decisions to refer patients with AUD to secondary-care still most commonly based on LFTs (Liver Function Tests) and alcohol history, rather than fibrosis measurement.

Moreover, the recommended pathway in the BSG guidelines shown above has not been validated in NHS primary-care settings, and the impact of using such strategies in primary-care in the UK is unknown. In particular, it is not known whether using the ELF test (or FibroScan) in primary-care to triage patients with AUD to secondary-care is effective in terms of increasing detection of advanced fibrosis and cirrhosis, reducing the proportion of 'unnecessary referrals' (those without advanced fibrosis) and whether it is practical and cost-effective.

This chapter describes the development and implementation of a new community pathway involving the ELF test in primary care in people identified to have AUD. Evaluation of the impact of the pathway is beyond the timeframe of this thesis.

5.3 The Camden and Islington (C&I) Alcohol pathway

5.3.1 Aim and objectives

I aimed to design and launch a new alcohol pathway in two London Boroughs (Camden and Islington) involving the use of the ELF test in patients with AUD, on the basis of the BSG guidelines.

The Camden and Islington (C&I) alcohol pathway has the potential to benefit:

1. Patients through a reduction in unnecessary hospital appointments and investigations; improved detection of liver damage due to alcohol; better understanding of the consequences of harmful drinking

2. The NHS through reduction in the unnecessary use of secondary-care resources; improved detection of liver disease at a point at which intervention can avoid harm; cost savings should accrue through the reduction in referrals and investigations and avoidance of harms from unnecessary investigations, and cost-utility should arise through the reduction in harms from advanced liver disease.

5.3.2 Reason for choosing ELF over FibroScan

Firstly, the proposed FibroScan thresholds in the BSG guidelines ≥8 kPa (1) have not been validated for use in ArLD, as I outlined in my systematic review in Chapter 3 (NICE do not suggest any particular FibroScan threshold in their guideline).

Secondly, the location of the pathway in Camden and Islington Clinical Commissioning Groups (CCGs) is the same location as the successful NAFLD pathway set up in 2014, which used the ELF test (14). Therefore, General Practitioners (GPs) in these boroughs are familiar with decision-making around the ELF score, and these CCGs have access to ELF testing. Thirdly, the ELF test is a simple blood test which can be performed at the same time as routine blood tests at the GP surgery and does not need any specialist equipment or training, and it takes less time than FibroScan, is less expensive, and has no associated failure rate (297). It is therefore more practical than FibroScan in the community setting.

5.3.3 Objectives

1. To review the evidence base surrounding the BSG guidelines, and select an appropriate ELF threshold for use in primary care for ArLD, as per best evidence.
2. To meet with local GP leads, public health professionals and hepatologists to scope current opinions on use of ELF in people with AUD, and ascertain if this is practical and achievable for uptake in primary care.
3. To collaborate with patients and the public on the design of the pathway
4. To design a new pathway incorporating BSG guidelines, in a way that is practical and applicable to the local clinical commissioning groups (CCGs)
5. To obtain the necessary permissions to allow the pathway to be officially approved for use in the NHS, and launched on the Camden and Islington GP website

6. To disseminate information about the pathway in primary and secondary care to maximise pathway uptake.
7. To design a future evaluation of the pathway, aiming to identify the proportion of 'unnecessary' referrals on the C&I pathway compared with standard care. The evaluation would also enable the calculation of the proportion of referred patients with advanced fibrosis in patients referred on the C&I pathway compared with standard-care, and a determination of the cost-effectiveness of the pathway.

5.3.4 Methods

5.3.5 Setting

Camden and Islington CCG

Camden and Islington are two north London boroughs serving a population of 430,000 patients, with 89 GP practices (35 in Camden, 34 in Islington), supported by secondary care provision from three hospitals – the Royal Free London NHS Foundation Trust, the Whittington Hospital NHS Trust, and University College London Hospital NHS Foundation Trust. When patients from primary care in Camden or Islington need referral to a hepatologist for suspected ArLD, the vast majority are referred to the Royal Free, which is the main hepatology service for North London (80% Camden hepatology referrals go to the Royal Free, with 13% going to St Mary's (Imperial), 2% to Bart's Health NHS Trust and 3% Other; 70% of Islington hepatology referrals also go to the Royal Free, with 30% going to UCLH or 'other'). Of the referrals to UCLH, the majority are viral hepatitis patients. Therefore, when planning an

evaluation of the pathway, this could be achieved by auditing the Royal Free ArLD referrals to hepatology, which would capture the vast majority of referred patients from Camden and Islington.

5.3.6 Designing the pathway

I initiated discussions surrounding the set-up of a new pathway in 2019, with the Camden CCG lead GP, who had been heavily involved in the previous successful NAFLD ELF pathway, and was enthusiastic to be involved in updating the alcohol pathway in a similar way. The Camden CCG shares its guidelines and pathways with neighbouring CCG Islington, so the CCG lead for Islington was also involved in the alcohol pathway discussions. In addition to primary care leads, the design and set-up of the pathway was a collaborative process with regular meetings held between myself and Professor Rosenberg with public health consultants, the alcohol-lead consultant hepatologist at the Royal Free, a health economist, statistical modeller, and expert patients.

Whilst there was currently an existing alcohol pathway in Camden that recommended conducting an AUDIT questionnaire, with referral to community alcohol services in hazardous or harmful drinkers, and performing LFTs, there was no mention of any fibrosis testing in the current local Camden pathways (which are shared with the neighbouring Islington CCG).

Over the course of 10 months in 2019, a series of regular meetings to discuss the design of the pathway was held at the Royal Free hospital and at the Hampstead Group Practice (the latter being the location of the Camden CCG lead GP). During

these meetings, the idea for the pathway and design process was discussed with the GP leads for Camden and Islington CCGs, the local hepatology alcohol lead, public health consultant, health economist and statistical modeller. There was a unanimous agreement that the alcohol pathway should be updated to incorporate the BSG guidelines, to reflect the use of the ELF test in women drinking ≥ 35 units per week and men ≥ 50 units per week, or who had AUDIT scores to reflect hazardous or harmful alcohol intake.

5.3.7 Patient and public involvement

I arranged two meetings at the Royal Free and follow up email communication with expert patients, to involve them in the planning of the pathway. These were three people who were identified by Professor Rosenberg, who each had previous experience of participating in studies of diagnostic tests in chronic liver disease, and also lived experience of liver disease and referral from primary care to the hepatology service.

I used the 'INVOLVE' guidelines (302) to arrange and conduct these meetings, which adhered to the specified standards for the involvement of members of the public in research.

Standard care was described, and the use of ELF testing in people with AUD in the community to stratify patients to secondary-care. Alternative approaches using blood and physical tests to assess liver fibrosis were discussed, and opinions regarding their use, interpretation of results and actions taken based on their use were explored. Their views on the acceptability of stratification on primary-care

and the consequent impact on referral to secondary care were discussed, and helped affirm the design of the study.

5.3.8 Ethics

The pathway was set-up to represent the current BSG guidelines on non-invasive fibrosis testing for ArLD. Thus, implementation and evaluation of the pathway was considered to be a 'service evaluation' using secondary anonymised data, not 'research', according to the Medical Research Council 'Is my study research' tool (IRAS ID 264420), and did not require ethics approval or patient consent.

Incorporating BSG guidelines into local guidelines

The BSG guidelines (Figure 5.1) were reviewed by me with input from the public health consultants and primary care leads. The BSG alcohol thresholds of 35 and 50 units per week (women and men) were agreed to be used in the pathway, with an ELF test to be advised in primary care in people drinking above these thresholds. The BSG guidelines also include recommendations for use of the AUDIT questionnaire in people drinking less than 35 units (women) or 50 units (men), to identify hazardous or harmful drinkers, with an AUDIT score > 19 representing higher risk drinkers and prompting ELF test. In those patients with intermediate AUDIT scores (8-19, representing 'hazardous drinking'), the BSG guidelines advise that for patients who 'continue to drink at hazardous levels, consideration should be given to fibrosis assessment as for the higher-risk category' (1).

Discussions with the primary care leads, however, revealed that full AUDIT questionnaires were not routinely performed in primary care because they were deemed to be too time-consuming within the constraints of the patient appointments. It was therefore decided that AUDIT would be included as an option in the pathway, along with assessment of amount of alcohol consumed per week, but with an alternative additional option for the GP to perform a shorter 'AUDIT-C' questionnaire. Details of the AUDIT-C can be found in Table 5.1.

Table 5.1: Alcohol use disorders identification test consumption (AUDIT C) (303)

Questions	Scoring system					Score
	0	1	2	3	4	
How often do you have a drink containing alcohol?	Never	Monthly or less	2 to 4 times per month	2 to 3 times per week	4 or more times per week	
How many units of alcohol do you drink on a typical day when you are drinking?	0 to 2	3 to 4	5 to 6	7 to 9	10 or more	
How often have you had 6 or more units if female, or 8 or more if male, on a single occasion in the last year	Never	Less than monthly	Monthly	Weekly	Daily or almost daily	
AUDIT-C Score						Total =
≥ 5 = positive screen 0 to 4 indicates low risk 5 to 7 indicates increasing risk 8 to 10 indicates higher risk 11 to 12 indicates possible dependence						

5.3.9 Pathway design process

The pre-existing alcohol pathway contained information about AUDIT scores, how to refer to local alcohol nurses and help with detoxification therapy, and prompted GPs to perform LFTs, but did not include any information about assessment for fibrosis.

I was able to use this existing local alcohol pathway as a template, and designed a new pathway such that it could incorporate some of the existing pathway elements, but which incorporated the new BSG guideline advocating the use of the ELF test in people drinking excess alcohol (Figure 5.2). This was approved by the GP leads for Camden and Islington, public health consultants, and the alcohol-lead for hepatology at the Royal Free. Figure 5.2 displays my flow-chart for the final pathway that was incorporated into the main alcohol pathway on the Camden CCG website (Figure 5.3).

Figure 5.2: Flow diagram illustrating new pathway for ELF testing, for incorporation into the Camden and Islington alcohol pathway.

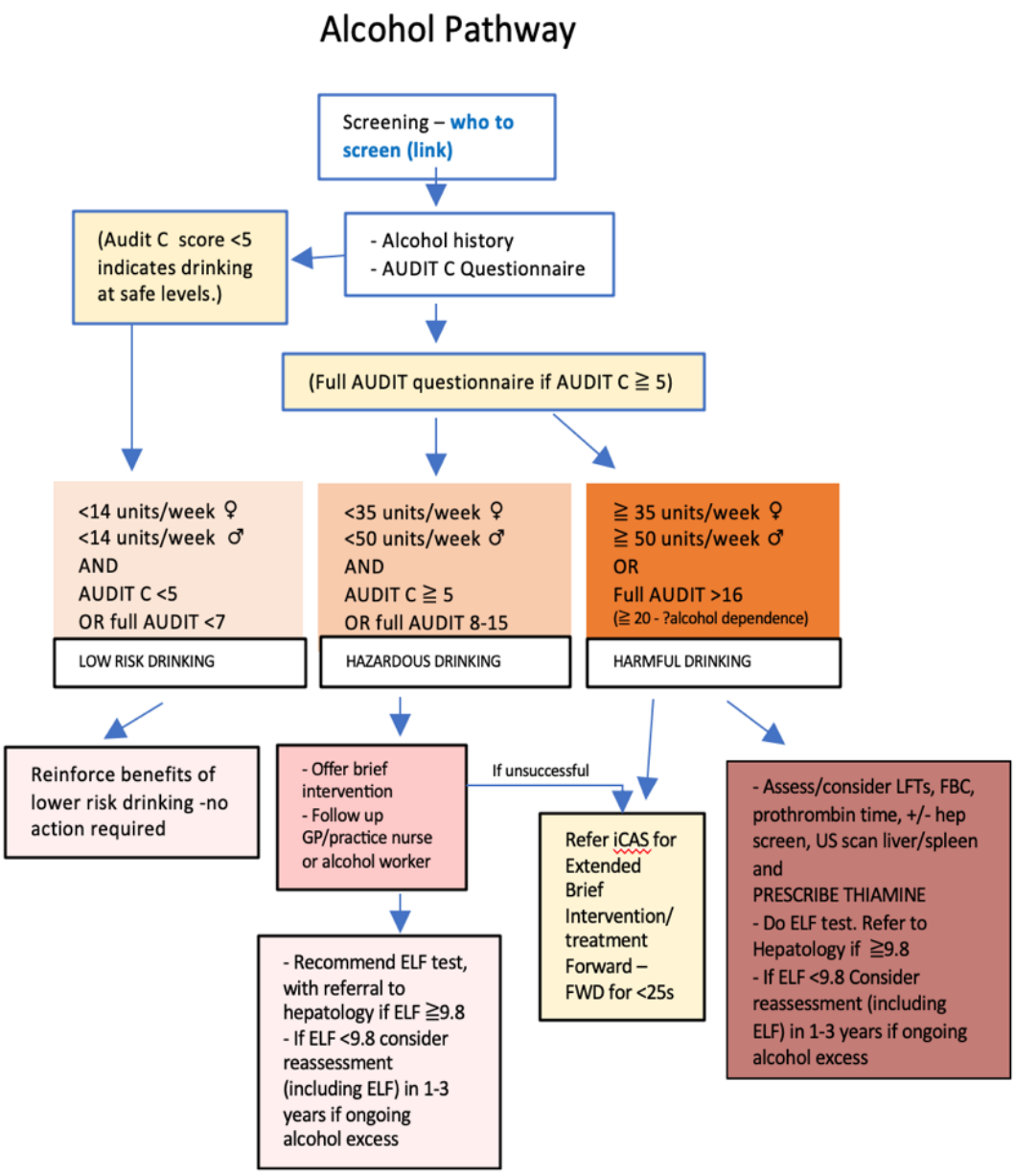
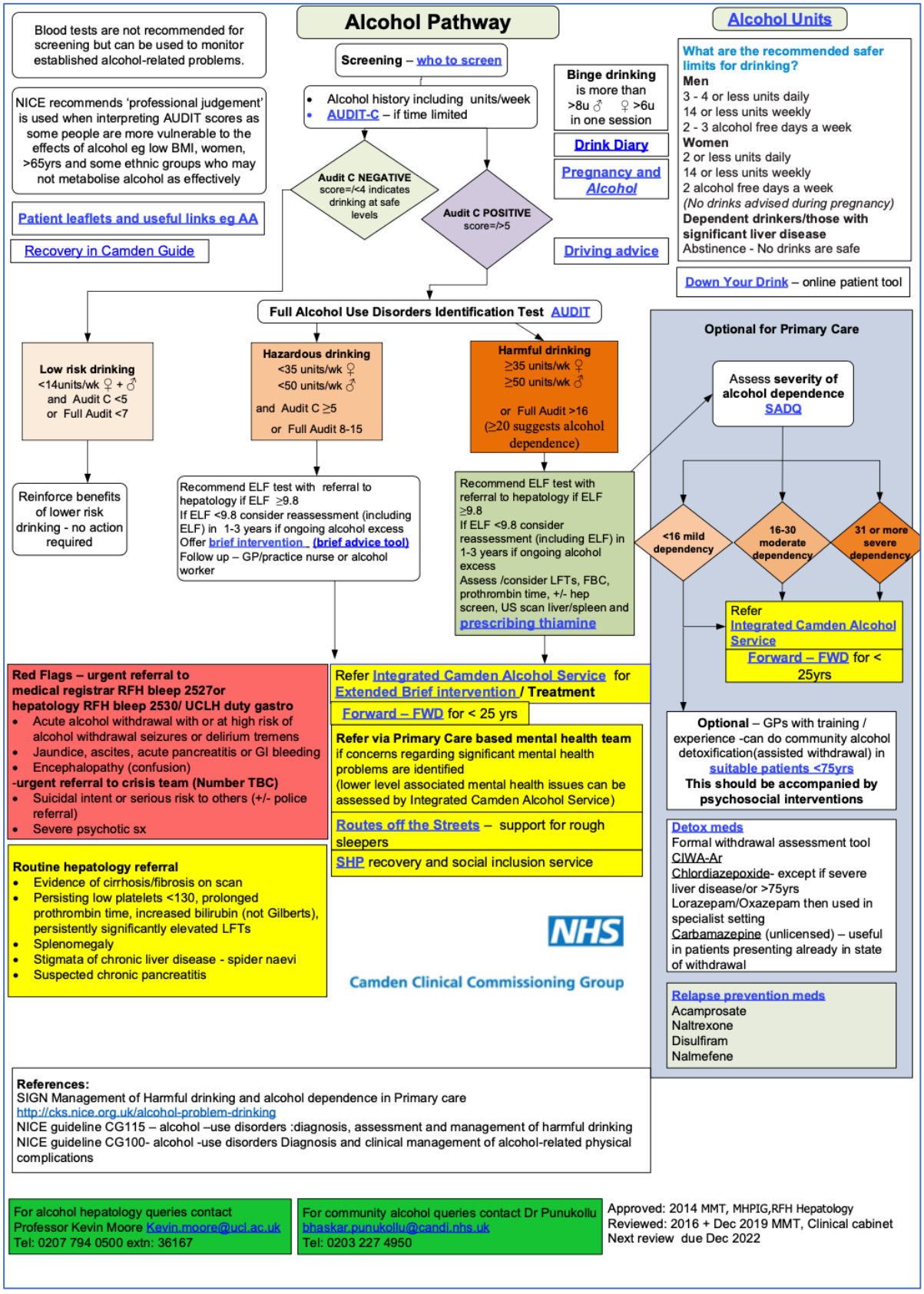


Figure 5.3: Final new alcohol pathway on Camden CCG website 2020:



5.4 Deciding on the ELF threshold for the pathway

The NAFLD pathway in Camden and Islington currently uses a 9.8 ELF threshold for advanced fibrosis. For evaluating ELF in ArLD, it has been suggested to be safe and effective in the Denmark study to use a referral threshold of 10.5 in AUD in primary care (23). However, a 9.8 threshold was also evaluated in this Danish study (23) which showed comparable performance (Table 5.2). This was a large prospective study of ArLD in 298 patients, all of whom had a liver biopsy on the same day as the ELF test (results included in my systematic review in Chapter 3). In their primary care cohort (n=128), although the 10.5 threshold was more specific, and with a higher PPV, the NPV was identical to the 9.8 threshold (98%), and sensitivity was equal at 75% between the two thresholds (Table 5.2). In the overall study cohort of n=298, whilst the 10.5 threshold was again more specific than 9.8, the NPV was higher with the 9.8 threshold at 96%, than the 10.5 threshold (94%), with higher sensitivity (89%) in the 9.8 threshold cohort than 10.5 threshold (79%).

As I discussed in Chapter 2, when designing a non-invasive diagnostic test, it is important to have the intended outcome in mind. For the purpose of a primary care pathway, where the pre-test probability of advanced fibrosis is relatively low, and, the prevalence of advanced fibrosis is also lower than that of secondary care, the purpose of an initial non-invasive test should be as a screening test to 'rule out' advanced fibrosis, with a more specific test to then be used later on to 'rule in' advanced fibrosis in those with a positive initial test. Therefore, the sensitivity and NPV are more important here, and these are identical or superior at the 9.8 threshold cohort in Thiele et al.'s study, than 10.5.

Another reason for deciding on a 9.8 threshold, is because the local NAFLD pathway also uses an ELF threshold of 9.8. Not only does this make things easier for GPs to have a uniform threshold for ArLD and NAFLD, but it is also recognised that many patients with AUD have co-morbid obesity or are overweight, (two-thirds in my retrospective study, Chapter 4). As such, by using a 10.5 threshold for AUD, it may have run the risk of missing advanced fibrosis in this cohort, as if the GP followed the Fatty liver pathway with a 9.8 threshold, there would likely be a proportion of people with ELF scores between 9.8 and 10.5 that would miss out on referral to hepatology if the NAFLD pathway were followed.

The final reason for deciding to proceed with a 9.8 threshold for ELF on the alcohol pathway is that as part of the evaluation of the pathway it would be possible to perform a post-hoc analysis to determine if a threshold of 10.5 would have avoided more unnecessary referrals without missing a clinically important number of cases. However, if 10.5 was the chosen threshold for the pathway *ab initio*, then it would preclude retrospective analysis of the 9.8 threshold.

Table 5.2 Diagnostic accuracy of ELF for advanced fibrosis in ArLD by Thiele et al. (23)

	Sensitivity % (95% CI)	Specificity % (95% CI)	PPV % (95% CI)	NPV % (95% CI)	LR+ (95% CI)	LR- (95% CI)	TP FP	FN TN
Overall cohort N=289)								
ELF 10.5	79 (67-88)	91 (86-94)	71 (59-81)	94 (89-96)	8.37 (5.46-12.81)	0.23 (0.15-0.37)	53 21	14 202
ELF 9.8	89 (79-96)	78 (72-83)	54 (44-64)	96 (92-98)	3.99 (3.08-5.16)	0.14 (0.07-0.28)	59 50	7 173
Primary care cohort N=128								
ELF 10.5	75 (35-97)	97 (92-99)	60 (26-88)	98 (94-100)	22.50 (7.93-63.87)	0.26 (0.08-0.86)	6 4	2 116
ELF 9.8	75 (35-97)	89 (82-94)	32 (13-57)	98 (94-100)	6.92 (3.61-13.27)	0.28 (0.08=0.93)	6 13	2 107

5.5.1 Camden 'Abnormal Liver Function Test (LFT) pathway'

The launch of the new pathway incorporating ELF in people with excess alcohol was complicated by the fact that Camden had a pre-existing pathway on their website entitled 'abnormal LFT pathway'. This also included information about alcohol excess that needed updating to include use of the ELF test. This 'abnormal LFT' pathway actually had a much higher 'hit-rate' on the Camden website than the alcohol pathway, and it was therefore important that I updated this as well introducing the new ArLD pathway.

This pathway starts with page 1 (Figure 5.4A) for patients with abnormal LFTs, and leads onto the fatty liver pathway (Figure 5.4B), which recommends the use of non-invasive tests in people with fatty liver and abnormal LFTs, with FIB4 to be used in the first instance. We then amended this pathway to specify if fatty liver is in the presence of alcohol excess (rather than suspected NAFLD) the ELF test should be used as the initial test, rather than FIB4 (which is recommended for NAFLD, by definition in the absence of alcohol excess).

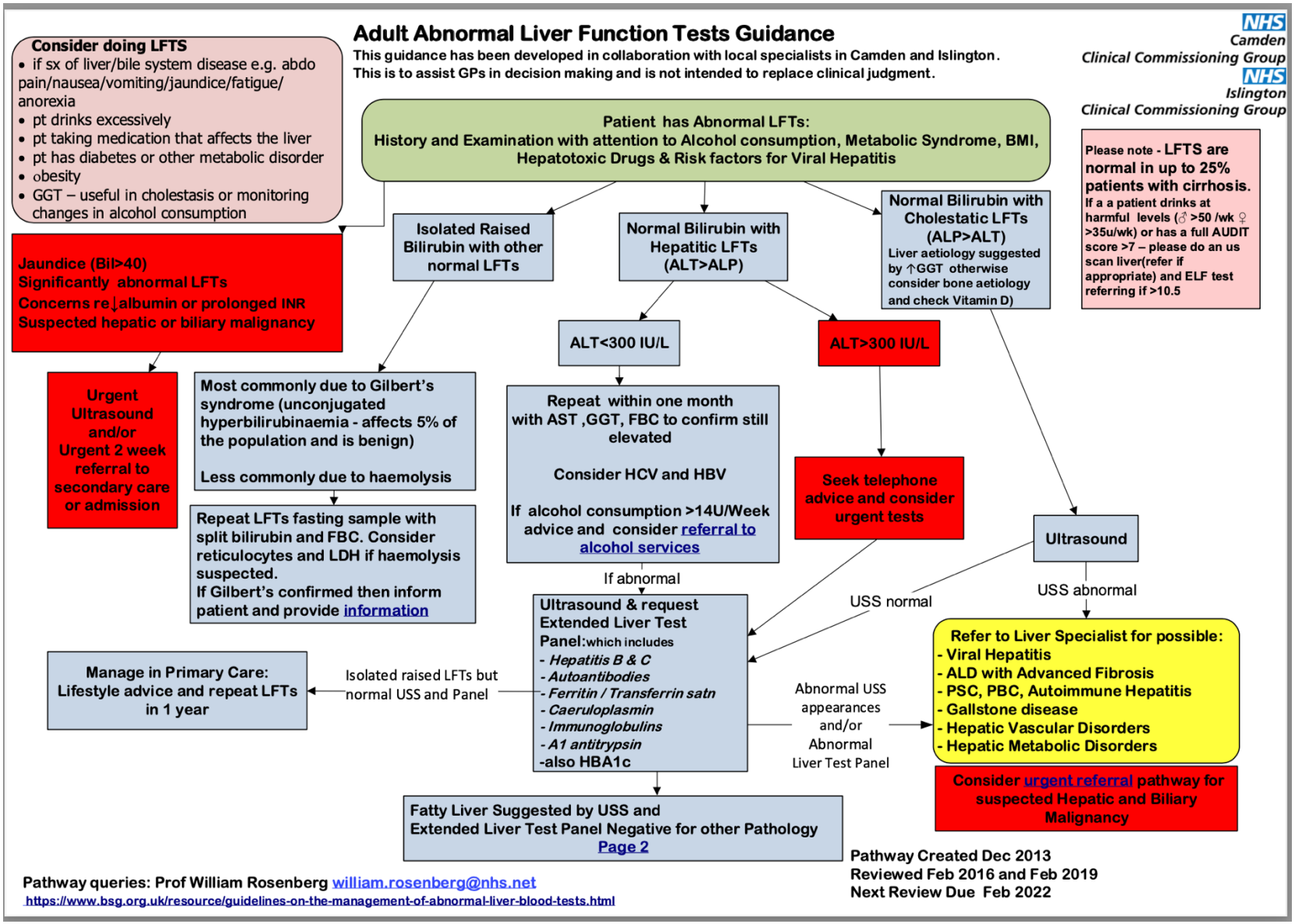


Fig 5.4A. First page of the updated 'abnormal liver function test' guidance
(Change made to the pink text box on the right-hand side, to highlight that LFTs can be normal in cirrhosis, and to consider ELF in excess alcohol.

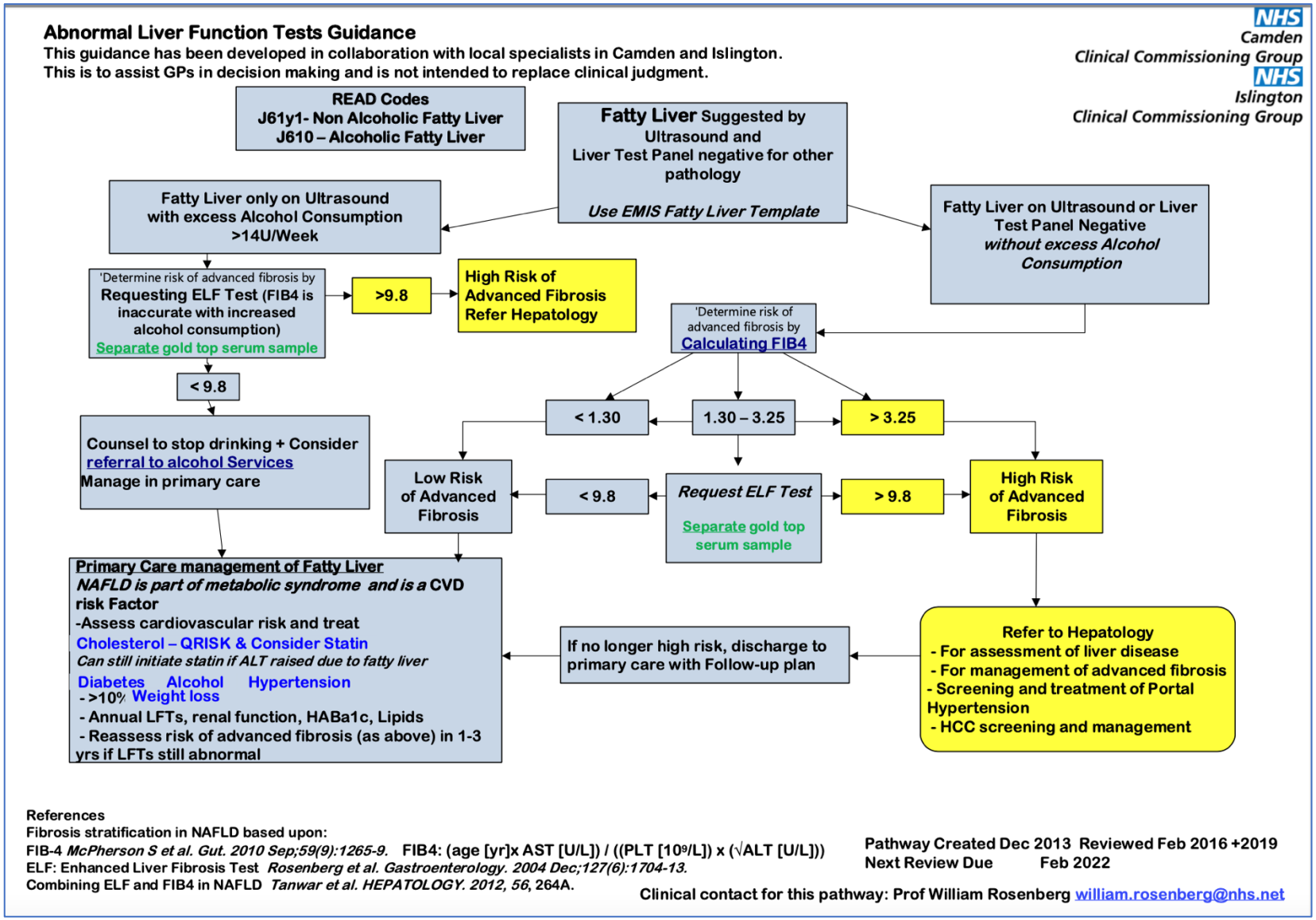


Fig 5.4B. Second page of the updated 'abnormal liver function test' guidance

This pathway flows on from the first page of the abnormal LFT guidance, so patients that have abnormal LFTs, fatty liver on imaging and drinking excess alcohol will be prompted to have an ELF test. If no alcohol excess, then the clinician should first do a FIB4.

5.5.2 Approval and launch of the pathways.

As there were two pathways that needed to be updated to reflect the use of ELF in people drinking excess alcohol, they both had to undergo separate approval processes.

The abnormal LFT pathway was taken to the Clinical Cabinet Group meeting first (Sept 2019), where we presented the proposed pathway changes in front of the commissioners. Once this had been reviewed and approved, it could then be prepared for launch on the Camden and Islington website (Islington did not require a separate approval process as they share guidelines and were happy to proceed).

Gaining approval for the alcohol pathway, however, was a lengthier process because it contained information about the prescribing of anti-craving medications, and the local addiction psychiatrist, along with the GPs, had recommended changes to their prescribing practice. This meant that that alcohol pathway then needed to be processed through the Medicines Management Committee (MMC), as well as the Clinical Cabinet Meeting, before it could be approved. This meant there was a delay between the launch of the LFT pathway, and the alcohol pathway. The LFT pathway went 'live' on the 3rd October 2020, but the alcohol pathway was not launched until the 13th January 2021.

As the 'LFT pathway website' was deemed to be the most popular, with a very high 'hit-rate' by service users, we were reassured by the GP leads that there should not be any significant loss of uptake of patients to the ELF pathway between the

October launch and the January launch, as they should be picked up on the LFT pathway.

5.5.3 Communicating the new pathway launch

The pathway changes and the new pathways were disseminated in a number of ways. Firstly, I wrote a news bulletin which was sent out to all Camden and Islington GPs, with a link to the new pathways. A similar message was sent out to hepatologists at the Royal Free.

We then had a 'launch event' where I delivered a talk about the new pathway to an invited audience of the Camden CCG GPs, followed up by a reminder email with a link to the pathway website. I followed this up by a talk at the Royal Free to the hepatology department, to advertise the new changes. The bulletin messages were then shared by the Islington CCG lead to all of the Islington GPs, and follow-up 'news flash' reminders were sent throughout the year.

5.6.1 Evaluation plan

The evaluation of the pathway was planned to compare before-and-after pathway data, with retrospective data collected for 12 months before the pathway introduction (Jan 2017 to January 2018), and then prospectively for 12 months after the pathway introduction (October 2019-September 2020). However, the COVID-19

pandemic impacted heavily as routine hepatology services were suspended from early Spring 2020 due to diversion of NHS resources into the COVID response. This significantly impacted the number of patients referred through the pathway so the evaluation will now be undertaken outside the timelines for this thesis.

However, in the following section I describe the planned method of analysis for the evaluation.

5.6.2 Evaluation populations and method

Evaluation design

The evaluation will collate data retrospectively for a 3-year period pre-pathway introduction, (Jan 2015-Jan 2018) (from Chapter 4), and then prospectively for 36 months after the pathway introduction (October 2019-September 2022).

Setting/population

The pathway implementation practices (35 in Camden and 34 in Islington) will be encouraged to use an ELF test to assess patients with AUD to aid decision making about whether to refer to specialist liver services in secondary-care. Patients with an ELF score $\geq$9.8, indicating advanced fibrosis in line with national guidance, will be referred (1). GPs opting not to use the C&I alcohol pathway will refer using standard-care.

5.6.3 Inclusion/exclusion criteria

Included patients will be aged ≥18, with hazardous or harmful alcohol use referred to the Royal Free London NHS foundation trust (RFL). Hazardous and harmful drinking will be defined as per the NICE guidance on AUD (CG115) (59), and not limited to people with overt evidence of liver disease.

Patients will be excluded from the evaluation if they are already under the care of a hepatologist, already known to have chronic liver disease or viral hepatitis, or are pregnant.

5.6.4 Primary Outcome measure

The primary outcome measure is the reduction in unnecessary referrals from primary to secondary-care attributable to the use of the C&I alcohol pathway compared with standard-care during the evaluation period.

Secondary Outcomes

Secondary outcomes include:

- The proportion of referrals with advanced fibrosis or cirrhosis in patients referred to secondary-care using the C&I alcohol pathway, compared with standard-care.

- The proportion of referrals with advanced fibrosis or cirrhosis in patients referred to secondary-care using the C&I alcohol pathway, compared with standard-care within C&I.

- A ‘before-and-after’ evaluation, comparing the proportion of unnecessary referrals and the number of referrals with advanced fibrosis or cirrhosis 2-years before and after the introduction of the C&I alcohol pathway.

- The cost-effectiveness of the C&I alcohol pathway will be compared to standard-care by determining the cost per case of cirrhosis detected. The costs associated with the C&I alcohol pathway will be compared to those of standard care by comparing the cost of (testing + referral and follow-up appointments + hospital investigations) for the C&I alcohol pathway versus (referral and follow-up appointments + hospital investigations) for standard care.

- A *post-hoc* comparison of the use of 9.8 ELF threshold versus 10.5, in terms of proportion of unnecessary referrals and detection of advanced fibrosis.

- A *post-hoc* comparison of the use of a two-step pathway using FIB4 for those patients who had been referred on the C&I alcohol pathway, in terms of reduction in need for ELF test, impact on number of unnecessary referrals and detection of advanced fibrosis, and cost-effectiveness.

- An evaluation of risk factors associated with diagnosis of advanced fibrosis in secondary-care, including documented current alcohol intake in U/w and duration of AUD, BMI, sex, and deprivation rank.

5.6.5 Definitions:

- *'Unnecessary'* will be defined as referral to secondary-care when there is no evidence of advanced fibrosis or cirrhosis and the patient could have been safely treated in primary-care.

- *'Standard care'* is defined as the GP's usual practice in referring patients with AUD to hepatology services, and referral decisions may be guided by the amount of alcohol consumed, LFTs, or ultrasound imaging of the liver. Whilst ELF/FibroScan is recommended by BSG guidelines, this is not in routine use by GPs for alcohol referrals (as found in Chapter 4).

- (*Advanced fibrosis* or *cirrhosis* will be decided using composite clinical judgment in secondary-care by a specialist hepatologist, and may include the use of liver biopsy, FibroScan, further non-invasive tests, and imaging to make the decision, depending on individual clinical circumstance.)

5.6.6 Sample size

I calculated the sample size powering for a 30 % expected decrease in unnecessary referrals to secondary care hepatology clinics in patients referred on the pathway compared to before the pathway (the primary outcome measure). This 30% was based on the fact that 64% of the patients referred to hepatology with suspected ArLD were deemed 'unnecessary referrals' in my retrospective analysis in the same patent population in Chapter 4, and a 22% reported false positive rate of ELF in the Thiele et al. study (23). A reduction from 60% to 20% decrease by adopting a

pathway using ELF would mean a 40% decrease in the percentage of unnecessary referrals. However, I decided to opt for a more conservative 30% expected decrease in unnecessary referrals. Applying 80% power, 95% confidence intervals, and percentage decrease from 64% to 34%, I calculated the sample size to be a minimum of 40 patients in each group (40 referred 'on the pathway' using ELF, and 40 pre-pathway) (304, 305). This power calculation was reviewed by a statistician who agreed with my calculation, and that this would be a sufficient size to detect a difference between groups for the primary outcome.

I estimated that participants would be referred on the pathway at a maximum rate of 15 patients per year per practice, based upon the number of patients seen at GP centres per week with AUD, and assuming a 50% uptake of the pathway (as per the NAFLD pathway in the same GP practices which saw a 48% uptake (168)) and allowing for 15% drop out rate. The overall recruitment was planned (pre COVID) for 12 months. Based on my study in Chapter 4, there were 231 alcohol referrals to the Royal Free in the 3 years from 2015-2018, which equated to 77 per year, so a target of 80 patients was expected to be achievable before the SARS-CoV-2 pandemic impacted on patient referrals through routine pathways.

5.6.7 Data collection

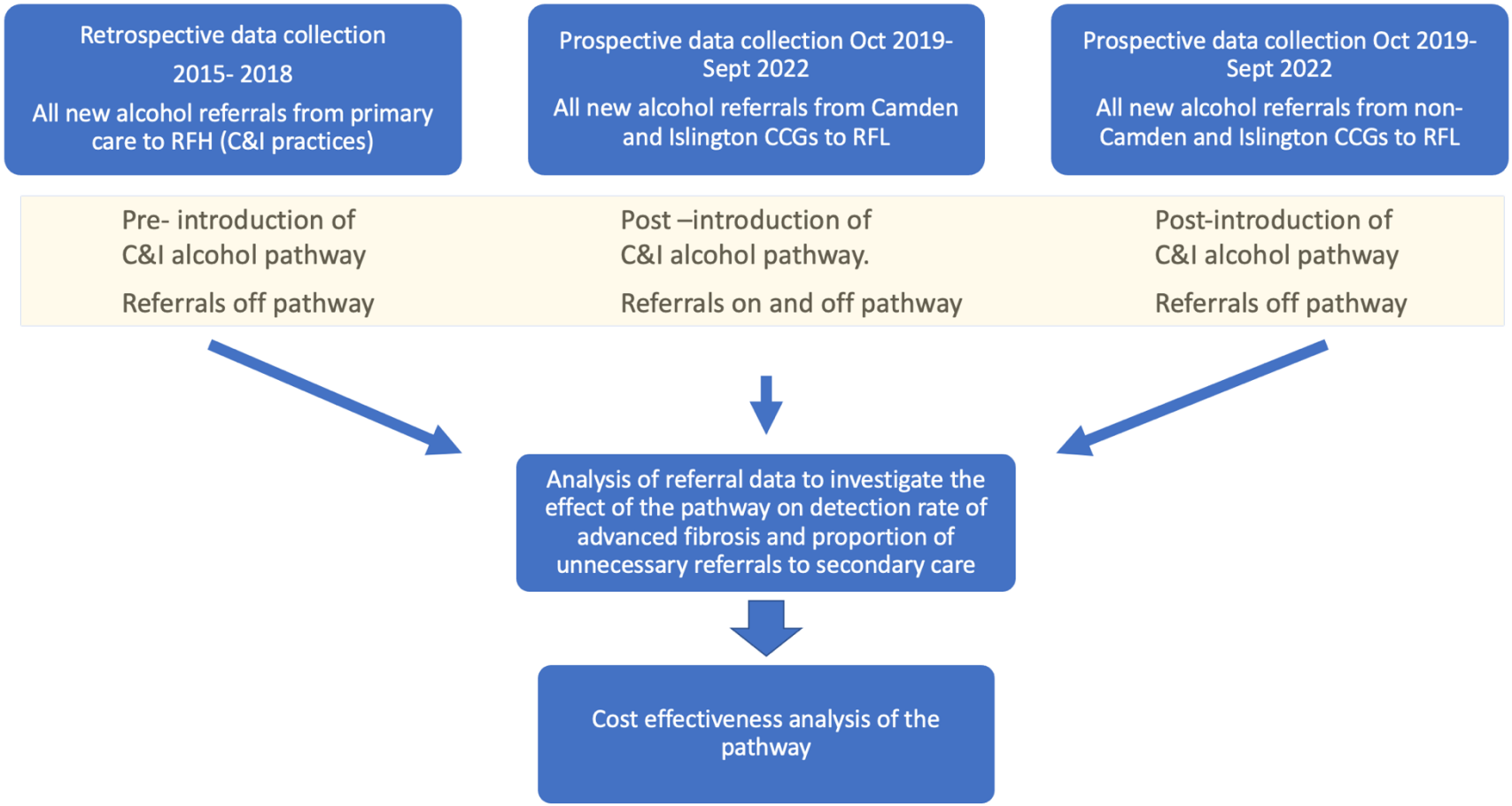
Referral data comparisons are planned to be pre-and post-pathway, (Retro-to-prospective) on-and-off pathway (prospective) and with comparison of the same with Camden and Islington practices versus other referring boroughs (Figure 5.5).

For both retrospective (Jan 2015-Jan 2018), and prospective (from start of pathway October 2019 to September 2022) data collection, all patients referred with suspected ArLD to hepatology clinics at Royal Free Hospital would be identified from electronic clinical records. This would enable identification of whether the patients are referred on the 'C&I alcohol pathway' or Standard of Care, and whether the referring practice is in or outside of the pathway referral practices (Camden and Islington).

Data would include: name of referring CCG, use of C&I alcohol pathway, results of investigations (including blood tests, FibroScan, imaging, liver biopsies, FIB4, APRI, AST:ALT ratio), number of appointments before discharge back to primary-care, diagnosis of advanced fibrosis or cirrhosis in secondary-care, socio-demographics including age, sex, ethnicity postcode for deprivation rank score, BMI and alcohol intake (including documented current intake in units per week, duration of alcohol excess, and if recent alcohol consumption in the 3 weeks before referral) and if any co-morbid hypertension or diabetes.

The anticipated methods of data synthesis are summarised in Figure 5.5.

Figure 5.5. Pathway evaluation data sources



5.6.8 Statistical analysis

Once evaluation data are collected, the proposed method of analysis by the Applied Research Collaboration (ARC) study statistician involved in this pathway is regression analysis and interrupted-time-series analysis on the collated dataset to determine the number of primary-care referrals with Advanced fibrosis using C&I alcohol pathway or standard-care. Using best-fit models, the Incidence Rate Ratio (and 95% CI) will be projected for the number of referrals using the C&I alcohol pathway or standard-care as well as the statistical significance of the difference between C&I alcohol pathway and standard-care. A difference-in-difference analysis will also be used to contrast the number of referrals between C&I alcohol pathway and standard care.

Economic evaluation

A comprehensive cost analysis plan will be developed with help from the NIHR Applied Research Collaboration health economist and statistician. It will involve an economic analysis to assess the impact of the C&I alcohol pathway in reducing the number of referrals and determining if there is cost-saving for NHS.

5.7.1 Impact of the SARS-CoV-2 Pandemic on data collection

I had initially planned to acquire some preliminary data to allow analysis of the impact of the pathway from 1-year post-launch (September 2020). However, the global SARS-CoV-2 pandemic arrived in the UK in the months following the launch of the pathway, which greatly impacted on the number of patients being referred and seen in hepatology clinics (which were closed for significant periods of time in the year following the launch). Despite extending the planned 1-year evaluation by 3 months, there were still not enough data for any meaningful analysis within this time frame.

By the time of the planned 12-month data analysis in September 2020, the SARS-CoV-2 pandemic had been ongoing for 9 months, and the UK was at the start of a 'second wave' of COVID-19 hospital admissions. Routine operations, procedures and outpatients' clinics were affected by the diversion of NHS resources into managing large numbers of COVID-19 cases in both Wave 1 (starting March 2020 and Wave 2 (starting August 2020), including redeployment of staff into covid wards areas, the ED and intensive care. Many clinicians (including myself) were moved to an 'emergency rota' covering additional clinical shifts to allow for the anticipated rota gaps due to staff becoming sick or having to self-isolate. Routine clinics were cancelled as outpatient departments and day units were converted into overflow intensive care units at the Royal Free. Routine hepatology clinics and investigation services were cancelled, and many patients had appointments delayed by several months. For example, the FibroScan service was closed for several months during the first 'wave' of COVID infections in Spring 2020, and again

in the Autumn and winter 2020-21, not re-opening until April 2021. Routine ultrasound appointments were also cancelled or delayed.

This meant significantly fewer referrals were captured in my 12-month analysis than expected in a 'normal' (pre-pandemic) year, and of the referrals I did capture, most were still either waiting to be seen in hepatology clinic, or had been seen once and were awaiting investigations such as FibroScan or ultrasound or other imaging. I continued data-collection beyond the 12 months, with a further census point 3 months later at the end of December 2020 (to include data from 3rd October 2019 to 30th November 2020). These extra months did not give me sufficiently more data, and I still had too few patients who had been seen in clinic and had investigations, to be able to determine if they had advanced fibrosis or not, and if they could be deemed 'unnecessary' referrals as per the primary outcome measure.

At my final census point at the end of Dec 2020, having scrutinised 1,609 referrals to hepatology clinics over 14 months, only 34 of these were referrals that met the inclusion criteria of having suspected ArLD. This was under half the number I would expect in a normal year, judging by the 77 alcohol referrals per year in my pre-covid evaluation (Chapter 4). Of the 34 who were referred with suspected ArLD, 12 had been referred using the new alcohol pathway with an ELF test, and 22 were referred without use of the pathway. This is roughly the same uptake of the pathway as the 48% uptake seen in the NAFLD pathway in Camden in 2016. Despite many of the 34 patients having been referred many months ago to the clinic, only 19 out of the 34 had been seen in hepatology clinic at this point (Dec 2020), and of these 10 had undergone the investigations that had been requested

at the clinic for assessment of liver fibrosis. The other 9 were still awaiting appointments for investigations. Having only 10 patients with available data, this was deemed not sufficient for any meaningful analysis. The FibroScan service was not due to be opened up again until April 2021, and a significant number of patients referred with suspected ArLD were having FibroScans requested by their clinicians for fibrosis assessment, as per the current NICE guidelines. In order to accrue the required ≥80 patients (40 off pathway, 40 on pathway), it was clear this would require a period of time beyond the end of the thesis. The decision was taken with my supervisors and thesis committee to write this chapter up as a pathway development and implementation chapter, with the evaluation to take place at a later date once sufficient number of patients had progressed through the pathway.

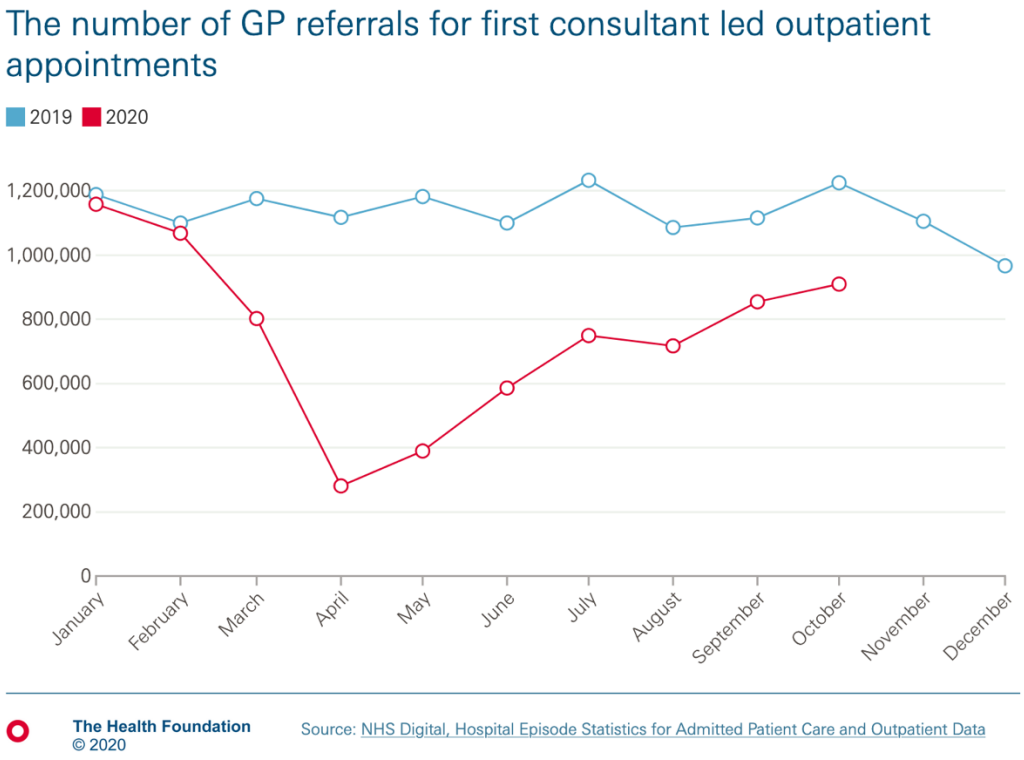
5.7.2 Additional reasons for low numbers of referrals

In addition to the closures of routine clinics, and postponements of clinic and investigation appointments in 2020, there were other factors due to the pandemic that are likely to have affected this small number of new patients referred to hepatology during this time period. There were fewer patients being referred to specialist clinics by their GPs during the pandemic, with NHS England data reporting that between January and September 2020, there were “4 million fewer referrals to outpatients than in the same period in 2019”(306) Figure 5.6). There was a 60% fall in the numbers of patients referred by their GP on a suspected cancer pathway in April 2020 compared with April 2019 in the UK (307), and whilst the same statistics are not reported for hepatology clinics in the UK in particular, my data from one hospital suggests that hepatology clinics were significantly affected by the

pandemic. It was also well reported during this time that people were avoiding seeing their GP for health problems, for fear of catching COVID, or for not wanting to 'bother health providers' when the NHS was under strain (307). As a consequence, there has been an observed increase in patients presenting with heart failure after 'missed' myocardial infarctions when they did not seek medical help (308). NHS England polling revealed that 4 in 10 people are not seeking help from their GP for health problems because they are afraid to be a burden on the NHS during the pandemic (307). This may be have contributed to the reason for the low number of new referrals for suspected ArLD during this year. The NHS subsequently launched a public information campaign to encourage people to seek help for non-COVID health conditions, but it is estimated that these knock-on effects to public health could take years to resolve (307).

Figure 5.6. Impact of SARS-CoV-2 pandemic on outpatient referral numbers

(Source: NHS Digital, licenced under the current version of the Open Government Licence)



5.7.3 Challenges of the pathway set-up and data collection

Developing a new pathway, whilst rewarding, has presented me with new challenges, even above those associated with the global pandemic described above.

These have included difficulties in making changes to the ELF threshold at the central laboratory, poor coding of alcohol referrals, and navigating the data protection laws to plan the future evaluation of the pathway. I will describe these now in more detail.

5.7.4 Coding

There is currently no system in place to filter GP referrals by 'alcohol-related liver disease' or 'alcohol referrals' to the Royal Free, due to lack of accurate coding in the referral system. Therefore, in order not to miss any relevant referrals, I reviewed every referral to any hepatology clinic (including NAFLD, autoimmune, HCC, transplant, viral hepatitis, along with general hepatology and alcohol nurse clinics) during the preliminary evaluation period (Sept 2019 to Nov 2020). This was time-consuming and meant reading many GP referral letters (1,609) to find only a small number of relevant referrals (n=34). This was similar to my experience for Chapter 4, when I looked through nearly 3,000 referrals of which only 231 were relevant to my study. However, this was the only way to make sure I did not miss any alcohol referrals that had been seen in NAFLD or viral hepatitis clinic etc. The SARS-CoV-2 pandemic enhanced these difficulties, because the way in which hepatology referrals were processed changed so that the hepatology clinicians could 'triage' GP referrals and, where appropriate, switch patients to telephone clinics (to minimise the risk to the patient of catching COVID by coming to the hospital), or deal with the referral by giving advice through correspondence without need for a clinic appointment, or to adjust the waiting time, delaying appointments that were not essential. Instead of being able to audit the referrals in one IT system called 'Cerner', I now needed to look in three systems, 'Cerner', 'ERS' and 'LUNA', which added to the time it took to sift through referrals.

5.7.5 ELF threshold laboratory changes

Prior to the initiation of the alcohol pathway, it had been agreed locally that the ELF threshold for NAFLD should be changed from 9.5 to 9.8. This process of raising the threshold had not yet taken effect. However, with the initiation of the alcohol pathway, which advises referral with an ELF result above the threshold of 9.8, it was vital to get the threshold changed at the laboratory. This is because the results are sent out to GPs with additional information advising referral to hepatology above the specified cut off. This could be very confusing if GPs were requested as part of the pathway to use a 9.8 threshold, but were receiving back results from the laboratory instructing them to refer to hepatology at a threshold of 9.5. I got in contact with the central laboratory that processes the ELF test (the 'HALO' building in central London, HSL (Health Service Laboratories). After confirming that they were indeed changing the ELF threshold to 9.8 for advanced fibrosis, we confirmed with them the appropriate information to include with the ELF results, equating 9.8 to advanced fibrosis, and recommending a hepatology referral. However, after initially being reassured it was all in hand, the changes did not happen. It then became a frustrating wait, requiring me to make regular phone calls to the HSL manager, with additional follow-up emails, where I was reassured that the changes would be taking effect, but then weeks passed, and we noticed the threshold was still set to 9.5. After a visit to the central HSL laboratory in August 2019 to meet with the head of the laboratories, to find out what was causing the delay, and see if there was anything we could do to help, we were again reassured the changes would be made. However, it took several further phone calls, emails and escalating

it as a patient safety issue before the threshold was finally changed to 9.8 in December 2019 but this was more than 6 months after the initial request. Fortunately, it was changed just prior to the launch of the new alcohol pathway.

5.7.6 Challenges with navigating data protection procedures

Whilst the primary outcome measure is a reduction in the proportion of unnecessary alcohol referrals to the Royal Free hepatology clinics and improvement in the detection of cases of advanced fibrosis and cirrhosis through use of the ELF pathway compared to not using the pathway, it would also be useful to know how many patients who had an ELF test on pathway in primary care were not referred. This would give us an idea of the impact of the pathway on the 'number of avoided referrals' to secondary care. The simpler way of accessing this data would be to ask the HSL laboratory to send us all the ELF test results and associated NHS number, and then we would be able to audit how many of the patients with a low (<9.8) ELF score remained in primary care. After submitting the necessary data-transfer information requests in July 2019, and being told by HSL it would be possible for them to send us monthly ELF results – in a similar vein to the efforts described above that were required to change the ELF threshold – multiple emails and phone calls over the best part of a year, with no result. I then began enquiries as to the possibility to access these data from the primary care end, rather than from the HSL laboratory. Ideally, I would be able to get access from primary care on the proportion of patients with suspected ArLD in whom the new pathway was

followed and they had an ELF test, compared to those who did not follow the pathway. Collecting patient data from within the Royal Free was straight forward because I have an NHS contract there, and the pathway was registered as a service evaluation, allowing the collection of anonymised data. However, in primary care, a different process would need to occur to get the appropriate permissions to access the data.

After enquiring through the Royal Free London NHS Foundation Trust Caldicott Guardian, I acquired the contact details of the relevant data protection officer and began the process of obtaining the relevant permissions to access the data. I completed multiple pages of data access request forms, along with a data-flow diagram (Figure 5.7). The forms included a DPIA (Data Protection Impact Assessment), which, along with the data-flow diagram, were used by the information governance team to decide if the request for data was appropriate and suitable. This decision is made by an IG (Information Governance) working group, including the responsible data protection officer – and was a lengthy process (4 months) before it was approved by the IG working group.

I then discovered that the 'data controllers' were the individual GP practices within Camden and Islington (69 practices) and so I would need to obtain separate permissions from each individual practice to access the data. The approved DPIA was uploaded by the IG working group to a central database accessible by the GP practices, and so for future analysis, it should now be relatively straight forward to get permissions from each practice at a time to access the data.

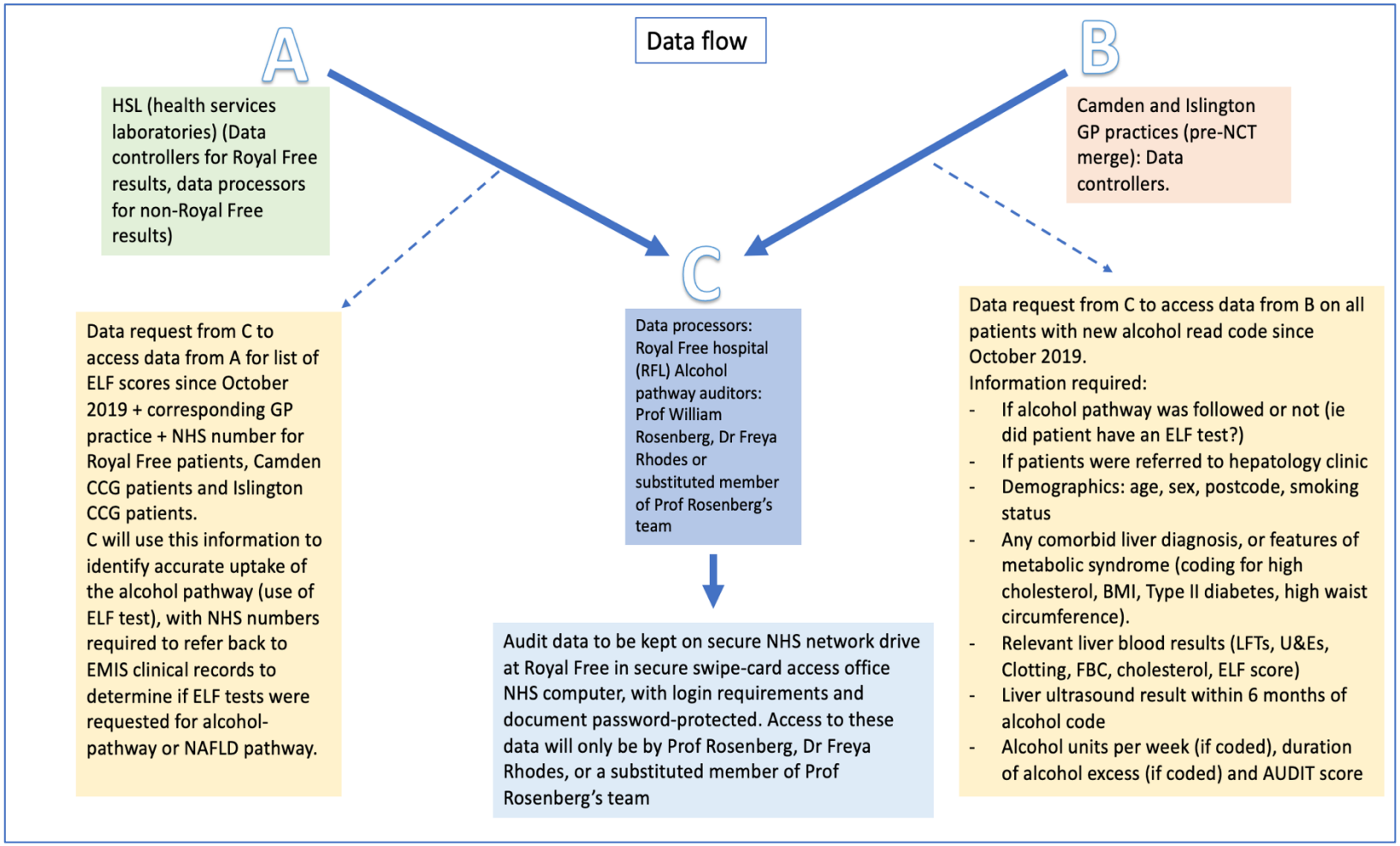


Fig. 5.7 Data flow diagram: illustrating the type of data requested to be accessed, and the expected direction of travel of the patient data

5.7.7 Camden and Islington 'merge'

To add to the challenges faced in setting up the pathway and planning its evaluation, I was informed after the launch of the pathway, that the Camden and Islington CCGs would be 'merging' with several other London CCGs to become 'North Central London' CCG in the spring of 2020 – half way through the planned pathway evaluation period. This process involved 5 separate CCGs (Camden, Islington, Barnet, Enfield, and Haringey) merging to form one 'mega CCG' called 'North Central London' (NCL) CCG. The 'merge' was activated on the 1st April 2020, and 'Camden and Islington' CCGs became 'NCL CCG'.

This then meant that there was a need for uniform guidelines for the whole NCL CCG, and the alcohol pathway that I had launched in Camden and Islington was suddenly under competition from other CCGs that were now within NCL, that had their own pathways that they wished to use.

Conversations occurred between hepatology department leads in the catchment area for the new NCL CCG, at the various hospital Trust sites, and it was decided that our Camden 'fatty liver pathway' (Figure 5.4B) would be scrapped in favour of an alternative NCL pathway (Figure 5.8). This alternative pathway advises that in the case of fatty liver on imaging, in patients who drink over 40 units of alcohol or more a week (men or women), that they should be referred to the alcohol pathway (Figure 5.9). This alcohol pathway was different to our Camden alcohol pathway, and advocates referral to hepatology in people drinking excess alcohol (over 40 units per week) if they have evidence of "organ damage/chronic liver disease or

enlarged liver/spleen e.g., alcoholic hepatitis or cirrhosis", but does not mention at all the use of non-invasive fibrosis tests, or how the GP would define 'evidence of chronic liver disease'.

Fortunately, it was then negotiated that Camden and Islington practices (within the new NCL CCG) could continue to use our alcohol pathway launched in October 2019 (Figure 5.3), and all other practices within the NCL CCG would follow this alternative alcohol pathway (Figure 5.9). All practices, however, would now follow the new fatty liver pathway (Figure 5.8).

It is unclear if this NCL merge may have had an impact on the number of ArLD referrals that had an ELF test (our alcohol pathway) prior to referral, or if it may have led to confusion, with some GPs following the new NCL guidelines.

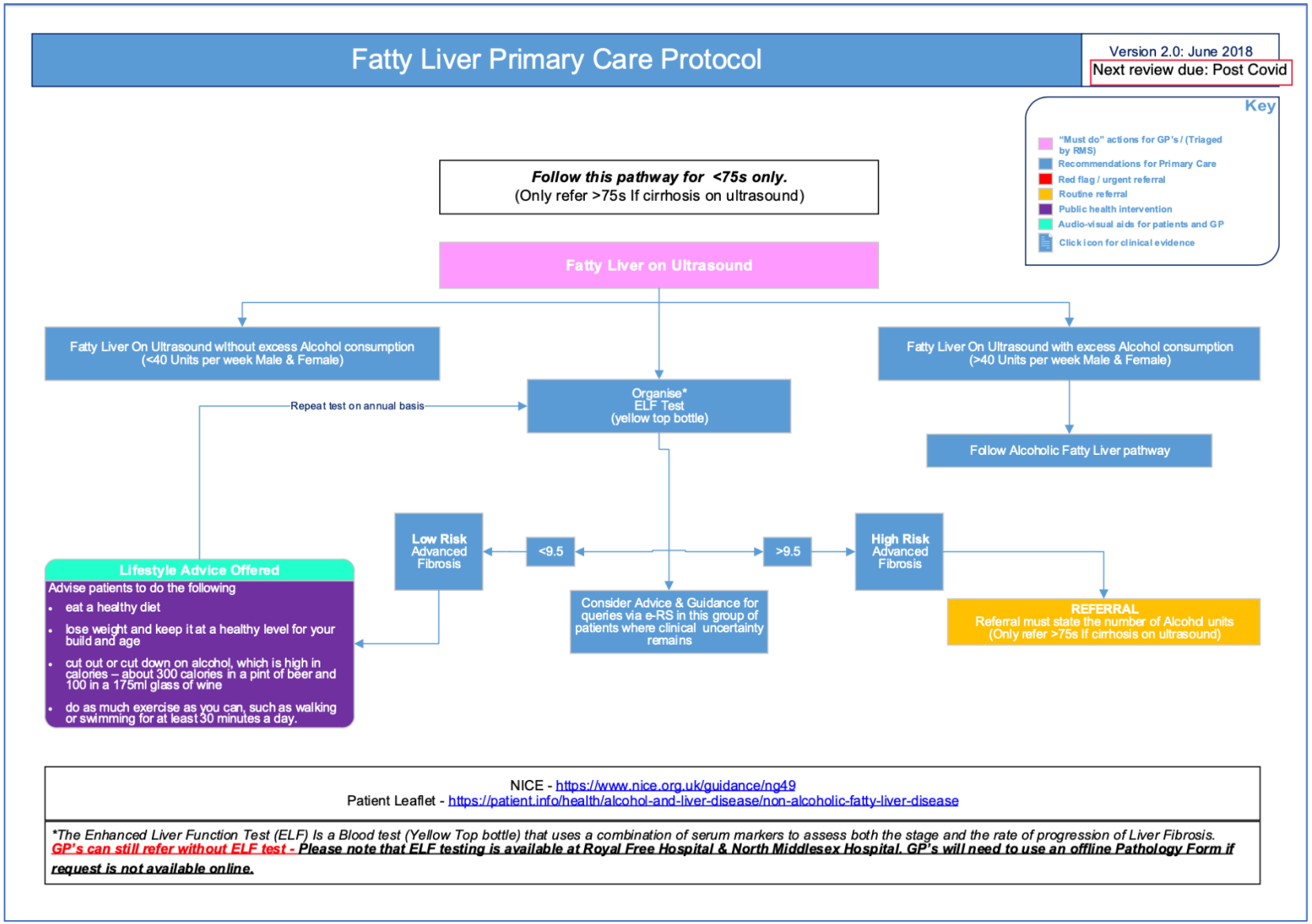


Figure 5.8: New NCL CCG fatty liver pathway

This was in effect for all NCL practices (including Camden and Islington) from 1st April 2020. Drinking > 40 units per week prompts pathway user to follow the alcoholic fatty liver pathway (Figure 5.9)

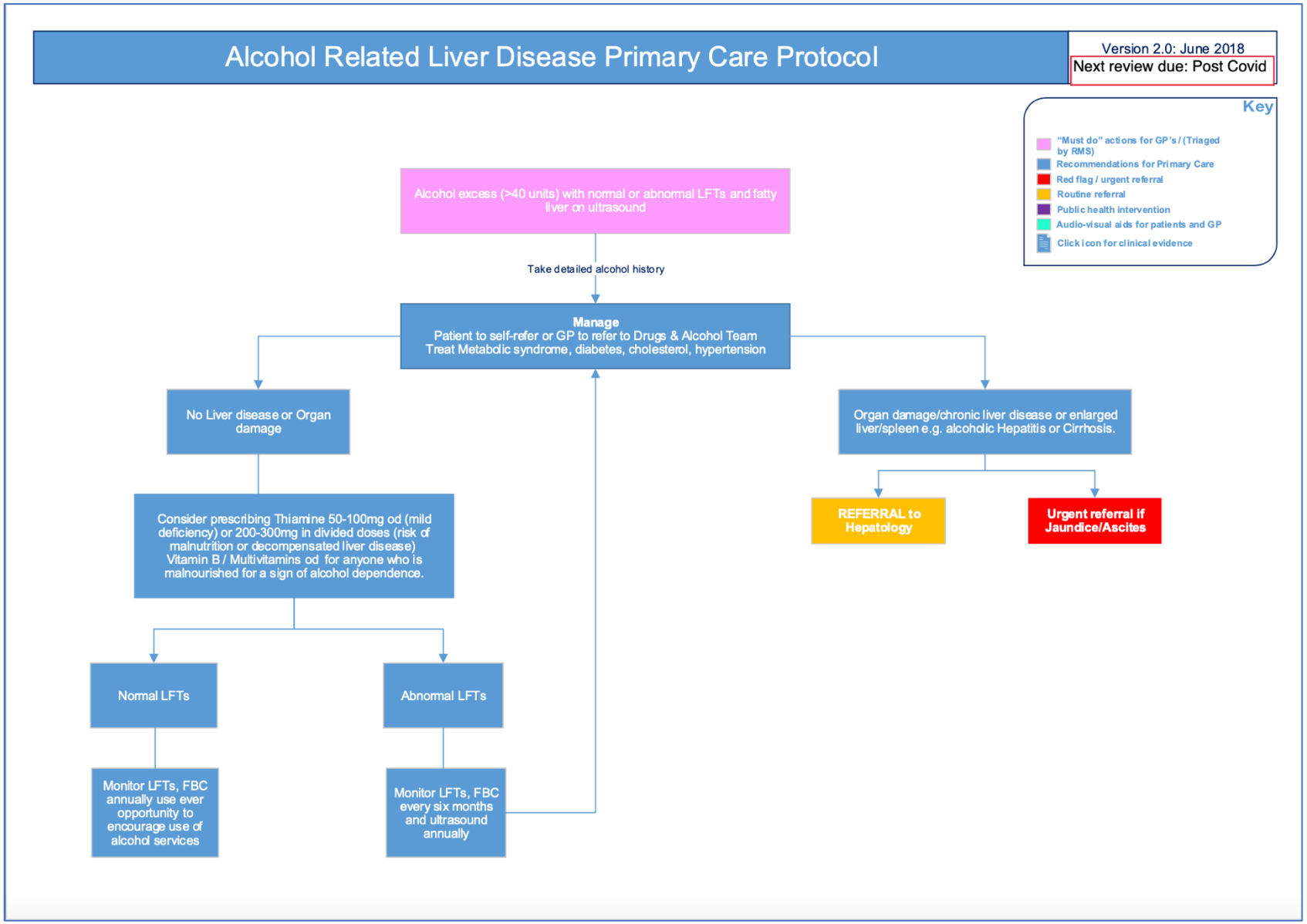


Figure 5.9: New NCL CCG ArLD pathway

This was in effect for all NCL practices from 1st April, but allowing original Camden and Islington CCG practices to continue to follow our alcohol pathway (Figure 5.3)

5.8 Lessons learned

Developing a new pathway has been a useful learning experience, which has provided me with a variety of new skills and knowledge of negotiating change across multiple NHS organisations, in particular around data governance processes.

The global SARS-CoV-2 pandemic, the CCG merger in North London, the unanticipated delays and issues with HSL including difficulties in trying to change the ELF thresholds, the challenges in accessing data from HSL and primary care, and the fact the pathway had to be launched in two stages with the fatty liver/LFT pathway and the alcohol pathway, all added to the overall challenge of pathway implementation and evaluation during the time period of my thesis.

As described above, the impact of the pandemic resulted in clinic closures and clinic delays, with less than half the expected number of alcohol referrals to hepatology services, and delays to patients accessing routine liver investigations after being seen in clinic, all of which impacted on the availability of patient data and therefore my ability to evaluate the pathway within the time constraints of my PhD.

However, on a positive slant, I successfully introduced a new pathway, incorporating best evidence and national guidelines on the use of non-invasive testing for liver fibrosis in people who drink excess alcohol.

This has potential for improving NHS care to the patient, through (i) potentially reducing unnecessary referrals to hospital; (ii) reducing the associated anxiety-provoking investigations; (iii) increasing the likelihood of detecting liver fibrosis

early in people at risk, and (iv) potentially allowing intervention to reduce alcohol intake and halt fibrosis progression, or allow for fibrosis regression. This would also mean that appropriate patients can be selected for referral to hepatology clinics, and can begin surveillance for portal hypertension and HCC, which can potentially alter their long-term outcomes (15, 138) This pathway also has potential for impact on the NHS, through reduction in the unnecessary use of secondary care resources, improved detection of liver disease at a point at which intervention can avoid harm; and cost savings should accrue through the reduction in referrals and investigations and cost-utility should arise through the reduction in harms from advanced liver disease.

My next chapter moves on to investigate the use of non-invasive fibrosis tests in people admitted to hospital with AUD.

CHAPTER 6

Uncovering unsuspected advanced liver fibrosis in patients referred to the alcohol nurse specialist using the ELF test

6.1 Abstract

Background: Alcohol Use Disorders (AUD) causes 7.2% of UK hospital admissions per year. Many of these patients who are presenting to hospital with manifestations of their alcohol use are not managed by hepatologists, but instead by general physicians/Emergency medicine doctors or surgical specialties, and is it therefore likely that liver disease may not be investigated for and opportunities for diagnosis of liver fibrosis and cirrhosis missed. Having reviewed the diagnostic performance of the ELF test in Chapters 2 to 4, and following current national guidelines (1), I aimed to use the ELF test to investigate prevalence and associations of occult advanced liver fibrosis in AUD patients not known to have liver fibrosis in a cohort of patients referred to the Royal Free London NHS foundation Trust (RFL) Alcohol service.

Methods: I used ELF as a marker of liver fibrosis in prospective sequential inpatients referred to the RFL Alcohol Specialist Nurse from November 2018 to December 2019. Known cases of liver disease (including ArLD) were excluded. I recorded data on patient demographics, blood tests, imaging data and alcohol histories. Advanced fibrosis was categorised as ELF $\geq$ 10.5 as per best evidence (23).

Results: 99 eligible patients were included (69% male, mean age 53.1 $\pm$ 14.4) with median alcohol intake 140 units/week (IQR 80.9-280), and a mean duration of harmful drinking of 15 years (IQR 10-27.5). The commonest reason for acute admission was symptomatic alcohol withdrawal (36%). The median ELF score was 9.62, range 6.87-13.78. An ELF score $\geq$10.5 was recorded in 28/99 (29%) patients, of whom 28.6% had normal liver tests (abnormal liver tests were defined as raised transaminases or ALP + GGT, not including isolated hyperbilirubinaemia (Gilbert's)).

Within the previous 5-years, 76% had attended A&E at least once without assessment of liver disease. The ELF score was not associated with recent alcohol intake ($r = -0.179$, $p = 0.081$), or inflammation as assessed by AST and ALT ($p = 0.574$).

Conclusion: Over a quarter of patients with AUD had previously undetected advanced liver fibrosis assessed by ELF testing. ELF was not associated with liver inflammation or recent alcohol intake. The majority had recent missed opportunities for investigating liver disease. This highlights the need to standardise the use of non-invasive tests by clinicians in to assess for liver fibrosis in all patients admitted to hospital with AUD.

Notice of publication

I have published a manuscript based on the data contained in this chapter in 'BMC Gastroenterology' in 2021 (Springer Nature Publishers):

"Rhodes F, Cococcia S, Panovska-Griffiths J, Tanwar S, Westbrook RH, Rodger A, et al. Uncovering unsuspected advanced liver fibrosis in patients referred to alcohol nurse specialists using the ELF test. BMC Gastroenterol. 2021;21(1):143." (274)

6.2 Introduction

As I highlighted in Chapter 2, one in five people in the UK drink alcohol at hazardous or harmful levels (309). While alcohol causes a wide array of health and social harms, the greatest morbidity and mortality are associated with alcohol-related liver disease (ArLD) with mortality rates increasing 400% since 1970 (12).

Hospital admissions related to alcohol are rising annually, with 350,000 alcohol related admissions per year in 2019, (an increase of 20% in a decade) (310) and with a cost to the NHS of £3.5 billion per year (4). This is likely to be due to a shift in drinking behaviours from low-strength beer in pubs to home consumption of higher strength beer, wine and spirits sold in supermarkets, as discussed in Chapter 1 (5). In addition, alcohol is now 64% more affordable than it used to be 30 years ago (309).

While a proportion of people admitted to hospital with harms arising from their drinking behaviour are recognised to have liver fibrosis and are managed by liver specialists, many are managed by a wide range of doctors and their liver disease may be missed, even if their alcohol use disorder (AUD) is recognised.

Moreover, it is estimated that up to 75% of people with Chronic Liver Disease (CLD) first present to healthcare when their liver disease is advanced often with decompensated cirrhosis, when it is too late for behaviour change or interventions to avert poor outcomes (140, 281, 309).

This work leads on from the evidence I explored in Chapter 2 that cirrhosis is often asymptomatic, and is not reliably detected using routine liver function blood tests (LFTs) or ultrasound (1). The last two decades have witnessed the development and validation of a number of non-invasive tests for liver fibrosis that are increasingly used in clinical practice. These tests create the possibility to quickly and easily detect advanced liver fibrosis in at-risk patients with AUD, in order to refer appropriately to hepatology services for regular follow up to avert, or detect and treat complications of portal hypertension including oesophageal varices (15, 16) and ascites (311), and screening for liver cancers (17, 18), as well as allowing timely assessment for transplantation where applicable, and enabling alcohol counselling from hospital or community services as appropriate.

Currently non-invasive testing for liver fibrosis with the Enhanced Liver Fibrosis test (ELF) is widely used for Non-Alcoholic-Fatty Liver Disease (NAFLD) to determine which patients with fatty liver have advanced fibrosis and warrant referral to hepatology, versus those with low ELF scores who can remain in primary care (312). However, although evidence-based (23) and recommended by BSG guidance, (1) I demonstrated in chapter 4 that this approach is not yet in widespread use in the NHS for people with AUD identified in the community. I wanted to know if there were further missed opportunities in secondary care to test for liver fibrosis in people with AUD.

In this chapter I aimed to investigate the prevalence of advanced liver fibrosis using the ELF test employing the literature-based ELF cut-off of 10.5 (23), in patients recognised to have AUD in secondary care but not recognised as having any liver

fibrosis. In addition, I aimed to examine the relationship between demographic factors (including age, Body Mass Index (BMI), deprivation score), baseline LFTs, and the ELF score.

6.3.1 Methods

6.3.2 Study Design and ethics

This was a prospective service evaluation conducted at the Royal Free Hospital from November 2018 to December 2019. It was reviewed by the Royal Free London NHS Foundation Trust Research and Development Office and deemed to be a service evaluation of an established pathway of care, using an established and CE marked diagnostic test (the ELF test). Using the Health Research Authority Tool (http://www.hra-decisiontools.org.uk/research/question1.html) the study did not meet the criteria that would require it to undergo external ethical review but rather to be registered as a service evaluation conducted within the NHS Trust. It was registered with the Royal Free London Audit and Service Evaluation Department on the 10th October 2018.

Based on this review of the study the Royal Free London NHS Foundation Trust Research and Development Office deemed that there was no necessity to obtain informed consent from participants in this service evaluation. Although not deemed to be “research”, the study was in compliance with the regulatory requirements for service evaluations and was conducted in accordance with the principles of the WMA Declaration of Helsinki.

6.3.3 Patient population:

I included consecutive referrals to the RFL alcohol specialist nurse (ASN) if they were aged ≥18, and excluded if the patients were already under the care of a hepatologist, if they had a known chronic liver condition, a diagnosis of acute alcoholic hepatitis, or acute liver injury secondary to a cause other than alcohol. Out of 100 consecutive referrals to the ASN 98 were in-patients or emergency department attendees. The vast majority of out-patient referrals were ineligible for inclusion because they had known ArLD, leaving 2/100 eligible out-patient referrals. One patient was excluded from the analysis as they were found to have an ALT of 1,023 following a pregabalin overdose, reducing the sample size from 100 to 99.

6.3.4 Clinical data

The data that I extracted from patients' electronic medical records included patient demographics, reason for presentation to hospital, BMI, alcohol intake (detailed in next section), postcode to enable deprivation score calculation, results of any imaging or fibrosis tests performed within 6 months of referral to ASN, blood test results to enable calculation of FIB4, AST:ALT ratio and APRI (23), and number of hospital presentations within the last five years. A research fellow in the research group assisted me in one aspect of the data collection through transcribing a small proportion of the blood results from the electronic patient records onto my predesigned spreadsheet.

Data were anonymised and entered into a password protected spreadsheet held on a secure NHS computer.

6.3.5 Alcohol data:

I recorded current alcohol consumption in units per week (U/w), and duration of 'excess alcohol consumption' in years. This was obtained from patients' self-reported consumption extracted from free text in the clinical notes. AUDIT scores or any other alcohol screening score were not available. I also noted if the patient had been actively drinking up to the point of presentation to hospital.

6.3.6 ELF score:

An ELF test was performed prospectively on consecutive eligible patients referred to the ASN. This was performed either by the ward phlebotomist or myself. Serum was extracted from 5ml blood per patient which was analysed at the Central ELF laboratory (iQur Limited, London). The samples were analysed for Hyaluronic acid (HA), Type III procollagen peptide (PIIINP) and Metalloproteinase inhibitor 1 (TIMP1) levels using the proprietary assays developed by Siemens Healthineers Inc (Tarrytown, New York, USA) for the ELF test, on a Siemens ADVIA centaur® immunoassay system. ELF Scores were calculated from test results using the manufacturer's published algorithm. An ELF threshold of 10.5 was pre-selected for detection of advanced fibrosis based on recommendations by Thiele et al. for use in ArLD (23).

6.3.7 Outcomes:

My primary outcome was the proportion of patients referred to the ASN at the Royal Free Hospital who had previously undetected advanced fibrosis (>F3) as determined by an ELF score of ≥10.5 (23). Secondary outcomes investigated potential risk factors for advanced fibrosis including alcohol consumption, BMI, age, sex, deprivation score and smoking status. In addition, I investigated missed opportunities by the RFL to previously diagnose advanced liver fibrosis by counting the number of attendances to the hospital within the previous five years without assessment for liver fibrosis being undertaken. This was done by a thorough case notes search for each patient, in order for me to ascertain if there had been any prior attempt at liver fibrosis assessment of any kind (including direct or indirect non-invasive fibrosis tests or biopsy).

6.3.8 Follow up:

In patients whose ELF scores were ≥10.5, I sent them a letter inviting them to attend a hepatology outpatient clinic to see a hepatologist with an interest in ArLD, and to have a FibroScan. Blood samples were also taken to screen for viral, immunological and metabolic causes of liver disease in accordance with current protocols if these tests had been omitted during their hospital admission.

6.3.9 Sample size:

As this is was an exploratory investigation, following statistical advice I accepted a precision of estimate at 0.1 which generated a sample size of between 62-89 using literature-based estimates of prevalence of advanced fibrosis (23). (A post-hoc sample size calculation for 0.29 prevalence of advanced fibrosis, using a precision around the estimate at 0.1 (10%) and 0.95% CI resulted in a minimum required sample size of n=80.)

6.4 Statistical analysis

I described demographic information using frequencies and percentages for categorical variables. For continuous data, I described these using means and SD or medians and IQR, depending on the normality of the data.

For the comparison of categorical variables, Chi-Squared or, if sample size was less than 5, Fisher's exact test was used as a conventional test, and for the continuous data the Mann Whitney U (for non-parametric data) or Student's t-test for normally distributed data.

I analysed alcohol 'units per week' both as continuous data, and in quartiles. After univariate analyses, to determine the variables associated with the presence of advanced fibrosis with the most significance, I used a multiple binary logistic regression analysis model, using the literature-based ELF threshold of 10.5 (23) as the binary value, and I used multiple linear regression for continuous ELF scores. Variables were selected if they were established in the literature as risk factors for

liver fibrosis, and if they had p values less than 0.25 in univariate analyses (either against ELF </≥10.5 or as a continuous variable). All p values were 2-sided and I considered them significant if $p<0.05$. In this chapter, I analysed all of the data using SPSS software (Version 25.0. Armonk, NY: IBM Corp).

6.5 RESULTS

6.5.1 Study demographics

My analysis included 99 patients (69% male) with a mean age of 53.1 years (SD 14.4) (Table 6.1). Mean BMI was 26.52 kg/m^2 (SD 5.94) and a high proportion (84%) were current or past smokers (Table 6.1). Alcohol intake was high with a median consumption of 140 U/w (80.9-280), and in this cohort men and women drank similar amounts ($p = 0.73$). The two patients seen in the ASN outpatient clinic had not been drinking alcohol within the past 3 months, and one inpatient had stopped drinking three weeks prior to hospital admission. All of the other 96/99 (97%) patients were drinking alcohol up to the point of presentation to hospital. The median duration of alcohol consumption was 15 years (IQR 10-27.5). This cohort of patients were from deprived areas, with 69% of them positioned within the lowest 4 deprivation deciles.

Table 6.1: Comparison of clinical characteristics of patients with and without advanced fibrosis (as determined by ELF score of ≥ 10.5).

Patient characteristics	Overall (n=99)	Advanced fibrosis ELF ≥ 10.5 N=28	Non-advanced fibrosis ELF <10.5 N=71	P value
Age mean sd	53.11 ± 14.37	55.7 ± 12.6	52.1 ± 15	0.27
Male sex n (%)	68/99 (69)	19 (68%)	49 (69%)	0.91
BMI mean sd	26.52 ± 5.94	26.4 ± 5.7	26.6 ± 6.1	0.90
T2DM diagnosis (%)	10/99 (10.1%)	4/28 (14.3%)	6/71 (8.5%)	0.46
Smoking status n (%)				
Non-smoker	15 (16)	6 (21)	9 (13)	0.35
Smoker	69 (73.4)	16 (57)	53 (75)	0.10
Ex-smoker	10 (10.6)	5 (18)	5 (7)	0.14
Unknown	5 (5)	1 (4)	4 (6)	
Ongoing active drinking n (%) ^	96/99 (97)	26/28 (93)	70/71 (99)	0.19
Current alcohol intake U/w, median (IQR)	140 (80.9-280)	112 (70-210)	150 (105-280)	0.03
Years of harmful drinking median (IQR)	15 (10-27.5)	20 (10-28)	15 (7.5-28)	0.36
Signs of CLD on exam				
Yes n (%)	4 (4%)	4 (14.3)	0 (0%)	<0.01
No n (%)	95 (96%)	24 (85.7)	71 (100%)	
Abnormal LFTs at referral $				
Yes	63 (66.3%)	18 (64.3)	45 (63.4%)	
No	32 (33.7%)	8 (28.6)	24 (33.8%)	0.71
N =	95	26	69	
ALT IU/L median (IQR)	39 (21 -73)	38 (14-76)	41 (22-71.75)	0.79
AST IU/L median (IQR)	43 (24-86.5)	52 (21-149)	40 (25.5-79.5)	0.49
MCV IU/L median (IQR)	96.9 (91.2-100.5)	97.8 (91.8-101.7)	96.8 (91.2-99.8)	0.52
Platelet count $x10^9$/L median (IQR)	206.5 (129-271)	203 (101-303)	206.5 (133.3-262.5)	0.65
Bilirubin μmol/L median (IQR)	10 (4-16)	10 (4-20)	10 (4-15.5)	0.62
FIB4 median (IQR)	2.00 (0.94-3.61)	2.04 (1.05-7.6)	1.96 (0.88-3.38)	0.30
APRI median (IQR)	0.64 (0.3-2.08)	0.64 (0.28-2.71)	0.63 (0.3-1.9)	0.55
AST:ALT ratio median (IQR)	1.3 (0.87-1.71)	1.5 (1.0-2.16)	1.26 (0.8-1.6)	0.07
HA median (IQR)	72.1 (35.1-144.5)			
PIIINP median (IQR)	8.18 (5.77-12.94)			
TIMP1 median (IQR)	265.7 (198.6-364)			
ELF median (IQR)	9.62 (8.93-10.6)			
ELF range (lowest to highest)	(6.87-13.78)			

^ At time of presentation to hospital or alcohol clinic

$ Abnormal LFTs defined as raised transaminases or ALP + GGT. (Not including isolated hyperbilirubinaemia (Gilbert's))

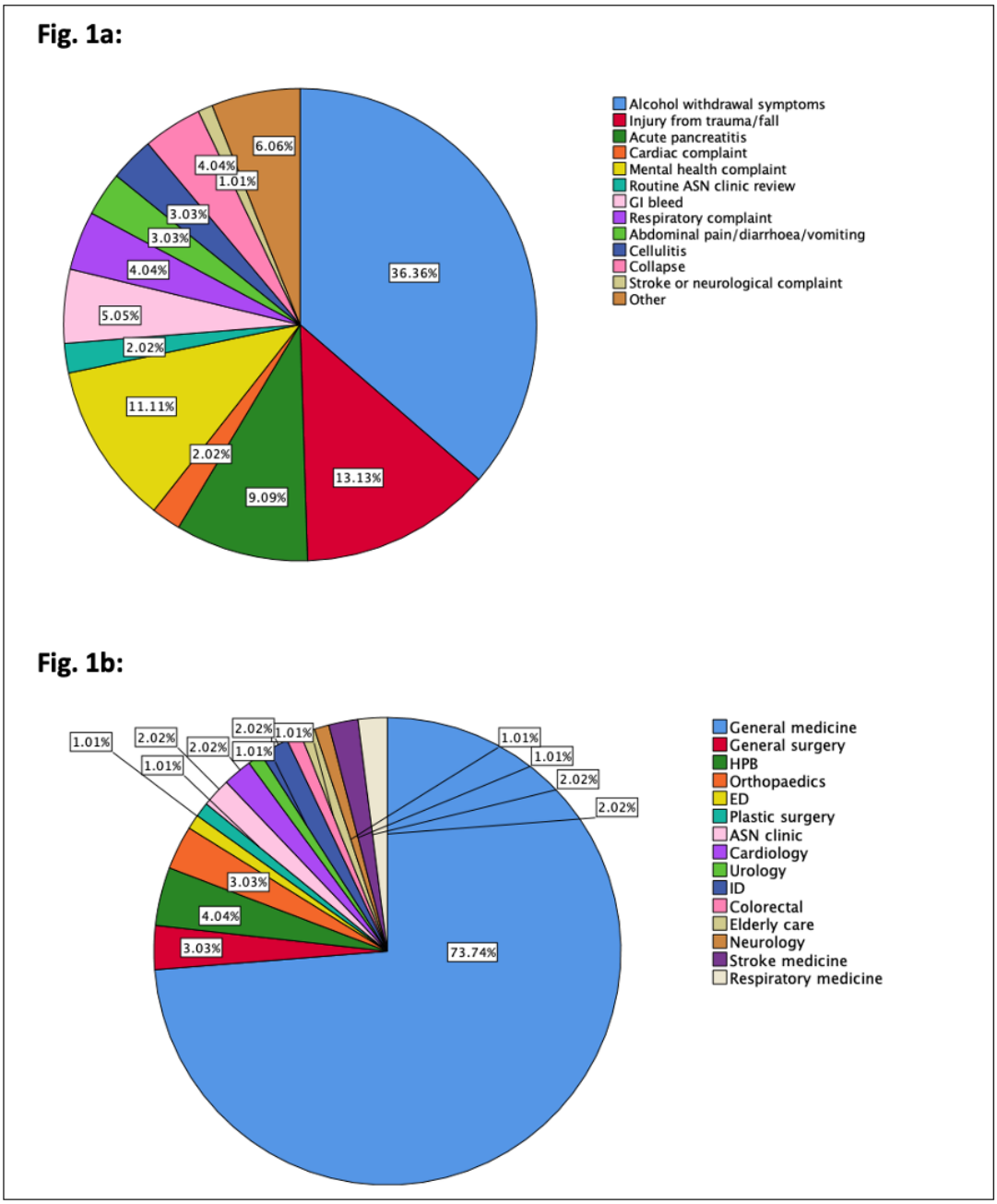
(sd = standard deviation, U/w = units per week, LFTs = Liver Function Tests, BMI =Body Mass Index, U/w = units per week, CLD = Chronic Liver Disease, T2DM = Type 2 Diabetes Mellitus, ALT = Alanine aminotransferase, AST = Aspartate aminotransferase, MCV = Mean Corpuscular Volume, APRI = AST to platelet ratio index, HA = Hyaluronic acid, PIIINP = Procollagen 3 N-terminal Peptide, ELF = Enhanced Liver Fibrosis score)

6.5.2 Reasons for presentation to healthcare

The vast majority (n=97/99, 98%) of patients were seen by the ASN as inpatients or in the emergency department. The most common reason for presentation to hospital to be symptomatic alcohol-withdrawal (36.4%) including seizures, followed by injuries from falling over (13.1%) and mental health presentations (11.1%) including overdose. The majority (73.7%) were under the care of a general medical team during their inpatient admission (Figures 6.1a and 6.1b). In the preceding 5 years 76% (75/99) of the patients had attended hospital (either inpatient admissions, or emergency department visits) without being diagnosed as having ArLD (aside from the current visit), with median number of hospital attendances being 4 (IQR 2-9).

Figure 6.1a: Pie chart of the principal recorded reasons for presentation to hospital

Figure 6.1b: Pie chart of the discharging hospital specialty team.



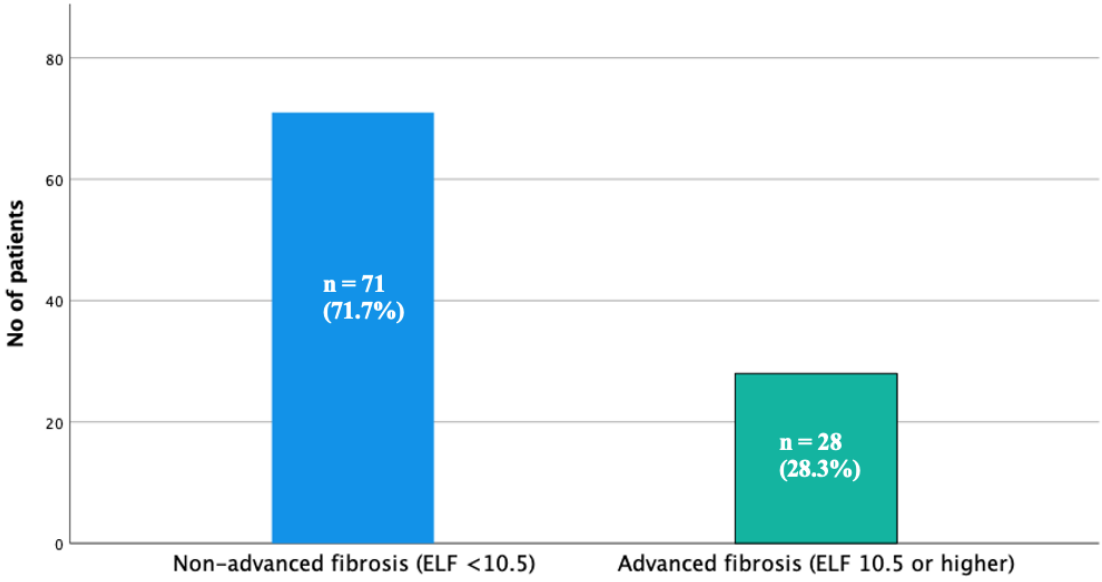
(AUD = Alcohol Use Disorder, ASN = Alcohol Specialist Nurse, GI = Gastro-Intestinal, HPB = Hepato-Pancreato-Biliary, ED = Emergency Department, ID = Infectious Diseases).

6.5.3 Results of non-invasive fibrosis tests and LFTs

LFTs were performed in 95/99 (96%) patients of which 63/95 (66.3%) of these patients had abnormal LFTs (raised transaminases or ALP + GGT), with median ALT of 39 IU/L (IQR 21-73) and AST of 43 IU/L (IQR 24-87 (Table 6.1). The median ELF score in the whole cohort was 9.62 (IQR 8.93-10.6, range 6.87-13.78). The ELF scores did not differ significantly between men and women (ELF score in men: 9.6 (IQR 8.8-10.6), and in women 9.8 (IQR 9-10.6), p = 0.435).

Twenty-eight participants (28.3%) had an ELF score of ≥ 10.5, indicating advanced fibrosis (Figure 6.2). Of the 28 patients with advanced fibrosis (ELF ≥ 10.5), 8 (28.6%) had normal LFTs (abnormal LFTs were defined as 'raised transaminases or ALP + GGT, not including isolated hyperbilirubinaemia (Gilbert's)').

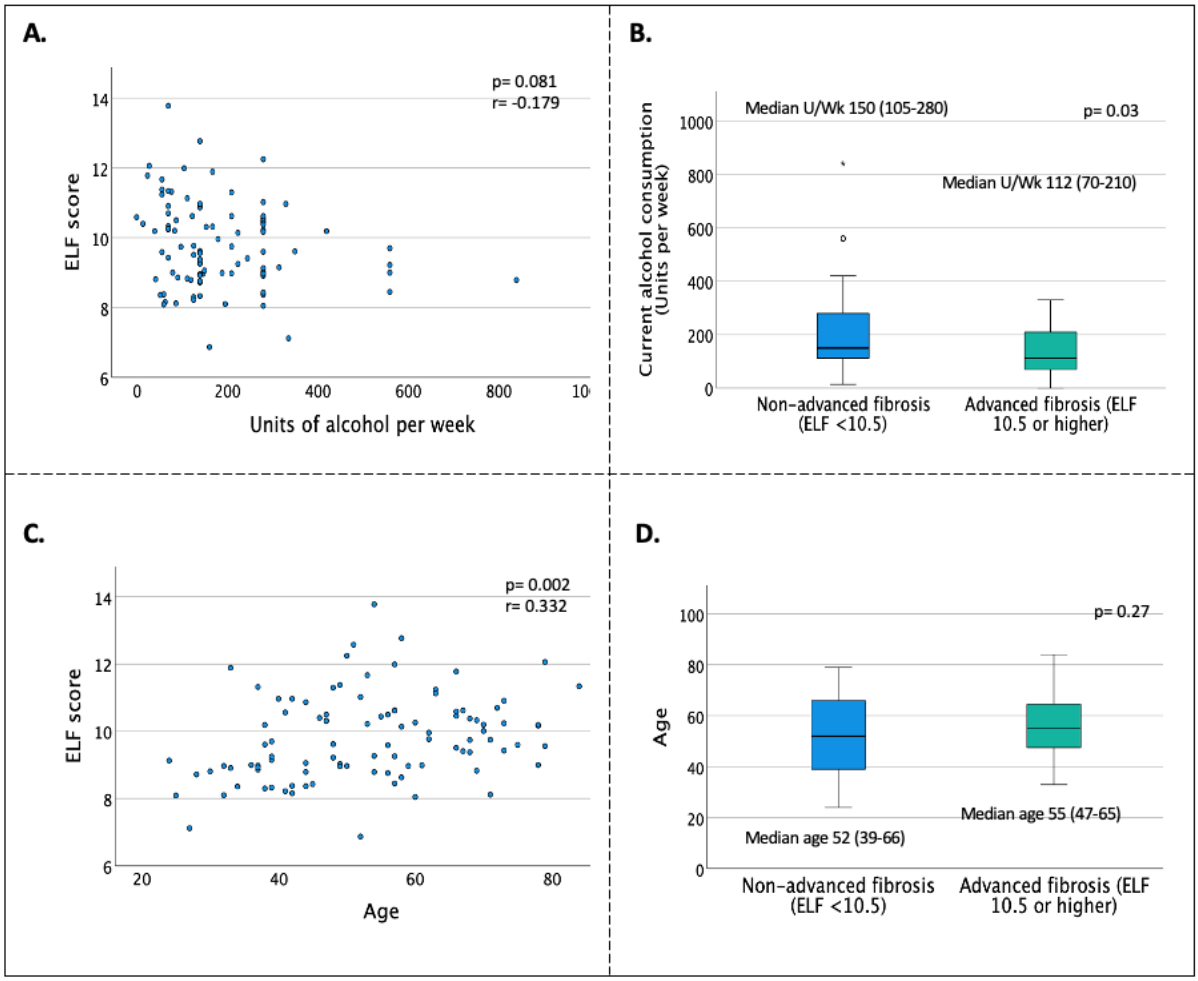
Figure 6.2: Proportion of patients in study cohort with advanced fibrosis as assessed by ELF ≥ 10.5



6.5.4 Comparison of clinical characteristics between patients with and without advanced fibrosis (ELF ≥ 10.5)

When comparing clinical characteristics with ELF score <10.5 versus ≥10.5, I could detect no significant difference in age, sex, or BMI between the two groups (Table 6.1, Figure 6.3D). There was also no significant difference in transaminase results, FIB4, or APRI scores between groups. Clinical signs of CLD (such as spider naevi, and palmar erythema), were only found to be documented in patients with ELF score ≥ 10.5 (n=4). Patients in the advanced fibrosis group (≥ 10.5) drank less alcohol than those with lower ELF scores (mean 112 U/w, compared with 150 U/w, p = 0.031, Figure 6.3B). However, there was no correlation observed between alcohol consumption and ELF score viewed as a continuous variable (Figure 6.3A). Furthermore, multivariable regression analysis revealed no association between alcohol consumption and ELF score (Table 6.2). There was no difference in the reported duration of alcohol excess in patients with ELF <10.5 compared to patients with ELF >10.5 (15 years (10 to 27.5 years) compared to 20 years (10 to 28 years); p = 0.357). Out of the three indirect biomarkers of fibrosis investigated (FIB4, APRI and AST:ALT ratio), AST:ALT ratios trended towards being higher in the advanced fibrosis group (median 1.5, IQR 1.0-2.16), than in the group with lower ELF scores (median 1.26, IQR 0.8-1.6; p = 0.074). On univariate analysis the AST:ALT ratio did significantly predict advanced fibrosis based on ELF (OR 2.081 (95% CI 1.145-3.779), p = 0.016 (Table 6.2). On multivariable regression analysis, increasing AST:ALT ratio was the only variable significantly associated with ELF scores indicative of advanced fibrosis, when adjusted for age, alcohol intake, bilirubin, MCV, and ALP (OR 1.984, 95%CI (1.014-3.884), p = 0.046 (see Table 6.2).

Figure 6.3 A, B, C, D: Influence of alcohol and age on binary and continuous ELF scores



A: Scatter plot of ELF by alcohol units per week (Spearman Rho correlation, with p value significance set at 0.05, r = correlation coefficient).
B: Boxplot of alcohol consumption (Units per week) by presence or absence of advanced fibrosis (ELF ≥ 10.5). Statistical test: Mann Whitney U, p value significance set at 0.05, median units per week displayed with IQR (interquartile range).
C: Scatter plot of ELF by age (Spearman Rho correlation, with p value significance set at 0.05, r = correlation coefficient).
D: Boxplot of age by presence or absence of advanced fibrosis (ELF ≥ 10.5). Statistical test: Mann Whitney U, p value significance set at 0.05, median age displayed with IQR (interquartile range).

Table 6.2: Factors associated with advanced fibrosis (ELF ≥ 10.5), as determined by univariable and multivariable regression analyses

Variable	B (unstandardized regression coefficient)	Univariable OR (95% CI)	p value	Multivariable OR (95% CI)	p value
Age $	0.018	1.018 (0.987-1.050)	0.264	1.010 (0.972-1.049)	0.609
Sex (male)	0.054	1.055 (0.412-2.699)	0.911		
BMI	-0.007	0.993 (0.891-1.107)	0.900		
Current alcohol intake (U/w)	- 0.005	0.995 (0.991-1.000)	0.041	0.995 (0.990-1.000)	0.070
Duration of alcohol excess	0.009	1.009 (0.975-1.044)	0.613		
Deprivation score	0.000	1 (1-1)	0.323		
Smoking (non-smoker)	-0.631	0.532 (0.170-1.667)	0.279		
Abnormal LFTs at referral*	0.182	1.20 (0.455-3.162)	0.712		
ALP	0.006	1.006 (0.999-1.012)	0.099	1.004 (0.998-1.011)	0.190
ALT	0.002	1.002 (0.995-1.010)	0.574		
MCV ^	0.005	1.005 (0.950-1.063)	0.856	0.971 (0.908-1.039)	0.399
Platelet count	0.000	1.000 (0.996-1.003)	0.786		
Bilirubin	0.025	1.026 (0.988-1.065)	0.187	0.999 (0.953-1.047)	0.966
AST §	0.003	1.003 (0.998-1.008)	0.215		
FIB4 §	0.076	1.079 (0.996-1.169)	0.063		
AST/ALT ratio	0.733	2.081 (1.145-3.779)	0.016	1.984 (1.014-3.884)	0.046
APRI §	0.082	1.085 (0.997-1.182)	0.059		

$ Although p value for age was above 0.25 in univariable logistic regression, it was <0.05 in correlation analysis with continuous ELF score, and is of clinical importance to investigate -so was included in this multivariable model.

^ Although p value for MCV was above 0.25 in univariable logistic regression, it was <0.05 in correlation analysis with continuous ELF score, and so was included in this multivariable model.

§ Left out of multivariable analysis as would be affected by multi-collinearity with AST:ALT ratio, which was more highly significant in the univariate analysis

* Abnormal LFTs defined as raised transaminases or ALP + GGT. (Not including isolated hyperbilirubinaemia (Gilbert's))

(BMI =Body Mass Index, U/w = units per week, CLD = Chronic Liver Disease, T2DM = Type 2 Diabetes Mellitus, ALT = Alanine aminotransferase, AST = Aspartate aminotransferase, MCV = Mean Corpuscular Volume, APRI = AST to platelet ratio index, ALP = Alkaline Phosphatase, OR = Odds Ratio).

6.5.5 Factors associated with increasing ELF score

When I analysed literature-derived risk factors for liver fibrosis against a continuous ELF score, there was no longer a significant association between the amount of alcohol consumption (U/w) and ELF score (p = 0.081), (Figure 6.3A) and this was confirmed in multivariable regression analysis, both using continuous ELF (Table 6.3) and binary ELF scores </≥ 10.5 (Table 6.2). Alcohol intake was also analysed by grouping the units consumed per week into quartiles (0-79 U/w, 80-140 U/w, 141-280 U/w and 281+ U/w. There was no significant difference in ELF score between the quartiles either when ELF was analysed as a continuous score or using the 10.5 threshold (Tables 6.4 A, B).

ELF scores increased with increasing age (patients' total age range 24-84) on univariate analysis (Figure 6.3C), (*r* = 0.33, p = 0.002), and this was confirmed in multivariable analysis, when adjusted for alcohol intake, AST:ALT ratio, ALP, MCV and bilirubin (p = 0.013, 95% CI 0.005-0.042) (Table 6.3). For every 10-years increase in age, the ELF score increased by 0.24.

ALT or AST were not associated with ELF score (either binary ELF of > or < 10.5, or continuous ELF (Figure 6.4A-D and Table 6.2). Whilst AST:ALT ratio predicted advanced fibrosis when assessed using the 10.5 ELF threshold, a significant correlation was not seen between AST: ALT and continuous ELF score (r = 0.12, p =0.27) (Figures 6.5A, 6.5B).

Table 6.3: Summary of multiple regression analysis of factors associated with continuous ELF score.

Model	Unstandardized Coefficients		Standardized Coefficients	Sig.	95% Confidence Interval for B	
	B	Std. Error	Beta		Lower Bound	Upper Bound
(Constant)	6.701	1.580		.000	3.554	9.849
Age	.024	.009	.278	.013	.005	.042
Current alcohol intake (U/w)	-.001	.001	-.105	.334	-.003	.001
ALP	.003	.002	.173	.111	-.001	.006
Bilirubin	.018	.012	.170	.146	-.006	.042
MCV	.014	.016	.094	.374	-.018	.046
AST:ALT ratio	.115	.170	.077	.502	-.224	.454

(Dependent Variable: ELF score)
(ALP = Alkaline Phosphatase, MCV = Mean Corpuscular Volume, AST = Aspartate Aminotransferase, ALT = Alanine Aminotransferase)

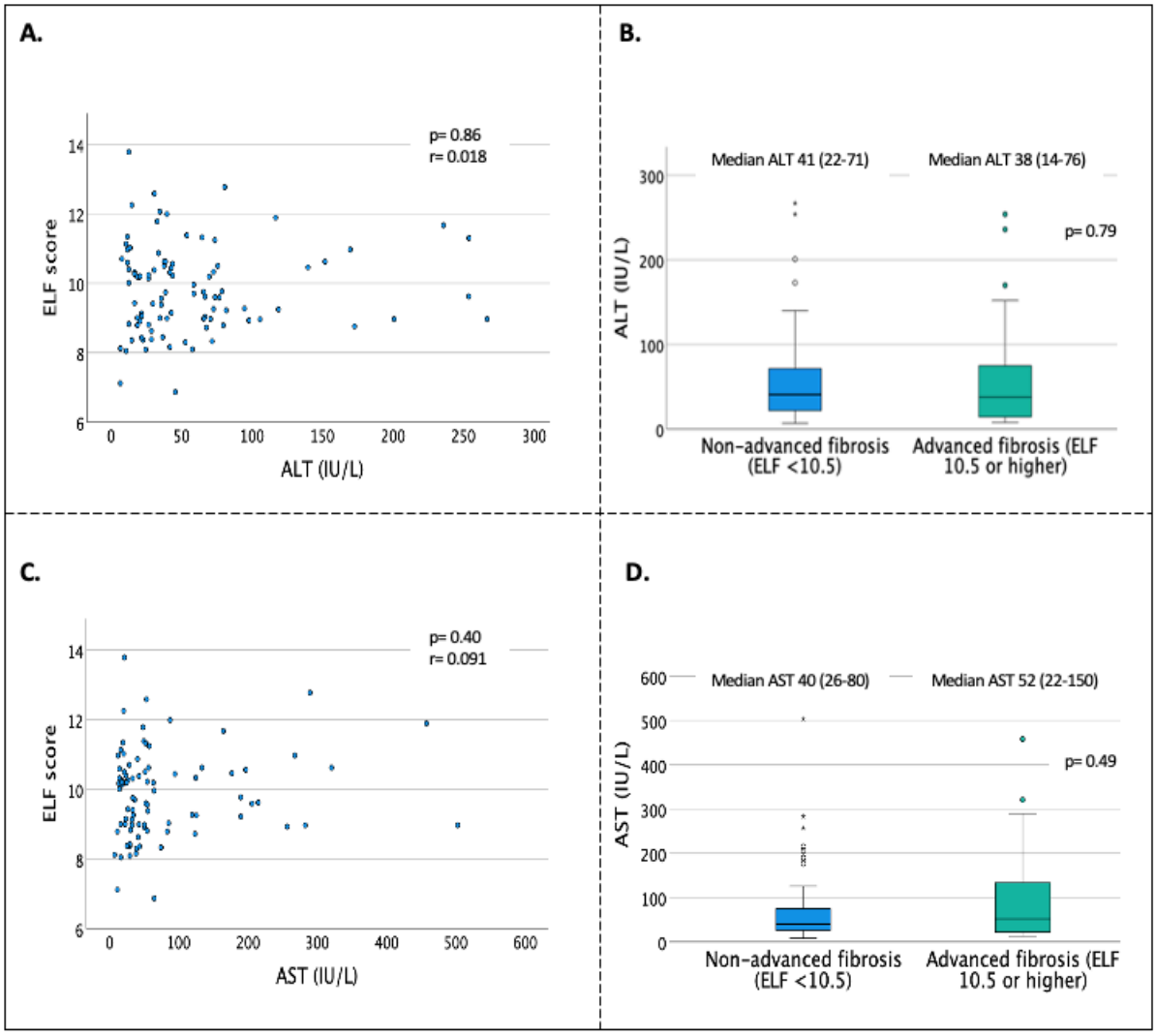
Table 6.4a: Multiple logistic regression analysis to investigate for effect of alcohol unit quartiles on presence or absence of advanced fibrosis (as per ELF ≥10.5)

Quartiles of alcohol units/week	B	S.E.	Wald	*df*	P value	OR	95% C.I. for OR	
							Lower	Upper
80-140 U/w (n=27)	-.963	.606	2.525	1	.112	.382	.116	1.252
141-280 U/w (n=35)	-1.129	.580	3.793	1	.051	.323	.104	1.007
281-840 U/w (n=10)	-2.110	1.134	3.464	1	.063	.121	.013	1.118
Constant (0-79 U/w) (n=23)	-.087	.417	.043	1	.835	.917		

Table 6.4b: Multiple linear regression analysis to investigate for effect of alcohol unit quartiles on continuous ELF score.

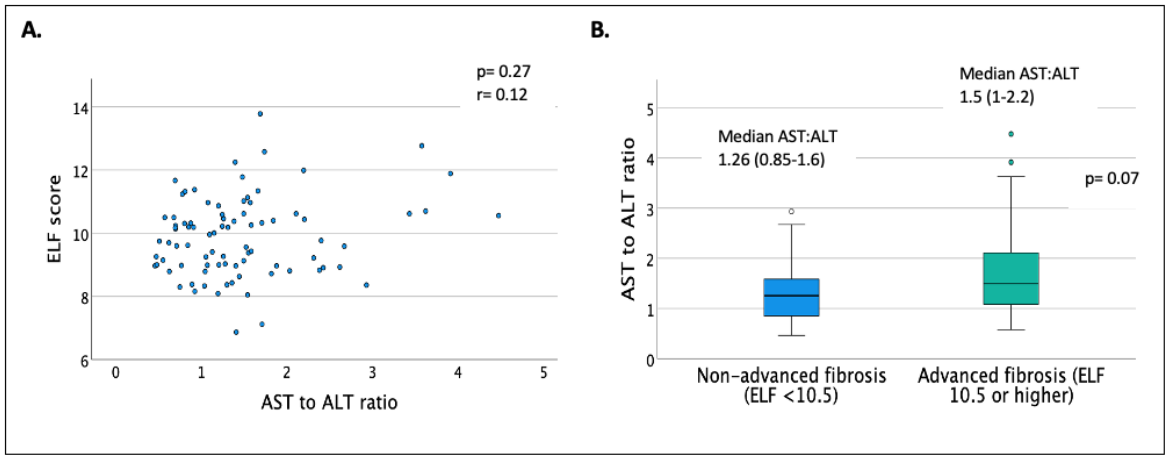
Model	Unstandardized Coefficients		Standardized Coefficients	Sig.	95.0% Confidence Interval for B	
	B	Std. Error	Beta		Lower Bound	Upper Bound
(Constant)	9.213	.700		.000	7.823	10.604
0-79 U/w (n=23)	1.270	.743	.435	.091	-.205	2.745
80-140 U/w (n=27)	.416	.738	.148	.575	-1.050	1.881
141-280 U/w (n=35)	.501	.730	.191	.494	-.948	1.950
281-840 U/w (n=10)	.007	.799	.002	.993	-1.579	1.592
Dependent Variable: ELF score						

Figure 6.4 A, B, C, D: Influence of ALT and AST on binary and continuous ELF scores



A: Scatter plot of ELF by ALT value (Spearman Rho correlation, with p value significance set at 0.05, r = correlation coefficient).
B: Boxplot of ALT by presence or absence of advanced fibrosis (ELF ≥ 10.5).
Statistical test: Mann Whitney U, p value significance set at 0.05, ALT displayed with IQR (interquartile range).
C: Scatter plot of ELF by AST value (Spearman Rho correlation, with p value significance set at 0.05, r = correlation coefficient).
D: Boxplot of AST by presence or absence of advanced fibrosis (ELF ≥ 10.5).
Statistical test: Mann Whitney U, p value significance set at 0.05, AST displayed with IQR (interquartile range).

Fig 6.5 A, B: Influence of AST:ALT ratio on binary and continuous ELF scores



A: Scatter plot of ELF by AST:ALT ratio (Spearman Rho correlation, with p value significance set at 0.05, r = correlation coefficient).
B: Boxplot of AST:ALT ratio by presence or absence of advanced fibrosis (ELF ≥ 10.5). Statistical test: Mann Whitney U, p value significance set at 0.05, AST:ALT ratio displayed with IQR (interquartile range).

6.5.6 FibroScan results

Of the 28 patients with ELF ≥ 10.5 who were offered a FibroScan appointment, only 18 attended (64%), one of whom did not have a valid FibroScan reading, leaving 17 valid results (failure rate of 6%). The mean FibroScan value was 10.9kPa (±7.1kpa), (total range of 4.2 kPa-25.3kPa). Using a literature-derived threshold of 9.5 kPa for advanced fibrosis (229), 10/17 (58.8%) with ELF ≥ 10.5 had a FibroScan value <9.5 kPa (range 4.2-9.1 kPa).

6.6 Discussion

I discovered that nearly a third (28.3%) of patients with AUD presenting to hospital for a variety of reasons had an ELF score of ≥10.5, indicating the presence of advanced liver fibrosis. None of these patients had been assessed previously for liver fibrosis or referred to a liver specialist. All of them were at high risk of liver damage, with a current median alcohol consumption of 140 U/w, and history of excess alcohol consumption lasting more than 15 years, and yet none had been investigated for liver disease during their index presentation or at any time previously. Moreover, 76% of the cohort had presented to hospital on a median of four times per person over the preceding five years without a diagnosis of ArLD, indicating missed opportunities for detection and treatment of liver fibrosis in a high-risk population. Missed opportunities for recognising and assessing liver damage in primary care were not investigated in this study, but none of the patients in this study had been referred to hepatology services for assessment of liver disease prior to diagnosis in this study.

I found that LFTs were not a reliable predictor of advanced fibrosis, with 28% of patients who had ELF ≥10.5 having normal LFTs, in concordance with previous reports (1). This highlights the need for education to secondary care clinicians to think about fibrosis testing in people drinking over recommended alcohol limits, even if LFTs are normal, in keeping with current guidelines (1). Moreover, the initiation of hospital alcohol pathways that prompt routine assessment and documentation of alcohol intake incorporating guidelines on when to perform

fibrosis assessments and refer to ASN, would standardise this practice and normalise fibrosis assessment in at-risk people.

Whilst I found that ELF scores were positively correlated with increasing age ($r = 0.303$, $p = 0.002$), I found no difference in the median ages of those with or without advanced fibrosis, as determined by ELF scores ≥ 10.5 or < 10.5 respectively. ELF score has been found to correlate with age in some (24, 25), but not all studies (26) and it is unclear how much of the reported correlation is due to the increased likelihood of advanced fibrosis being present in older patients (27). McPherson et al. (28) studied the impact of age on the performance of a range of NIT (NFS, FIB4, AST:ALT ratio, but not ELF) in detecting advanced fibrosis (compared with biopsy) in 634 patients with Non-Alcoholic-Fatty-Liver-Disease (NAFLD), and found that all the tests performed less well in people over the age of 65 (with an increase in false positive rates in this age group). They suggested the use of adjusted thresholds for diagnosing advanced fibrosis in this age range. Fagan et al. (24) found increased risk of false positives with ELF above the age of 45 in a cohort of 329 patients (mixed aetiology liver disease), concluding that caution needs to be taken in interpreting ELF scores in older age groups. Thiele et al. (3) also reported increased false positive ELF results in people over 60 in a cohort of 289 patients investigated for ArLD, and advised caution in interpreting ELF in the over 60s. In contrast Parkes et al. (26) found no influence of age on ELF score in a cohort of 347 patients with chronic hepatitis C (CHC) raising the possibility that age influences ELF score in NAFLD and ArLD but not in CHC.

The evidence discussed indicates that ELF may be less accurate in older patients, with risk of more false positive results. It would be useful in future to specifically investigate in a prospective cohort if there is a predictable rise in ELF with age or if the test is merely less accurate in older people, and to evaluate the performance of age-based cut-offs for ELF.

The amount of alcohol consumed in U/w or duration of heavy drinking was not associated with ELF score in this cohort, and this was also the case in a large biopsy-controlled study of ELF in AUD (n=289) by Thiele *et al.*(23) The same study also found that ALP was associated with increased ELF score, as observed in my univariate analysis in this study, although not when adjusting for other factors in multivariable analysis.

Increasing AST:ALT ratio was the only other marker significantly associated with advanced fibrosis (ELF $\geq$10.5) in this study (OR 1.984, 95%CI 1.014-3.884, p = 0.046). Thiele et al (23) found that AST:ALT had a Negative Predictive Value of 91% in a large biopsy-paired study and it may be that AST:ALT ratio could be used as a simple direct fibrosis test in addition to ELF in the assessment of advanced liver fibrosis in ArLD in a manner analogous to the combination of FIB4 and ELF in NAFLD,(14) but this would require validation.

Whilst it has previously been reported that ELF scores may be influenced by inflammation (314, 318), I did not find any correlation between ELF score and ALT or AST, as markers of hepatic inflammation in this study, suggesting ELF was not influenced by inflammation in this cohort. In a large (n=289) biopsy-paired

prospective study on ArLD in 2018, Thiele. et al found that ELF was associated with histological inflammation but, like in this study, not with AST or ALT values (23). Similarly, Connoley et al 2021 found no correlation between ELF and ALT. It must be noted, however, that patients with acute alcoholic hepatitis or acute liver injury from non-alcohol-related causes were excluded in my study, and in the other two published studies (23, 236), and ELF is not currently validated in these settings.

The fact that ELF does not appear to be influenced by AST or ALT in this study nor in the only two other published studies on ELF in ArLD (23, 236), suggests that ELF could confer advantage over FibroScan in this regard (319-321). (As I discovered in Chapter 3, FibroScan is influenced by transaminases and Bilirubin levels with increased false positive readings in the presence of elevated transaminases/Bilirubin (132, 321, 322)). I think, however, there does need to be further prospective investigation, for any potential association between ELF and histological inflammation.

Limitations of this study include the lack of paired biopsies that would have provided a more robust reference standard assessment of liver fibrosis, and also help answer the question about association between ELF and histological inflammation. However the use of non-invasive tests to assess liver fibrosis in this study is representative of current clinical practice within the NHS and in many other countries, where patients presenting to hospital with AUD are not routinely biopsied, partly due to increasing recognition of the imperfections of biopsy as a test for liver fibrosis due to sampling error, inter and intra observer variability and the costs and hazards associated with biopsy (171, 218). FibroScan was offered to

all participants with ELF scores ≥ 10.5, but only 18/28 attended, of which valid readings were obtained for 17/18. Whilst FibroScan and ELF were discordant in 10/17 cases, FibroScan cannot be considered a robust reference standard measurement of fibrosis in ArLD, due to the recognised impact of alcohol and inflammation on the accuracy of elastography. The small number of patients attending for FibroScan means that it is not possible for me to draw robust conclusions about the performance of FibroScan in this cohort. Furthermore, the poor attendance rate illustrates both the need to assess patients 'opportunistically' while they are inpatients, and the greater reliability of using a blood test to assess fibrosis that can be incorporated in routine investigations.

In common with routine practice, I relied on patients' self-reported alcohol intake extracted from clinical records, an approach that is likely to be inaccurate. Unfortunately, it is not local routine practice to obtain AUDIT scores (either by admitting clinician or the ASN) but these would provide additional valuable information about drinking behaviour. Fibrosis was assessed using a single ELF test at the start of the patients' hospital admission. Although liver stiffness as measured by FibroScan reduces significantly on withdrawal of alcohol (206, 267, 323), a study of ten patients found that there was no significant difference in the ELF scores recorded from intoxicated patients when re-tested two weeks after alcohol withdrawal (324). Similarly, Thiele et al found no association between alcohol intake and ELF score (23). However, the impact of drinking on ELF score needs further investigation. One way of addressing this would be to design a prospective study (sufficiently powered for the primary outcome measure), whereby ELF tests

are performed sequentially on patients withdrawing from alcohol, as has been performed in the FibroScan studies referenced in the previous paragraph.

Overall, in this study I have highlighted the missed opportunities for detecting liver fibrosis in at-risk patients in a hospital setting. Alcohol use disorder must be viewed as a multimorbid condition with psycho-social morbidity and the potential to damage every organ in the body. However, ArLD accounts for much of the mortality and costs of drinking and accurate and relatively inexpensive blood tests are now available that permit detection of liver fibrosis in those at risk. It could be argued that there is no longer any excuse to miss the diagnosis of liver fibrosis in patients presenting to hospital with AUD. Whilst people with AUD encompass some of the more socially disadvantaged members of society that may find engaging with routine health services difficult, it is imperative that all opportunities to detect liver fibrosis should be taken especially on those occasions when they present to hospital with complications of AUD or other conditions.

BSG guidance now recommends non-invasive fibrosis testing for people with high-risk alcohol intake (>35 U/w in women, >50 U/w in men) (1) with either FibroScan or ELF.

This study emphasises the importance of implementing this guidance and incorporating it into hospital guidelines in emergency departments and in alcohol care teams (325) to improve the detection of advanced fibrosis in people with AUD.

Having so far examined the diagnostic performance of non-invasive tests in ArLD, and their current use in primary and secondary care and opportunities for

improving practice, in the next chapter I will go on to investigate their prognostic performance in ArLD.

Chapter 7

Investigating the prognostic performance of four non-invasive tests in Alcohol-related Liver Disease: a systematic review

7.1 ABSTRACT:

Background/Aims:

Mortality of Alcohol-related-Liver-Disease (ArLD) is increasing, and liver fibrosis stage is the best predictor of mortality. Non-invasive-tests (NIT) are increasingly used to detect fibrosis, but their value as prognostic tests in chronic liver disease (CLD), and in particular in ArLD is less well recognized. After having evaluated the diagnostic performance of four widely used NITs in ArLD in Chapter 3 (FIB4, ELF test, FibroScan and FibroTest), I now aim to investigate their prognostic performance.

Methods:

Applying systematic-review methodology, I searched four databases from inception to May 2020. Inclusion/exclusion criteria were applied to search using MeSH terms and keywords. Both myself and a second reviewer independently screened the search results, extracted data and performed risk-of-bias assessment using Quality-In-Prognostic-Studies (QUIPS) tool.

Results:

Searches identified 25,088 articles. After initial screening, 1,020 articles were reviewed independently by myself and the second reviewer. Eleven articles remained after screening for eligibility: one on ELF, four on FibroScan, four on FIB4, one on FIB4 and FibroScan and one on FibroTest and FIB4. I found few studies focused on NIT performance in ArLD compared to NAFLD, viral hepatitis and mixed-aetiologies. Area-Under-Receiving-Operator-Characteristics-curves (AUROCS) for outcome-prediction ranged from: 0.65-0.76 for FibroScan, 0.64-0.83 for FIB4, 0.69-0.79 for FibroTest and 0.72-0.85 for ELF. Studies scored low-moderate risk of bias for most domains, but high-risk in confounding/statistical reporting domains. The results

were heterogeneous for outcomes and reporting, making pooling of data unfeasible. However, where a study reported direct comparisons between tests, FIB4 performed better than MELD in prognosis prediction in ArLD, FibroTest and ELF performed at least as well as histology.

Conclusions:

This systematic-review returned eleven papers, six of which were conference-abstracts. Whilst the heterogeneity of studies precluded direct comparisons of NITs, each NIT performed well in predicting prognosis in ArLD (AUROCs >0.7 in each NIT category) in individual studies, and each may add value to prognostication in clinical practice.

Notice of publication

I have published a version of this chapter in the Journal of Gastroenterology and Hepatology in 2021 (224): *(Rhodes FA, Trembling P, Panovska-Griffiths J, Tanwar S, Westbrook RH, Rodger A, et al. Systematic review: Investigating the prognostic performance of four non-invasive tests in alcohol-related liver disease. J Gastroenterol Hepatol. 2020.)*

7.2 Introduction

Mortality rates from cirrhosis have increased by 400% over the last 30 years, largely attributed to alcohol (326). The degree of liver fibrosis is the strongest predictor of mortality in chronic liver disease,(327) and thus it is important for clinicians to have information about fibrosis in order to predict clinical outcomes and guide individualised treatment decisions.

Liver biopsy is the traditional modality for detecting and quantifying fibrosis in alcohol-related liver disease (ArLD) and the current reference standard against which other tests for fibrosis are evaluated. However, liver biopsy is considered an imperfect test owing to its invasive nature with associated risks to the patient, as well as sampling error and reporting bias (171, 218). Therefore, there has been a drive to develop non-invasive tests (NIT) for liver fibrosis over the last two decades to assess fibrosis severity and to determine prognosis. These NIT largely comprise blood tests that measure direct and indirect markers of liver fibrosis, of which the most widely studied are The Enhanced Liver Fibrosis (ELF) test, FibroTest, HepaScore, Fibrometer, FIB4, Forns' Index, APRI, AST:ALT ratio, and age-platelet index (328). There are also "physical" techniques assessing liver stiffness, including FibroScan, sheer-wave elastography and MR elastography but these are less generalisable due to operator performance and availability.

I conducted a scoping exercise to identify NIT that had been investigated for both their prognostic and diagnostic performance, and were established enough that they could be readily translated into clinical practice for routine prognostic

assessment. The markers selected on these criteria are: FibroScan, FIB4, ELF test and FibroTest. The selection of prognostic markers is of particular importance in the practice of stratified or personalised medicine where they can support clinicians and patients in making decisions about management such as initiating treatments, and initiating enhanced monitoring for complications of cirrhosis.

Whilst there is an increasing number of studies on prognostic markers, few have been externally validated for use in clinical practice (329). Moreover, as I discovered in Chapter 3, the majority of validation studies have been performed in patients with either viral hepatitis or unselected chronic liver disease, rather than specifically in ArLD. It has been shown in cholangiopathies and all-cause CLD that NIT can out-perform histology in predicting clinical outcomes (30, 330), and therefore it is of great clinical importance to know if NIT can also reliably predict outcomes in ArLD, the commonest aetiology of cirrhosis.

In this systematic review I aim to determine the prognostic performance of four commonly used NIT for liver fibrosis in ArLD, specifically in predicting mortality, and liver related events (LRE) resulting in decompensated cirrhosis and death.

7.3.1 Patients and methods

I conducted this systematic review using the guidance provided in the Cochrane Handbook (331). My aim in this study was to identify the prognostic performance of four non-invasive tests for liver fibrosis in alcohol-related liver disease – FibroScan, ELF test, FibroTest and FIB4. I followed the PICO structure (participants, interventions, comparators, outcomes and study design) and followed PRISMA guidance (Preferred Reporting Items for Systematic reviews and Meta-Analyses) (See Table 7.1). I prospectively registered the protocol for this review with PROSPERO (Registration ID: CRD42020175605).

Table 7.1: Databases searched

I updated the search on 26/05/2020 using the same search strategy and methodologies.

Databases	**Date initial search performed**	**Date repeat search performed**
MEDLINE (Ovid) (1946 to date of search)	10/10/2019	26/05/2020
EMBASE (Ovid) (1974 to date of search)	10/10/2019	26/05/2020
Web of Science (1900 to date of search)	10/10/2019	26/05/2020
Cochrane Database of Systematic Reviews	10/10/2019	26/05/2020

7.3.2 PICO

Participants:

All adult humans (age 18+ years) with Alcohol related Liver Disease

Intervention:

Studies that included FIB4, FibroTest, FibroScan or ELF test as prognostic markers in ArLD were included.

Comparisons:

Each of the above interventions were compared to one another.

Outcomes:

1. The ability of ELF, FibroTest, FibroScan and FIB4 to predict all cause and liver-related mortality
2. The ability of ELF, FibroTest, FibroScan and FIB4 to predict liver-related cirrhotic decompensation events including ascites, variceal bleeding, encephalopathy, need for liver transplantation and development of hepatocellular carcinoma (HCC)

7.3.3 Ethics and patient consent

Ethical approval was not required for this systematic review, since it used data from previous studies which had their own ethics and patient consent.

7.3.4 Search strategy

I searched four databases systematically, using search strategies which can be found in Tables 7.2, 7.3 and 7.4.

Search themes related to my study PICO, including a combination of MeSH terms and keywords. I conducted pilot searches in order to refine the search strategy. Firstly, Web of Science, Ovid Medline, Embase and Cochrane Library were systematically searched (see Table 7.1). Secondly, reference lists of included studies and relevant review articles were hand-searched by myself to identify any further potentially relevant publications. Thirdly, where information from abstracts or full texts was not sufficient for us to include the study, I contacted relevant authors by email to request data.

To make the search as comprehensive as possible, the key words searched were a combination of MeSH terms and free text words. I used different combinations of similar words for example for ELF I used 'ELF' or 'elf adj score' or 'hyaluronic acid' or 'hyalauronate' or 'hyaluronan' or 'procollagen' or 'piiinp' or 'p3np' or 'ppcp' or 'tissue adj inhibitor adj metalloproteinase$' or 'timp$'. For a more general search of non-invasive liver fibrosis tests, I included 'biological marker$' or 'biomarker$' or 'algorithm$' or 'non adj invasive adj test' or 'non adj invasive', and these terms were included using 'AND' along with terms for alcohol-related liver disease including MeSH terms. I did not limit the cirrhosis to alcohol-related in these searches, so as to be able to explore studies on cirrhosis in case they included sub-analyses on alcohol-related liver disease. These terms were finally combined with 'AND' for the prognostic terms, e.g., 'predict$', 'prognos$', 'mortality', 'prediction', 'risk', 'follow adj up', 'prediction and forecasting', 'adverse outcome', 'predictive value', 'prognos$, 'outcome$', 'treatment adj outcome', 'disease progression', 'course', 'mortal$', 'death', 'cancer', 'neoplas$', 'malignan$', 'transplant (Tables 7.2 a, b, c).

I imported the references into Endnote web basic reference manager, and then the selection of articles for both myself and the second reviewer to review was imported into Rayyan systematic review manager (225), which enabled independent, blinded review of each article and documentation of reasons for exclusion.

TABLE 7.2a: DETAILS OF SEARCH STRATEGY: Cochrane database

No.	Searches	Search type	Total number of results
1	(enhanced adj liver fibrosis) OR (elf adj test) OR (Elf and prognos$) OR (elf) OR (elf adj score) OR (FibroTest) OR (FibroScan) OR (transient adj elastography$) OR (elastography$ and liver) OR (hyaluronic adj acid) OR (hyalauronate) OR (hyaluronan) OR (procollagen) OR (piiinp) OR (p3np) OR (ppcp) OR (tissue adj inhibitor adj metalloproteinase$) OR (timp$) OR (FIB4) OR (FIB 4) OR (biological marker$) OR (biomarker$) OR (algorithm$) OR (non adj invasive adj test) OR (non adj invasive)	Advanced	27742
2	MeSH descriptor: [Fatty Liver, Alcoholic] explode all trees	Advanced	21
3	MeSH descriptor: [Liver Diseases, Alcoholic] explode all trees	Advanced	468
4	(predict$ or prognos$) OR (mortality) OR (Prediction) OR (risk) OR (follow adj up) OR (prediction and forecasting) OR adverse outcome) OR (predictive value) OR (prognos$) OR (outcome$) OR (treatment adj outcome) OR (disease progression) OR (predictive value of test$) OR (course) OR (mortal$) OR (death) OR (cancer) OR (neoplas$) OR (malignan$) OR (transplant$)	Advanced	784625
5	#2 or #3	Advanced	468
6	#1 and #4 and #5	Advanced	11

TABLE 7.2b: DETAILS OF SEARCH STRATEGY: Ovid MEDLINE and EMBASE

No.	Searches	Search type	Total number of results
1.	(enhanced adj liver adj fibrosis).tw.	Advanced	
2.	(elf adj test$).tw.	Advanced	
3.	(elf and prognos$).tw.	Advanced	
4.	elf.tw.	Advanced	
5.	(elf adj score).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	Advanced	
6.	FibroTest.tw.	Advanced	
7.	FibroScan.tw.	Advanced	
8.	(transient adj elastograph$).tw.	Advanced	
9.	(elastograph$ and liver).tw.	Advanced	
10.	(hyaluronic adj acid).mp. or (hyalauronate or hyaluronan).tw. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	Advanced	
11.	(procollagen or piiinp or p3np or ppcp).tw.	Advanced	
12.	((tissue and inhibitor and metalloproteinase$) or timp*).tw.	Advanced	
13.	FIB 4.tw.	Advanced	
14.	FIB4.tw.	Advanced	
15.	biological markers/	Advanced	
16.	biomarker$.tw.	Advanced	
17.	algorithm$.tw.	Advanced	
18.	(non adj invasive adj test).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	Advanced	
19.	(non adj invasive).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	Advanced	
20.	exp liver cirrhosis/ or exp liver diseases, alcoholic/	Advanced	
21.	((liver adj fibros*s) or cirrhos*s or (hepatic adj fibros*s)).tw.	Advanced	
22.	(predict* or prognos*).mp.	Advanced	
23.	mortality/ or prediction/ or risk/ or follow up/	Advanced	
24.	"prediction and forecasting"/ or exp adverse outcome/ or exp predictive value/ or exp prognosis/	Advanced	
25.	outcome*.mp. or treatment outcome/	Advanced	
26.	disease progression/ or predictive value of tests/	Advanced	
27.	course/ or mortal*/ or death*/	Advanced	
28.	cancer/ or neoplas$/ or malignan$/ or transplant$/	Advanced	

29.	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19	Advanced	
30.	20 or 21	Advanced	
31.	22 or 23 or 24 or 25 or 26 or 27 or 28	Advanced	
32.	29 and 30 and 31	Advanced	
33.	limit 32 to human	Advanced	15,424

TABLE 7.2c: DETAILS OF SEARCH STRATEGY: Web of Science

No.	Searches	Search type	Total number of results
1.	TS = enhanced adj liver adj fibrosis	Advanced	
2.	TS = elf adj test$	Advanced	
3.	TS = elf	Advanced	
4.	TS = elf adj score	Advanced	
5.	TS = FibroTest	Advanced	
6.	TS = FibroScan	Advanced	
7.	TS = transient adj elastograph$	Advanced	
8.	TS = (elastograph$ and liver)	Advanced	
9.	TS = (hyaluronic adj acid OR hyalauronate OR hyaluronan)	Advanced	
10.	TS = (procollagen OR piiinp or p3np or ppcp)	Advanced	
11.	TS = (procollagen OR piiinp or p3np or ppcp)	Advanced	
12.	TS = (tissue adj inhibitor adj1 metalloproteinase$) OR TS = timp$	Advanced	
13.	TS = FIB4	Advanced	
14.	TS = FIB 4	Advanced	
15.	TS = biological adj marker$	Advanced	
16.	TS = biomarker$	Advanced	
17.	TS = algorithm$	Advanced	
18.	TS = non adj invasive	Advanced	
19.	TS = non adj invasive adj test$	Advanced	
20.	TS = cirrhosis	Advanced	
21.	TS = liver adj fibrosis	Advanced	
22.	TS = (predict$ OR prognos$)	Advanced	
23.	TS = (mortality OR prediction OR risk OR follow adj up)	Advanced	
24.	TS = ("prediction and forecasting" OR adverse outcome OR predictive value OR prognosis)	Advanced	
25.	TS = (outcome$ OR treatment adj outcome$)	Advanced	
26.	TS = (disease adj progression OR predictive value of tests)	Advanced	
27.	TS = (course OR mortal$ OR death$)	Advanced	
28.	TS = (cancer OR neoplas$ OR malignan$ OR transplant$)	Advanced	
29.	#19 OR #18 OR #17 OR #16 OR #15 OR #14 OR #13 OR #12 OR #11 OR #10 OR #9 OR #8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1	Advanced	
30.	#21 OR #20	Advanced	
31.	#28 OR #27 OR #26 OR #25 OR #24 OR #23 OR #22	Advanced	
32.	#31 AND #30 AND #29	Advanced	9,653

7.3.5 Process for reviewing articles

I compiled the search strategy, performed the first search, and conducted the first sift of journal articles by title and abstract. The abstracts of the remaining 1,020 articles were then reviewed again by the second reviewer and me independently using Rayyan systematic review manager (225). Articles were selected using my pre-defined inclusion and exclusion criteria (Table 7.3). Reasons for exclusion were documented, and where there was any discrepancy in decision, this was able to be resolved by discussion between us, or by input from a third reviewer when consensus not achieved. The resulting articles were then reviewed by full text independently by the second reviewer and me, and a final list of articles for inclusion was created (Figure 7.1).

7.3.6 Selection criteria

The full list of inclusion and exclusion criteria is presented in Table 7.3. All levels of evidence were included apart from descriptive review articles and opinion pieces. Non-human and pre-clinical studies were excluded. No restriction was made on language. Grey literature (conference abstracts and unpublished work) was not excluded, in line with Cochrane guidance (331). Due to the paucity of prognostic biomarker data on ArLD, relevant studies on chronic liver disease in general were included, as long as they incorporated at least 10 patients where the primary aetiology was alcohol, and that these alcohol data could be extracted separately. Studies were required to have reported either relative risk (RR), hazard ratio (HR) or Area Under the Receiver Operating Characteristic curve (AUROC), with

corresponding confidence intervals (CI), for data extraction in order to address prognosis.

Table 7.3: Selection criteria

Inclusion criteria	**Exclusion Criteria:**
All adult humans (age 18+)	Review articles and opinion pieces
Participants have ArLD	Non-human studies
Where studies investigate chronic liver disease of mixed aetiology, they are only to be included if they comprise ≥ 10 patients where the primary aetiology is alcohol, and that these alcohol data are able to be extracted separately.	Pre-clinical and biological studies
Study relates to at least one of the four non-invasive tests of interest (FIB4, ELF, FibroTest, FibroScan)	Aetiology of liver disease other than alcohol
RR, HR or AUROC with corresponding 95% CI must be able to be extracted from the data	Alcoholic hepatitis

RR: Relative Risk, HR: Hazard Ratio, AUROC: Area Under Receiver Operating Characteristic curve, CI: Confidence Interval

7.3.7 DATA EXTRACTION STRATEGY

Data extraction was undertaken by both second reviewer and me independently using a pre-defined data-entry form. Any disagreements were resolved through discussion between ourselves, or with a third reviewer if persisting uncertainty. Information collected included journal title, year, type of publication, type of non-invasive test investigated, number of patients in study cohort, number of patients with alcohol as primary aetiology, patient demographics, alcohol consumption data, statistical methods used and test performance characteristics (See Tables 7.5 and

7.6). Where studies had included a comparison of non-invasive test with histology, this was recorded and evaluated.

If more than one publication included data from the same cohort of patients, the data from the most recent and comprehensive report were included to avoid duplication in line with Cochrane methodology.

Where data were not clear, or where data were reported for chronic liver disease in general but not specifically for those patients with alcohol as the aetiology, I contacted authors by email to request clarification or access to their original data. If the author failed to reply, the study was excluded.

7.3.8 Quality assessment

The second reviewer and I assessed the quality of the included prognostic studies independently using the QUIPS (Quality In Prognosis Studies) tool (332).

This allowed grading of each publication for risk of bias as being at “low”, “medium” or “high” based on six domains: study participation, study attrition, prognostic factor measurement, outcome measurement, study confounding and statistical analysis and reporting. Disagreements between us were resolved through discussion. This process identified the need to clarify the prognostic factor measurement. Originally a ‘low’ risk of bias was assigned only to those studies in which the NIT failed in less than 10%. After discussion it was recognised that blood biomarkers have a result reporting success rate of 100% and this failure rate only applied to studies using FibroScan. The QUIPS tool can be found in Table 7.4.

7.3.9 Data synthesis and analysis

Prognostic outcomes were reported as HR or AUROC with 95% confidence intervals. Where RR were reported, these were taken to be equivalent to HR. Due to the heterogeneity and small number of final included studies, I adopted a descriptive approach to analyse the results.

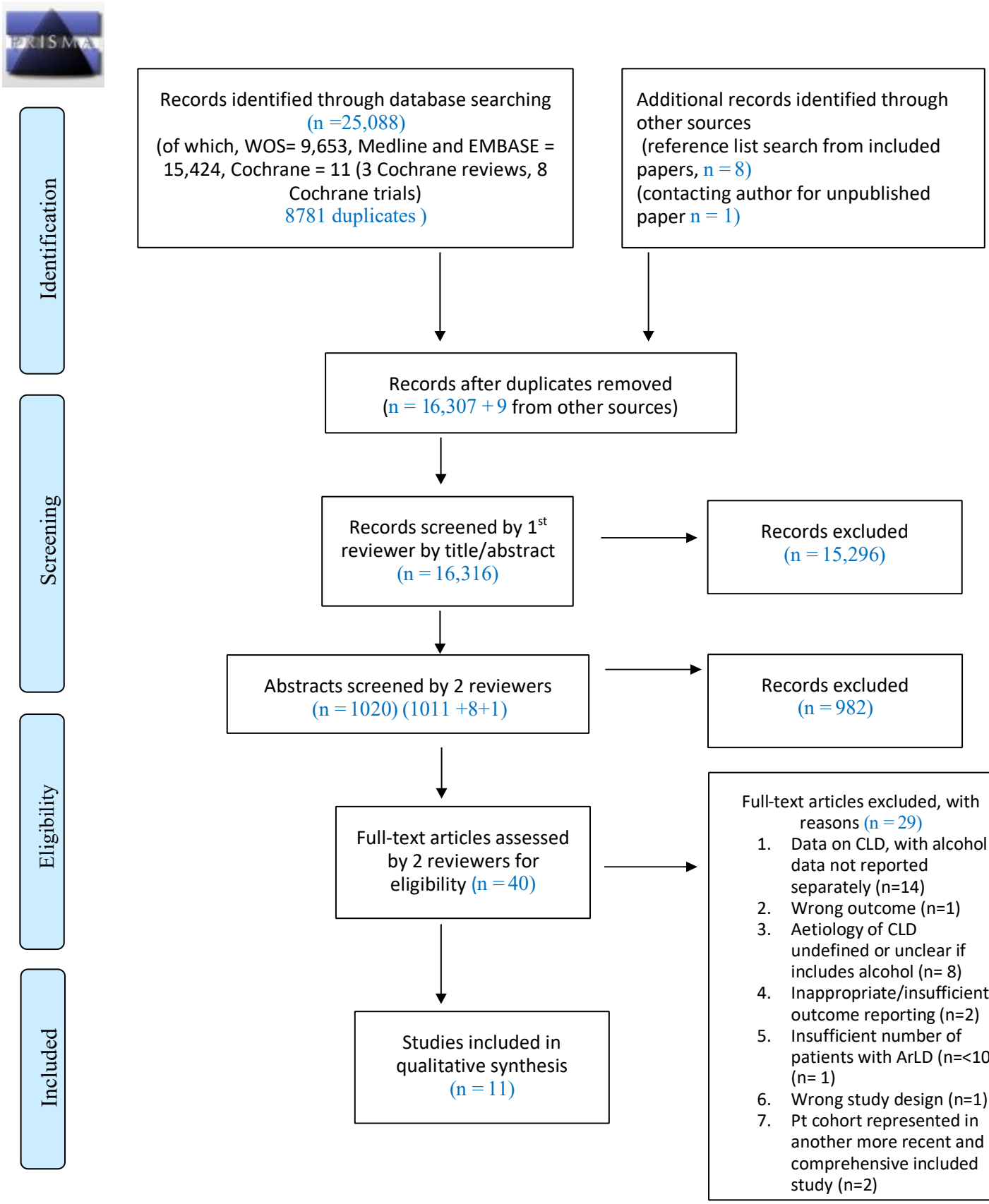
Table 7.4 Summary of the Bias Domains, Rating of "Risk of bias" and Prompting Items and Considerations using an adapted QUIPS Tool, based on the following references: (332-334)

Bias Domains	Rating of "Risk of bias"	Prompting Items and Considerations
1.Study participation	Low	Clearly defined the sampling frame, period and place of recruitment, description of population of interest, as well as baseline study sample; ensured adequate participation of eligible subjects; and clearly reported inclusion and exclusion criteria
	Moderate	All of the earlier-described criteria were met except insufficient description of inclusion and exclusion criteria
	High	The study failed to clearly define the sampling frame, period and place of recruitment; there was an inadequate description of the population of interest, as well as the baseline study sample; was not able to confirm adequate participation of eligible subjects, and did not report inclusion and exclusion criteria
2. Study Attrition	Low	Reported a 100% follow-up rate or less than a 20% attrition rate at the end of the study, or in case of more than 20% attrition a clear statement that patients compliant with follow-up evaluation were not significantly different from those lost to follow-up evaluation
	Moderate	Did not report any attrition rate or an attrition rate of more than 20% but with no description of any systematic differences between those followed up and those lost to follow-up evaluation

	High	The attrition rate was higher than 20% with reported systematic differences between those followed up and those lost to follow-up evaluation
3. Prognostic Factor Measurement	Low	For FibroScan studies: clearly described elastographic technique using a valid and reliable method, and failure rate of test occurred in less than 10% of the sample. For all other prognostic marker studies- used appropriate cut-off values from previous experience or published literature.
	Moderate	Where applicable, the failure rate was not reported or was between 10% and 25% Or cut off values not clearly defined or referenced
	High	Where applicable, failure rates of greater than 25%
4.Outcome Measurement	Low	Clearly and appropriately defined liver-related events (hepatic decompensation based on ascites, variceal bleeding, hepatic encephalopathy, need for liver transplantation, HCC; OR liver-related mortality or all-cause mortality based on medical record review), and used a valid and reliable method of ascertainment
	Moderate	Inappropriately reported the presence of oesophageal varices or the development of sepsis as suggestive of hepatic decompensation
	High	No clear report of which outcomes were measured or how they were measured
5. Study Confounding	Low	Clearly defined and adequately measured relevant confounders, in particular, markers of hepatic synthetic function such as MELD or its components, Child–Pugh score, as well as type of treatment for cohort members

	Moderate	Adjusted for at least 3 other confounding variables, but not including markers of hepatic synthetic dysfunction
	High	The reported adjusted analysis not clearly described (undefined confounding variables) OR the study adjusted for fewer than 3 confounding variables OR the study reported unadjusted analysis
6. Statistical Analysis and Reporting	Low	Performed a multivariable Cox proportional hazard model without overfitting, reporting hazard ratios with corresponding confidence intervals
	Moderate	Reported a multivariable regression analysis instead of a time to event analysis, or reports AUROC as only measure of prediction of mortality
	High	Just reported a univariate analysis or if there was selective reporting of results, or reported hazard ratios or AUROCs without corresponding confidence intervals

FIGURE 7.1: PRISMA FLOW DIAGRAM ('Preferred Reporting Items for Systematic Reviews and Meta-Analyses')



7.4 Results

7.4.1 Study Selection

Searching the four databases returned 25,088 results, of which 8,781 were duplicates (detected and removed using Endnote). An additional 8 results were found by searching reference lists of included papers and relevant review articles. Three of the results (two conference abstracts (40, 233) and one full paper (335)) reported data from patients with ArLD from the same patient cohort. The full paper (335) did not detail the prognostic performance of ELF in the ArLD cohort separately from the mixed-aetiology liver patients, and so was excluded. The senior author of the most recent abstract (40) was contacted, and permission gained to use unpublished data from this abstract that had been written up as a manuscript under review for publication (236). As this was the most recent and comprehensive of the articles, this was included even though it was not yet published at the time of conducting the review, and the older conference abstract (233) reporting the same cohort was excluded. Since this time, the manuscript has now been published (236). I found several articles investigating the prognostic performance of non-invasive tests in mixed aetiology chronic liver disease that did not separately specify the performance in ArLD patients, and therefore I excluded these studies from this review (336-358). One systematic review (359) was excluded as it was also investigating prognostic performance of FibroScan in mixed aetiology liver disease – and although it referenced studies which included patients with ArLD (358, 360-363), these studies either had unrelated outcomes, the sample size of the alcohol

patients was less than n=10, or they did not report data on patients with ArLD separately.

This resulted in 16,316 articles that I screened by title or abstract, with 15,296 being excluded, leaving 1,020 publications for review of abstract by the second reviewer and me independently. The full texts of the 40 selected articles were assessed for eligibility by us both independently, and, after resolving discrepancies, 11 articles remained for inclusion in the data analysis. This comprised 5 full-text published papers and a further 6 conference abstracts (see Figure 7.1). I found no systematic reviews that specifically reported the prognostic performance in ArLD of any of the selected four non-invasive tests.

7.4.2 Study characteristics

Of the eleven studies included, four evaluated FIB4 (three full papers, one abstract), four FibroScan (all abstracts), one evaluated both FIB4 and FibroScan (abstract), one study evaluated both FIB4 and FibroTest (full paper), and one study evaluated ELF (full paper). The total number of patients with ArLD included in the analyses of these eleven studies was 20,412, with a median number of participants of 218 (range 64 - 17,300). Seven studies were prospective, two were retrospective and two were unspecified. The general characteristics for each study, with references, are detailed in Tables 7.5 and 7.6.

Studies were conducted between 2009 and 2019. The median age of participants was 48.5 (range 41.6-60), and 74% were male (range 62.9-100%). The median background prevalence of cirrhosis or advanced fibrosis was heterogeneous, with one study excluding patients with known cirrhosis from the outset (recruiting

patients with fatty liver on imaging and a significant alcohol history), six reporting 100% with cirrhosis, one with 31% cirrhosis (biopsy-proven) and one with 77.8% cirrhosis (biopsy proven), and a further two did not specify.

Outcomes were also heterogeneous, and included liver-related events (LRE), development of HCC, index variceal bleed, liver-related mortality (LRM) and all-cause mortality (ACM). Eight of the eleven articles exclusively investigated patients with alcohol-related liver disease, and three investigated people with chronic liver disease of mixed aetiology, but included details of sub-group analyses, specifying results for patients with alcohol-related liver disease within their cohorts.

The significant heterogeneity of these studies precluded meta-analysis or pooling of results.

Table 7.5: Baseline characteristics of included studies

Study author, year, location (reference)	Publication type Retrospective, prospective, or retro-to-prospective	Aetiology	Time period; Median follow up period (IQR)	Total no pts in study	Total no pts with ArLD included	NIT of interest investigated	Additional prognostic marker assessed	Outcomes assessed (mortality or LRE)	Time point for recorded outcome	Statistical analysis (HR/RR/ AUROC + CI)
Chang et al 2019, South Korea(364)	Full paper Retro-to-prospective	mixed	2002-2015 (5.2 years, IQR 2.8-8.8)	437,828	17,300	FIB4	APRI	Liver related mortality	End of study period (14 years)	HR + 95%CI
Chaudhari et al 2017, India(365)	Abstract Not stated	Alcohol only	From Jan 2015 7.5 months (5-21)	158	158	FIB4	FIBRO-Q, MELD, APRI, AST:ALT ratio	Mortality (unspecified)	unspecified	AUROC + 95% CI
Raker et al 2016, UK(366)	Abstract Retrospective	mixed	2008-2014 26 months (max 83.6)	408	98	FibroScan	-	All-cause mortality	3 years	AUROC, HR + CI
Bertrais et al 2012, France(367)	Abstract Prospective	Alcohol only	2004-2009 3.4 years (no IQR)	302	302	FibroScan	FibroScan, Fibrometer, Hepascore, CP, Quanti-meter	All-cause mortality and liver-related mortality	1 year	AUROC +95% CI

Study author, year, location (reference)	Publication type Retrospective, prospective, or retro-to-prospective	Aetiology	Time period; Median follow up period (IQR)	Total no pts in study	Total no pts with ArLD included	NIT of interest investigated	Additional prognostic marker assessed	Outcomes assessed (mortality or LRE)	Time point for recorded outcome	Statistical analysis (HR/RR/ AUROC + CI)
Mueller et al 2019, Germany(368)	**Retrospective, prospective, or retro-to-prospective** Prospective	Alcohol only	**Median follow up period (IQR)** 3.7 years (mean)	943	675	FibroScan	Albumin, bilirubin, ALP, Hb	All-cause mortality	1,3,5 years	AUROC, HR +95% CI
Gomez et al 2018, Spain(369)	Abstract Prospective	Alcohol only	Not specified 29.2 years (mean, SD 17.3)	276	276	FibroScan	Child Pugh Score, AST, ALT, platelet count	Liver related event	'outcomes during mean follow-up of 29.2 months (SD 17.3)'	AUROC, OR +95% CI
Hyun Kim et al 2018, South Korea(370)	Full paper Retrospective	Alcohol only	2007-2015 58 months (IQR 31-94)	924	924	FIB4	Modified FIB4, APRI, eLIFT score	Development of HCC	3 years	AUROC +95% CI
Cho E et al 2013, South Korea(39)	Abstract Not specified	Alcohol only	Not specified Not specified	195	195	FIB4, FibroScan	APRI, Child Pugh score	Liver-related death and all-cause death	Not specified	AUROC

Study author, year, location (reference)	Publication type Retrospective, prospective, or retro-to-prospective	Aetiology	Time period; Median follow up period (IQR)	Total no pts in study	Total no pts with ArLD included	NIT of interest investigated	Additional prognostic marker assessed	Outcomes assessed (mortality or LRE)	Time point for recorded outcome	Statistical analysis (HR/RR/ AUROC + CI)
Naveau et al 2009, France(148)	**Retrospective, prospective, or retro-to-prospective** Retro-to-prospective	Alcohol only	**Median follow up period (IQR)** 8.2 years (range 5 days to 11.8 years)	218 change to 292	218	FibroTest, FIB4	Fibrometer, Hepascore, APRI, Forns'	Liver-related death and all-cause death	Survival at 5 and 10 years	AUROC + 95% CI
Connoley et al, UK, 2021 (236)	Full paper Prospective	Mixed	Not specified 6.4 years (IQR 2.8-8.5)	786	64	ELF	Liver biopsy	LRE and all-cause mortality	6 years	HR +95% CI
Kothari et al 2019 (371)	Full paper Prospective	Alcohol only	2016-2017 6 months	202	202	FIB4	APRI, MELD, Child Pugh	Variceal bleed	6 months	AUROC + 95% CI

ALT: Alanine Aminotransferase, AST: Aspartate transaminase, ALP: Alkaline Phosphatase, CP: Child Pugh, OR: Odds Ratio, LRE: Liver-related event, SD: Standard deviation, Hb: Haemoglobin, APRI: AST-to-platelet ratio index, MELD: Model for end-stage liver disease

Table 7.6: Baseline participant characteristics

Study author, year, location (reference)	Recruitment details (where reported)	Alcohol consumption required for inclusion	Age year mean (SD)	%male	BMI (SD)	ALT (IQR)	% cirrhosis (advanced fibrosis)	Number of events Event rate (incidence rate per 1000 person years)					
								Liver-related mortality	All-cause mortality	HCC	ascites	HE	Variceal bleed
Chang et al 2019(364)	Pt cohort nested in existing multicentre health study. Included pts: those attending employment-related screening clinics with either NAFLD/AFLD based on US and alcohol history. Excluded pts: with evidence of cirrhosis at start of study.	≥30g/day men ≥ 20g/day women	41.6 (9.3)	94.4	26.5 (2.9)	33 (24-48)	0	19	-	-	-	-	-

Study author, year, location (reference)	Recruitment details (where reported)	Alcohol consumption required for inclusion	Age year mean (SD)	%male	BMI (SD)	ALT (IQR)	% cirrhosis (advanced fibrosis)	Number of events Event rate (incidence rate per 1000 person years)					
								Liver-related mortality	All-cause mortality	HCC	ascites	HE	Variceal bleed
Chaudhari et al 2017(365)	Inpatients with decompensated alcohol-related cirrhosis	-	43.86 (9.03)	-	-	-	100 % cirrhosis		12†	-	-	-	-
Raker et al 2016 (366)	Inpatients with compensated cirrhosis or advanced fibrosis of mixed aetiology in one UK hospital	-	53.5‡	63‡	-	-	100% either advanced fibrosis or cirrhosis		41 (3yrs) (3%, 6%, 10% at 1, 2, and 3 yrs.)‡ -	-	-	-	-
Bertrais et al 2012(367)	Patients with alcohol-related cirrhosis with no history of HCC	-	60	69.9	-	-	100% cirrhosis	50 -	91 -	-	-	-	-
Mueller et al 2019,(368)	Caucasian heavy drinkers, presenting for alcohol detox (6 days)	presented for alcohol detox) (Mean consumption 178 g/d)	-	-	-	-	-	16 -	106 -	-	-	-	-

Study author, year, location (reference)	Recruitment details (where reported)	Alcohol consumption required for inclusion	Age year mean (SD)	%male	BMI (SD)	ALT (IQR)	% cirrhosis (advanced fibrosis)	Number of events Event rate (incidence rate per 1000 person years)					
								Liver-related mortality	All-cause mortality	HCC	ascites	HE	Variceal bleed
Gomez et al 2018, Spain (369)	Patients with Child Pugh A/B alcohol-related cirrhosis without HCC or decompensation at time of enrolment	-	56.5 (8.4)	82	-	-	100% (of which 80% child A, 20% child B)	-	-	13	29	14	17
Hyun et al 2018(370)	Inpatients and outpatients with alcohol-related cirrhosis, excluding 'active alcoholism' and excluding decompensation or HCC at enrolment	Alcohol >10yrs, >60g/day for males, 40g/day females, but no alcohol for past 2 yrs	59	62.9	-	19	100%	-	-	- -	-	-	-
Cho E et al 2013 (39)	'patients with alcohol related liver disease'	-	-	-	-	-	-	-					

Study author, year, location (reference)	Recruitment details (where reported)	Alcohol consumption required for inclusion	Age year mean (SD)	%male	BMI (SD)	ALT (IQR)	% cirrhosis (advanced fibrosis)	Number of events Event rate (incidence rate per 1000 person years)					
								Liver-related mortality	All-cause mortality	HCC	ascites	HE	Variceal bleed
Naveau et al 2009 (148)	'patients with heavy alcohol consumption and available liver biopsy and FibroTest results	Patients had to have consumed at least 50g of alcohol per day over past year	47 (0.7)	78	-	65 (SD 5)	31% cirrhosis (biopsy)	42 -	85 -	7	-	-	4
Connoley et al 2021 (236)	Patients aged between 18-75 yrs undergoing a planned liver biopsy	-	50 (IQR 41.5-57.5)	67.9	-	36 (23-66)	77.8% ≥F3, 66.7% ≥F5	32 at 6 yrs, 34 at 7 yrs, 35 at 8 yrs	23 at 6 yrs, 26 at 7 yrs, 26 at 8 years	-	-	-	-
Kothari et al 2019 (371)	Male aged 18-70 with clinical diagnosis of Alcohol-related cirrhosis, absence of TIPS/previous variceal bleed	'clinically significant alcohol intake'	43.77 (9.95)	100	-	29 (21-50)	100% (clinical/ imaging-based diagnosis)	-	-	-	-	-	61

† Presumed all-cause mortality not liver-related, but not actually specified in study. ‡ Of whole study cohort. (Data on this not reported for alcohol cohort separately)
HE: Hepatic Encephalopathy, HCC: Hepatocellular carcinoma, RR: relative risk, HR: Hazard ratio, BMI: body mass index, ALT: alanine aminotransferase, NAFLD: non-alcoholic fatty liver disease, AFLD: alcoholic fatty liver disease

Table 7.7: Prognostic performance of non-invasive tests (NIT) in each study

Study	Cut off or continuous NIT value	outcome	AUROC	95% CI		HR	95% CI		Other analysis	Adjustment for confounding factors
				Lower limit	Upper limit		Lower limit	Upper limit		
FIB4										
Chang et al (364)	Low FIB 4 = <1.3 Intermediate FIB4 = 1.3 to <2.67 High FIB4 = ≥ 2.67	LRM (14 years)	- - -	- - -	- - -	1.14 4.48 32.9	0.34 1.91 15.04	3.85 10.5 71.96		Yes (sex, yr of screening exam, center, BMI, smoking status, regular exercise, educational level, diabetes, HTN)
Chaudhary et al (365)	Not stated	Mortality (unspecified if LRM, unspecified time point)	0.825	0.71	0.93	-	-	-		Not stated
Hyun et al 2018 (370)	Continuous FIB4 and cut offs – low FIB4 ≤3.25, high FIB4 >3.25	HCC (3 years)	0.69	0.63	0.75	-	-	-	Fib4 high versus low HR 1.71 (95% CI 1.08-2.71)	Yes (age, albumin, platelets, modified FIB4, APRI, eLIFT score)
Cho E et al 2013 (39)	Continuous FIB4	LRM (unspecified time point)	0.78	-	-	1.11	-	-	-	Yes (age)
Naveau et al (148)	Continuous FIB4	ACM (unspecified time point)	0.64	0.55	0.71	-	-	-	-	No
Kothari et al 2019 (371)	Continuous FIB4	Index variceal bleed (6 months)	0.74	0.66	0.81	-	-	-	-	No

Study	Cut off or continuous NIT value	outcome	AUROC	95% CI		HR	95% CI		Other analysis	Adjustment for confounding factors
				Lower limit	Upper limit		Lower limit	Upper limit		
FibroScan										
Cho E et al 2013 (39)	Continuous liver stiffness (kPa)	LRM (unspecified time point)	0.73	-	-	-	-	-	-	Yes (Age)
Raker et al (366)	Continuous LSM + Threshold of <20kpa vs >20kpa	ACM (3 yrs)	0.74	-	-	-†	-	-	Highest vs lowest – 3% incidence of death with LSM <20kpa, VS 15% deaths with LSM >20kpa (p = <0.004)	No
Bertrais et al (367)	Continuous LSM	ACM (1-yr) LRM (1yr)	0.65 0.73	0.51 0.64	0.79 0.83	-	-	-	-	no
Mueller et al (368)	Continuous LSM	ACM (1-yr) ACM (3-yr) ACM (5-yr)	0.76 0.74 0.73	-	-	1.013	1.003	1.023	-	Yes (Age, ALP, albumin)
Gomez et al (369)	Continuous LSM +threshold of <25kpa vs >25kpa	LRE (during follow up period) (unspecified time point)	0.675	0.607	0.743	-	-	-	Highest vs lowest (TE <25kPa = mean incidence of 4.5% for LRE versus 15.5% for >25kPa	Yes (sex, Child Pugh score)

Study	Cut off or continuous NIT value	outcome	AUROC	95% CI		HR	95% CI		Other analysis	Adjustment for confounding factors
				Lower limit	Upper limit		Lower limit	Upper limit		
FibroTest										
Naveau et al (148)	Continuous FibroTest value Continuous FibroTest value	LRM (unspecified time point) ACM (unspecified time point)	0.79 0.69	0.68 0.61	0.86 0.76	23.2‡ 3.7‡	3.2 1.2	167.3 11.7		Yes (liver biopsy, fibrometer, hepascore, abstinent vs non-abstinent)
	FibroTest cut offs:								5-yr SNLRD	
	0-0.31 (no or minimal fibrosis)	SNLRD (5-yr)							98.7% (96-1)	
	0.21-0.58 (moderate fibrosis)	SNLRD (5-yr)							92.1% (83.5-100)	
	0.59-1 (severe fibrosis)	SNLRD (5-yr)							68.3% (79.5-89.4)	
	FibroTest cut offs:								10-yr SNLRD	
	0-0.31 (no or minimal fibrosis)	SNLRD (10-yr)							92% (84.9-99)	
	0.21-0.58 (moderate fibrosis)	SNLRD (10-yr)							87.5% (75.5-99.5)	
	0.59-1 (severe fibrosis)	SNLRD (10-yr)							78.5% (72.4-84.6)	

Study	Cut off or continuous NIT value	outcome	AUROC	95% CI		HR	95% CI		Other analysis	Adjustment for confounding factors
				Lower limit	Upper limit		Lower limit	Upper limit		
ELF										
Connoley et al 2021 (236)	ELF as continuous	LRE (6-yr)	-	-	-	1.82§	1.169	2.83	-	Yes (age and sex)
	ELF as continuous	LRE (6-yr)	0.816	0.713	0.920	-	-	-	-	no
	ELF as continuous	LRE (7-yr)	0.844	0.750	0.938					no
	ELF as continuous	LRE (8-yr)	0.847	0.754	0.940					no
	ELF as continuous	ACM (6-yr)	0.733	0.645	0.861	-	-	-	-	no
	ELF as continuous	ACM (7-yr)	0.722	0.591	0.852	-	-	-	-	no
	ELF as continuous	ACM (8-yr)	0.722	0.591	0.852	-	-	-	-	no
	ELF cut offs in 4 categories (compared to <9.8)		-	-	-				-	Yes (age and sex)
	9.8-10.49	LRE (6-yr)				1.49	0.287	7.74		
	10.5-11.29	LRE (6-yr)				3.84	0.9	16.39		
	≥11.3	LRE (6-yr)				10.24	2.97	35.27		
	ELF cut offs in two categories		-	-	-				-	Yes (age and sex)
	<10.5 and ≥10.5	LRE (6-yr)				6.42	2.63	15.24		

† Study reported HR with corresponding 95% CI for LSM cut offs above and below 20kPa, and for ArLD versus non-ArLD, but not specifically for LSM in the ArLD cohort.

‡ RR (risk ratio) § OR (odds ratio). LRM = Liver Related Mortality, LRE: Liver related event, ACM: all-cause mortality, SNLRD: Survival or Non-liver related Death, LSM: Liver Stiffness Measurement, HTN: hypertension, TE: transient elastography

7.4.3 Risk of bias within studies

On review of the six bias domains in the QUIPS tool, the majority of the eleven included studies were assessed to be at low or moderate risk of bias in study attrition, prognostic factor measurement and outcome measurement. However, there were some studies that were at high risk of bias in the other three domains. In the first domain (study participation), 4/11 studies scored 'high risk'. Three of these were conference abstracts, and one was the unpublished manuscript which incorporated pooled data from three patient cohorts, but did not clearly specify the time period for each of the studies.

In the study confounding domain, 6/11 (55%) of the studies scored 'high risk of bias', due either to not defining the confounding variables, adjusting for fewer than three confounding variables, or reporting an unadjusted analysis. Four out of six of these articles were conference abstracts and so may have omitted this information because of restrictions on word count. The other two (40, 371) had either only adjusted the analysis for two variables or no adjustment was documented, and thus were both graded as being at high risk of bias using the QUIPS tool. In the final domain 'statistical analysis and reporting', three of the eleven articles were graded as being at high risk because they either did not report multivariable analysis, or they reported HRs or AUROCs without corresponding confidence intervals. These three were all conference abstracts.

Overall, 78.8% of the six domains across ten studies were rated 'Low' or 'Moderate' risk of bias (See Figure 7.2 and Table 7.8). Cohen's kappa (κ) was measured to

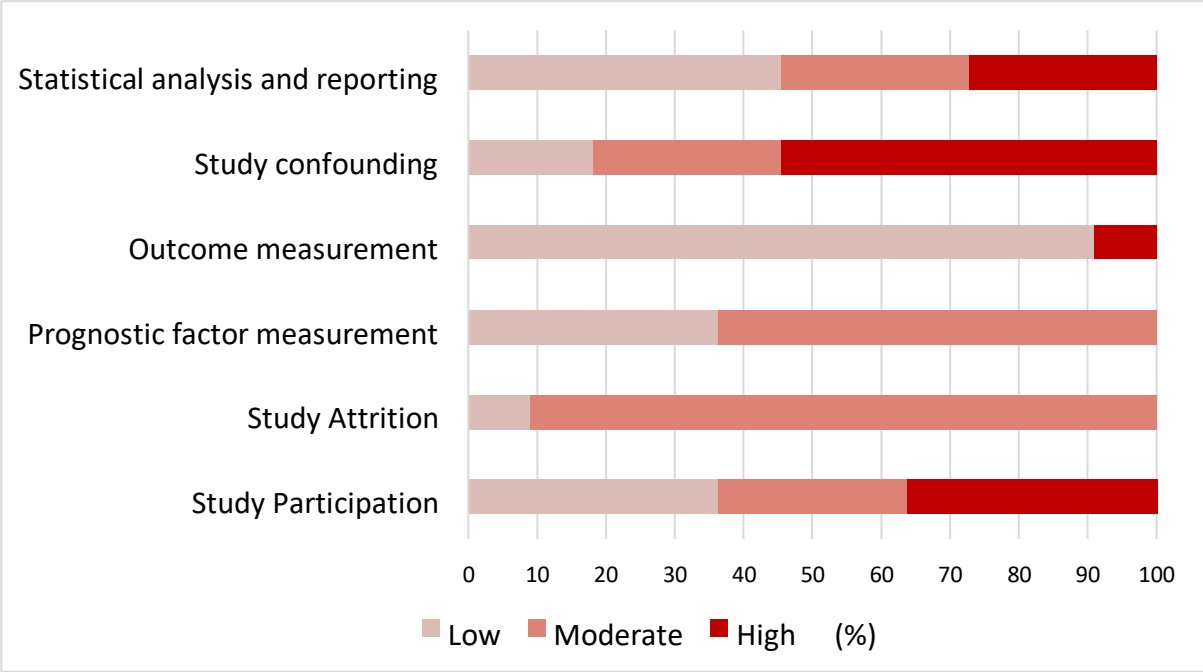
determine if there was agreement between myself and the second reviewer on the grading of low, moderate and high risk of bias across six domains over the 11 articles. This showed moderate-to-good agreement (259), with 74.2% of all grades (6 domains x 11 papers) being the same between our responses, and k = 0.59 (95%CI, 0.426 to 0.75), p = <0.0005.

Table 7.8: Study–Level Quality Assessment using the Quality In Prognosis Studies (QUIPS) Tool

Study	Study Participation	Study Attrition	Prognostic Factor Measurement	Outcome Measurement	Study Confounding	Statistical Analysis and Reporting
Chang et al 2019	L	M	L	L	M	L
Chaudhari et al 2017	M	M	M	H	H	H
Raker et al, 2016	M	M	M	L	H	L
Bertrais et al, 2012	H	M	M	L	H	M
Hyun et al 2018	L	M	L	L	L	L
Mueller et al, 2019	M	M	M	L	M	H
Gomez et al 2018	H	M	M	L	L	M
Cho et al 2013	H	M	M	L	H	H
Naveau et al 2009	L	L	L	L	M	L
Connoley et al 2021	H	M	L	L	H	L
Kothari et al 2019	L	M	M	L	H	M

L = Low M= Moderate H = High

Figure 7.2: QUIPs tool results: Quality assessment of included studies using the quality in prognosis studies tool.



7.5 Prognostic performance of each of the four non-invasive tests

(See table 7.7, Figure 7.3)

7.5.1 FIB4

Six studies examined the prognostic performance of FIB4 in ArLD, three of which used Liver Related Mortality (LRM) as the outcome. The fourth study used 'development of HCC at 3 years' as the outcome, the fifth used index variceal bleed within 6 months, and the sixth used 'mortality' – which did not specify if liver-related or all-cause.

AUROCS for mortality were recorded in 3 studies, and Hazard ratio (HR) in the other mortality study. AUROCS were 0.64 (95%CI 0.55-0.71), 0.78 (no CI reported) and 0.825 (95%CI 0.71-0.93). The study which reported HR for LRM at 14 years showed a

significant difference in mortality based on FIB4 thresholds, with the higher threshold of >2.67 giving a HR of 32.9 (95%CI 15.04-71.96) compared with a low FIB4 threshold of <1.3 (HR 1.14, 95% CI 0.34-3.85).

Four studies used continuous FIB4 score in their analysis, and two used FIB4 thresholds, which were different in each study. One study identified three categories of FIB4 score: low (1.3), intermediate (1.3-2.67) and high >2.67 and the other used a single threshold categorising results above or below 3.23. The latter study used development of HCC at 3 years as the main outcome, and FIB4 was able to predict this with AUROC of 0.69 (0.63-0.75).

Two studies examined FIB4 along with another non-invasive test (FibroTest in one and FibroScan in the other). Cho et al. (39) found no reported difference between AUROC for FIB4 (AUROC 0.78) and FibroScan (AUROC 0.73) in predicting LRM (although p values were not stated for this comparison). However, when using a multivariable cox proportional hazard model, FIB4 was able to predict LRM (HR 1.11, p= 0.03) but FibroScan was not. FIB4 was also able to independently predict ACM (p = <0.001) whereas FibroScan was not (confidence intervals were not reported).

Naveau et al. (148) compared FIB4 with FibroTest. Although there was a statistically significant difference between the AUROCs for the tests for prediction of liver related death (FibroTest AUROC 0.79 (95%CI 0.68-0.86), FIB4 AUROC 0.65 (95% CI 0.54-0.74); p=0.004) there was no significant difference in AUROCS for predicting overall survival (FibroTest AUROC 0.69 (95%CI 0.61-0.76), FIB4 AUROC 0.64 (0.55-0.71); p = 0.20).

7.5.2 FibroScan

Five studies investigated the prognostic performance of FibroScan. One of them (39) as described above, compared FibroScan (continuous liver stiffness measurement) with FIB4, and found that FIB4 was able to predict mortality in multivariable cox proportional hazard analysis, but FibroScan was not, although AUROCS were not significantly different between FibroScan (0.73) and FIB4 (0.78).

Four other studies (366-369) reported AUROCs for predicting mortality or liver-related events using continuous liver stiffness measurements, with AUROCs of 0.65, 0.675, 0.7 and 0.76. Two of these four studies (366, 369) also reported liver stiffness thresholds, with one using a threshold of 25kPa, finding a significant difference in mean incidence of LRE of 4.5% in the <25kPa cohort compared with 15.5% in the >25kpa cohort (369). The other study (366) reported a liver stiffness threshold of 20kPa, with incidence of death at 3% in patients with liver stiffness measurement (LSM) <20kPa, compared to 15% deaths in those with LSM >20kPa (p = <0.004).

7.5.3 FibroTest

Only one study reported the prognostic performance of FibroTest in ArLD (148). This study of 218 people with ArLD compared FibroTest with liver biopsy, Hepascore, Fibrometer, FIB4, APRI and Forns' Index in predicting LRM and ACM. FibroTest performed better than FIB4 (p= 0.004) in predicting LRM (FibroTest AUROC 0.79 (95% CI 0.68-0.86) compared to FIB4 AUROC of 0.65 (95% CI 0.54-0.74) (although there was

no difference between FIB4 and FibroTest in predicting ACM). When compared with other markers of fibrosis in this study, the prognostic values of FibroTest (AUROC 0.79 $\pm$ 0.04 for survival or non-liver disease related death), Hepascore (0.78 $\pm$ 0.04), and Fibrometer (0.80 $\pm$ 0.04) did not differ from that of liver biopsy fibrosis staging (0.77 $\pm$ 0.04). In multivariate analysis, they found that the best performing tests were FibroTest (p = 0.004) and liver biopsy (p = 0.03).

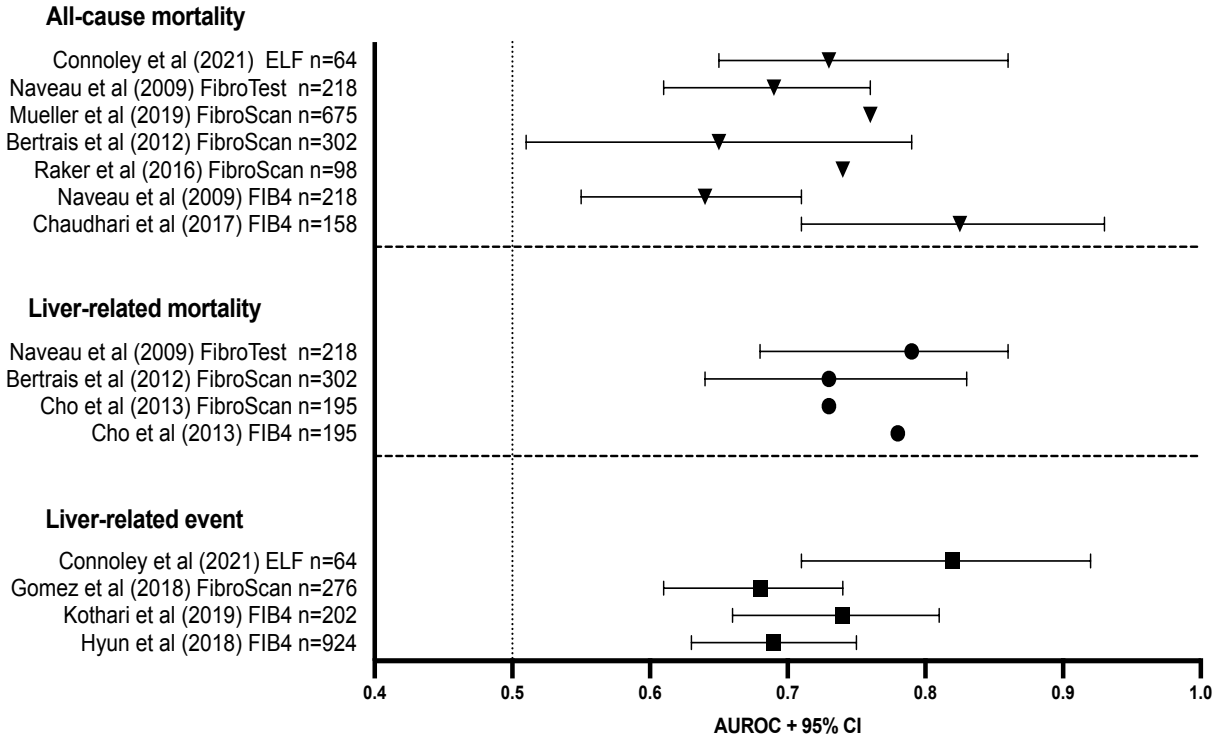
7.5.4 ELF

I found only one study (236) that investigated the prognostic performance of the ELF test in ArLD. Data from this study have been published by the same authors as a conference abstract (40).

This study comprised 64 people with ArLD taken from three different study cohorts with a total sample size (of mixed aetiology liver disease) of 786.

ELF was analysed both as a continuous value and using thresholds of 9.8, 10.5 and 11.3. When analysed as a continuous value, ELF predicted LRE at 6, 7 and 8 years with AUROCs of 0.816, 0.844 and 0.847 respectively. Risk ratio for the prediction of LRE at 6 years was 1.82 (1.169-2.83). ELF also predicted All-Cause Mortality (ACM) at 6,7 and 8 years with AUROCs of 0.733, 0.722 and 0.722 respectively. When analysed using a cox proportional hazard model (adjusted for age and gender), Connoley et al. found that each unit increase in ELF was associated with a 1.44 times increased risk of LRE (95% CI 1.25-1.66, p <0.001). When analysed, the HR for ELF scores between 10.5-11.29 was 3.84 (95% CI 0.90-16.39), HR for ELF $\geq$11.3 was 10.24 (95% CI 2.97-35.27), compared to a low ELF threshold of <9.8 where HR was 1.49 (95% CI 0.287-7.74).

Figure 7.3: Forest plot of AUROCs + 95% CI for outcome prediction*



* Chang et al 2019 study on FIB4 not represented in this forest plot as study did not report AUROCs, however, this study had largest sample size of 17,300. HR for liver-related mortality at 14 years = Low FIB4 threshold <1.3: HR 1.14 (95% CI 0.34-3.85), Intermediate FIB4 1.3-2.67: HR 4.48 (95% CI 1.91-10.5), High FIB4 ⍰ 2.67: HR 32.9 (95% CI 15.04-71.96).

7.6 Performance of non-invasive tests compared with histology and other prognostic scores:

Only 2 out of 11 studies were biopsy-paired (148, 236). Whilst Naveau et al found that FibroTest performed equally as well as histology, (148) Connoley et al found that ELF was superior to histology in predicting prognosis in ArLD (ELF AUROC for all-cause mortality 0.733 (95% CI:0.645-0.861), compared to 0.600 (95% CI:0.470-0.730) for liver biopsy, p = <0.05) (236).

Where one of the four NITs of interest were directly compared with other more traditional prognostic scores such as MELD and Child Pugh (CTP), FibroTest outperformed CTP (FibroTest AUROC for survival 0.79 (95%CI 0.68-0.86), CTP AUROC 0.69 (95%CI 0.58-0.77), p=0.02) (148), and FIB4 outperformed MELD in two separate studies: In Chaudhari et al.'s study, FIB4 AUROC for mortality was 0.83 (95%CI 0.71-0.93), MELD 0.70 (0.53-0.87) p=0.001 (365), and in Kothari et al.'s study, FIB4 AUROC for predicting variceal bleed was 0.74 (95%CI 0.66-0.81), MELD AUROC 0.54 (95%CI 0.46-0.62) (372).

Whilst MELD and CTP are very much still a part of current clinical practice for prognosticating in CLD, these findings suggest that non-invasive fibrosis markers may be a better choice in predicting prognosis in ArLD.

7.7 Discussion

7.7.1 Main findings

Whilst there is now good evidence for the use of non-invasive tests for prognosticating in chronic liver disease of mixed aetiologies (336, 338-358, 373), this systematic review has found fewer studies on ArLD. This corroborates with my findings in Chapter 3 in the systematic review of diagnostic performance of non-invasive tests. This is important, as each aetiology of liver disease behaves differently in terms of pathophysiology, clinical presentation and complications and I have found that there is evidence suggesting that the performance of some non-invasive tests varies with disease aetiology (203, 204, 221, 222, 374, 375). Whilst non-invasive tests are becoming increasingly widely used both for the diagnostic staging of liver fibrosis as well as prognosis, it is imperative that they are evaluated in representative populations with specific aetiologies.

Mortality rates for cirrhosis have increased by 400% in the last thirty years, with alcohol-related liver disease identified as the predominant cause (326). There is thus an international growing recognition of the importance of research in alcohol-related liver disease, and hopefully this will be reflected in forthcoming published works over the next decade. This study has highlighted gaps in current knowledge of commonly used non-invasive liver fibrosis tests as prognostic markers in this condition.

Nevertheless, despite the relative paucity of studies, all the four non-invasive tests investigated in this review show promising prognostic performance, with AUROCS above 0.7 in some studies for each test, and AUROCS above 0.8 for one study of FIB4

(365) and the single ELF study (236). While heterogeneity of the results prevented me from performing statistical comparisons of test performance between studies, there were two studies which directly compared two non-invasive tests. The study which compared FibroTest to FIB4 (148) found a significantly better prognostic performance of FibroTest when compared to FIB4. The other study which directly compared FIB4 with FibroScan (39), found that FIB4 could predict Liver Related Death (HR 1.11; p=0.003) while FibroScan could not.

However, it is not possible to identify a single non-invasive test that performs better than the rest based on this systematic review due to the heterogeneity and lack of data.

The ELF test did appear to be superior to histology in predicting outcomes in one study in ArLD (236), and FibroTest performed equally well to histology (148). FibroTest also outperformed Child Pugh score (148) and FIB4 outperformed MELD in two studies in ArLD (365, 372). Biopsy is not routinely required in the management of ArLD, and the findings from this review indicate that these commonly available non-invasive tests can perform a useful role when predicting prognosis in clinical practice for patients with ArLD, avoiding the need for liver biopsy and may even be superior to more traditional prognostic scores like MELD and Child-Turcotte-Pugh.

7.7.2 Strengths

The strengths of this study are based on a comprehensive search, with rigorous screening of texts. The screening of texts, extraction of data, and assessment of risk of bias were all conducted by myself and a second reviewer independently, to minimise

reporting bias. Where it was not possible to extract data from certain papers, efforts were made to contact authors by email. In addition, reference lists of relevant articles were hand-searched to maximise inclusion for study, and grey literature was included. Where data were found to be using the same patient cohort – only the most recent and comprehensive study was used, to prevent risk of duplication of results.

7.7.3 Limitations

I recognise there are limitations of this study. The heterogeneity and low number of included studies prevented meta-analysis, or pooling of hazard ratios or AUROCS.

For example, of the four studies on FIB4 that all reported AUROCs for mortality, one reported Liver Related Mortality, and the three that reported All-Cause Mortality had different end points (between 1 and 5 years), and two out of these three did not report corresponding confidence intervals. This variety in outcomes and reporting made it impossible to pool results, and is in keeping with existing wider literature on prognostic studies in all fields that recognises the variable quality of prognostic studies. The Cochrane methods group has acknowledged the challenges in systematic reviews on prognosis due to “low quality of primary studies, poor reporting, and difficulties in combining results across different research designs, analyses and presentations of results”(376). Other studies have commented on the barriers to synthesis of prognostic study data ranging from poor reporting, lack of consistency in statistical analysis across prognostic studies and often prognostic model studies are based on relatively small sample sizes leading to overfitting and poor generalisability of

results (377, 378). A clear finding arising from the conduct of this review is the necessity for larger rigorous studies of non-invasive tests in ArLD.

Six out of eleven of these studies were conference abstracts, limiting the data that could be extracted, leading to higher scores on the risk of bias assessment.

As all of the studies reported significant results for the performance of non-invasive tests, it is possible there was publication bias. However, it is difficult to the be certain as the number of publications in this area is so small. The inclusion of grey literature in this systematic review may have reduced publication bias.

7.7.4 Conclusions

In this study I have demonstrated that all four of the examined non-invasive tests (FIB4, FibroTest, ELF, FibroScan) can perform well in predicting prognosis in ArLD with AUROCs >0.7. Of the two included studies that were biopsy-paired, they found that FibroTest and ELF performed as well and better than histology in predicting outcomes, respectively. Whilst the heterogeneity of studies precluded pooling of results and direct comparisons, those studies that did include direct comparisons of non-invasive tests with other 'traditional' prognostic scores showed non-invasive tests for liver fibrosis to be superior to MELD and CTP. With easy availability of FIB4/FibroTest/FibroScan/ELF, it therefore may be preferable to use one of these fibrosis markers when prognosticating in ArLD in situations where biopsy is not necessary. Due to the small number of included studies, further validation studies of these non-invasive tests as prognostic scores are warranted.

Chapter 8

Investigating the prognostic performance of ELF in alcoholic hepatitis

8.1 Abstract

Background

Alcohol-related hepatitis (AH) carries a high mortality - up to 60% in patients with GAHS (Glasgow Alcoholic Hepatitis Score) of ≧ 9. Tests that predict mortality could improve management of critically ill patients with AH. The STOPAH trial team analysed the performance of static scores in predicting mortality in patients with AH, finding these scores performed sub-optimally and recommended further research. The static scores included Maddrey's Discriminant Function (DF), Model for End-stage Liver Disease (MELD), Age-Bilirubin-INR-creatinine (ABIC) and GAHS, with 90-day AUCs of 0.670, 0.704, 0.726 and 0.713 respectively. The ELF test is a good prognostic marker for CLD but has not been studied in AH.

Methods

The individual components of the ELF test and composite scores were measured in a sample of 179 baseline sera from STOPAH trial participants. ELF tests were performed using Siemen's reagents on an Advia Centaur XP. Other biomarkers were calculated from participants' STOPAH data. Prognostic performance of ELF, ELF components, Model for End-stage Liver Disease (MELD), Glasgow Alcoholic Hepatitis Score (GAHS), Age-Bilirubin-INR-Creatinine (ABIC) and Neutrophil to lymphocyte ratio (NLR) for predicting mortality at 28, 90, and 120 days was assessed using logistic regression and Kaplan Meier analysis, and scores compared using log-rank test, and de-long method to compare AUROCs. FIB4 and Lille were examined in a sub-analysis.

Results

All scores were available for 162/179 (96.4%) participants. ELF scores ranged from 11.97-21.68, with median of 15.9 (IQR 14.9-16.9). Median Hyaluronic acid (HA) level was 3971 (IQR 2298-12246). ELF was able to predict mortality at 28, 90 and 120 days, with 90-day OR of 1.7 (95% CI 1.3-2.2, p <0.001). ELF positively correlated with CRP (rho = 0.51, p <0.001) and AST (rho = 0.35, p <0.001), but there was no association between ELF and infection episodes (p = 0.55) or variceal bleeding (p=0.56). Logistic regression revealed that an algorithm combining ELF and ABIC outperformed all other variables (including GAHS, MELD, NLR) in predicting mortality at 90 days. ELF+ABIC as a single marker had the highest numerical AUROC for 90-day mortality prediction at 0.81 (95% CI 0.73-0.89) compared to all the other variables, and significantly so (p =0.01)

when compared with ABIC alone. Patients with both low-ABIC (<6.71) and low ELF score (<16.78) had a 100% 90-day survival rate, compared with 45% survival in those with high ABIC/high ELF ($p < 0.01$).

Those with High ABIC/high ELF and Lille <0.45 had a 90-day survival rate of 90% (n=9/10 survived), compared to 6/13 survival (46% survival) in the group with Lille ≥0.45 ($p = 0.047$). In those with high ABIC and low ELF (n=100), those with baseline sepsis (n =17/100) had a significantly worse outcome (survival rate 59%) compared to those without sepsis (83/100, survival rate 89%), $p = 0.001$.

Conclusion

For prognosis-prediction in AH, this represents the first study to discover a marker with AUROC >0.8 in ABIC-ELF. The addition of Lille allows further differentiation of survival rates in those with high ABIC/ELF. ABIC-ELF could be easily adopted as a clinically useful prognostic score in AH, but first requires further validation in a larger cohort.

8.2 Introduction

Alcohol-related hepatitis (AH) is a severe clinical syndrome characterised by new onset jaundice in patients with ongoing alcohol excess, commonly after a recent period of heavy drinking. Specifically, patients present with a bilirubin >50 μMol/L, an elevated AST (50-400 U/L) and AST:ALT ratio >1.5 with no other cause for hepatitis (120). Risk factors include drinking more than 30-60g/day of alcohol (284), although only 35% of excessive drinkers will develop AH (379) suggesting that other factors such as genetics, (380) microbiome (381), and nutrition (382), also play a role. In addition, the pattern of drinking can influence risk, with binge-drinking and drinking outside of meal-times conferring enhanced risk of AH (379, 383). Women at are greater risk of AH at lower-alcohol intake than men (383-385) and malnutrition is also strongly associated with AH (379, 382).

The pathogenesis of AH is complex, and still being explored (122). It involves multiple processes that result in an inflammatory cascade, leading to steatohepatitis, resulting from both the intra-hepatic and extra-hepatic effects of alcohol. Intra-hepatically, the metabolites of alcohol (including acetaldehyde) directly induce hepatocyte cell death through apoptosis and necrosis, and this hepatic cell injury leads to the release of DAMPs (Damage-associated molecular patterns) (122). These DAMPs bind to TLRs (Toll-like receptors) in Kupffer cells in the liver, leading to an inflammatory cascade (122). The extra-hepatic effects include the effects of alcohol on the intestine – which in turn worsen hepatic failure. This is thought to be through the direct toxic effect of alcohol on intestinal epithelial cells which reduces the expression of 'tight-junction proteins' leading to increased permeability of intestinal mucosa, or 'leaky gut', allowing bacterial translocation across the gut wall, leading to further 'switching on' of

inflammatory cytokines and further inflammation of the liver, ultimately resulting in hepatic failure (122-124). The hallmark morphological features of AH include steatosis, lobular inflammation (usually with heavy neutrophilic infiltration), ballooning of hepatocytes, and necrosis (77). It should, however, be noted that steatosis may only be present in less than 5% of the parenchyma, or even absent in some cases with severe AH, or following periods of abstinence or in cirrhosis (77).

AH carries a high mortality – around 30% at 90 days, with currently no universally effective pharmacotherapeutic option that improves medium or long-term outcomes (125). Nutritional support and abstinence of alcohol are the mainstays of management, with corticosteroids being the only pharmacotherapeutic option currently (122), although trials are ongoing to investigate new therapies (122, 129, 130). However, almost half of patients cannot tolerate or do not respond to steroid therapy, and any positive survival benefit they may have to steroids is lost beyond 28 days (122, 386). Recognising those patients with the poorest prognosis is essential to guiding clinical management. This is becoming even more of a necessity in recent times with options for transplantation for Acute-on-Chronic-Liver Failure (ACLF) and AH increasingly considered in selected patients (387-389). The selection criteria for transplantation in AH needs further refining, but evidence to date suggests that appropriate patients for selection are likely to be those with a) severe disease (currently defined as Maddrey's Discriminant Function (DF) $\geq$ 32) (390), b) Lille score 0.5-0.82 (391), c) lack of co-morbid conditions (390), d) adequate psychosocial support favouring the ongoing abstinence to alcohol post-transplant (387), and with careful consideration of those at higher risk of relapse (younger age, and more than 10 drinks a day prior to transplant) (390). Being able to select out patients with either a very

high chance of survival (and therefore unlikely to need intensive-care management or transplantation) or with very poor prognosis is therefore fundamental. It would be ideal if there was a prognostic score that could even predict those patients with a <10% chance of survival, as they would have a worse outcome without liver transplant, which has a 10% reported mortality rate (392).

Multiple prognostic scores for AH have been developed; the most widely used up until recent years was Maddrey's DF, mentioned in the previous paragraph. However, later studies on AH, including the STOPAH trial (Steroids OR Pentoxifylline in Alcoholic Hepatitis) (41) have shown improved outcomes in the current standard care of alcoholic hepatitis, with improvements in 28-day survival. The knowledge that these improved outcomes will lead to a further reduction in the specificity of DF, led to a further prognostic biomarker study by Forrest, E and the STOPAH trial Management group, et al in 2018 which identified that GAHS, MELD and ABIC all performed better than DF in predicting mortality at 28 and 90 days (and similar to each other) (131), with 90-day AUCs of 0.70, 0.73 and 0.71 and 0.67 respectively. Combining the use of one of these 'static' scores with a 'dynamic' score assessment at day 7 led to an improved (yet still 'modest') ability to prognosticate 90-day outcomes. The authors suggested that those with low baseline 'static' scores could avoid prednisolone treatment, and those with high static scores with favourable day-7 response (dynamic assessment), in the absence of sepsis or Gastro-Intestinal bleeding, could benefit from prednisolone (90-day mortality 19.2% compared with 28.2% standard approach of treating all patients with prednisolone with a DF over 32, (p= 0.026, 95% CI 0.63-14.72)

(131). However, the study concluded that there remains to be a need to predict mortality and response to steroids more accurately with further biomarker research.

Fibrosis-staging is recognised to be an important tool in the prognostication of liver disease, with advanced fibrosis (not clinical or biochemical factors) being the only independent predictor of prognosis in alcohol-related liver disease in a recent study (327). Whilst AH is thought of as an inflammatory condition, it is interesting to note that in a landmark study from 2014 where liver biopsy was used in AH, the severity of fibrosis/presence of cirrhosis was the strongest predictor of mortality, whereas neutrophilic inflammation was associated with improved survival (393). In clinical practice, however, liver biopsy is not routinely performed in AH, and is currently only recommended by international guidelines to be carried out where there the clinical diagnosis of AH is in debate (77). Other non-invasive tests (NIT) for fibrosis have not been validated in AH. FibroScan, the most commonly studied NIT for prognostication in ARLD (224), has been shown to be influenced by inflammation (394-396), cholestasis (397) and alcohol intake/withdrawal (132, 133, 206), and is therefore likely to give inaccurate fibrosis staging readings in AH.

The Enhanced Liver Fibrosis (ELF) test has proven to be a good prognostic marker for Chronic Liver Disease (CLD) (335, 339, 341), with AUROCs >0.7 for mortality-prediction in alcohol-related liver disease (224), with no reported influence of alcohol (23, 274) so far, and undetermined influence of inflammation (23, 274, 398). However, it has not yet been studied in AH.

My aim in this study is to investigate the performance of the ELF test and its components in the prediction of mortality in AH, compared with the static scores GAHS, MELD and ABIC. My hypothesis was that ELF, as a direct marker of liver fibrosis, would carry prognostic information in AH and that either alone or in combination with established static prognostic markers it would outperform the established prognostic models for AH.

8.3 Methods

8.3.1 STOPAH Methodology

I studied participants recruited to the STOPAH trial (41). STOPAH was a multicentre, randomised, double-blinded trial which evaluated the effect of prednisolone and pentoxifylline versus placebo and each other in the management of alcoholic hepatitis. A total of 1092 patients were included in the analysis, with inclusion criteria specifying a clinical diagnosis of AH, average alcohol consumption of >80g per day for men and >60g per day for women, serum bilirubin >80 μmol/L, and DF of 32 or higher. Exclusion criteria were: jaundice for >3 months, absence of alcohol consumption for >2 months prior to randomisation, alternative causes of liver disease, AST >500 Iu/L or ALT >300. Standard supportive care and nutrition were given to each patient. Patients were randomised to receive either prednisolone and placebo or pentoxifylline and placebo or prednisolone and pentoxifylline or double placebo, with analysis on an intention-to-treat basis. The primary endpoint was mortality at 28 days, with secondary end points including 90-day outcomes.

8.3.2 Ethics

ELF was analysed in samples from the STOPAH trial (41) which had ethical approval in place and all patients were consented for the trial (Multicentre Research Ethics Committee reference number 09/MRE09/59). The samples for ELF were analysed as part of the Medical Research Council (MRC) stratified medicine initiative 'MIMAH' (Minimising Mortality from Alcoholic Hepatitis) (130) (reference code MR/R014019/1).

8.3.3 Patients and data collection

A total of 179 serum samples and their associated STOPAH dataset were made available by the STOPAH working group for inclusion in this study. These samples were selected as, of the patients with serum remaining for analysis, these 179 were chosen randomly by computer, and according to the STOPAH trial group comprised a cohort typical of the whole STOPAH cohort (n = 1092). I went on to formally compare characteristics between my cohort of 179 samples and the main STOPAH cohort of 1092, and this is described later in the results section.

8.3.4 Measurement of key outcomes

Mortality was analysed at 28 and 90 days, as in the published STOPAH trial, and for this sub-study, 120-day mortality data were also available and so I included these in the analyses.

8.3.5 Selection of prognostic models

Results for 'static scores' MELD, GAHS, ABIC and NLR were available from the baseline blood results at the time of starting treatment, and the ELF scores were calculated

from the measurement of the constituent analytes (HA, PIIINP, TIMP1) in the corresponding baseline serum samples. I selected these 'static scores' MELD, GAHS, and ABIC for comparison with ELF in this study because they all performed the best in predicting mortality (and equally to each other) in the biomarker study published by the STOPAH team on the main cohort of 1,068 patients with alcoholic hepatitis, outperforming Maddrey's DF (131) with AUROCs of 0.704 (MELD), 0.713 (GAHS), 0.726 (ABIC), and 0.670 (DF) for 90-day mortality. In addition, I included Neutrophil-to-Lymphocyte Ratio (NLR) in the analysis, as this was shown to be predictive of prognostic outcomes in a further published study by the STOPAH trial team (399).

8.3.6 ELF scores

The ELF test was performed on the serum samples which were analysed at the Central ELF laboratory (iQur Limited, London). The samples were analysed for HA, PIIINP and TIMP1 levels using the proprietary assays developed by Siemens Healthineers Inc. (Tarrytown, New York, USA) for the ELF test, on an ADVIA centaur® immunoassay system. Results were entered into the manufacturer's published algorithm for the Adviar Centaur XP ELF test to generate each ELF score by combining PIIINP, HA and TIMP1 results. (Published algorithm: 2.278 + 0.851 ln(HA) + 0.751 ln(P3NP) + 0.394 ln(TIMP1)).

8.3.7 Outcome measures

The primary outcome measure was the performance of ELF, in comparison to and in combination with ABIC, GAHS, and MELD, in the prediction of survival at 90 days.

Secondary outcomes included survival at 28 and 120 days, investigating for correlation between ELF and inflammation (assessed by AST and CRP), and for association between ELF and episodes of infection or Gastro-Intestinal bleeding.

8.3.8 Statistical analyses

I performed the analyses using IBM SPSS statistics for Mac, version 27 (Armonk, NY: IM corp, https://www.ibm.com/support/pages/spss-statistics-27; 2020), apart from sensitivity, specificity, likelihood ratios and NPV/PPV which I calculated using MedCalc statistical software (MedCalc Statistical Software version 19.2.6 (MedCalc Software by Ostend, Belgium; https://www.medcalc.org; 2020).

I created a database which combined the ELF, HA, PIIINP, and TIMP1 results alongside the associated anonymised trial data for each participant. I used the existing data to calculate ABIC scores (Age, Bilirubin, INR, Creatinine) using the standard equation: (**A**ge, years x 0.1) + (Serum **B**ilirubin, mg/dL x 0.08) + (**I**NR x 0.8) + (Serum **C**reatinine, mg/dL x 0.3) (400).. MELD and GAHS scores were already available on the dataset. The formula for MELD is: MELD = 3.78×ln[serum bilirubin (mg/dL)] + 11.2×ln[INR] + 9.57×ln[serum creatinine (mg/dL)] + 6.43 ; and the formula for GAHS is: Age (<50 years = 1, ≥50 years = 2) + Leucocytes (10^9/L) (<15 = 1, ≥15 = 2) + Urea (mmol/L) (<5 = 1, ≥5 = 2) + PT ratio (<1.5 = 1, 1.5–2.0 = 2, >2.0 = 3) + bilirubin (mmol/L) (<125 = 1, 125–50 = 2, <250 = 3).

I summarised the patient characteristics and outcomes using standard summary statistics (n % for binary measures and median (IQR) or mean (SD) for continuous measures). The Chi-squared test was used to compare categorical variables, and

Mann-Whitney-U test for non-parametric continuous data. The rates of missing data for any variable are reported, and analyses were performed only on available data.

To identify the best-performing predictors of mortality, I analysed ELF and the ELF components (HA, PIIINP, TIMP1) individually along with the established prognostic scores (MELD, GAHS, ABIC, NLR) in a backwards stepwise logistic regression model. Comparison of scores was performed by Area Under the Receiver Operative Curve (AUC) analysis, with pairwise comparison of AUC results using DeLong analysis. This was performed for the individual variables, and also for the combination of the ELF with the established prognostic scores.

Optimal thresholds for the static scores were taken from the published STOPAH biomarker study of >1,000 patients with alcoholic hepatitis (131). I derived the optimal ELF threshold from AUROC analysis by calculating the maximum Youden Index (J) using the following equation: J = Sensitivity + specificity – 1 for each value of ELF, and selecting the highest scoring value (J_{max}), which equates to the optimum sensitivity and specificity, relating to the top left corner of the AUROC curve.

I used Kaplan-Meir analysis to assess survival, with application of the log-rank test to compare survival curves. Results are presented with 95% confidence intervals (95% CI) and significance set at $p < 0.05$.

I conducted *post-hoc* exploratory analyses of the addition of FIB4 and Lille Score to ELF in prognosis prediction. These were performed *post-hoc*, as available data were limited: Only 99/162 patients in this cohort had available Lille Score data, and 120/162 a FIB4 score. Upon liaising with the STOPAH-Trial management group, the missing Lille data were deemed to be due to either patients being lost to follow-up or discharged

prior to day 7 (day-7 bilirubin value is necessary for the Lille Score calculation). In these patients the STOPAH team were able to use MRIS data to record mortality outcomes.

The FIB4 data were missing in 42/162 patients because AST was not measured in these patients.

I compared the baseline characteristics of both the reduced cohort with FIB4 scores, and the reduced cohort with available Lille Scores, with those of the overall cohort of n=162, using Chi-squared for categorical data.

Logistic regression and AUROC were used to assess the performance of FIB4 and Lille in predicting prognosis, using the literature-derived Lille threshold of 0.45 (401), and log-rank test to compare Kaplan Meir survival curves with high ($\geq$ 0.45) versus low (<0.45) Lille score.

8.4 Results

A total of 179 serum samples and their associated STOPAH dataset were provided in two batches by the STOPAH working group for inclusion in this study.

A total of 7 patients did not have complete associated data for prognostic scores or mortality and were excluded. The Advia Centaur automatically dilutes samples that exceed the dynamic range of the assays by 1/5. Samples exceeding the dynamic range after automated dilution were subsequently diluted 1/10 manually. In the first batch of samples, despite manual dilution 10 samples yielded HA levels, and in one case PIIINP

levels, that exceeded the dynamic range of the assay. Insufficient sample volume remained for further dilutions and so these 10 subjects were also excluded from the analyses. Subsequently, for the second batch of samples, to avoid this issue a minimum sample volume to permit manual dilutions, should they be required, was first extracted prior to the initial measurement of the ELF analytes.

Thus, a total of 162 patients were included in the analysis.

8.4.1 Patient characteristics:

Patient characteristics were largely representative of the overall STOPAH cohort, although the mortality rate was lower, and infection rate higher in this subset (Table 8.1). The average age was 48.3 $\pm$ 10.2 and 64% were male. The 28/90/120-day mortality rate was 13/21/25% respectively. GAHS, MELD and ABIC scores were reflective of those in the original STOPAH cohort (n=1,092). Gastro-Intestinal bleeding (GI bleed) was a feature in 13/162 (8%) and sepsis in 27 (17%) of the cohort, leaving 122/162 (75.3%) without these complications. There was no difference in the proportion of patients who received prednisolone (50%) in this cohort when compared with the original STOPAH study cohort.

Table 8.1: Patient characteristics: Comparison of sub-study with STOPAH original cohort

	ELF sub-study cohort (n=162)	STOPAH original cohort (n=1092)
Age (years)	48.3 ± 10.2	48.7 ± 10.2
% male	64	63
Had prednisolone n (%)	81 (50)	547 (50.1)
Day 28 mortality n (%)	21/162 (13)	174 (16)
Day 90 mortality n (%)	34/162 (21)	285/968 (29)
Day 120 mortality n (%)	41/162 (25)	318/1077 (29.5)
Sepsis on admission n (%)	27 (17)	110 (10)
Renal failure on admission n (%)	2 (1)	2 (<0.5)
GI bleeding on admission n (%)	13 (8)	67 (6)
GAHS mean SD	8.3 ± 1.3	8.4 ±1.3
MELD mean SD	20.6 ± 6	21.2 ±6.2
ABIC mean SD	7.9 ± 1.3	7.95 (1.54)
NLR mean SD	3.1 ± 2.7	6.27 (4.64)

SD (±) = Standard Deviation, IQR = Interquartile range, , GI = Gastrointestinal, GAHS = Glasgow Alcoholic Hepatitis Score, MELD = Model for End-stage Liver Disease, ABIC= age/bilirubin/INR/creatinine, NLR = neutrophil to Lymphocyte ratio.

8.4.2 ELF test

ELF score and its components were high, with a median ELF score of 15.9 (IQR 14.9-16.9) (ELF score total range 11.97-21.68), HA 3971 mg/ml (IQR 2298-12246), PIIINP 16 mg/ml (IQR 83-249), TIMP1 1198 mg/ml (IQR 767-1638) (Table 8.1).

Associated mortality data were available for 28, 90 and 120 days. Following discussion with the STOPAH steering group, the 90-day value was taken to be the most clinically useful, as per the STOPAH biomarker study (131) as beyond this time mortality is thought to be influenced by possible return to alcohol consumption (402).

On univariate analysis, ELF was able to predict mortality at 28, 90 and 120 days (Figure 8.1), with 90-day OR of 1.7 (95% CI 1.3-2.2, p <0.001). ELF positively correlated with CRP (rho = 0.51, p <0.001) and AST (rho = 0.35, p <0.001) (Figure 8.2 A/B), with every CRP increase by 10mg/l increasing the ELF by 0.25, and with every AST increase by 100 increasing the ELF by 0.5. There was no association between ELF and infection episodes (p = 0.55) or variceal bleeding (p=0.56).

Figure 8.1 Boxplot of ELF against mortality status at 28, 90 and 120 days

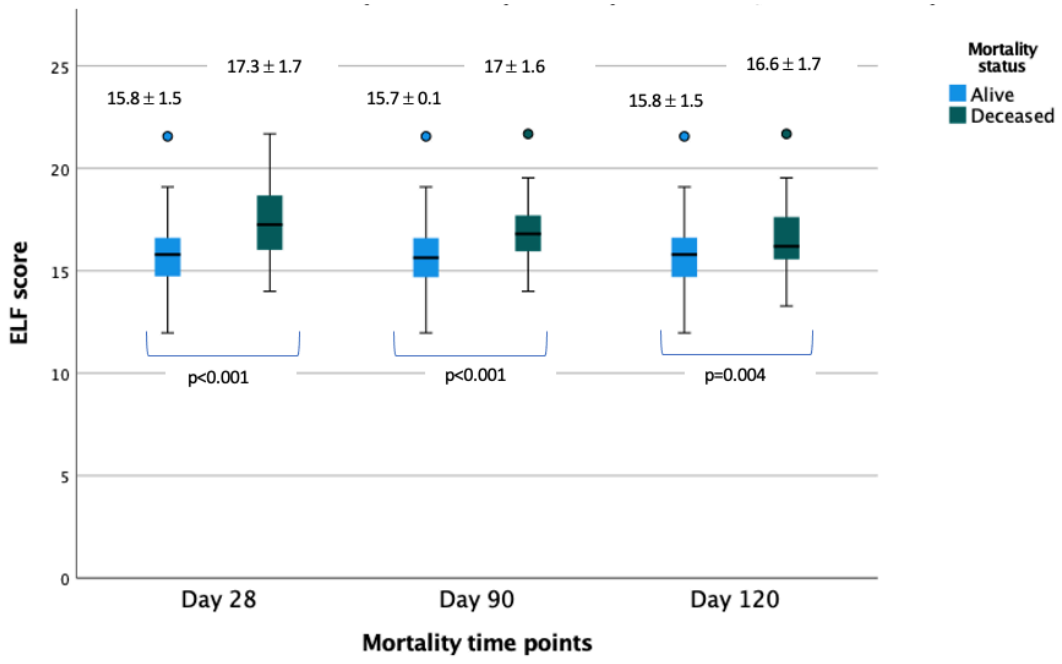
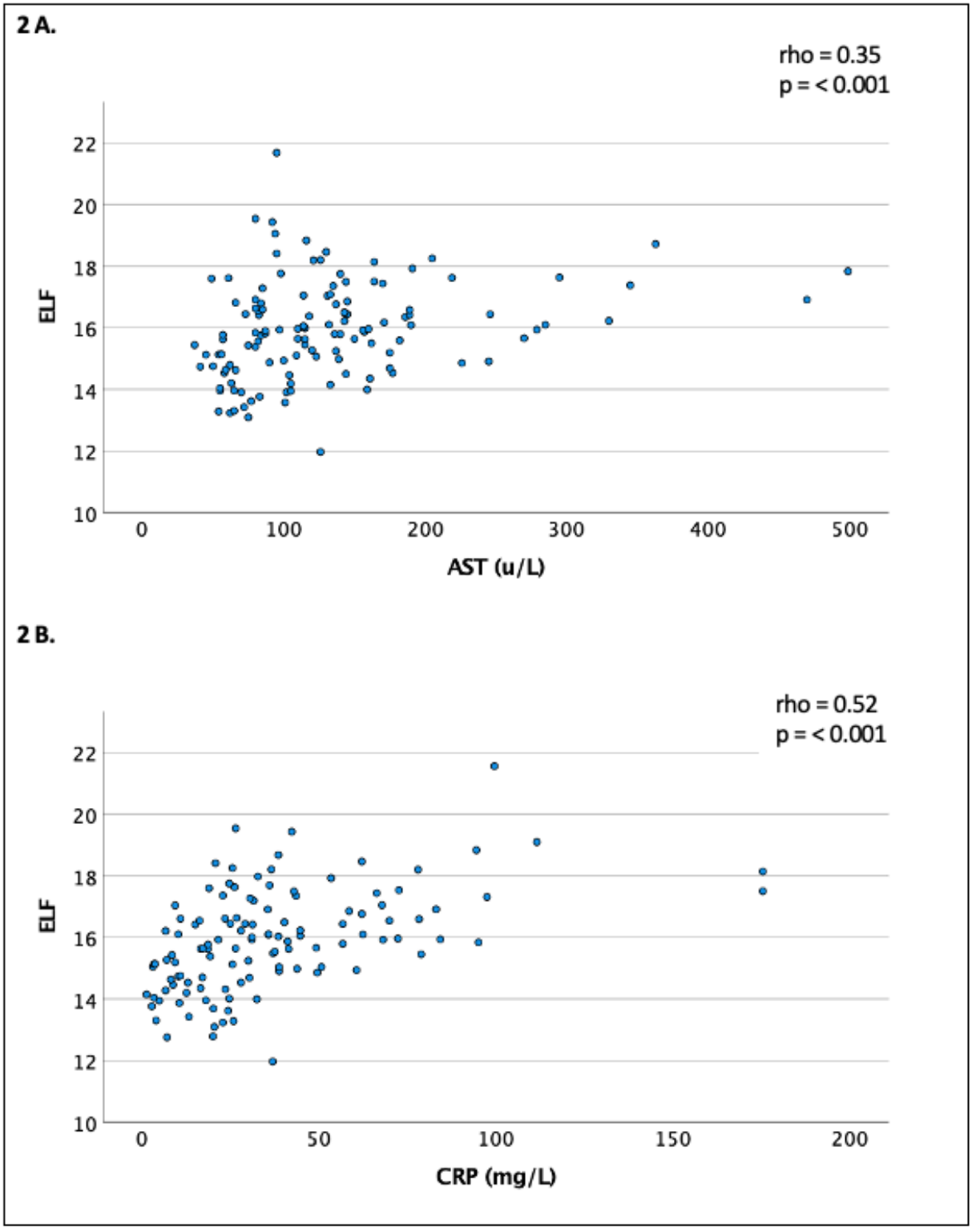


Figure 8.2 A/B: Scatter plot of A) ELF by AST (n= 125) and B) ELF by CRP (n=118)



8.4.3 Analysis of ELF compared with baseline 'static scores' GAHS, MELD, ABIC.

When examining individual performance in AUROC, ELF performed similarly to ABIC, GAHS and MELD (no significant difference). At 28-days ELF AUC= 0.77 (0.66-0.88), at 90-days AUC=0.72 (0.63-0.81) and 120-days AUC=0.64 (0.54-0.74) (Table 8.2).

A multivariable backwards logistic regression model was used to compare ELF with established static scores at 90 days, and repeated at 28 and 120 days. Variables entered at step 1 were: NLR, MELD, GAHS, ABIC, ELF, HA, PIIINP, TIMP1. This resulted in just ABIC and ELF remaining in the last step as the best performing biomarkers in predicting 90-day mortality (ABIC O.R.=2.21 (95% CI 1.5-3.2), ELF O.R.=1.64 (1.2-2.2). (Table 8.3).

Table 8.2: Areas Under Receiver Operating Characteristic curves for variables at 28, 90 and 120-day mortality time points. (N=162)

Blood biomarker	28-day AUROC (95%CI)	90-day AUROC (95% CI)	120-day AUROC (95% CI)
PIIINP	0.64 (0.52-0.75)	0.63 (0.53-0.72)	0.55 (0.45-0.65)
TIMP1	0.73 (0.62-0.85)	0.69 (0.58-0.79)	0.62 (0.51-0.72)
HA	0.76 (0.65-0.88)	0.72 (0.63-0.82)	0.66 (0.57-0.76)
ELF	0.77 (0.66-0.88)	0.72 (0.63-0.81)	0.64 (0.54-0.74)
NLR	0.78 (0.68-0.89)	0.70 (0.60-0.80)	0.63 (0.53-0.74)
ABIC	0.81(0.71-0.91)	0.74 (0.64-0.83)	0.71(0.62-0.80)
MELD	0.83 (0.75-0.91)	0.74(0.65-0.83)	0.70 (0.60-0.79)
GAHS	0.86 (0.79-0.93)	0.75 (0.66-0.84)	0.71 (0.62-0.80)
ABIC+ELF	0.87 (0.80-0.95)	0.81 (0.73-0.89)	0.75 (0.66-0.84)
MELD+ELF	0.87 (0.80-0.93)	0.78 (0.70-0.86)	0.71 (0.62-0.80)
GAHS+ELF	0.89 (0.83-0.95)	0.79 (0.70-0.87)	0.72 (0.63-0.82)
NLR+ELF	0.80 (0.70-0.90)	0.74 (0.66-0.83)	0.65 (0.56-0.75)

CI = Confidence Interval, AUROC = Area Under Receiver Operating Characteristic curve, PIIINP = Type III procollagen peptide, TIMP1 = Tissue Inhibitor of Metalloproteinase-11, HA = Hyaluronic acid, ELF = Enhanced Liver Fibrosis score, NLR = Neutrophil to Lymphocyte ratio, ABIC = Age/bilirubin/INR/creatinine, MELD = Model for End-stage Liver Disease, GAHS = Glasgow Alcoholic Hepatitis Score

Table 8.3: Multivariable backwards stepwise logistic regression for 28, 90, 120-day mortality end-points

(N=162)	Variable	Odds ratio	95% C.I. for Odds Ratio		p value
			Lower	Upper	
28 days	**ABIC**	2.02	1.13	3.61	0.017
	ELF	1.75	1.196	2.545	0.004
	GAHS	1.99	1.008	3.914	0.048
90 days	**ABIC**	2.21	1.513	3.218	<0.001
	ELF	1.64	1.233	2.188	0.001
120 days	**ABIC**	2.03	1.442	2.862	<0.001
	ELF	1.34	1.053	1.707	0.017

(Variables entered on step 1: NLR, MELD, GAHS, ABIC, ELF, HA, PIIINP, TIMP1.)
C.I = Confidence Interval, ELF = Enhanced Liver Fibrosis score, ABIC = Age/Bilirubin/INR/Creatinine score

8.4.4 Combining ELF with ABIC in AUROC analysis

My initial hypothesis was that ELF as a direct biomarker of liver fibrosis would either outperform or augment the prognostic information provided by the established "static" prognostic scores. Having established that the combination of ABIC+ELF performed optimally in the logistic regression illustrated in Table 8.3, equations from the logistic regression were applied to combine ABIC and ELF, and this new marker 'ABIC+ELF' was then compared with the single markers (ELF, ABIC, GAHS, MELD) in AUROC analysis. The same method was then applied to combine GAHS+ELF and MELD+ELF and compare performance at predicting mortality with the simple static markers.

Table 8.2 and Figure 8.3 demonstrate that the addition of ELF to ABIC improves the performance of both ELF and ABIC at predicting mortality, with an AUROC of 0.81 (95% CI 0.73-0.89) at 90-days, higher than all the other variables.

Pairwise comparisons of AUCs (Table 8.4) show that the combination of ABIC+ELF significantly outperforms ABIC alone at 28 and 90 days.

Figure 8.3 A/B/C ROC curves for ABIC+ELF, ABIC alone, and ELF alone at 28/90/120-day mortality end points.

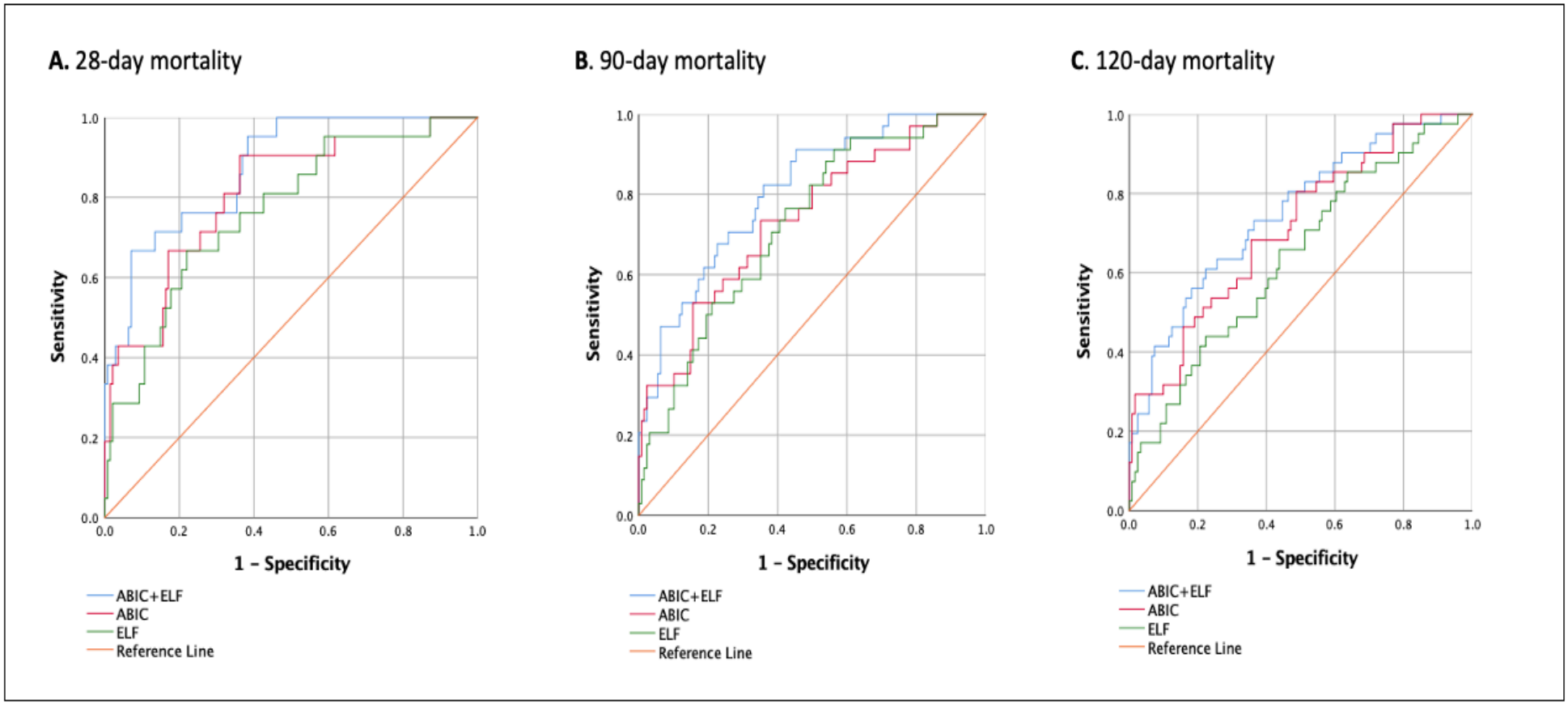


Table 8.4: Comparison of AUCs (whole cohort): combined direct fibrosis marker (ELF) with indirect fibrosis marker (GAHS, MELD or ABIC) versus indirect fibrosis marker alone.

			28-day outcome	90-day outcome	120-day outcome
ELF+ABIC	**vs ABIC**	95% CI	0.00035 to 0.128	0.02 to 0.12	-0.09-0.11
		Z statistic	1.97	2.6	-1.5
		p value	**0.049**	**0.01**	0.13
	vs MELD	95% CI	-0.143 to 0.052	-0.02 to 0.16	-0.14-0.04
		Z statistic	-0.92	1.5	-1.19
		p value	0.36	0.15	0.24
	vs GAHS	95% CI	-0.104 to 0.076	-0.02 to 0.13	-0.12-0.04
		Z statistic	-0.31	1.4	-1.04
		p value	0.758	0.17	0.3
	vs ELF	95% CI	2.01	-0.007-0.176	0.025-0.2
		Z statistic	0.003-0.214	1.8	2.5
		p value	**0.044**	0.07	**0.012**
ELF+GAHS	**vs ABIC**	95% CI	-0.19 to 0.027	-0.14 to 0.04	-0.1-0,08
		Z statistic	-1.47	-1.06	-0.27
		p value	0.14	0.29	0.79
	vs MELD	95% CI	-0.14 to 0.01	-0.11 to 0.02	-0.1-0.04
		Z statistic	-1.68	-1.4	-0.79
		p value	0.094	0.16	0.43
	vs GAHS	95% CI	0.003 to 0.06	0.005 to 0.064	-0.04-0.02
		Z statistic	2.16	2.3	-0.9
		p value	**0.03**	**0.02**	0.37
	vs ELF	95% CI	0.02-0.23	-0.16-0.03	-0.17-0.001
		Z statistic	-2.3	-1.36	-1.9
		p value	**0.02**	0.17	0.054
ELF+MELD	**vs ABIC**	95% CI	-0.17 to 0.06	-0.15 to 0.06	-0.1-0.97
		Z statistic	-.94	-0.86	-0.04
		p value	0.35	0.39	0.97
	vs MELD	95% CI	-0.09 to 0.02	-0.09 to 0.004	-0.06-0.03
		Z statistic	-1.3	-1.8	-0.7
		p value	0.18	0.07	0.47
	vs GAHS	95% CI	-0.08 to 0.07	-0.098 to 0.038	-0.07-0.07
		Z statistic	-0.15	-0.86	-0.12
		p value	0.88	0.39	0.9
	vs ELF	95% CI	0.008-0.19	-0.14-0.02	-0.15-0.002
		Z statistic	-2.12	-1.46	-1.9
		p value	0.034	0.15	0.057

ELF – Enhanced Liver Fibrosis test, ABIC = Age/bilirubin/INR/creatinine, MELD = Model for End-stage Liver Disease, GAHS = Glasgow Alcoholic Hepatitis Score, CI = Confidence Interval

8.4.5 Applying thresholds

Established literature-derived thresholds were used for ABIC (403), GAHS (125), and MELD (131). The ELF thresholds established for the diagnosis of liver fibrosis and prognosis in CLD were exceeded by the levels measured in this study of AH. Thus, new prognostic thresholds for mortality were established in this STOPAH cohort. For ELF, the Youden's (J) index was calculated from AUROC analysis for the optimum threshold as described above. This resulted in an optimum ELF threshold of 16.78 yielding a specificity of 79% (95% CI 71-86) and NPV of 87% (95% CI 82-90) for the prediction of 90-day mortality. (Table 8.5).

Using AUROC values, I then explored the use of an upper threshold of ELF, with high specificity of 90% for the prediction of mortality, and a lower ELF threshold with high sensitivity of 90% for the prediction of survival.

The resulting high ELF threshold that produced a specificity of 90% for predicting 90-day mortality was 17.8, and this threshold was consistent at 28- and 120-day mortality end points. At 90 days, applying a high ELF threshold of 17.8 results in a specificity of 91% (95% CI 84-95%), sensitivity of 24% (95% CI 11-41%), PPV of 42 (95% CI 24-62%) and an NPV of 82% (95% CI 79-85%) which was similar to the performance of MELD. However, this only identified 19/62 (12%) of the study cohort.

Applying a low ELF threshold of 15.5 resulted in a sensitivity of 88% (95% CI 73-97%) in predicting 90-day survival (or 'ruling out' mortality), with NPV of 94% (95% CI 85-97%). This again only applied to a minority (62/162 or 38%) of the cohort.

The Youden-index derived ELF of 16.78 was therefore more clinically useful, and in the next section I demonstrate its performance in prognostication when used in conjunction with ABIC at the established threshold of 6.71.

Table 8.5: Performance of ELF, GAHS, MELD, ABIC in predicting 28, 90 and 120-day mortality

Biomarker + threshold N=162	AUROC (95% CI)	TP	FP	TN	FN	Sensitivity % (95% CI)	Specificity % (95% CI)	PPV % (95% CI)	NPV % (95% CI)	LR+ (95% CI)	LR- (95% CI)	Prevalence (mortality) (%)
28-day mortality												
ELF <15.5	0.77 (0.66-0.88)	19	81	60	2	91 (70-99)	43 (34-52)	19 (16-23)	97 (89-99)	1.6 (1.3-1.9)	0.2 (0.06-0.85)	13.1
ELF ≥17.8	0.77 (0.66-0.88)	6	13	128	15	29 (11-52)	91(85-95)	32 (17-52)	89 (86-92)	3.1 (1.3-7.3)	0.8 (0.6-1.04)	13.1
ELF ≥16.78	0.77 (0.66-0.88)	14	31	110	7	67 (43-85)	78 (70-85)	31 (23-41)	94 (89-97)	3 (1.97-4.7)	0.4 (0.2-0.8)	13.1
GAHS ≥9	0.86 (0.79-0.93)	19	50	91	2	91 (70-99)	65 (56-72)	28 (23-33)	98 (92-99)	2.6 (2-3.3)	0.15 (0.04-0.55)	13.1
MELD ≥25	0.83 (0.75-0.91)	10	19	122	11	48 (26-70)	87 (80-92)	35 (22-50)	92 (88-94)	3.6 (1.9-6.5)	0.6 (0.4-0.9)	13.1
ABIC ≥6.71	0.81 (0.71-0.91)	20	118	23	1	95 (76-99.9)	16 (11-24)	15 (13-16)	96 (76-99)	1.1 (1-1.3)	0.3 (0-2.1)	13.1
90-day mortality												
ELF <15.5	0.72 (0.63-0.81)	30	70	58	4	88 (73-97)	45 (37-54)	30 (26-34)	94 (85-97)	1.6 (1.3-2)	0.3 (0.1-0.7)	20.8
ELF ≥17.8	0.72 (0.63-0.81)	8	11	117	26	24 (11-41)	91 (84-95)	42 (24-62)	82 (79-85)	2.7 (1.2-6.3)	0.8 (0.7-1)	20.8
ELF ≥16.78	0.72 (0.63-0.81)	18	27	101	16	53 (35-70)	79 (71-86)	40 (29-51)	87 (82-90)	2.5 (1.6-4)	0.6 (0.4-0.9)	20.8
GAHS ≥9	0.75 (0.66-0.84)	23	46	82	11	68 (49-83)	64 (55-72)	33 (26-41)	88 (82-93)	1.9 (1.4-2.6)	0.5 (0.3-0.8)	20.8
MELD ≥25	0.74 (0.65-0.83)	13	16	112	21	38 (22-56)	88 (81-93)	45 (30-60)	84 (80-88)	3.1 (1.7-5.7)	0.7 (0.5-0.9)	20.8
ABIC ≥6.71	0.74 (0.64-0.83)	33	105	23	1	97 (85-99.9)	18 (12-26)	24 (22-26)	96 (77-99)	1.2 (1-1.3)	0.16 (0-1.17	20.8
120-day mortality												
ELF <15.5	0.64 (0.54-0.74)	31	69	52	10	76 (60-88)	43 (34-52)	31 (26-36)	84 (75-90)	1.3 (1.1-1.7)	0.6 (0.3-1)	25
ELF ≥17.8	0.64 (0.54-0.74)	8	11	110	33	20 (9-35)	91 (84-95)	42 (24-62)	77 (74-80)	2.2 (0.9-5)	0.9 (0.8-1)	25
ELF ≥16.78	0.64 (0.54-0.74)	18	27	94	23	44 (28-60)	78 (69-85)	40 (29-52)	81 (76-85)	2 (1.2-3.2)	0.7 (0.5-1)	25
GAHS ≥9	0.71 (0.61-0.80)	26	43	78	15	63 (47-78)	65 (55-73)	37 (30-45)	84 (78-89)	1.8 (1.3-2.5)	0.6 (0.4-0.9)	25
MELD ≥25	0.70 (0.6-0.79)	14	15	106	27	34 (20-51)	88 (80-93)	48 (33-63)	80 (76-83)	2.8 (1.5-5.2)	0.8 (0.6-1)	25
ABIC ≥6.71	0.71 (0.62-0.80)	40	98	23	1	98 (87-99.9)	19 (12-27)	29 (27-31)	96 (77-99)	1.2 (1.1-1.3)	0.1 (0-0.9)	25

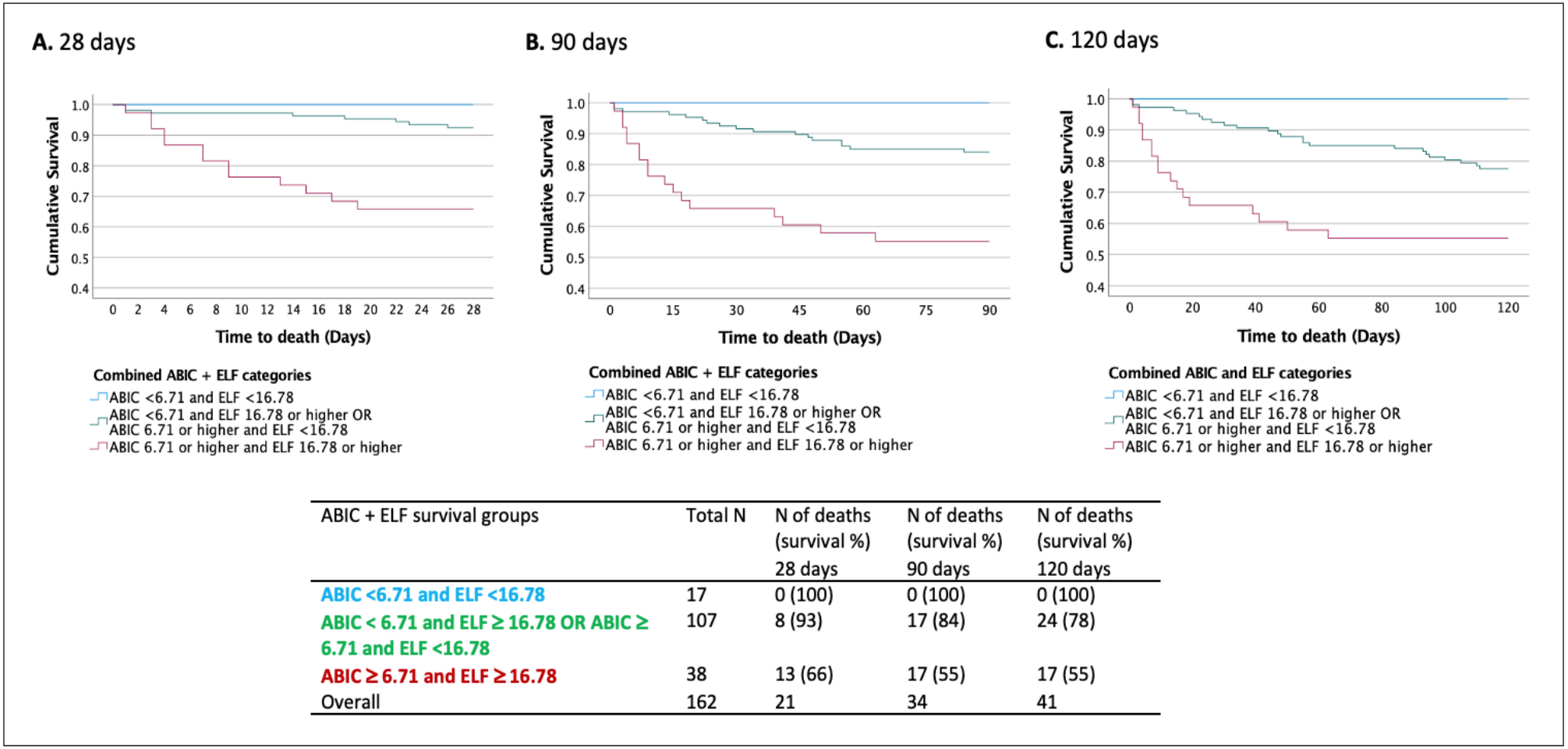
(ABIC/MELD/GAHS: published thresholds used from STOPAH biomarker study (131). ELF threshold of 16.78 derived from AUROC by calculating Youdens (J) index. Alternative low 'rule out' and high 'rule in' mortality ELF thresholds also displayed. AUROC= Area Under Receiver Operator Characteristic curve, CI = Confidence Interval, TP = True Positive, FP = False Positive, TN = True Negative, FN = False Negative, PPV = Positive Predictive Value, NPV = Negative Predictive Value, LR+ = Positive Likelihood Ratio, LR- = Negative Likelihood Ratio.

8.4.6 Combining ELF and ABIC thresholds for prognosis prediction

I separated the ELF and ABIC results into three threshold-categories for Kaplan Meier analysis: Low ELF/low ELF, high ABIC/high ELF, and an intermediate discordant category of low ABIC/high ELF or high ABIC/low ELF. (Figure 8.3 A/B/C). This resulted in good separation of survival curves between these three groups, and log-rank test found a significant difference ($p < 0.01$) between the highest and lowest curves.

At 28, 90 or 120-day end-points, having a low-ABIC score <6.71 and a low ELF score <16.78 guaranteed survival in 100% of the patients. (This applies to 17/162 or 10.5% of the sample). Conversely, those patients with a high ABIC (≥ 6.71) and high ELF (≥ 16.78) had a mortality rate of 45% at 90 or 120 days, and 34% at 28 days (applying to 38/162 or 23.5%). In total, the highest and lowest curves on the graph accounted for 55/162 (34%) of the cohort. The mortality rate in the intermediate group (low ABIC/high ELF or high ABIC/low ELF) was 7%, 16% and 22% at 28, 90, and 120 days respectively.

Figure 8.4 A/B/C: Kaplan Meier Survival curves for combined ABIC+ELF at 28, 90, and 120-day mortality end-points.



ABIC + ELF survival groups	Total N	N of deaths (survival %) 28 days	N of deaths (survival %) 90 days	N of deaths (survival %) 120 days
ABIC <6.71 and ELF <16.78	17	0 (100)	0 (100)	0 (100)
ABIC < 6.71 and ELF ≥ 16.78 OR ABIC ≥ 6.71 and ELF <16.78	107	8 (93)	17 (84)	24 (78)
ABIC ≥ 6.71 and ELF ≥ 16.78	38	13 (66)	17 (55)	17 (55)
Overall	162	21	34	41

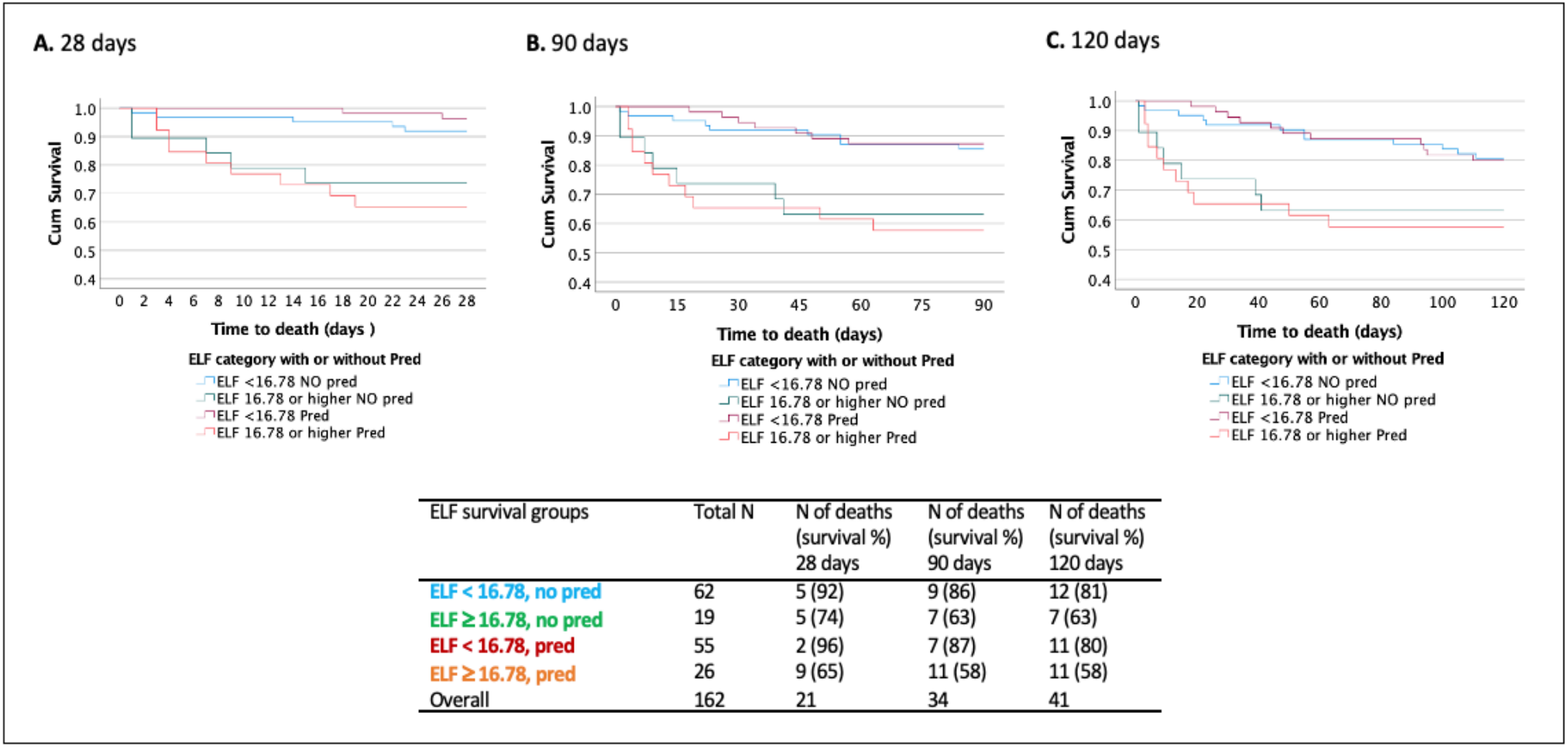
Fig 8.4. Kaplan-Meier survival probability for patients stratified by a combination of ELF with ABIC at A) 28-day, B) 90-day, and C) 120-day censored study end-points. ABIC threshold of 6.71 taken from published optimum threshold (131). ELF threshold of 16.78 derived from Youden's (J) index. ABIC = Age, Bilirubin, INR, Creatinine.

8.5 Application to management strategies

When examining ELF alone as a prognostic marker, using a threshold of 16.78, prednisolone did not confer a significant survival benefit in either the low-ELF (<16.78) or high ELF (≥ 16.78) groups when tested using the log-rank test. (Figure 8.5 A/B/C) When ELF and ABIC were grouped into three categories as described above (group 1: low ABIC/low ELF, group 2: low ABIC/high ELF OR high ABIC/low ELF, group 3: High ABIC/high ELF), there was also no significant survival benefit seen with prednisolone between the category groups. (Figure 8.6 and 8.7 A/B/C)

As liver transplantation is currently being explored as a treatment for patients with alcoholic hepatitis, and carries with it a mortality of approximately 10%, it would be useful to be able to predict which patients with alcoholic hepatitis have a survival rate of <10%, who are unlikely to get better with current treatment options and might be appropriate for selection for transplantation. Higher thresholds of ABIC at 9 (literature-derived (131)) and ELF of 17.8 (AUROC-derived) were therefore explored via Kaplan Meier and log-rank test analysis to see if <10% survival could be predicted from these scores. Applying the higher ABIC threshold of 9 with ELF of 16.78 (Youden-index derived), resulted in a survival rate of 36% at 90 days but only affected 11/162 patients (6.8%). Applying an even higher ELF threshold of 17.8 (AUROC-derived) with the ABIC of 9 reduced the survival rate above this threshold to 25% but only affected 4/162 (2.5%) patients. It was therefore not possible to identify biomarkers capable of predicting <10% survival in this cohort. Details of this analysis can be found in Figure 8.8/8.9 A/B/C.

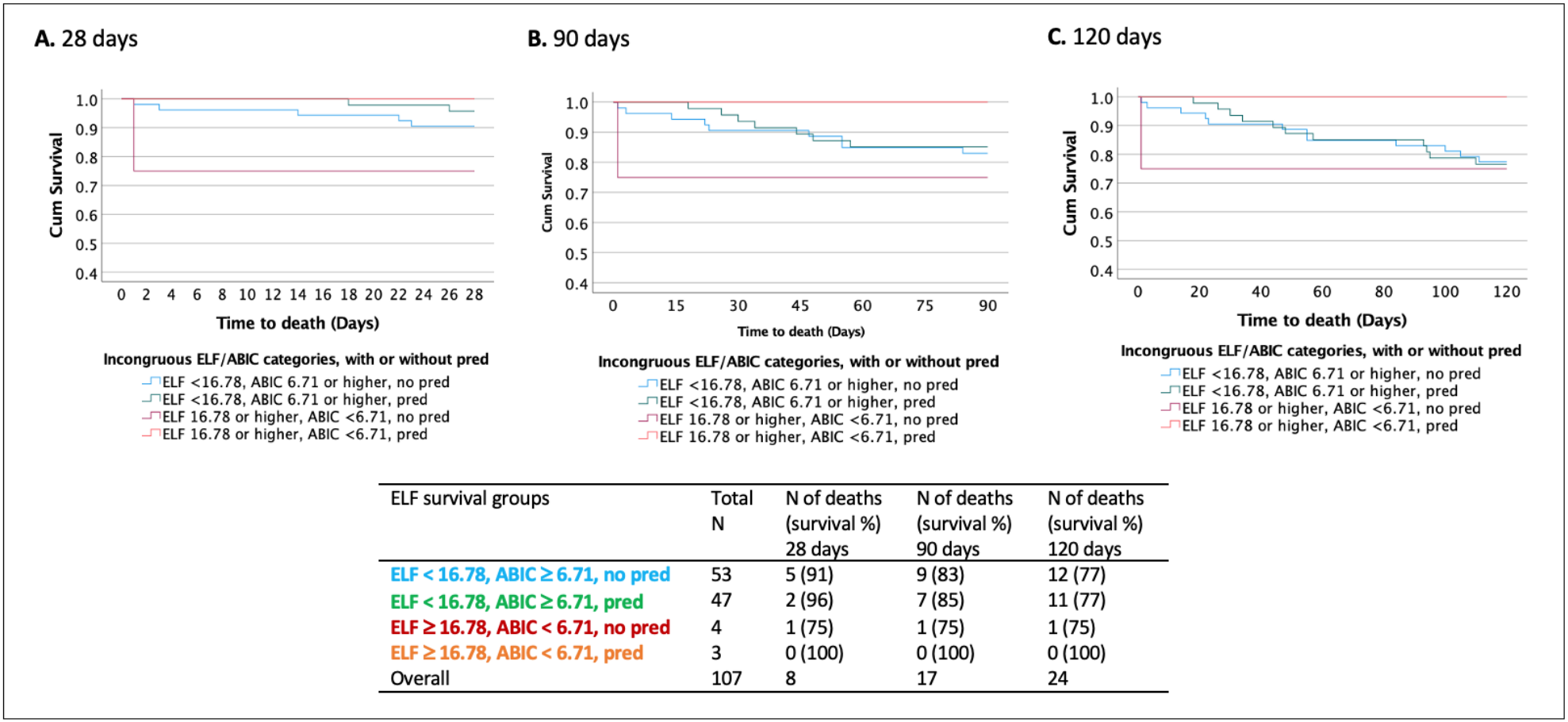
Figure 8.5 A/B/C Kaplan Meier survival curves for high or low ELF with and without prednisolone at 28/90/120 days.



ELF survival groups	Total N	N of deaths (survival %) 28 days	N of deaths (survival %) 90 days	N of deaths (survival %) 120 days
ELF < 16.78, no pred	62	5 (92)	9 (86)	12 (81)
ELF ≥ 16.78, no pred	19	5 (74)	7 (63)	7 (63)
ELF < 16.78, pred	55	2 (96)	7 (87)	11 (80)
ELF ≥ 16.78, pred	26	9 (65)	11 (58)	11 (58)
Overall	162	21	34	41

Fig 8.5 A/B/C. Kaplan-Meier survival probability for patients stratified by high or low ELF, with or without prednisolone exposure at A) 28-day, B) 90-day, and C) 120-day censored study end-points. ELF threshold of 16.78 derived from Youden's (J) index.

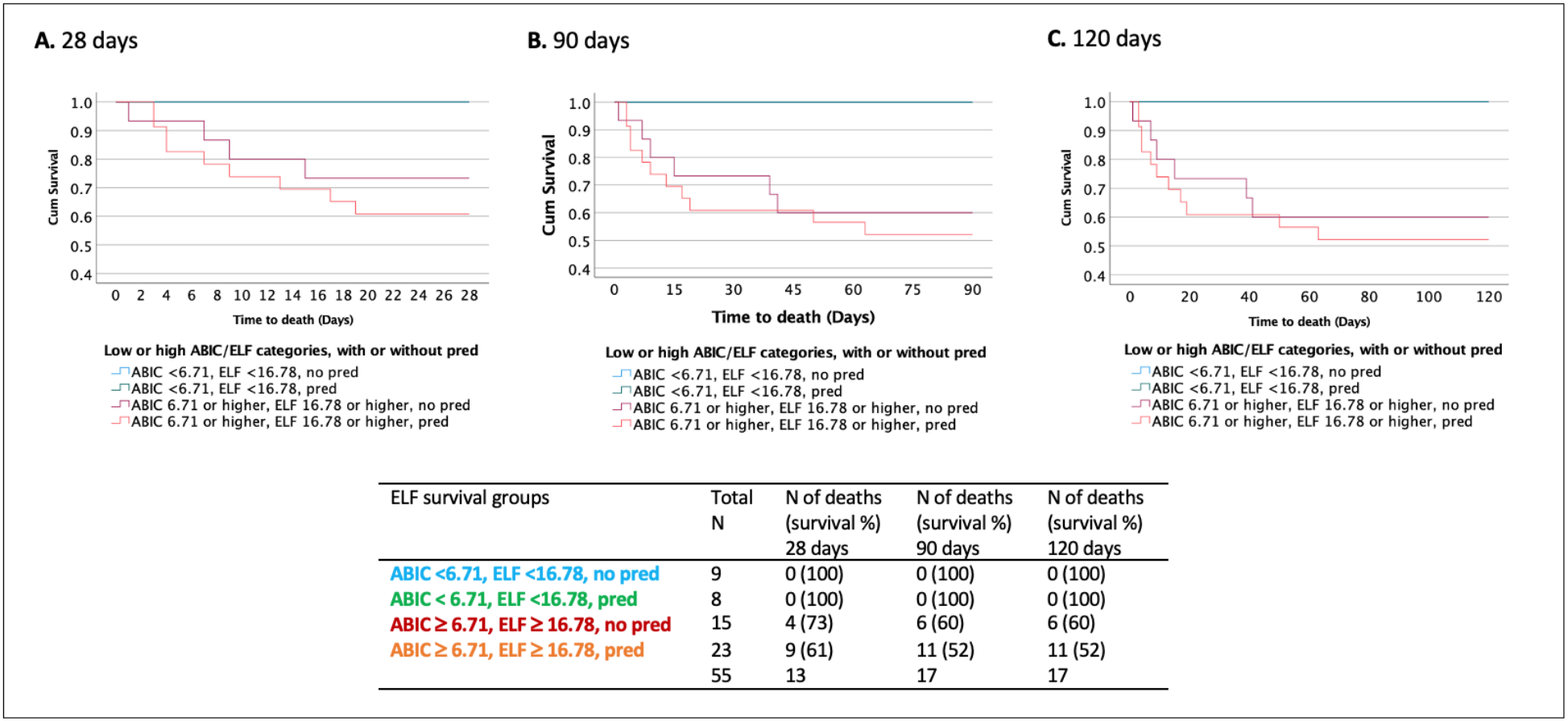
Figure 8.6 A/B/C: Kaplan Meier Survival curves for incongruous ABIC+ELF categories, with or without prednisolone, at 28, 90, and 120-day mortality end-points.



ELF survival groups	Total N	N of deaths (survival %) 28 days	N of deaths (survival %) 90 days	N of deaths (survival %) 120 days
ELF < 16.78, ABIC ≥ 6.71, no pred	53	5 (91)	9 (83)	12 (77)
ELF < 16.78, ABIC ≥ 6.71, pred	47	2 (96)	7 (85)	11 (77)
ELF ≥ 16.78, ABIC < 6.71, no pred	4	1 (75)	1 (75)	1 (75)
ELF ≥ 16.78, ABIC < 6.71, pred	3	0 (100)	0 (100)	0 (100)
Overall	107	8	17	24

Fig 8.6 A/B/C. Kaplan-Meier survival probability for patients stratified by incongruous combinations of ABIC and ELF, with or without prednisolone exposure at A) 28-day, B) 90-day, and C) 120-day censored study end-points. ELF threshold of 16.78 derived from Youden's (J) index.

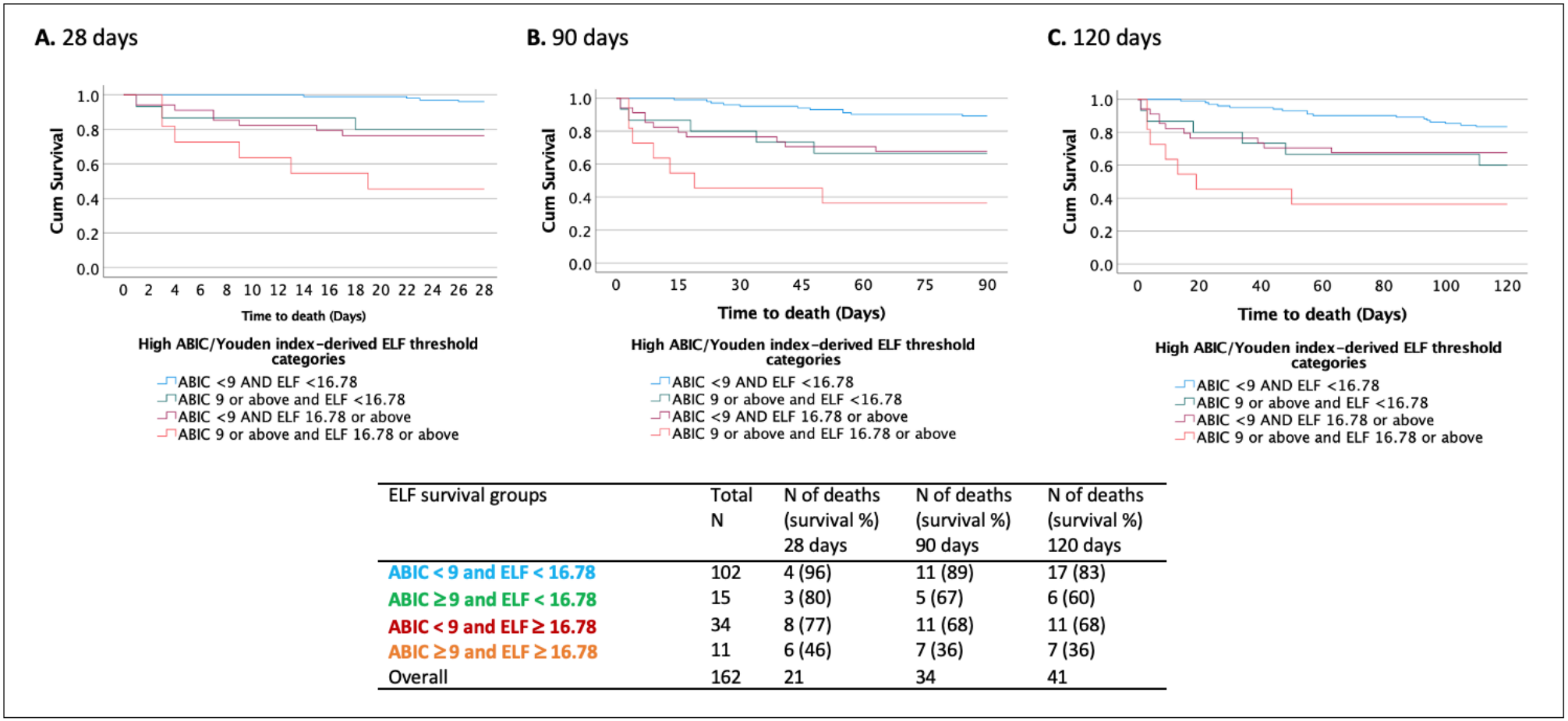
Figure 8.7 A/B/C: Kaplan Meier Survival curves for combinations of low ABIC/ELF, high ABIC/ELF, with or without prednisolone, at 28, 90, and 120- day mortality end-points.



ELF survival groups	Total N	N of deaths (survival %) 28 days	N of deaths (survival %) 90 days	N of deaths (survival %) 120 days
ABIC <6.71, ELF <16.78, no pred	9	0 (100)	0 (100)	0 (100)
ABIC < 6.71, ELF <16.78, pred	8	0 (100)	0 (100)	0 (100)
ABIC ≥ 6.71, ELF ≥ 16.78, no pred	15	4 (73)	6 (60)	6 (60)
ABIC ≥ 6.71, ELF ≥ 16.78, pred	23	9 (61)	11 (52)	11 (52)
	55	13	17	17

Fig 8.7 A/B/C. Kaplan-Meier survival probability for patients stratified by high or low ABIC/ ELF categories, with or without prednisolone exposure at A) 28-day, B) 90-day, and C) 120-day censored study end-points. ELF threshold of 16.78 derived from Youden's (J) index.

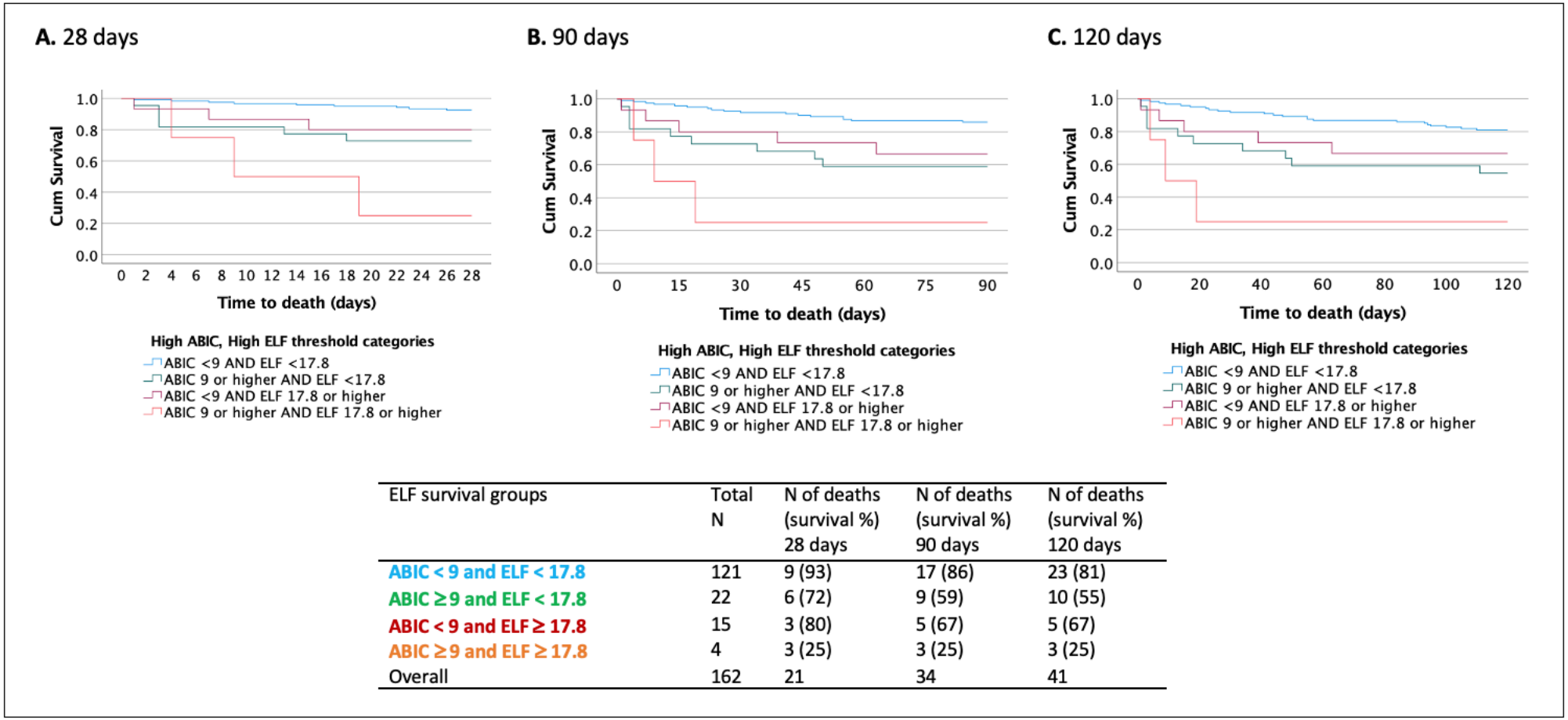
Figure 8.8 A/B/C Kaplan Meier Survival curves for combined ABIC+ELF, using high ABIC threshold of 9, at 28, 90, and 120-day mortality end-points.



ELF survival groups	Total N	N of deaths (survival %) 28 days	N of deaths (survival %) 90 days	N of deaths (survival %) 120 days
ABIC < 9 and ELF < 16.78	102	4 (96)	11 (89)	17 (83)
ABIC ≥ 9 and ELF < 16.78	15	3 (80)	5 (67)	6 (60)
ABIC < 9 and ELF ≥ 16.78	34	8 (77)	11 (68)	11 (68)
ABIC ≥ 9 and ELF ≥ 16.78	11	6 (46)	7 (36)	7 (36)
Overall	162	21	34	41

Fig 8.8 A/B/C. Kaplan-Meier survival probability for patients stratified by high or low ABIC/ ELF categories using high ABIC threshold of 9 at A) 28-day, B) 90-day, and C) 120-day censored study end-points. ELF threshold of 16.78 derived from Youden's (J) index.

Figure 8.9 A/B/C Kaplan Meier Survival curves for combined ABIC+ELF, using high ABIC threshold of 9, at 28, 90, and 120-day mortality end-points.



ELF survival groups	Total N	N of deaths (survival %) 28 days	N of deaths (survival %) 90 days	N of deaths (survival %) 120 days
ABIC < 9 and ELF < 17.8	121	9 (93)	17 (86)	23 (81)
ABIC ≥ 9 and ELF < 17.8	22	6 (72)	9 (59)	10 (55)
ABIC < 9 and ELF ≥ 17.8	15	3 (80)	5 (67)	5 (67)
ABIC ≥ 9 and ELF ≥ 17.8	4	3 (25)	3 (25)	3 (25)
Overall	162	21	34	41

Fig 8.9 A/B/C. Kaplan-Meier survival probability for patients stratified by high or low ABIC/ ELF categories using high ABIC threshold of 9 (literature-derived (131)) and high ELF threshold of 17.8 (AUROC-derived) at A) 28-day, B) 90-day, and C) 120-day censored study end-points

8.6 *Post-hoc* analyses: Investigating the performance of FIB-4 and Lille in predicting 90-day mortality.

8.6.1 FIB-4

Data were available to calculate FIB-4 for 120 patients (42 lacked AST values), and there was no significant difference in baseline characteristics of the n=120 with FIB4, compared to the total n=162 cohort. Applying binary logistic regression, FIB4 did not reliably predict mortality, with a 90-day OR of 1.01 (95% CI 0.996-1,027, p = 0.14), and similar results were obtained for 28 and 120 days.

8.6.2 Lille Score

Only 99/162 patients had available Lille scores. Baseline characteristics of these 99 patients were not significantly different to the total cohort of 162.

Lille score predicted 90-day mortality (OR 9.3, 95% CI 1.7-51, p = 0.01), with an AUROC of 0.7 at 90 days (95% CI 0.57-0.83).

8.6.3 Addition of Lille to ABIC/ELF for prognosis prediction

In the cohort of patients with available Lille (n=99), the Lille score (using a threshold of 0.45) was applied to three categorised groups: 1) Low ABIC/low ELF, 2) Discordant ABIC/ELF (i.e., low ABIC/high ELF or high ABIC/low ELF, 3) High ABIC, High ELF.

In groups 1 and 2, the Lille score did not provide additional survival information to the ABIC/ELF scores. However, in group 3, the addition of Lille score to those patients with high ABIC and high ELF scores allowed their separation into two

survival groups. Those with High ABIC/high ELF and Lille <0.45 had a 90-day survival rate of 90% (n=9/10 survived), compared to 6/13 survival (46% survival) in the group with Lille ≥0.45 (p = 0.047). (Figure 8.9.1).

Although the numbers were small, I investigated the outcomes in the group of patients with high ELF/ABIC scores (ELF ≥ 16.78 and ABIC ≥ 6.71) that were taking prednisolone. Survival was 100% (n=8/8) in those with a Lille Score < 0.45, compared with 25% survival (n= 1/4) in those taking prednisolone who had Lille ≥ 0.45, p = 0.004. (Figure 8.9.2). It was not possible to examine the effect of Lille score on survival benefit with prednisolone in those with discordant ABIC/ELF (high ABIC, low ELF and vice versa) or Low ABIC/Low ELF, due to the numbers being too few in these groups.

8.6.4 Impact of infection on survival in those in differing ABIC/ELF categories.

In the patients with high ABIC plus high ELF, (n=38/162), or low ABIC plus Low ELF (n=17) there was no additional mortality impact in those who had reported baseline sepsis compared to those without sepsis. In those with high ABIC and low ELF (n=100), however, those with baseline sepsis (n =17/100) had a significantly worse outcome (survival 59%) compared to those without baseline sepsis (83/100, survival 89%), p = 0.001.

Figure 8.9.1: Kaplan Meier survival curve for high or low ELF/ABIC with and without Lille score at threshold of 0.45, at 90 days.

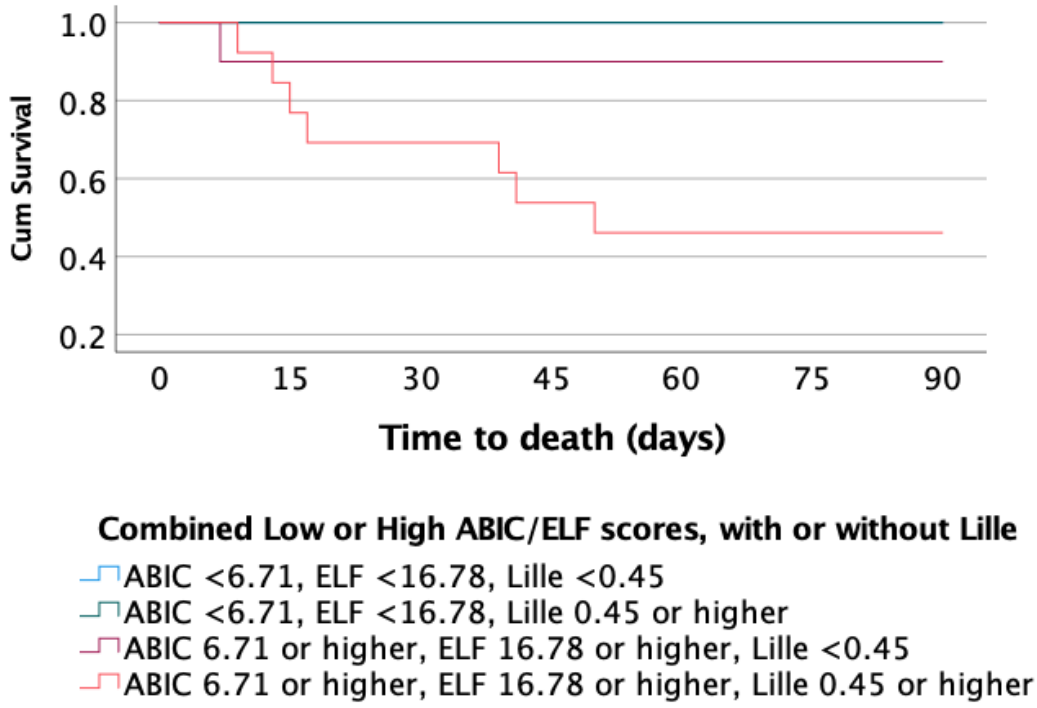
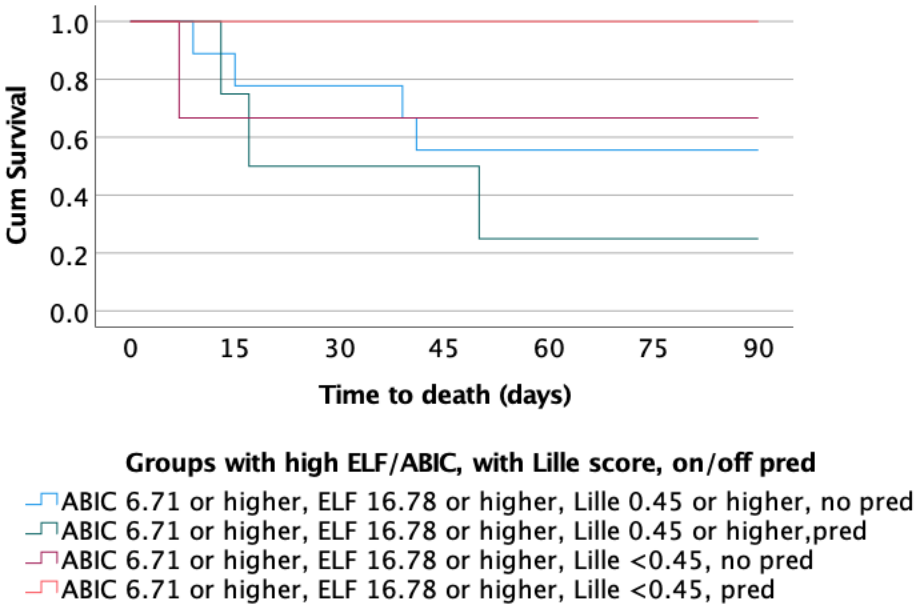


Figure 8.9.2: Kaplan Meier survival curve for high ABIC/ELF groups, with Lille score, on vs off prednisolone at 90 days.



8.7.1 Discussion

AH is the most severe presentation of ArLD, associated with high mortality rates, and currently lacking effective pharmacological treatment. With the recent promising trials on transplantation for AH, it is now even more important to be able to have a reliable prognostic score that can predict which patients have the poorest prognosis and might potentially benefit from access to early liver transplantation. Whilst pre-existing scores that are currently used widely in clinical practice such as MELD, GAHS, and ABIC have all been shown to perform equally to each other, and outperform the traditionally used DF score in AH (131), their performance is still suboptimal, with AUROCs all less than 0.8 for the prediction of 90-day mortality.

This is the first study of ELF in alcoholic hepatitis. I have discovered that ELF is a useful prognostic marker in AH, having been shown previously to be superior to histology in predicting outcomes in CLD (335) and on par with other non-invasive fibrosis tests such as FibroTest, FibroScan and FIB4 in predicting prognosis in alcohol-related liver disease (224). This is also the first study where ELF has been directly compared with currently used prognostic scores in AH -MELD, GAHS and ABIC. I have shown that not only is ELF non-inferior to MELD/GAHS/ABIC in predicting prognosis in AH, but that whilst previous studies of prognostic markers in AH have shown AUROCs >0.7, this is the first study to identify a combination of biomarkers yielding a prognostic AUROC >0.8 for 90-day mortality in AH, when combining ELF with ABIC.

My finding that the best-performing marker in combination with ELF was ABIC, rather than GAHS or MELD, is in keeping with previous literature findings: A study by Dominguez et al (403) found that ABIC was the best independent predictor of 90-day mortality when compared to MELD, Maddrey's DF, and GAHS (HR 2.78, 95% CI 1.9-4.09, p = 0.0001) (403). ABIC was also found to be the only independent predictor of 1-year mortality (HR 2.49, 95% CI 1.77-3.52, P = 0.0001) when compared to other prognostic scores, and a 7-day ABIC was found to have a better 6-month mortality prediction than Lille (403).

It is important to note, however, that the Dominguez et al. study was the derivation cohort. Whilst the more recent STOPAH biomarker study (131) of n=1092 patients found ABIC to also have the numerically highest 90-day HR of 1.82 out of ABIC/GAHS/MELD, this was not statistically distinct from the other markers, with authors concluding that ABIC, GAHS and MELD all perform similarly at 90 days.

It was not possible in this study to select out a sub-group of patients with survival rate of <10% (the mortality threshold below which their chance of survival would likely to be greater with liver transplantation) - perhaps due to the better-than-expected survival rates in this cohort, and the relatively small sample size in this sub-study. For example, I highlighted in Figure 9 that only 4 patients had a high-enough ABIC and ELF score in order to have a 90-day survival rate of 25% (applying to only 1 out of 4 patients), therefore it was not possible to seek out a predictive score for <10% survival in this cohort. Similarly, in the main STOPAH cohort of 1092 patients, the highest-reported 90-day mortality was 42.2% in those patients with

the highest baseline static scores and the authors did not report a <10% survival group (131).

However, using ABIC/ELF thresholds, I was able to find that those with low ABIC (<6.71) and low ELF (<16.78) had a 100% survival rate at 120 days, whether they were taking prednisolone or not (47% (8/17) were on prednisolone, compared with 53% (9/17) not on prednisolone, with 100% survival in all). This is clinically useful to know as these patients have a good prognosis and are likely to do well with standard care comprising nutritional/psychological support. Further prospective analysis is needed to determine conclusively if there is any potential survival benefit of prednisolone in this low-risk group.

Conversely, I found that those patients with high ABIC (≥ 6.71) and high ELF (≥ 16.78) had a 45% mortality rate at 90 days. The addition of the Lille score to the latter group allowed further differentiation of survival rates, with those patients with high ABIC/high ELF and Lille ≥0.45 having a survival rate of 46% compared to 90% survival in those with high ABIC/high ELF and Lille <0.45. This is in keeping with findings from the STOPAH biomarker study (131) that the addition of a 'dynamic' score such as Lille to a baseline 'static score' improves the prognostic ability of the score (131).

Furthermore, I was able to identify that in the group of patients with high ELF/ABIC scores (ELF ≥ 16.78 and ABIC ≥ 6.71) that were taking prednisolone, 100% (n=8/8) survived who had a Lille response (< 0.45), compared with a 25% survival (n= 1/4) in those taking prednisolone who showed no Lille response (≥ 0.45), p = 0.004. Whilst

the sample size of patients in this sub-analysis was clearly very small, these results are in keeping with those from other studies that have shown no benefit of prednisolone in patients with Lille $\geq$ 0.45 (131, 401).

In patients with discordant ABIC/ELF (high ABIC and low ELF, n=100), I discovered that those with presence of sepsis at baseline (n =17/100) had a significantly worse outcome (90-day survival 59%) compared to those without sepsis (83/100, survival 89%), p = 0.001. The presence of sepsis thus further helps the risk stratification in this 'intermediate' group with discordant ABIC/ELF scores. (Baseline sepsis was defined in the STOPAH cohort by Nikhil Vergis et al. (404): the diagnosis was made prospectively by clinicians blinded to the use of steroids or pentoxifylline, and diagnostic criteria followed the guidelines for diagnosing infection in patients with cirrhosis published by Bajaj et al (405)).

Whilst histological data were lacking in this patient cohort, it is reflective of real-world practice, where liver biopsy is not routinely indicated in AH. Therefore, it is not possible to know whether the prognostic ability of ELF in AH in this study is related to its association with severity of fibrosis, or whether ELF is a marker of inflammation severity in this cohort, or a combination of both. I did find positive correlations between ELF and AST/CRP in this study, and ELF scores were much higher than reported in patients with established cirrhosis without alcoholic hepatitis. The manufacturer's 'high' ELF threshold for detecting the presence of cirrhosis is 11.3, and the median ELF score in this cohort was 15.9 (IQR 14.9-16.9), with scores ranging up to 21.68. While ELF scores of up to 16 or 17 may be encountered in the assessment of patients with established cirrhosis (W. Rosenberg

personal communication) the levels recorded in this study were uniformly very high. Therefore, it is possible that the high level of inflammation that occurs in AH is the cause of this, but this requires further evaluation in a biopsy-paired study.

Previous studies have reported conflicting results on the potential correlation between ELF and inflammation, with one study in viral hepatitis finding correlation with inflammation (398), and two on ArLD that did not find any correlation between ELF and AST/ALT as markers of inflammation (23, 274), although one of these two latter studies did find a correlation between ELF and histological inflammation (23). As AST/ALT are less reliable markers of inflammation than histological changes are, there is a need for further studies to prove a link between ELF and histological inflammation in ArLD. No studies up until now, however, have investigated ELF in alcoholic hepatitis, which by definition is a highly inflammatory state, although with mortality rates strongly linked to fibrosis severity (393).

Table 8.2 interestingly shows that PIIINP produced lower AUROCs than TIMP1 and HA at all three time points, although the greatest difference was between PIIINP and HA. (PIIINP AUROC 0.63 at 90 days, compared to HA AUROC of 0.72). This is a significant difference, with p value 0.031 (95% CI 0.009 to 0.188). The difference between PIIINP and TIMP1 was not significant (p = 0.26, 95% CI -0.16 to 0.04). As these biomarkers have not been evaluated before in alcoholic hepatitis, it is difficult to draw any firm conclusions from this.

However, PIIINP has previously been shown to be associated with severity of inflammation in NAFLD, independently from histological fibrosis stage or ELF (406). HA levels have also been observed to reflect the severity of liver inflammation

(407), and in a 1997 study by Murawaki et al., TIMP1 levels were closely correlated with the histological degrees of portal inflammation (although the author's conclusion of this study was that TIMP1 increases with progression of liver disease and is key in the development of liver fibrosis (408).

In my study, whilst all ELF components were elevated compared to previously published ELF cohorts, the HA levels in particular were the most markedly raised. I observed HA levels that were 'off the scale' with some samples being required to undergo 1/10 dilutions in order to achieve a readable result on the adviar centaur analyser. As far as I am aware, levels this high (median HA 3971 mg/ml (IQR 2298-12246), have not been observed before. In this cohort, the CRP correlated with HA ($p = 0.006$, rho 0.25), PIIINP ($p < 0.001$, rho 0.677) and TIMP1 ($p < 0.001$, rho 0.703), as well as with ELF ($p < 0.001$, rho 0.52).

A hypothesis could be that as HA levels are so significantly elevated in this cohort, that the reason HA performs better than PIIINP in predicting 90-day mortality is that it is reflective of a higher level of inflammation.

An alternative hypothesis would be that HA is a better marker of liver fibrosis than PIIINP. Without paired histology samples, however, it is not possible to come to any firm conclusion, and future prospective studies on biomarkers in AH with paired histology samples in order to correlate these biomarkers with CPA/fibrosis stage and histological inflammation severity scores would be informative.

8.7.2 Strengths

This study is the first to report on ELF in AH. Whilst it lacks histologically-matched samples, it is reflective of 'real-world' clinical practice, where liver biopsy is not routinely performed in AH, and the diagnosis is made clinically. All data for comparison with ELF scores were prospectively collected under STOPAH trial regulations in multiple centres. My primary-outcome was focussed on 90-day survival, which is now recognised to be the optimal end-point for survival outcomes in AH, as beyond this is likely to be affected by return to drinking (402). Where available, I applied pre-defined prognostic score thresholds from published sources, and where this was not available (for ELF in AH), I used an evidenced-based approach, calculating the Youden Index, to derive the optimal threshold from AUC.

8.7.3 Limitations

The overall sample size for this study was restricted by the availability of stored sera from the STOPAH trial. Whilst this cohort of 162 patients is representative of the 1092-patient STOPAH study in terms of key sociodemographic and clinical factors, it did have a significantly higher baseline sepsis rate (17% compared to 10%) and a lower 90-day mortality rate (21% compared to 29%). Taking this into consideration, in the knowledge of how prevalence impacts on PPV/NPV (409), one might expect the PPV of ELF, and therefore its ability to predict mortality, to improve with increasing prevalence of mortality and lower sepsis rates, so ELF might have performed better in the whole 1092-patient cohort. Conversely, the NPV might be expected to be lower in the higher-prevalence mortality population

in the 1092-patient STOPAH cohort, and therefore survival-prediction not as good. Arguably, the unmet need is for a test that better-predicts mortality in AH (rather than survival), as it would permit the selection of patients who may benefit from additional treatments such as intensive-care support or liver transplantation.

The sub-analyses were also restricted in sample-size by the lack of available AST (for FIB4 calculation) in 25% of the cohort, and lack of available Lille score in 40%. Further analyses for Lille in a larger cohort are required.

8.7.4 Overall conclusion

To conclude, I have discovered that ELF is a useful prognostic marker in AH, performing similarly to GAHS/MELD/ABIC. When combined with ABIC, I found that the ELF-ABIC marker outperformed ABIC alone in the prediction of 90-day mortality with AUROC >0.8, the highest reported AUROC of a prognostic marker in AH for 90-day mortality thus far. The addition of Lille is useful to further stratify the patients with high ELF/high ABIC scores into those who are very likely to survive (survival rate of 90%) and those with poor survival rate of 46%.

Further validation of ELF-ABIC is required in a larger sample of patients, and it would be interesting to include histology data where possible, in order to learn more about the prognostic ability of ELF in AH.

Chapter 9

Summary of main findings

and

Discussion

9.1 Outline of thesis aims and methodology

My overall aims for the thesis were to investigate the performance and current practice of non-invasive testing to risk-stratify liver fibrosis, and predict prognosis in ArLD. The methods I used to address these aims included two systematic reviews (one with meta-analysis), a retrospective cross-sectional analysis of GP referrals with suspected ArLD, a prospective evaluation of advanced fibrosis prevalence in people presenting to hospital with AUD, and finally a biomarker study using prospectively collected data to investigate prognostic performance of ELF in AH.

In this final section, I will highlight the key findings from each chapter. I will then summarise how these findings meet my original aims and objectives, and how they add to the existing evidence base in this area. I will discuss the limitations of my work, and finally consider what are the remaining unanswered questions in this field, and make recommendations for future research.

9.2 Summary of key points emerging from the new work presented in this thesis

9.2.1 CHAPTER 3: Investigating the diagnostic performance of four non-invasive tests in alcohol-related liver disease: a systematic review with meta-analysis.

- A search through 11,000 publications produced 16 articles with a total of 2,280 participants for inclusion in an analysis of the diagnostic accuracy of FIB4, FibroTest, FibroScan and ELF in ArLD.
- Despite alcohol being the commonest cause of cirrhosis, there were fewer studies of non-invasive tests in ArLD than in other aetiologies such as viral hepatitis and NAFLD.

- FibroScan was the most extensively studied non-invasive test in ArLD.
- Due to heterogeneity between studies, FibroScan was the only one of the 4 tests for which meta-analysis was possible. The pooled AUROC for studies of FibroScan was 0.91 (95% CI 0.89-0.94) for F3, and for F4: pooled sensitivity was 88% (95% CI 0.84-0.92), and pooled specificity was 84% (95% CI 0.81-0.87).
- For ELF the AUROC ranges for F3 was 0.82-0.92 and that for FibroTest was 0.80-0.90. Both were higher than AUROCs for FIB4 (0.70-0.85), although direct comparisons were not possible.
- Whilst ELF, FibroScan, FibroTest and FIB4 all performed well (AUROCs > 0.80) for the detection of advanced fibrosis (F3), thresholds for all these 4 tests still need further validation for use in ArLD.
- ELF was associated with histological inflammation in one study on ArLD, but not with AST or ALT in either of the two included studies. Alcohol did not appear to influence ELF.
- FibroScan was influenced by recent alcohol consumption or withdrawal in 3 out of 4 studies examining this (increased false positives).
- FibroScan was influenced by inflammation (histological or transaminases) in 5 studies out of 8 that investigated this.
- FibroScan was the only one out of ELF, FibroScan, FibroTest and FIB4 to have a failure rate, ranging from 1-22%.

9.2.2 Chapter 4: Diagnosing advanced fibrosis in alcohol-related liver disease in practice – examining current referral strategies from primary to secondary care, and risk factors associated with advanced fibrosis

- Over two-thirds (147/229) of referrals to secondary with suspected ArLD were 'unnecessary' in that they had no advanced fibrosis and could be discharged back to primary care.
- Despite national guidelines, 84% of alcohol referrals to secondary care did not have fibrosis testing of any kind prior to referral, reflecting the need for clear, easy-to-follow local pathways for primary care physicians.
- Drinking over 14 units per week doubled the odds of advanced fibrosis in overweight/obesity (OR 2.11; CI 1.44-to-3.09; $p<0.001$).
- Alcohol unit thresholds of $\geq$35 units per week for women, and $\geq$50 units per week for men were associated with a diagnosis of advanced fibrosis, validating these thresholds that were adopted in NICE and BSG guidelines without clear evidence.
- FIB4 had an AUROC of 0.80 (95% CI 0.74-0.86) for the composite clinical diagnosis of advanced fibrosis. This AUC is consistent with the findings of the systematic review presented in Chapter 3.
- Modelling demonstrated that use of FIB4 in primary care could halve the number of unnecessary alcohol referrals to secondary care (OR 0.50; CI 0.32-to-0.79, $p = 0.003$).

9.2.3 Chapter 5: Implementing a community referral pathway involving the ELF test in patients with alcohol use disorder – the 'Camden and Islington alcohol pathway'

- In Camden and Islington CCGs, the pre-existing local alcohol pathway did not contain any information about non-invasive fibrosis testing for ArLD.
- In collaboration with GPs, hepatologists and public health professionals, I designed a new alcohol pathway that incorporated BSG guidelines, recommending the use of the ELF test to detect liver fibrosis in women drinking $\geq$ 35 units per week, and men $\geq$ 50 units per week, or in patients considered to be hazardous or harmful drinkers (WHO classification) as per AUDIT score.
- I participated in gaining approval for this new pathway that was launched in the Camden and Islington CCGs in October 2019.
- In promoting the pathway, I helped raise awareness amongst GPs and hospital physicians at the Royal Free about the potential benefits and need to proactively perform non-invasive testing in people at risk of ArLD.
- This pathway has the potential to improve care for patients by reducing unnecessary referrals to hospital (and associated anxiety-provoking investigations) and increasing the likelihood of early detection of liver fibrosis, potentially allowing intervention to reduce alcohol intake, detect and treat complications of fibrosis, and improve outcomes.
- This pathway has potential to benefit the NHS, through better resource utilization with a higher proportion of referred patients expected to have advanced fibrosis/cirrhosis, and improved detection of liver disease at a

point when intervention could avoid harm, with anticipated cost-savings through a reduction in referrals and investigations, and cost-utility though reduction in harms from advanced liver disease.

9.2.4 Chapter 6: Uncovering unsuspected advanced liver fibrosis in hospitalized patients admitted with alcohol related problems, referred to the alcohol nurse specialist but not recognised to have liver disease using the ELF test

- Ninety-nine consecutive hospitalised patients referred to the alcohol liaison service but not thought to have liver disease were investigated for the presence of liver fibrosis using the ELF test.
- None of the 99 patients in this study had undergone prior investigation for liver fibrosis, despite consuming a median of 140 units of alcohol per week, and with three-quarters (75/99) having presented to hospital a median of 4 times in the previous 5 years.
- A third of patients in hospital with AUD (28/99) over a 13-month period had previously undetected advanced liver fibrosis as assessed by an ELF score $\geq$10.5.
- Twenty-eight percent (8/28) of those with ELF score $\geq$10.5 indicating advanced fibrosis, had normal LFTs.
- ELF was found to correlate with age (r = 0.303, p = 0.002), but not with the amount of recent alcohol consumed in this cohort.
- ELF was not found to correlate with AST or ALT in this cohort of 99 patients with AUD.

9.2.5 Chapter 7: Investigating the prognostic performance of four non-invasive tests in alcohol-related liver disease: a systematic review

- I found 25,088 articles that were screened, identifying 11 articles for inclusion in the systematic review.
- Fewer studies investigated prognostic performance of non-invasive tests in ArLD than in viral hepatitis or NAFLD.
- FIB4, FibroTest, ELF and FibroScan were all able to predict prognosis in ArLD with AUROCs > 0.7 for each test.
- FIB4 may perform better than MELD in prognosis prediction in ArLD, (FIB4 AUROC for mortality prediction 0.83, compared to MELD AUROC 0.70 (p = 0.001) in 1 study (365) and FIB4 AUROC for prediction of variceal bleed 0.74 (0.66-0.81) in a second study (372), compared to MELD AUROC 0.54 (95% CI 0.46-0.62).
- FibroTest and ELF performed at least as well as histology in predicting prognosis in ArLD.

9.2.6 Chapter 8: Investigating the prognostic performance of ELF in alcoholic hepatitis (AH)

- In the STOPAH study cohort ELF was found to predict mortality in AH at 28, 90 and 120 days. The Odds Ratio for predicting mortality at 90-days was 1.7 (95% CI 1.3-2.2, p <0.001)
- In this cohort of patients with AH (n=162), ELF positively correlated with CRP (rho = 0.51, p <0.001) and AST (rho = 0.35, p <0.001)
- ELF was not associated with infection episodes (p = 0.55) or variceal bleeding (p=0.56).
- Logistic regression revealed an algorithm combining ELF and ABIC outperformed all other variables (including GAHS, MELD, NLR) in predicting mortality at 90 days. ELF+ABIC as a single marker had the highest numerical AUROC for 90-day mortality prediction at 0.81 (95% CI 0.73-0.89) compared to all other variables. When combined with ABIC, ELF significantly enhanced the performance of ABIC alone (p =0.01). ELF+ABIC is the first prognostic marker with AUROC > 0.8 in AH.
- Patients with both low-ABIC (<6.71) and low ELF score (<16.78) (n=17/162, 10.5%) had a 100% 90-day survival rate, compared with 55% survival in those with high ABIC/high ELF (n= 38/162, 23.5%) (p <0.01).
- In those with high ABIC and low ELF (n=100), those with baseline sepsis (n =17/100) had a significantly worse outcome (survival rate 59%) compared to those without sepsis (83/100, survival rate 89%), p = 0.001.

- Those with High ABIC/high ELF and Lille <0.45 had a 90-day survival rate of 90% (n=9/10 survived), compared to 6/13 survival (46% survival) in the group with Lille ≥0.45 (p = 0.047).

9.3 Discussion

9.3.1 Trends in alcohol use since the SARS-CoV2 pandemic

In Chapter 2, I noted that whilst there has been a global reduction in the prevalence of people who drink alcohol (from 47.6% to 43% between 2000 and 2016), the total alcohol per capita (APC) consumption has increased during this time, suggesting that people who drink alcohol, although fewer in number, are drinking more (42). This trend is predicted by the WHO to continue with an expected global increase in total APC from 6.4 litres to 7 litres by 2025 (410).

Since the start of my PhD in 2019, these trends have continued in the wrong direction, enhanced by the SARS-CoV-2 pandemic (411). Mortality from alcohol has continued to rise in the UK since 2019, with deaths from alcohol reaching a 'record high' in 2020 during the pandemic, with 5460 recorded alcohol-related deaths between January and September 2020, which was a 16.4% increase compared with the same period in 2019 (412). Deaths from ArLD increased by 20.8% in the same period, and accounted for 80% of the alcohol-specific deaths (413). In addition to increased alcohol consumption, a lack of access to healthcare during the pandemic may have contributed to this (414). The full impact of the SARS-CoV-2 pandemic on alcohol trends is unclear, but it is recognized to have led to a shift to more home

drinking in lockdown, with an increase in alcohol sales by 67% as people reacted to the closure of bars and pubs restrictions (415). Emerging data also suggests an increase in domestic violence since the first lockdown in March 2020, with the WHO reporting a 60% increase in emergency calls from women subjected to violence by their partners during the pandemic in 2020 (416). Three-quarters of domestic violence perpetrators had been drinking alcohol at the time of the assault (415).

When reviewing trends of alcohol consumption during the pandemic, a survey in 55,811 adults in 2020 from 11 countries found that overall, whilst 42% did not change their drinking patterns in response to the pandemic, in those who did, a larger proportion of them increased their alcohol intake (36% increased alcohol compared to 22% who decreased) (411). Risk factors for increased drinking included being young, female, or the presence of stress or anxiety (417), suggesting the need to target support towards these groups. This demographic group is not the same as the demographic impacted by increased mortality from ArLD – which tends to be older males, aged between 45-69 (418). However, the development of liver fibrosis and cirrhosis usually takes several years, and it is important to provide support to younger people with risky alcohol intake in order to help reduce their risk of long-term liver damage.

ArLD carries the highest burden of mortality and morbidity from alcohol, with over three-quarters of the alcohol-related deaths in the UK in the last year being caused by ArLD (412). In the vast majority of cases (75%) of ArLD liver fibrosis and early cirrhosis are asymptomatic and patients first present to healthcare with

decompensated disease when prognosis can be poor, and options to intervene to improve outcomes at this point are limited. With the worrying alcohol trends described in this thesis, it is imperative that more is done to address the mortality rates in ArLD. One important angle to tackle this is to improve the support available to patients with AUD, and address the amount of alcohol consumed on a population level (an example of this is the Minimum Unit Pricing (MUP) policies introduced in Scotland, a public health policy not yet introduced in England). Whilst long-term impacts of MUP on health outcomes will only become apparent over the next decade, this public health initiative has proven to be successful in reducing alcohol consumption, and therefore this may be expected to positively impact on ArLD morbidity and mortality trends. Another vital method is addressing the early detection of liver fibrosis in people at risk from AUD. For those patients identified prior to the development of cirrhosis this approach gains sufficient time to allow interventions such as "brief advice" on harm minimization and advice to support abstinence and permit fibrosis regression. For those patients with established cirrhosis, surveillance for portal hypertension and HCC can permit the instigation of interventions that have been shown to improve outcomes in randomized controlled trials (116, 419). My thesis has focused on this latter tactic of early detection of liver fibrosis, exploring the current practice in the detection of chronic liver disease and investigating if diagnostic testing could lead to earlier detection of alcohol-related liver disease in primary and secondary care. I have also investigated the performance of non-invasive tests in diagnosing fibrosis and prognosticating in ArLD, the practicalities of incorporating non-invasive tests into a primary care pathway, and the ability of non-invasive fibrosis tests to predict prognosis in AH.

9.4 Current practice of non-invasive fibrosis testing in ArLD- scope for change

In this thesis, I aimed to investigate the performance of current methods used by GPs to select patients with AUD for referral to liver specialists for the investigation of ArLD, what proportion of referred patients had advanced fibrosis or cirrhosis and could be considered 'necessary referrals', the performance of available simple indirect markers for fibrosis in this population, and if there were potentially missed opportunities for detecting advanced fibrosis in primary and secondary care.

In Chapter 4, using real-world data, I discovered that out of the patients referred from primary care with suspected ArLD over 3 years, 147/229 (64%) had no evidence of advanced fibrosis when assessed in secondary care, and could be discharged back into primary care for management of their alcohol use, representing unnecessary referrals. The vast majority of GPs appeared to have based their decision to refer on factors such as alcohol intake, presence of steatosis on ultrasound, and abnormal LFTs, with only 38/231 (16%) having had any kind of fibrosis assessment prior referral. Most of the patients who did have a fibrosis assessment pre-referral (16%) were overweight or living with obesity in addition to having concerning alcohol intakes, and so it was likely these GPs were following the local NAFLD pathway because these patients had a FIB4 test followed by ELF where indicated. A local alcohol pathway specifying use of non-invasive fibrosis tests for patients with suspected alcohol related liver disease prior to referral had not been implemented at the time of this study. My evaluation in Chapter 4 is also the first to report a threshold effect that matches the NICE and BSG guidelines for non-invasive fibrosis testing, of ≥35 units in women and ≥50 units in men.

In Chapter 6, my work demonstrates that opportunistic use of ELF could uncover occult liver disease – in 100% of patients referred to the alcohol specialist nurse from hospital over 13 months, none had undergone prior fibrosis testing despite consuming a median of 140 units per week of alcohol and having presented to hospital a median of 4 times in the previous 5 years to health care services. Despite national (BSG) guidelines (1) advocating use of ELF to detect advanced fibrosis in ArLD, these guidelines are clearly not being followed.

By performing ELF on these patients and detecting a third (28/99) with ELF $\geq$10.5 which were referred on to hepatology, I have likely improved the early diagnosis of liver disease in many of these patients that would have otherwise gone untested.

In Chapter 5 I describe how I presented evidence about alcohol use and liver disease data to GPs, public health consultants, and hospital clinicians at the Royal Free hospital, highlighting the BSG guidelines that should be followed to improve early detection of liver fibrosis in ArLD. I successfully helped develop and launch a local pathway in Camden and Islington, recommending that the ELF test should be performed in people consuming excess alcohol ($\geq$35 units in females, and $\geq$50 units in males), which has potential to significantly improve the way patients are selected for referral to secondary care, enabling those with low ELF scores to remain in primary care and be supported in the community for their alcohol use, and those with elevated ELF scores to be referred for further investigation and management in secondary care. A similar pathway for NAFLD was launched in 2014 and resulted

in 88% reduction in unnecessary referrals to secondary care, a 5-fold increase in the detection of advanced fibrosis, 3-fold increase in cirrhosis-detection, and significant cost savings for the NHS. In Chapter 4, I demonstrated that the use of FIB4 alone would halve the proportion of unnecessary alcohol referrals to secondary care (OR 0.50; CI 0.32-to-0.79, p = 0.003). As I found in Chapter 3 that ELF is superior to FIB4 for the detection of advanced fibrosis, and by analogy with the Camden and Islington NAFLD pathway evaluation, the use of ELF in my primary care alcohol pathway is likely to improve both the detection of advanced fibrosis and the reduction in unnecessary referrals to secondary care.

9.5 Diagnostic tests for fibrosis in ArLD- which to use?

The development of non-invasive tests for liver fibrosis has not displaced the role of liver biopsy in the investigation of patients with ArLD. Histological examination of a liver biopsy provides insight into the cellular pattern of inflammation as well as architectural disruption due to fibrosis. Biopsy may be particularly informative in situations where aetiology is uncertain, where there is comorbidity or if the clinician does not have satisfactory conclusion of fibrosis stage from the available non-invasive tests. However, it is an invasive procedure, not without risk, and the inaccuracies associated with sampling bias and intra- and inter-observer variability have been well reported. With the availability of many non-invasive tests, and the increasing mortality rates in ArLD which have been associated with delayed presentation, there is arguably now no reason to not proactively perform a non-

invasive fibrosis test in people at risk from chronic liver disease by drinking excess alcohol.

In Chapter 3 I investigated the most commonly used and reported non-invasive tests in ArLD in a systematic review and meta-analysis.

Compared to other published systematic reviews (204, 248, 321, 420), my systematic reviews in Chapters 3 and 7 have the largest sample of included participants, with all included studies having more than 30 patients each, and a combined sample of 1,268 for advanced fibrosis in Chapter 3 – larger than any previous meta-analysis on FibroScan in ArLD. My literature searches were comprehensive, with rigorous screening of texts, following Cochrane guidelines for conducting systematic reviews, and including use of a second reviewer to minimize reporting bias. In addition, efforts were made to contact authors where clarification of data was needed, or for request of more data, and references were hand-searched in addition to the database searches to minimize the chance of missing any relevant studies.

Whilst alcohol is the commonest cause of cirrhosis, it was perhaps surprising to find it was the least explored aetiology in terms of non-invasive testing, with the majority of studies either looking at viral hepatitis, NAFLD, or mixed aetiology cirrhosis. Mixing disease aetiologies introduces spectrum bias between fibrosis staging of different causes of liver disease, as I discussed in Chapter 2.

I found that ELF, FibroScan, FibroTest and FIB4 all perform well (AUROCs >0.7) for the detection of advanced fibrosis (≥F3) in ArLD. For F3 fibrosis the AUROC ranges for ELF (0.82-0.92), FibroScan (0.77-0.94) and FibroTest (0.80-0.90) were higher

than AUROCs for FIB4 (0.70-0.85), and significantly higher when directly compared by Thiele et al. in a large biopsy-paired non-invasive test study in ArLD. However, although FIB4 was the least well performing test, it still produced good AUROCs of between 0.70-0.85 in this systematic review (and AUROC 0.80 in my study reported in chapter 4). As it is a cheap and readily available test based on variables that can be calculated from simple routine blood tests, its potential use as a first-step rule-out test in primary care in ArLD, as has been investigated in NAFLD, warrants further exploration in prospective studies. Thresholds for all tests still need further validation in ArLD. FibroTest, ELF, and FIB4 all had only one study each that reported thresholds for use in advanced fibrosis and so even these studies require validation. FibroScan had thresholds reported for more studies, but still requires further validation, as different studies report different thresholds. A recent set of guidelines published by the Baveno working group in 2021 suggested a threshold of 8kPa to rule out advanced fibrosis, and 12kPa to rule in advanced fibrosis in ArLD (266). However, for several reasons this needs further exploration. The authors acknowledge that the guidelines are based on studies where information about alcohol consumption prior to the FibroScan was not documented, whilst it is known that alcohol consumption affects FibroScan results (206). In addition, only 17% of patients in this study had ArLD, with the majority of the included patients having viral hepatitis or NAFLD. The data for this study were collected retrospectively using previously published results from individual studies (all of which were included in my systematic review). In addition, Genesca et al., responding to this article (275), questioned the validity of the results, given lack of availability of the XL probe for obese patients for studies included in this summary study. Approximately one third

of the total cohort were obese, and as it is known that liver stiffness measurements can be inaccurate in obesity if the XL probe is not used (276), then the optimum derived thresholds may not be valid. Ideal FibroScan thresholds remain in debate for ArLD as evidenced in my systematic review (Chapter 3) that found different reported optimum threshold for all 11 different included FibroScan studies, preventing pooling of the data. I suggest that authors of future studies evaluate a uniform threshold in order for easier validation. Perhaps the best threshold to adopt for this until further validation would be the one recommended by Nguyen-Khac et al.'s meta-analysis on FibroScan in ArLD (243), which had access to individual study data of 1,026 patients, and recommended bilirubin and AST be factored in to the interpretation. Their suggested thresholds were 9kPa for F2, 12.1kPa for F3, and 18.6 kPa for F4 (243).

9.6.1 Impact of inflammation and alcohol on non-invasive test results

In my systematic review in Chapter 3, I found only one study (by Thiele et al.(23)) that compared ELF with histological inflammation in ArLD, with the authors finding a positive correlation (coefficient 0.18, $p < 0.001$ between ELF and NAS-CRN activity score (NAFLD Activity Score). Thiele et al. also found a positive correlation between NAS-CRN and TE (coefficient 1.93), and FibroTest (coefficient 0.02), both $p < 0.001$). However, in the two studies (including Thiele et al.) that compared transaminases as a surrogate marker for inflammation, no association was found between ELF and transaminases (23, 134). This is in keeping with my findings from Chapter 6, in my

prospective study of ELF in people with AUD in hospital: ELF score did not correlate with ALT or AST.

In AH, there is only surrogate inflammation data available at present, with my results from Chapter 8. In this chapter I explore the use of ELF as a prognostic marker in AH, and found ELF to significantly correlate with CRP (rho = 0.51, p <0.001) and AST (rho = 0.35, p <0.001). However, it is important to note that this was not a biopsy-paired study so I did not have access to histological results with which to compare ELF. I therefore cannot conclude from this study alone that ELF is influenced by inflammation in AH– as it is possible that the patients who had the highest AST and ALT values also had the highest levels of fibrosis and cirrhosis but I was not able to assess this.

To summarise, whilst ELF does not appear to be influenced by surrogate markers of inflammation (AST/ALT) in ArLD (not including AH), Thiele et al. found correlation with histological inflammation and this needs further validation in a biopsy-paired cohort of patients with AH.

The influence of inflammation on FibroScan performance has been better studied than in ELF. In my Chapter 3 systematic review, I found FibroScan to be influenced by inflammation (histological or transaminases) in ArLD in 5 studies out of 8 that investigated this (23, 133, 258, 322, 421).

9.6.2 Alcohol

In Chapter 6, in my prospective study of ELF in people with AUD in hospital, I found no correlation between ELF and amount of recent alcohol consumption. This is in keeping with findings from a 2021 study by Connoley et al. (134) and 2018 study by Thiele et al., (23) neither of which found any association of ELF with quantity of alcohol consumed. The fact that ELF does not appear to be influenced by alcohol consumption may confer an advantage over FibroScan, which has shown to produce false positive results in the context of excess alcohol and alcohol withdrawal (206, 422, 423).

9.7 Which non-invasive fibrosis tests can be used to predict prognosis in ArLD?

With the stage of fibrosis known to be the best predictor of mortality in ArLD, but biopsy not being recommended for routine use in all patients, I wanted to explore if non-invasive fibrosis tests have a role in predicting clinical outcomes in patients with ArLD. Having shown the performance for diagnosing liver fibrosis severity to be good for commonly used fibrosis tests: ELF, FibroScan, FIB4 and FibroTest, I detailed in Chapter 7 the results from a systematic review (aiming to investigate prognostic performance of non-invasive fibrosis tests in ArLD) finding 11 articles on these tests. All four tests could predict prognosis with AUROCs for outcome-prediction (including liver-related mortality, all-cause mortality and variceal bleeding) ranging from: 0.65-0.76 for FibroScan, 0.64-0.83 for FIB4, 0.69-0.79 for FibroTest and 0.72-0.85 for ELF. Whilst heterogeneity between studies precluded meta-analysis, where studies

reported direct comparisons between tests, FIB4 performed better than MELD in prognostication in two studies in ArLD (365, 424), and FibroTest outperformed Child-Pugh score (148). The ELF test was superior to histology in predicting outcomes in one study in ArLD (236), and FibroTest performed equally well to histology in another (148). Biopsy is not routinely performed in the management of ArLD, and my findings indicate that these commonly available non-invasive fibrosis tests can perform a useful role for assessing prognosis in clinical practice for patients with ArLD, avoiding the need for liver biopsy. They may even be superior to more traditional prognostic scores like MELD and Child-Pugh.

Finally, in Chapter 8, I investigated the role of ELF in prognosticating in AH, a condition that carries a mortality up to 60% for which there is currently no effective treatment that impacts on medium to long-term outcomes (41). Currently used prognostic scores do not reflect underlying fibrosis severity, which is known to be an important factor affecting mortality – in fact fibrosis stage was the only independent predictor of prognosis in ArLD in a recent study (327). Whilst AH is thought of as an inflammatory condition, I noted in Chapter 8 that in a landmark study from 2014 where liver biopsy was used in AH, the severity of fibrosis/presence of cirrhosis was the strongest predictor of mortality, whereas neutrophilic inflammation was associated with improved survival (393). Current commonly used scores perform sub-optimally, with AUROCs below 0.8 (Maddrey's DF, MELD, ABIC, GAHS with 90-day AUCs of 0.670, 0.704, 0.726 and 0.713 respectively) (131).

My study on ELF presented in Chapter 8 is the first to report the use of ELF in prognosticating in AH. In a cohort of 162 patients with AH from the 'STOPAH' trial I

found not only that ELF is able to predict 90-day mortality (AUROC 0.72), but that when ELF is combined with ABIC, MELD or GAHS, it improves the performance of both the ELF score and the ABIC/MELD/GAHS score. The combination of ABIC with ELF performed the best, with AUROC 0.81 (95% CI 0.73-0.89). This was significantly higher than ABIC alone.

This is the first prognostic biomarker study I am aware of that has achieved AUROC of >0.80 for the prediction of 90-day mortality in AH, with other published biomarker studies achieving AUROCs of between 0.70-0.80 (131, 425, 426).

I also found that patients with both low-ABIC (<6.71) and low ELF score (<16.78) had a 100% 90-day survival rate, compared with 45% survival in those with high ABIC and high ELF ($p < 0.01$). The addition of use of the Lille score allowed further stratification such that patients with high ABIC and high ELF, with a Lille score <0.45 had a 90-day survival rate of 90% (n=9/10 survived), compared to 6/13 survival (46% survival) in the group with Lille $\geq$0.45 ($p = 0.047$). In those with high ABIC and low ELF (n=100), those with baseline sepsis (n =17/100) had a significantly worse outcome (survival rate 59%) compared to those without sepsis (83/100, survival rate 89%), $p = 0.001$.

If validated, ELF+ABIC as a single biomarker could be readily incorporated into clinical practice and used as a prognostic tool with superior performance to the existing available scores.

9.8.1 Limitations

I recognize that there are limitations in my studies. Three of my studies (Chapters 4, 6 and 8), would have benefitted from examination of paired biopsy samples. In Chapter 4, I relied on the composite clinical judgement of expert hepatologists for the diagnosis of advanced fibrosis and cirrhosis, rather than being able to have access to liver biopsy results. However, this clinic-based sample of real-world cases reflects the current practice where biopsy is not routinely indicated in the investigation of patients with ArLD. In Chapter 6, where I discovered missed opportunities for fibrosis testing in people with AUD who present to hospital, I found that a third of the patients had ELF scores ≥10.5 indicating advanced fibrosis. However, it would have been useful to have paired liver biopsy samples in this study as a more robust reference standard. The lack of biopsies reflected current NHS clinical practice where biopsy is not routinely performed. Whilst FibroScan was offered to all patients with elevated ELF scores (≥10.5), only 18/28 attended their appointments, with 1/18 not having a valid FibroScan reading. The small number of FibroScan results meant it was not possible to draw robust conclusions about the performance of FibroScan. although I found discordance between FibroScan and ELF of 58% in this cohort. The failure of many patients to attend their FibroScan appointments highlights even further the need for 'opportunistic' assessment of fibrosis in this patient group when they present to hospital or GP practices, as people suffering with AUD may find clinic attendance challenging. A test that can be readily incorporated in routine laboratory blood tests may be preferrable to a "stand alone" specialist tests such as FibroScan that requires dedicated equipment, a skilled operator and time.

In Chapter 8, in my study of ELF in AH in the STOPAH cohort it would have been useful to examine whether ELF correlated with fibrosis stage, degree of inflammation, or both in AH, especially as histological fibrosis and inflammation have been shown to be associated with clinical outcomes in AH and so may be considered to be the reference prognostic test in this context.

I also relied on self-reported alcohol intake in my studies, which may not always be accurate, with patients sometimes underestimating their alcohol intake (427). An additional complication is that the concept of 'one unit' or 'one drink' can vary from person to person, and even amongst the scientific community there is controversy around how many units there are in a standard alcoholic 'drink'. Whilst the AUDIT questionnaire is recommended for use in NICE guidelines, and by the WHO, through my work with local GPs I found that this was not something that was routinely completed either in primary or secondary care in the assessment of AUD, in my study sites.

9.8.2 Unanswered questions and areas for future research

There remains to be several unanswered questions, which I think would be interesting and useful to answer in future studies. I discuss them below:

1) What are the optimal thresholds for fibrosis staging in ArLD for FibroScan, FibroTest, and ELF, and can FIB4 be used at a low-threshold to rule out advanced fibrosis in primary care?

As I discussed in Chapter 3, it is difficult to compare the performance of non-invasive tests when the majority of studies seem to be evaluating different thresholds, or investigating the test in mixed-aetiology liver disease, when it is recognized that thresholds need to differ between aetiologies (266). In order for validation, future studies should adopt uniform disease-specific pre-defined thresholds, which for advanced fibrosis, I think should be: 9.8 (for ELF) (23), 12.1 kPa for FibroScan (321), 0.58 for FibroTest (23). FIB4 should be evaluated at a threshold of 1.45 in primary care (low-prevalence population) (23), and it would be useful to know if FIB4 could perform well in a two-step pathway with ELF in ArLD, as it has been demonstrated in NAFLD (14). An Evaluation of my Camden and Islington Alcohol pathway would be ideal for this as all patients on the pathway will have an ELF test, and available blood tests with which to calculate FIB4.

2) Is the reason that ELF is able to predict prognosis in AH because of its correlation with fibrosis stage?

I found that ELF correlated with AST and ALT in AH (Chapter 8), but also predicted 90-day mortality. In order to understand whether the performance of ELF in this situation is due to its ability to stage fibrosis, or predict inflammation, or both is unclear. As the study in Chapter 8 was not biopsy-paired this question was not able to be answered in my thesis, and needs further exploration of ELF in a biopsy-paired cohort with AH.

3) Is ELF influenced by inflammation in ArLD?

As I have detailed above, only one study (by Thiele et al.) (23) investigated this with biopsy-paired cohort with ArLD, in which there was a low-prevalence of inflammation (48% of patients (139/289) had none or minimal hepatic inflammation, with steatohepatitis being present in 28%). ELF therefore needs further evaluation in a cohort with available biopsy results, and in a population of patients with ArLD with higher expected levels of inflammation than in Thiele et al.'s paper.

4) Is ELF influenced by alcohol in ArLD?

Whilst several studies, mine included in Chapter 6, do not suggest an association, ideally this would be investigated as a primary outcome in a study, where ELF tests could be taken sequentially from patients during an extended period following presentation with alcohol withdrawal symptoms.

5) Can the use of ELF in a community alcohol pathway improve the early detection of advanced fibrosis and cirrhosis, reduce unnecessary referrals to hepatology and save costs? Would pathways such as this result in an improvement in liver related morbidity and mortality long-term?

The adoption of non-invasive fibrosis tests in NAFLD pathways (FIB4 and ELF) has been demonstrated to lead to 88% reduction in unnecessary referrals to hepatology and five-fold increase in the detection of advanced fibrosis, with significant cost savings. However, such impact of non-invasive fibrosis tests in an alcohol pathway has not yet been evaluated. I have described in Chapter 5 an optimal evaluation plan to answer these questions from the Camden and Islington alcohol pathway.

6) What is the optimal threshold of alcohol intake in the population, above which liver fibrosis testing should be performed?

This is a challenging question to address as ultimately the risk to the individual is compounded by factors such as genetic susceptibility, microbiome, sex, and drinking pattern (Chapter 2). With the adoption of 35 units per week (females) and 50 units per week (males) in BSG and NICE guidelines, above which to use non-invasive

fibrosis tests, it would be useful to additionally compare the performance of tests at these thresholds, with testing at a 14-35-unit threshold (in terms of false negative and false positive rates, impact on referrals to secondary care and cost-effectiveness.

7) Should people who have BAFLD (NAFLD + ArLD risk factors) have a lower threshold of alcohol intake above which to test for advanced fibrosis?

It would also be important to evaluate if alcohol thresholds (above which to test for advanced fibrosis) need reducing in people with overweight or obesity, as it has recently been described that people with a BMI >35 have double the risk to the liver with any given alcohol intake (278).

8) In AH, is there the possibility of finding a biomarker that can predict <10% survival chance?

This may enable better selection of patients for liver transplant which has a mortality rate of around 10%. The combination of ELF + ABIC also needs validating as prognostic marker in AH.

Whilst this is not an exhaustive list of questions, answering these would enable further progress in the diagnosis and management of the increasing population of patients with ArLD.

9.9.1 Final remarks

In this thesis I have explored the use of non-invasive fibrosis tests in ArLD, evaluating their current use in primary and secondary care, modelling the impact of simple non-invasive tests (APRI and FIB4) on the proportion of unnecessary referrals to hepatology, investigating their diagnostic and prognostic performance in ArLD, and in AH. Whilst I have highlighted some remaining unanswered questions about their performance, I think it is important to note that firstly, any of the non-invasive tests I evaluated would be preferable to none in the stratification of patients from primary to secondary care. In my studies I found that currently 84% of patients in primary care are not getting any fibrosis testing before referral with ArLD, and 100% of patients with AUD were not tested outside of hepatology clinics. This highlights the lack of diffusion and adoption of non-invasive fibrosis testing that has been recommended in national and international guidance for three years (1). Before attempting to identify the perfect non-invasive test, greater effort must be expended on trying to encourage the use of 'any' non-invasive test in the first instance.

Secondly, identifying the 'perfect' non-invasive test, that does not produce false positive or false negative results is not possible. This is because the current reference standard is liver biopsy, which is imperfect itself, with AUROCs for fibrosis staging when compared to liver resections not reaching > 0.9, and estimated sensitivities between 80-90%, therefore it does not make sense to be striving to each AUROCs > 0.90 for non-invasive tests for fibrosis.

It is important, instead, for non-invasive tests to be designed and interpreted with the clinical context in mind, and also bearing in mind factors that may influence the result – for example recent alcohol use in FibroScan should prompt repeat FibroScan 2 to 3 weeks after the first to check the result. In the context of inflammation, FibroScan results may be falsely high, and again need repeating if the inflammation is anticipated to resolve, or needs interpreting with caution, and in conjunction with other tests. Most importantly, as with any diagnostic test (for example radiological tests, or d-dimers for detecting thromboses), the result is more accurate when evaluated with knowledge of other clinical parameters and an estimated pre-test probability. When designing a non-invasive fibrosis test for use in a community setting where prevalence is lower, and pre-test probability is lower, emphasis should be placed on selecting a test and threshold of the test that ensure the test performs with high sensitivity to screen out the condition (e.g. advanced fibrosis), and if wanting to 'rule in' advanced fibrosis in the context of a higher pre-test probability and prevalence, as in a hepatology clinic, then a test with higher specificity, and thresholds, is preferable. The test also needs to be accessible and acceptable to GPs and patients, and blood-based tests are more practical for use in primary care than FibroScan.

9.9.2 Overall conclusions

In this thesis, I have highlighted significant shortcomings in the care of people with AUD, with zero patients presenting to hospital with AUD over 13 months having had a fibrosis test despite having had multiple prior hospital attendances. In addition, I have highlighted inefficiencies in the way patients with suspected ArLD

are referred to hepatology clinics from primary care, with the vast majority of referrals based on steatosis on ultrasound and abnormal LFTs, neither of which are sensitive nor specific for liver fibrosis. During the course of this PhD, I have analysed the diagnostic (Chapter 3) and prognostic (Chapter 7) performance of non-invasive tests in ArLD. I have demonstrated (in Chapter 4) that non-invasive tests can improve the detection of advanced fibrosis and reduce the proportion of unnecessary referrals to hepatology by 50%, and I have successfully helped design and set-up a new primary care pathway incorporating a non-invasive test (ELF) in alcohol referrals.

I hope that the pathway that I have created will improve patient care in ArLD and allow earlier detection of liver fibrosis that could positively influence outcomes, and that in the process of launching the pathway I have increased awareness to GPs and hospital clinicians about the need for pro-active non-invasive testing in people with risk factors. I also hope that the data I have presented in this thesis may contribute to initiatives to support hospital clinicians – for example on the acute medical units, to create pathways to increase the opportunistic detection of advanced fibrosis in people presenting to hospital.

Finally, with all the recognized limitations of liver biopsy, and the fact it is not widely used in AH, I have presented novel data demonstrating that ELF can accurately predict 90-day mortality in AH, with a similar performance to currently used scores MELD, ABIC, and GAHS. This is the first study on ELF in AH, and I have demonstrated that where a fibrosis test (ELF) is combined with a simple score based on blood tests and age (ABIC) to predict mortality in AH, the combined ELF-

ABIC score performs superiorly to traditional prognostic scores. This is the first study where the performance of a biomarker in predicting mortality in AH has reached an AUROC above 0.8. If validated, this could be readily adopted in routine clinical practice and used to more accurately predict prognosis in this cohort.

References

1. Newsome PN, Cramb R, Davison SM, Dillon JF, Foulerton M, Godfrey EM, et al. Guidelines on the management of abnormal liver blood tests. Gut. 2018;67(1):6-19.
2. Burton R, Henn C, Lavoie D, O'Connor R, Perkins C, Sweeney K. Public Health England: The Public Health burden of alcohol and the effectiveness and cost-effectiveness of alcohol control policies, an evidence review 2016 [Available from: https://www.gov.uk/government/publications/the-public-health-burden-of-alcohol-evidence-review].
3. British-Liver-Trust. Alcohol Related Liver Disease: Statistics 2019 [Available from: https://britishlivertrust.org.uk/information-and-support/living-with-a-liver-condition/liver-conditions/alcohol/].
4. Rice P, Drummond C. The price of a drink: the potential of alcohol minimum unit pricing as a public health measure in the UK. Br J Psychiatry. 2012;201(3):169-71.
5. Hydes T, Gilmore W, Sheron N, Gilmore I. Treating alcohol-related liver disease from a public health perspective. J Hepatol. 2019;70(2):223-36.
6. Health-And-Social-Care-Information-Centre. Statistics on alcohol in England 2016 [Available from: http://www.hscic.gov.uk/catalogue/PUB20999/alc-eng-2016-rep.pdf].
7. 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.
8. Gao B, Bataller R. Alcoholic liver disease: pathogenesis and new therapeutic targets. Gastroenterology. 2011;141(5):1572-85.
9. Teli MR, Day CP, Burt AD, Bennett MK, James OF. Determinants of progression to cirrhosis or fibrosis in pure alcoholic fatty liver. Lancet. 1995;346(8981):987-90.
10. Bruha R, Dvorak K, Petrtyl J. Alcoholic liver disease. World J Hepatol. 2012;4(3):81-90.
11. British-Liver-Trust. British Liver Trust 'statistics - liver disease crisis' 2019 [Available from: https://britishlivertrust.org.uk/about-us/media-centre/statistics/].
12. Williams R, Aspinall R, Bellis M, Camps-Walsh G, Cramp M, Dhawan A, 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. The Lancet. 2014;384(9958):1953-97.
13. Guss D, Sherigar J, Mohanty SR. Missed Diagnosis of Liver Cirrhosis Leads to Disparities in Care for Older Patients. Gastroenterology Res. 2018;11(5):333-9.
14. Srivastava A, Gailer R, Tanwar S, Trembling P, Parkes J, Rodger A, et al. Prospective evaluation of a primary care referral pathway for patients with non-alcoholic fatty liver disease. J Hepatol. 2019;71(2):371-8.
15. Lo GH, Lai KH, Cheng JS, Lin CK, Hsu PI, Chiang HT. Prophylactic banding ligation of high-risk esophageal varices in patients with cirrhosis: a prospective, randomized trial. J Hepatol. 1999;31(3):451-6.

16. Sarin SK, Guptan RK, Jain AK, Sundaram KR. A randomized controlled trial of endoscopic variceal band ligation for primary prophylaxis of variceal bleeding. Eur J Gastroenterol Hepatol. 1996;8(4):337-42.
17. Trinchet JC, Chaffaut C, Bourcier V, Degos F, Henrion J, Fontaine H, et al. Ultrasonographic surveillance of hepatocellular carcinoma in cirrhosis: a randomized trial comparing 3- and 6-month periodicities. Hepatology. 2011;54(6):1987-97.
18. Santi V, Trevisan F, Gramenzi A, Grignaschi A, Mirici-Cappa F, Del Poggio P, et al. Semianual surveillance is superior to annual surveillance for the detection of early hepatocellular carcinoma and patient survival. . J Hepatol. 2010;53(2):291-7.
19. Patel PJ, Connoley D, Rhodes F, Srivastava A, Rosenberg W. A review of the clinical utility of the Enhanced Liver Fibrosis test in multiple aetiologies of chronic liver disease. Ann Clin Biochem. 2019:4563219879962.
20. Sumida Y, Nakajima A, Itoh Y. Limitations of liver biopsy and non-invasive diagnostic tests for the diagnosis of nonalcoholic fatty liver disease/nonalcoholic steatohepatitis. World J Gastroenterol. 2014;20(2):475-85.
21. Nallagangula KS, Nagaraj SK, Venkataswamy L, Chandrappa M. Liver fibrosis: a compilation on the biomarkers status and their significance during disease progression. Future Sci OA. 2018;4(1):FSO250.
22. Wai CT, Greenson JK, Fontana RJ, Kalbfleisch J, Marrero JA, Conjeevaram H, et al. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. Hepatology. 2003;38:518-26.
23. Thiele M, Madsen BS, Hansen JF, Detlefsen S, Antonsen S, Krag A. Accuracy of the Enhanced Liver Fibrosis Test vs FibroTest, Elastography, and Indirect Markers in Detection of Advanced Fibrosis in Patients With Alcoholic Liver Disease. Gastroenterology. 2018;154(5):1369-79.
24. Nyblom H, Berggren U, Balldin J, Olsson R. High AST/ALT ratio may indicate advanced alcoholic liver disease rather than heavy drinking. Alcohol and Alcoholism. 2004;39:336-9.
25. Li Q, Lu C, Li W, Huang Y, Chen L. The gamma-glutamyl transpeptidase to platelet ratio for non-invasive assessment of liver fibrosis in patients with chronic hepatitis B and non-alcoholic fatty liver disease. Oncotarget. 2017;8(17):28641-9.
26. Lombardi R, Buzzetti E, Roccarina D, Tsochatzis EA. Non-invasive assessment of liver fibrosis in patients with alcoholic liver disease. World J Gastroenterol. 2015;21(39):11044-52.
27. Srinivasa Babu A, Wells ML, Teytelboym OM, Mackey JE, Miller FH, Yeh BM, et al. Elastography in Chronic Liver Disease: Modalities, Techniques, Limitations, and Future Directions. Radiographics. 2016;36(7):1987-2006.
28. Irvine KM, Wockner LF, Shanker M, Fagan KJ, Horsfall LU, Fletcher LM, 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):370-7.
29. Rosenberg WM, Voelker M, Thiel R, Becka M, Burt A, Schuppan D, et al. Serum markers detect the presence of liver fibrosis: a cohort study. Gastroenterology. 2004;127(6):1704-13.
30. Mayo MJ, Parkes J, Adams-Huet B, Combes B, Mills AS, Markin RS, et al. Prediction of clinical outcomes in primary biliary cirrhosis by serum enhanced liver fibrosis assay. Hepatology. 2008;48(5):1549-57.

31. Vesterhus M, Hov JR, Holm A, Schrumpf E, Nygard S, Godang K, et al. Enhanced liver fibrosis score predicts transplant-free survival in primary sclerosing cholangitis. Hepatology. 2015;62(1):188-97.
32. Guha IN, Parkes J, Roderick P, Chattopadhyay D, Cross R, Harris S, 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):455-60.
33. Tanwar S, Trembling PM, Hogan BJ, Parkes J, Harris S, Grant P, et al. Biomarkers of Hepatic Fibrosis in Chronic Hepatitis C: A Comparison of 10 Biomarkers Using 2 Different Assays for Hyaluronic Acid. J Clin Gastroenterol. 2017;51(3):268-77.
34. Trembling PM, Lampertico P, Parkes J, Tanwar S, Vigano M, Facchetti F, et al. Performance of Enhanced Liver Fibrosis test and comparison with transient elastography in the identification of liver fibrosis in patients with chronic hepatitis B infection. Journal of Viral Hepatitis. 2014;21(6):430-8.
35. Day J, Patel P, Parkes J, Rosenberg W. Derivation and Performance of Standardized Enhanced Liver Fibrosis (ELF) Test Thresholds for the Detection and Prognosis of Liver Fibrosis. The Journal of Applied Laboratory Medicine. 2019;3(5):815.
36. Srivastava A, Jong S, Gola A, Gailer R, Morgan S, Sennett K, et al. Cost-comparison analysis of FIB-4, ELF and fibroscan in community pathways for non-alcoholic fatty liver disease. Bmc Gastroenterology. 2019;19.
37. Guidance NIfHaCE. Cirrhosis in Over 16s: Assessment and Management. Cirrhosis in Over 16s: Assessment and Management. National Institute for Health and Care Excellence: Guidance. London2016.
38. NICE-Guidance. Cirrhosis in Over 16s: Assessment and Management. National Institute for Health and Care Excellence: Guidance. 2016 [Available from: https://www.nice.org.uk].
39. Cho EJ, Lee JH, Kim MY, Yoo JJ, Choi WM, Cho YY, et al. Which noninvasive hepatic fibrosis test most effectively predicts portal pressure and survival in alcoholic liver disease patients? Hepatology. 2013;58(4):987A.
40. Connoley D, Patel P, Hogan B, Tanwar S, Rhodes F, Parkes J, et al. The utility of the enhanced liver fibrosis test in alcoholic liver disease. J Hepatol. 2019;70 (1):e815.
41. Thursz MR, Richardson P, Allison M, Austin A, Bowers M, Day CP, et al. Prednisolone or pentoxifylline for alcoholic hepatitis. N Engl J Med. 2015;372(17):1619-28.
42. WHO. Global Status Report on Alcohol and Health 2018 2018 [Available from: https://apps.who.int/iris/handle/10665/274603].
43. The Lancet Public H. Failing to address the burden of alcohol. Lancet Public Health. 2020;5(6):e297.
44. Collaborators GBDA. Alcohol use and burden for 195 countries and territories, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet. 2018;392(10152):1015-35.
45. Buykx P, Li J, Gavens L, Hooper L, Lovatt M, Gomes de Matos E, et al. Public awareness of the link between alcohol and cancer in England in 2015: a population-based survey. BMC Public Health. 2016;16(1):1194.

46. NHS-Digital. Statistics on alcohol, England 2020 [Available from: https://www.gov.uk/government/statistics/statistics-on-alcohol-england-2020.
47. Rehm J, Baliunas D, Borges GL, Graham K, Irving H, Kehoe T, et al. The relation between different dimensions of alcohol consumption and burden of disease: an overview. Addiction. 2010;105(5):817-43.
48. Rehm J, Gmel GE, Sr., Gmel G, Hasan OSM, Imtiaz S, Popova S, et al. The relationship between different dimensions of alcohol use and the burden of disease-an update. Addiction. 2017;112(6):968-1001.
49. NICE. Alcohol Use Disorders: The NICE clinical guideline on diagnosis, assessment, and management of harmful drinking and alcohol dependence 2020 [Available from: https://www.nice.org.uk/guidance/cg115/evidence/full-guideline-136423405.
50. Pisinger VS, Bloomfield K, Tolstrup JS. Perceived parental alcohol problems, internalizing problems and impaired parent - child relationships among 71 988 young people in Denmark. Addiction. 2016;111(11):1966-74.
51. Wolfe JD. Maternal alcohol use disorders and depression in emerging adulthood: Examining the relevance of social ties, childhood adversity, and socioeconomic status. Psychiatry Res. 2017;257:441-5.
52. Mangiavacchi L, Piccoli L. Parental alcohol consumption and adult children's educational attainment. Econ Hum Biol. 2018;28:132-45.
53. Finan LJ, Simpson E, Schulz J, Ohannessian CM. Parental Problem Drinking and Emerging Adult Problem Behavior: The Moderating Role of Parental Support. J Child Fam Stud. 2018;27(4):1175-85.
54. Anda RF, Whitfield CL, Felitti VJ, Chapman D, Edwards VJ, Dube SR, et al. Adverse childhood experiences, alcoholic parents, and later risk of alcoholism and depression. Psychiatr Serv. 2002;53(8):1001-9.
55. British-Liver-Trust. Alcohol-related Liver Disease: Statistics 2021 [Available from: https://britishlivertrust.org.uk/about-us/media-centre/statistics/#alcohol.
56. Saunders J, Degenhardt L, Reed G, Poznyak V. Alcohol Use Disorders in ICD-11: Past, Present, and Future. Alcoholism: Clinical and Experimental Research. 2019;43:1617-31.
57. Brandish E. Alcohol Use Disorders. Medicine. 2020;48(12):754-6.
58. American-Psychiatric-Association. Diagnostic and Statistical Manual of Mental Disorders: DSM-V (5th ed.) 2013.
59. NICE. NICE clinical guideline [CG115]: Alcohol-use disorders: diagnosis, assessment and management of harmful drinking (high-risk drinking) and alcohol dependence 2011 [Available from: https://www.nice.org.uk/guidance/CG115].
60. Bradley KA, Bush KR, McDonell MB, Malone T, Fihn SD, Ambulatory Care Quality Improvement P. Screening for problem drinking : Comparison of CAGE and AUDIT. J Gen Intern Med. 1998;13(6):379-88.
61. Hodgson R, Alwyn T, John B, Thom B, Smith A. The FAST Alcohol Screening Test. Alcohol Alcohol. 2002;37(1):61-6.
62. Thursz M, Gual A, Lackner C, Mathurin P, Moreno C, Spahr L, et al. EASL Clinical Practice Guidelines: Management of alcohol-related liver disease. J Hepatol. 2018;69(1):154-81.

63. Pimpin L, Cortez-Pinto H, Negro F, Corbould E, Lazarus JV, Webber L, et al. Burden of liver disease in Europe: Epidemiology and analysis of risk factors to identify prevention policies. J Hepatol. 2018;69(3):718-35.
64. Shield K, Rylett M, Rehm J. Public health successes and missed opportunities. Trends in alcohol consumption and attributable mortality in the WHO European Region, 1990-2014. World Health Organization. Regional Office for Europe. 2016 [Available from: https://apps.who.int/iris/handle/10665/329489.
65. World-Health-Organization-regional-office-for-Europe. European health for all database (HFA-DB) 2014 [Available from: http://data.euro.who.int/hfadb/.
66. Sheron N, Gilmore I. Effect of policy, economics, and the changing alcohol marketplace on alcohol related deaths in England and Wales. BMJ. 2016;353:i1860.
67. Williams R, Ashton K, Aspinall R, Bellis MA, Bosanquet J, Cramp ME, et al. Implementation of the Lancet Standing Commission on Liver Disease in the UK. Lancet. 2015;386(10008):2098-111.
68. Williams R, Alexander G, Armstrong I, Baker A, Bhala N, Camps-Walsh G, et al. Disease burden and costs from excess alcohol consumption, obesity, and viral hepatitis: fourth report of the Lancet Standing Commission on Liver Disease in the UK. Lancet. 2018;391(10125):1097-107.
69. Williams R, Aithal G, Alexander GJ, Allison M, Armstrong I, Aspinall R, et al. Unacceptable failures: the final report of the Lancet Commission into liver disease in the UK. Lancet. 2020;395(10219):226-39.
70. Roscoe S, Pryce R, Buykx P, Gavens L, Meier PS. Is disinvestment from alcohol and drug treatment services associated with treatment access, completions and related harm? An analysis of English expenditure and outcomes data. Drug Alcohol Rev. 2021.
71. Burton R, Henn C, Lavoie D, O'Connor R, Perkins C, Sweeney K, et al. A rapid evidence review of the effectiveness and cost-effectiveness of alcohol control policies: an English perspective. Lancet. 2017;389(10078):1558-80.
72. Stockwell T, Zhao J, Giesbrecht N, Macdonald S, Thomas G, Wettlaufer A. The raising of minimum alcohol prices in Saskatchewan, Canada: impacts on consumption and implications for public health. Am J Public Health. 2012;102(12):e103-10.
73. Anderson P, O'Donnell A, Kaner E, Llopis EJ, Manthey J, Rehm J. Impact of minimum unit pricing on alcohol purchases in Scotland and Wales: controlled interrupted time series analyses. Lancet Public Health. 2021.
74. Wagenaar AC, Salois MJ, Komro KA. Effects of beverage alcohol price and tax levels on drinking: a meta-analysis of 1003 estimates from 112 studies. Addiction. 2009;104(2):179-90.
75. Wagenaar AC, Tobler AL, Komro KA. Effects of alcohol tax and price policies on morbidity and mortality: a systematic review. Am J Public Health. 2010;100(11):2270-8.
76. Corrao G, Rubbiati L, Bagnardi V, Zambon A, Poikolainen K. Alcohol and coronary heart disease: a meta-analysis. Addiction. 2000;95(10):1505-23.
77. European Association for the Study of the Liver. Electronic address eee, European Association for the Study of the L. EASL Clinical Practice Guidelines: Management of alcohol-related liver disease. J Hepatol. 2018;69(1):154-81.

78. Bagnardi V, Rota M, Botteri E, Tramacere I, Islami F, Fedirko V, et al. Alcohol consumption and site-specific cancer risk: a comprehensive dose-response meta-analysis. Br J Cancer. 2015;112(3):580-93.
79. Burton R, Sheron N. No level of alcohol consumption improves health. Lancet. 2018;392(10152):987-8.
80. UK-Department-of-Health. UK Chief Medical Officers' Low Risk Drinking Guidelines. (Accessed May 2020) 2016 [Available from: https://www.gov.uk/government/uploads/].
81. Glyn-Owen K, Bohning D, Parkes J, Roderick P, Buchanan R. The combined effect of alcohol and body mass index on risk of chronic liver disease: A systematic review and meta-analysis of cohort studies. Liver Int. 2021;41(6):1216-26.
82. Hart CL, Morrison DS, Batty GD, Mitchell RJ, Davey Smith G. Effect of body mass index and alcohol consumption on liver disease: analysis of data from two prospective cohort studies. BMJ. 2010;340:c1240.
83. Corrao G, Bagnardi V, Zambon A, Torchio P. Meta-analysis of alcohol intake in relation to risk of liver cirrhosis. Alcohol Alcohol. 1998;33(4):381-92.
84. Murray CJ, Richards MA, Newton JN, Fenton KA, Anderson HR, Atkinson C, et al. UK health performance: findings of the Global Burden of Disease Study 2010. Lancet. 2013;381(9871):997-1020.
85. Lelbach W. Quantitative aspects of drinking in alcoholic liver cirrhosis. . Alcoholic Liver Pathology. 1975;1:1-18.
86. Rehm J, Taylor B, Mohapatra S, Irving H, Baliunas D, Patra J, et al. Alcohol as a risk factor for liver cirrhosis: a systematic review and meta-analysis. Drug Alcohol Rev. 2010;29(4):437-45.
87. Roerecke M, Vafaei A, Hasan OSM, Chrystoja BR, Cruz M, Lee R, et al. Alcohol Consumption and Risk of Liver Cirrhosis: A Systematic Review and Meta-Analysis. Am J Gastroenterol. 2019;114(10):1574-86.
88. Simpson RF, Hermon C, Liu B, Green J, Reeves GK, Beral V, et al. Alcohol drinking patterns and liver cirrhosis risk: analysis of the prospective UK Million Women Study. Lancet Public Health. 2019;4(1):e41-e8.
89. Kamper-Jorgensen M, Gronbaek M, Tolstrup J, Becker U. Alcohol and cirrhosis: dose--response or threshold effect? J Hepatol. 2004;41(1):25-30.
90. Corrao G, Bagnardi V, Zambon A, La Vecchia C. A meta-analysis of alcohol consumption and the risk of 15 diseases. Prev Med. 2004;38(5):613-9.
91. Fairfield B, Schnabl B. Gut dysbiosis as a driver in alcohol-induced liver injury. JHEP Rep. 2021;3(2):100220.
92. Crabb DW, Im GY, Szabo G, Mellinger JL, Lucey MR. Diagnosis and Treatment of Alcohol-Associated Liver Diseases: 2019 Practice Guidance From the American Association for the Study of Liver Diseases. Hepatology. 2020;71(1):306-33.
93. Singal AK, Bataller R, Ahn J, Kamath PS, Shah VH. ACG Clinical Guideline: Alcoholic Liver Disease. Am J Gastroenterol. 2018;113(2):175-94.
94. Askgaard G, Gronbaek M, Kjaer MS, Tjonneland A, Tolstrup JS. Alcohol drinking pattern and risk of alcoholic liver cirrhosis: a prospective cohort study. J Hepatol. 2015;62(5):1061-7.
95. Dam MK, Flensborg-Madsen T, Eliasen M, Becker U, Tolstrup JS. Smoking and risk of liver cirrhosis: a population-based cohort study. Scand J Gastroenterol. 2013;48(5):585-91.

96. Chen Y, Yang F, Lu H, Wang B, Chen Y, Lei D, et al. Characterization of fecal microbial communities in patients with liver cirrhosis. Hepatology. 2011;54(2):562-72.
97. Addolorato G, Ponziani FR, Dionisi T, Mosoni C, Vassallo GA, Sestito L, et al. Gut microbiota compositional and functional fingerprint in patients with alcohol use disorder and alcohol-associated liver disease. Liver Int. 2020;40(4):878-88.
98. Jones L, Bates G, McCoy E, Bellis MA. Relationship between alcohol-attributable disease and socioeconomic status, and the role of alcohol consumption in this relationship: a systematic review and meta-analysis. BMC Public Health. 2015;15:400.
99. Jones L, McCoy E, Bates G, Bellis M, Sumnall H. Understanding the alcohol harm paradox in order to focus the development of interventions. Centre for public Health. [accessed 02 June 2021] 2015 [Available from: http://www.cph.org.uk/publication/understanding-the-alcohol-harm-paradox-in-order-to-focus-the-development-of-interventions/.]
100. Morleo M, Dedman D, O'Farrell I, Cook P, Michael B, Tocque K, et al. Alcohol-attributable hospital admissions; segmentaion series report 3. [Accessed 01.May. 2021] 2010 [Available from: http://www.ias.org.uk/uploads/pdf/Health%20Impacts%20docs/Alcohol%20-%20Series%203%20-%20May%202010-2.pdf.
101. Klatsky AL, Armstrong MA. Alcohol, smoking, coffee, and cirrhosis. Am J Epidemiol. 1992;136(10):1248-57.
102. Stroffolini T, Cotticelli G, Medda E, Niosi M, Del Vecchio-Blanco C, Addolorato G, et al. Interaction of alcohol intake and cofactors on the risk of cirrhosis. Liver Int. 2010;30(6):867-70.
103. Wadhawan M, Anand AC. Coffee and Liver Disease. J Clin Exp Hepatol. 2016;6(1):40-6.
104. Kennedy OJ, Roderick P, Buchanan R, Fallowfield JA, Hayes PC, Parkes J. Systematic review with meta-analysis: coffee consumption and the risk of cirrhosis. Aliment Pharmacol Ther. 2016;43(5):562-74.
105. Liangpunsakul S, Beaudoin JJ, Shah VH, Puri P, Sanyal AJ, Kamath PS, et al. Interaction between the patatin-like phospholipase domain-containing protein 3 genotype and coffee drinking and the risk for acute alcoholic hepatitis. Hepatol Commun. 2018;2(1):29-34.
106. Lackner C, Tiniakos D. Fibrosis and alcohol-related liver disease. J Hepatol. 2019;70(2):294-304.
107. Betts J, DeSaix P, Johnson E, Johnson J, Korol O, Kruse D, et al. Anatomy and Physiology: OpenStax; 2017.
108. Bedossa P, Paradis V. Liver extracellular matrix in health and disease. J Pathol. 2003;200(4):504-15.
109. Arriazu E, Ruiz de Galarreta M, Cubero FJ, Varela-Rey M, Perez de Obanos MP, Leung TM, et al. Extracellular matrix and liver disease. Antioxid Redox Signal. 2014;21(7):1078-97.
110. Friedman SL. Mechanisms of hepatic fibrogenesis. Gastroenterology. 2008;134(6):1655-69.
111. Friedman SL. Liver fibrosis: from mechanisms to treatment. Gastroenterol Clin Biol. 2007;31(10):812-4.

112. Theise ND. Histopathology of alcoholic liver disease. Clin Liver Dis (Hoboken). 2013;2(2):64-7.
113. Fleming KM, Aithal GP, Card TR, West J. The rate of decompensation and clinical progression of disease in people with cirrhosis: a cohort study. Aliment Pharmacol Ther. 2010;32(11-12):1343-50.
114. Pericleous M, Sarnowski A, Moore A, Fijten R, Zaman M. The clinical management of abdominal ascites, spontaneous bacterial peritonitis and hepatorenal syndrome: a review of current guidelines and recommendations. Eur J Gastroenterol Hepatol. 2016;28(3):e10-8.
115. Pedersen JS, Bendtsen F, Moller S. Management of cirrhotic ascites. Ther Adv Chronic Dis. 2015;6(3):124-37.
116. Roccarina D, Best LM, Freeman SC, Roberts D, Cooper NJ, Sutton AJ, et al. Primary prevention of variceal bleeding in people with oesophageal varices due to liver cirrhosis: a network meta-analysis. Cochrane Database Syst Rev. 2021;4:CD013121.
117. Vilstrup H, Amodio P, Bajaj J, Cordoba J, Ferenci P, Mullen KD, et al. Hepatic encephalopathy in chronic liver disease: 2014 Practice Guideline by the American Association for the Study of Liver Diseases and the European Association for the Study of the Liver. Hepatology. 2014;60(2):715-35.
118. Fattovich G, Stroffolini T, Zagni I, Donato F. Hepatocellular carcinoma in cirrhosis: incidence and risk factors. Gastroenterology. 2004;127(5 Suppl 1):S35-50.
119. European Association for the Study of the Liver. Electronic address eee, European Association for the Study of the L. EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma. J Hepatol. 2018;69(1):182-236.
120. Mitchell MC, Friedman LS, McClain CJ. Medical Management of Severe Alcoholic Hepatitis: Expert Review from the Clinical Practice Updates Committee of the AGA Institute. Clin Gastroenterol Hepatol. 2017;15(1):5-12.
121. Crabb DW, Bataller R, Chalasani NP, Kamath PS, Lucey M, Mathurin P, et al. Standard Definitions and Common Data Elements for Clinical Trials in Patients With Alcoholic Hepatitis: Recommendation From the NIAAA Alcoholic Hepatitis Consortia. Gastroenterology. 2016;150(4):785-90.
122. Rachakonda V, Bataller R, Duarte-Rojo A. Recent advances in alcoholic hepatitis. F1000Res. 2020;9.
123. Yin H, Hu M, Zhang R, Shen Z, Flatow L, You M. MicroRNA-217 promotes ethanol-induced fat accumulation in hepatocytes by down-regulating SIRT1. J Biol Chem. 2012;287(13):9817-26.
124. Parlesak A, Schafer C, Schutz T, Bode JC, Bode C. Increased intestinal permeability to macromolecules and endotoxemia in patients with chronic alcohol abuse in different stages of alcohol-induced liver disease. J Hepatol. 2000;32(5):742-7.
125. Forrest EH, Evans CD, Stewart S, Phillips M, Oo YH, McAvoy NC, et al. Analysis of factors predictive of mortality in alcoholic hepatitis and derivation and validation of the Glasgow alcoholic hepatitis score. Gut. 2005;54(8):1174-9.
126. Singh S, Murad MH, Chandar AK, Bongiorno CM, Singal AK, Atkinson SR, et al. Comparative Effectiveness of Pharmacological Interventions for Severe Alcoholic Hepatitis: A Systematic Review and Network Meta-analysis. Gastroenterology. 2015;149(4):958-70 e12.

127. Louvet A, Thursz MR, Kim DJ, Labreuche J, Atkinson SR, Sidhu SS, et al. Corticosteroids Reduce Risk of Death Within 28 Days for Patients With Severe Alcoholic Hepatitis, Compared With Pentoxifylline or Placebo-a Meta-analysis of Individual Data From Controlled Trials. Gastroenterology. 2018;155(2):458-68 e8.
128. Pavlov CS, Varganova DL, Casazza G, Tsochatzis E, Nikolova D, Gluud C. [Glucocorticosteroids for people with alcoholic hepatitis (Cochrane review)]. Ter Arkh. 2019;91(8):52-66.
129. Nguyen-Khac E, Thevenot T, Piquet MA, Benferhat S, Goria O, Chatelain D, et al. Glucocorticoids plus N-acetylcysteine in severe alcoholic hepatitis. N Engl J Med. 2011;365(19):1781-9.
130. MIMAH-Consortium. Minimising Mortality from Alcoholic hepatitis: Developing innovative strategies 2021 [Available from: https://www.mimah.org/.
131. Forrest EH, Atkinson SR, Richardson P, Masson S, Ryder S, Thursz MR, et al. Application of prognostic scores in the STOPAH trial: Discriminant function is no longer the optimal scoring system in alcoholic hepatitis. J Hepatol. 2018;68(3):511-8.
132. Mueller S, Millonig G, Sarovska L, Friedrich S, Reimann FM, Pritsch M, et al. Increased liver stiffness in alcoholic liver disease: differentiating fibrosis from steatohepatitis. World J Gastroenterol. 2010;16(8):966-72.
133. Voican CS, Louvet A, Trabut JB, Njike-Nakseu M, Dharancy S, Sanchez A, et al. Transient elastography alone and in combination with FibroTest((R)) for the diagnosis of hepatic fibrosis in alcoholic liver disease. Liver Int. 2017;37(11):1697-705.
134. Connoley D, Patel PJ, Hogan B, Tanwar S, Rhodes F, Parkes J, et al. The Enhanced Liver Fibrosis test maintains its diagnostic and prognostic performance in alcohol-related liver disease: a cohort study. BMC Gastroenterol. 2021;21(1):268.
135. Day J, Patel P, Parkes J, Rosenberg W. Derivation and Performance of Standardized Enhanced Liver Fibrosis (ELF) Test Thresholds for the Detection and Prognosis of Liver Fibrosis. J Appl Lab Med. 2019;3(5):815-26.
136. Verrill C, Markham H, Templeton A, Carr NJ, Sheron N. Alcohol-related cirrhosis--early abstinence is a key factor in prognosis, even in the most severe cases. Addiction. 2009;104(5):768-74.
137. Kaner EF, Beyer F, Dickinson HO, Pienaar E, Campbell F, Schlesinger C, et al. Effectiveness of brief alcohol interventions in primary care populations. Cochrane Database Syst Rev. 2007(2):CD004148.
138. Santi V, Trevisan F, Gramenzi A, Grignaschi A, Mirici-Cappa F, Del Poggio P, et al. Semianual surveillance is superior to annual surveillance for the detection of early hepatocellular carcinoma and patient survival. J Hepatol. 2010;53(2):291-7.
139. Hydes T, Gilmore W, Sheron N, Gilmore I. Treating alcohol-related liver disease from a public health perspective. J Hepatol. 2019;70(2):223-36.
140. Sheron N, Moore M, O'Brien W, Harris S, Roderick P. Feasibility of detection and intervention for alcohol-related liver disease in the community: the Alcohol and Liver Disease Detection study (ALDDeS). Br J Gen Pract. 2013;63(615):e698-705.
141. Eisenberg E, Konopniki M, Veitsman E, Kramskay R, Gaitini D, Baruch Y. Prevalence and characteristics of pain induced by percutaneous liver biopsy. Anesth Analg. 2003;96(5):1392-6, table of contents.

142. Filingeri V, Francioso S, Sforza D, Santopaolo F, Oddi FM, Tisone G. A retrospective analysis of 1.011 percutaneous liver biopsies performed in patients with liver transplantation or liver disease: ultrasonography can reduce complications? European Review for Medical and Pharmacological Sciences. 2016;20(17):3609-17.
143. Rockey DC, Caldwell SH, Goodman ZD, Nelson RC, Smith AD, American Association for the Study of Liver D. Liver biopsy. Hepatology. 2009;49(3):1017-44.
144. Kalambokis G, Manousou P, Vibhakorn S, Marelli L, Cholongitas E, Senzolo M, et al. Transjugular liver biopsy--indications, adequacy, quality of specimens, and complications--a systematic review. J Hepatol. 2007;47(2):284-94.
145. Theodossi A, Skene AM, Portmann B, Knill-Jones RP, Patrick RS, Tate RA, et al. Observer variation in assessment of liver biopsies including analysis by kappa statistics. Gastroenterology. 1980;79(2):232-41.
146. Bejarano PA, Koehler A, Sherman KE. Second opinion pathology in liver biopsy interpretation. Am J Gastroenterol. 2001;96(11):3158-64.
147. Hahm GK, Niemann TH, Lucas JG, Frankel WL. The value of second opinion in gastrointestinal and liver pathology. Arch Pathol Lab Med. 2001;125(6):736-9.
148. Naveau S, Gaude G, Asnacios A, Agostini H, Abella A, Barri-Ova N, et al. Diagnostic and Prognostic Values of Noninvasive Biomarkers of Fibrosis in Patients with Alcoholic Liver Disease. Hepatology. 2009;49(1):97-105.
149. Goodman ZD. Grading and staging systems for inflammation and fibrosis in chronic liver diseases. J Hepatol. 2007;47(4):598-607.
150. Bedossa P, Poynard T. An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group. Hepatology. 1996;24(2):289-93.
151. Brunt EM, Janney CG, Di Bisceglie AM, Neuschwander-Tetri BA, Bacon BR. Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. Am J Gastroenterol. 1999;94(9):2467-74.
152. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology. 2005;41(6):1313-21.
153. Scheuer PJ. Classification of chronic viral hepatitis: a need for reassessment. J Hepatol. 1991;13(3):372-4.
154. Ishak K, Baptista A, Bianchi L, Callea F, De Groote J, Gudat F, et al. Histological grading and staging of chronic hepatitis. J Hepatol. 1995;22(6):696-9.
155. Tiniakos D, Anstee Q, Burt A. MacSweens Pathology of the Liver: Elsevier; 2017.
156. Yip WW, Burt AD. Alcoholic liver disease. Semin Diagn Pathol. 2006;23(3-4):149-60.
157. Lackner C, Tiniakos D. Fibrosis and alcohol-related liver disease. J Hepatol. 2019;70(2):294-304.
158. Janssens F, De Suray N, Piessevaux H, Horsmans Y, De Timary P, Starkel P. Can transient elastography replace liver histology for determination of advanced fibrosis in alcoholic patients: A real-life study. Journal of Clinical Gastroenterology. 2010;44(8):575-82.
159. Young Kim M, Koo Baik S, YeonPark, RaKim, Heon Hong J, WonJo. The comparison of the usefulness in prediction of advanced fibrosis between the

ultrasonographic scoring system and liver stiffness measurement in alcohol related liver disease. Hepatology. 2011;54:1233A.
160. Salavrakos M, Piessevaux H, Komuta M, Lanthier N, Starkel P. Fibroscan Reliably Rules Out Advanced Liver Fibrosis and Significant Portal Hypertension in Alcoholic Patients. Journal of Clinical Gastroenterology. 2019;53(10):772-8.
161. Nguyen-Khac E, Chatelain D, Tramier B, Decrombecque C, Robert B, Joly JP, et al. Assessment of asymptomatic liver fibrosis in alcoholic patients using fibroscan: prospective comparison with seven non-invasive laboratory tests. Alimentary Pharmacology & Therapeutics. 2008;28(10):1188-98.
162. Naveau S, Essoh BM, Ghinoiu M, Marthey L, Njike-Nakseu M, Balian A, et al. Comparison of Fibrotest and PGAA for the diagnosis of fibrosis stage in patients with alcoholic liver disease. European Journal of Gastroenterology & Hepatology. 2014;26(4):404-11.
163. Hien LTT. Assessment of liver fibrosis using fibroscan in patients with alcoholic liver disease. Gut. 2018;67 (Supplement 2):A91-A2.
164. Fern, ez M, Trepo E, Degre D, Gustot T, Verset L, et al. Transient elastography using Fibroscan is the most reliable noninvasive method for the diagnosis of advanced fibrosis and cirrhosis in alcoholic liver disease. European Journal of Gastroenterology & Hepatology. 2015;27(9):1074-9.
165. Lefkowitch JH. Morphology of alcoholic liver disease. Clin Liver Dis. 2005;9(1):37-53.
166. Tannapfel A, Denk H, Dienes HP, Langner C, Schirmacher P, Trauner M, et al. [Histopathological diagnosis of non-alcoholic and alcoholic fatty liver disease. Grade 2 consensus-based guidelines]. Pathologe. 2010;31(3):225-37.
167. Pavlov CS, Casazza G, Nikolova D, Tsochatzis E, Burroughs AK, Ivashkin VT, et al. Transient elastography for diagnosis of stages of hepatic fibrosis and cirrhosis in people with alcoholic liver disease. Cochrane Database Syst Rev. 2015;1:CD010542.
168. Srivastava A, Gailer R, Tanwar S, Trembling P, Parkes J, Rodger A, et al. Prospective evaluation of a primary care referral pathway for patients with non-alcoholic fatty liver disease. J Hepatol. 2019;71(2):371-8.
169. Bedossa P, Dargere D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. Hepatology. 2003;38(6):1449-57.
170. Ratziu V, Charlotte F, Heurtier A, Gombert S, Giral P, Bruckert E, et al. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. Gastroenterology. 2005;128(7):1898-906.
171. Regev A, Berho M, Jeffers LJ, Milikowski C, Molina EG, Pyrsopoulos NT, et al. Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection. Am J Gastroenterol. 2002;97(10):2614-8.
172. Rousselet MC, Michalak S, Dupre F, Croue A, Bedossa P, Saint-Andre JP, et al. Sources of variability in histological scoring of chronic viral hepatitis. Hepatology. 2005;41(2):257-64.
173. Afdhal NH, Curry M. Technology evaluation: a critical step in the clinical utilization of novel diagnostic tests for liver fibrosis. J Hepatol. 2007;46(4):543-5.
174. Mehta SH, Lau B, Afdhal NH, Thomas DL. Exceeding the limits of liver histology markers. J Hepatol. 2009;50(1):36-41.
175. Papastergiou V, Tsochatzis E, Burroughs AK. Non-invasive assessment of liver fibrosis. Ann Gastroenterol. 2012;25(3):218-31.

176. Poynard T, Halfon P, Castera L, Munteanu M, Imbert-Bismut F, Ratziu V, et al. Standardization of ROC curve areas for diagnostic evaluation of liver fibrosis markers based on prevalences of fibrosis stages. Clin Chem. 2007;53(9):1615-22.
177. Suk KT, Kim DY, Sohn KM, Kim DJ. Biomarkers of liver fibrosis. Adv Clin Chem. 2013;62:33-122.
178. Nguyen-Khac E, Chatelain D, Tramier B, Decrombecque C, Robert B, Joly JP, et al. Assessment of asymptomatic liver fibrosis in alcoholic patients using fibroscan: Prospective comparison with seven non-invasive laboratory tests. Alimentary Pharmacology and Therapeutics. 2008;28(10):1188-98.
179. Stickel F, Poeschl G, Schuppan D, Conradt C, Strenge-Hesse A, Fuchs FS, et al. Serum hyaluronate correlates with histological progression in alcoholic liver disease. European Journal of Gastroenterology & Hepatology. 2003;15(9):945-50.
180. Lieber CS, Weiss DG, Paronetto F. Value of fibrosis markers for staging liver fibrosis in patients with precirrhotic alcoholic liver disease. Alcoholism: Clinical and Experimental Research. 2008;32(6):1031-9.
181. Gabrielli GB, Faccioli G, Casaril M, Capra F, Bonazzi L, Falezza G, et al. Procollagen III peptide and fibronectin in alcohol-related chronic liver disease: correlations with morphological features and biochemical tests. Clinica Chimica Acta. 1989;179(3):315-22.
182. Deng H, Qi X, Guo X. Diagnostic Accuracy of APRI, AAR, FIB-4, FI, King, Lok, Forns, and FibroIndex Scores in Predicting the Presence of Esophageal Varices in Liver Cirrhosis: A Systematic Review and Meta-Analysis. Medicine (Baltimore). 2015;94(42):e1795.
183. Imbert-Bismut F, Ratziu V, Pieroni L, Charlotte F, Benhamou Y, Poynard T, et al. Biochemical markers of liver fibrosis in patients with hepatitis C virus infection: a prospective study. Lancet. 2001;357(9262):1069-75.
184. Adams LA, Bulsara M, Rossi E, DeBoer B, Speers D, George J, et al. Hepascore: an accurate validated predictor of liver fibrosis in chronic hepatitis C infection. Clin Chem. 2005;51(10):1867-73.
185. Patel K, Gordon SC, Jacobson I, Hezode C, Oh E, Smith KM, et al. Evaluation of a panel of non-invasive serum markers to differentiate mild from moderate-to-advanced liver fibrosis in chronic hepatitis C patients. J Hepatol. 2004;41(6):935-42.
186. Koda M, Matunaga Y, Kawakami M, Kishimoto Y, Suou T, Murawaki Y. FibroIndex, a practical index for predicting significant fibrosis in patients with chronic hepatitis C. Hepatology. 2007;45(2):297-306.
187. Cales P, Halfon P, Batisse D, Carrat F, Perre P, Penaranda G, et al. Comparison of liver fibrosis blood tests developed for HCV with new specific tests in HIV/HCV co-infection. J Hepatol. 2010;53(2):238-44.
188. Stickel F, Datz C, Hampe J, Bataller R. Pathophysiology and Management of Alcoholic Liver Disease: Update 2016. Gut Liver. 2017;11(2):173-88.
189. Srivastava A, Jong S, Gola A, Gailer R, Morgan S, Sennett K, et al. Cost-comparison analysis of FIB-4, ELF and fibroscan in community pathways for non-alcoholic fatty liver disease. BMC Gastroenterology. 2019;19(1):122.
190. Johnston DE. Special considerations in interpreting liver function tests. Am Fam Physician. 1999;59(8):2223-30.

191. Diehl A, Potter J, Boitnott JK, Van Duyn MA, Herlong HF, Mezey E. Relationship between pyridoxal 5'-phosphate deficiency and aminotransferase levels in alcohol hepatitis. . Gastroenterology. 1984;85:632-6.
192. Nyblom H, Berggren U, Balldin J, Olsson R. High AST/ALT ratio may indicate advanced alcoholic liver disease rather than heavy drinking. . Alcohol and Alcoholism. 2004;39:336-9.
193. Moreno C, Mueller S, Szabo G. Non-invasive diagnosis and biomarkers in alcohol-related liver disease. J Hepatol. 2019;70(2):273-83.
194. Weissmann B, Meyer K, Sampson P, Linker A. Isolation of oligosaccharides enzymatically produced from hyaluronic acid. J Biol Chem. 1954;208(1):417-29.
195. Patel PJ, Connoley D, Rhodes F, Srivastava A, Rosenberg W. A review of the clinical utility of the Enhanced Liver Fibrosis test in multiple aetiologies of chronic liver disease. Ann Clin Biochem. 2020;57(1):36-43.
196. Gudowska M, Gruszewska E, Panasiuk A, Cylwik B, Swiderska M, Flisiak R, et al. High serum N-terminal propeptide of procollagen type III concentration is associated with liver diseases. Prz Gastroenterol. 2017;12(3):203-7.
197. Benyon RC, Arthur MJ. Extracellular matrix degradation and the role of hepatic stellate cells. Semin Liver Dis. 2001;21(3):373-84.
198. Li J, Rosman AS, Leo MA, Nagai Y, Lieber CS. Tissue inhibitor of metalloproteinase is increased in the serum of precirrhotic and cirrhotic alcoholic patients and can serve as a marker of fibrosis. Hepatology. 1994;19(6):1418-23.
199. Sherman KE, Abdel-Hameed EA, Ehman RL, Rouster SD, Campa A, Martinez SS, et al. Validation and Refinement of Noninvasive Methods to Assess Hepatic Fibrosis: Magnetic Resonance Elastography Versus Enhanced Liver Fibrosis Index. Dig Dis Sci. 2020;65(4):1252-7.
200. Poynard T, Imbert-Bismut F, Munteanu M, Messous D, Myers RP, Thabut D, et al. Overview of the diagnostic value of biochemical markers of liver fibrosis (FibroTest, HCV FibroSure) and necrosis (ActiTest) in patients with chronic hepatitis C. Comp Hepatol. 2004;3(1):8.
201. Sharma S, Khalili K, Nguyen GC. Non-invasive diagnosis of advanced fibrosis and cirrhosis. World J Gastroenterol. 2014;20(45):16820-30.
202. Castera L, Forns X, Alberti A. Non-invasive evaluation of liver fibrosis using transient elastography. J Hepatol. 2008;48(5):835-47.
203. Tsochatzis EA, Gurusamy KS, Ntaoula S, Cholongitas E, Davidson BR, Burroughs AK. Elastography for the diagnosis of severity of fibrosis in chronic liver disease: a meta-analysis of diagnostic accuracy. J Hepatol. 2011;54(4):650-9.
204. Pavlov CS, Casazza G, Nikolova D, Tsochatzis E, Gluud C. Systematic review with meta-analysis: diagnostic accuracy of transient elastography for staging of fibrosis in people with alcoholic liver disease. Aliment Pharmacol Ther. 2016;43(5):575-85.
205. Voican CS, Louvet A, Trabut JB, Njike-Nakseu M, Dharancy S, Sanchez A, et al. Transient elastography alone and in combination with FibroTest^R for the diagnosis of hepatic fibrosis in alcoholic liver disease. Liver International. 2017;37(11):1697-705.
206. Gelsi E, Dainese R, Truchi R, Marine-Barjoan E, Anty R, Autuori M, et al. Effect of detoxification on liver stiffness assessed by Fibroscan(R) in alcoholic patients. Alcohol Clin Exp Res. 2011;35(3):566-70.

207. Castera L, Pinzani M. Biopsy and non-invasive methods for the diagnosis of liver fibrosis: does it take two to tango? Gut. 2010;59(7):861-6.
208. NICE-Guidance. FibroScan for assessing liver fibrosis and cirrhosis in primary care: Medtech innovation briefing 2020 [Available from: www.nice.org.uk/guidance/mib216.
209. Imajo K, Kessoku T, Honda Y, Tomeno W, Ogawa Y, Mawatari H, et al. Magnetic Resonance Imaging More Accurately Classifies Steatosis and Fibrosis in Patients With Nonalcoholic Fatty Liver Disease Than Transient Elastography. Gastroenterology. 2016;150(3):626-37.e7.
210. Kiani A, Brun V, Laine F, Turlin B, Morcet J, Michalak S, et al. Acoustic radiation force impulse imaging for assessing liver fibrosis in alcoholic liver disease. World J Gastroenterol. 2016;22(20):4926-35.
211. Zhang DK, Li P, Chen M, Liu LP, Liu Y, Zhao YY, et al. Non-invasive assessment of liver fibrosis in patients with alcoholic liver disease using acoustic radiation force impulse elastography. Abdominal Imaging. 2015;40(4):723-9.
212. Herrmann E, de Ledinghen V, Cassinotto C, Chu WC, Leung VY, Ferraioli G, et al. Assessment of biopsy-proven liver fibrosis by two-dimensional shear wave elastography: An individual patient data-based meta-analysis. Hepatology. 2018;67(1):260-72.
213. Beyer FR, Campbell F, Bertholet N, Daeppen JB, Saunders JB, Pienaar ED, et al. The Cochrane 2018 Review on Brief Interventions in Primary Care for Hazardous and Harmful Alcohol Consumption: A Distillation for Clinicians and Policy Makers. Alcohol Alcohol. 2019;54(4):417-27.
214. Subhani M, Knight H, Ryder S, Morling J. Does advice based on biomarkers of liver injury or non-invasive tests of liver fibrosis impact high-risk drinking behaviour: A systematic review with meta-analysis. Alcohol and Alcoholism. 2021;56:185-200.
215. Kaner EF, Beyer FR, Garnett C, Crane D, Brown J, Muirhead C, et al. Personalised digital interventions for reducing hazardous and harmful alcohol consumption in community-dwelling populations. Cochrane Database Syst Rev. 2017;9:CD011479.
216. C. A, Gillespie D, Ally A, Brennan A. Modelling the impact of Minimum Unit Price and Identification and Brief Advice polices using the Sheffield Alcohol Policy Model Version 3. (A Public Health England Commissioned Report). . 2015.
217. Ryder S. Does knowledge of liver fibrosis affect high risk drinking behaviour (KLIFAD)? A feasibility randomised controlled trial protocol lodged at: doi: 10.1186/ISRCTN16922410. . 2021.
218. Maharaj B, Maharaj RJ, Leary WP, Cooppan RM, Naran AD, Pirie D, et al. Sampling variability and its influence on the diagnostic yield of percutaneous needle biopsy of the liver. Lancet. 1986;1(8480):523-5.
219. Pavlov CS, Casazza G, Nikolova D, Tsochatzis E, Gluud C. Systematic review with meta-analysis: diagnostic accuracy of transient elastography for staging of fibrosis in people with alcoholic liver disease. Alimentary Pharmacology & Therapeutics. 2016;43(5):575-85.
220. Crossan C, Tsochatzis EA, Longworth L, Gurusamy K, Papastergiou V, Thalassinos E, et al. Cost-effectiveness of noninvasive liver fibrosis tests for

treatment decisions in patients with chronic hepatitis B in the UK: systematic review and economic evaluation. Journal of Viral Hepatitis. 2016;23(2):139-49.
221. Salkic NN, Jovanovic P, Hauser G, Brcic M. FibroTest/Fibrosure for significant liver fibrosis and cirrhosis in chronic hepatitis B: a meta-analysis. American Journal of Gastroenterology. 2014;109(6):796-809.
222. Chrostek L, Panasiuk A. Liver fibrosis markers in alcoholic liver disease. World Journal of Gastroenterology. 2014;20(25):8018-23.
223. Stevenson M, Lloyd-Jones M, Morgan MY, Wong R. Non-invasive diagnostic assessment tools for the detection of liver fibrosis in patients with suspected alcohol-related liver disease: a systematic review and economic evaluation. Health Technology Assessment (Winchester, England). 2012;16(4):1-174.
224. Rhodes FA, Trembling P, Panovska-Griffiths J, Tanwar S, Westbrook RH, Rodger A, et al. Systematic review: Investigating the prognostic performance of four non-invasive tests in alcohol-related liver disease. J Gastroenterol Hepatol. 2020.
225. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev. 2016;5(1):210.
226. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane Handbook for Systematic Reviews of Interventions version 6.0 Cochrane2019 [updated July 201917/03/2020]. Available from: www.training.cochrane.org/handbook].
227. Parkes J, Guha IN, Harris S, Rosenberg WMC, Roderick PJ. Systematic review of the diagnostic performance of serum markers of liver fibrosis in alcoholic liver disease. Comparative Hepatology. 2012;11(1).
228. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529-36.
229. Pavlov CS, Casazza G, Nikolova D, Tsochatzis E, Gluud C. Systematic review with meta-analysis: Diagnostic accuracy of transient elastography for staging of fibrosis in people with alcoholic liver disease. Alimentary Pharmacology and Therapeutics. 2016;43(5):575-85.
230. Review-Manager(RevMan). [Computer programme] Version 5.4. The Cochrane Collaboration. 2020.
231. Reiberger T, Ferlitsch A, Payer BA, Pinter M, Schwabl P, Stift J, et al. Noninvasive screening for liver fibrosis and portal hypertension by transient elastography--a large single center experience. Wiener Klinische Wochenschrift. 2012;124(11):395-402.
232. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557-60.
233. Trembling PM, Parkes J, Tanwar S, Burt AD, Rosenberg WM, Investigators ELF. ENHANCED LIVER FIBROSIS TEST ACCURATELY IDENTIFIES LIVER FIBROSIS AND PREDICTS CLINICAL OUTCOMES IN ALCOHOLIC LIVER DISEASE. J Hepatol. 2012;56:S424-S.
234. Connoley D, Patel P, Hogan B, Tanwar S, Rhodes F, Parkes J, et al. The utility of the enhanced liver fibrosis test in alcoholic liver disease. J Hepatol. 2019;70 (1 Supplement):e815.

235. Rosenberg WMC, Voelker M, Thiel R, Becka M, Burt A, Schuppan D, et al. Serum markers detect the presence of liver fibrosis: A cohort study. Gastroenterology. 2004;127(6):1704-13.
236. Connoley D, Patel P, Hogan B, Tanwar S, Rhodes F, Parkes J, et al. The Enhanced Liver Fibrosis test maintains its diagnostic and prognostic performance in alcohol-related liver disease. . (Manuscript in preparation). 2020.
237. Xie QS, Zhou XH, Huang PF, Wei JF, Wang WL, Zheng SS. The Performance of Enhanced Liver Fibrosis (ELF) Test for the Staging of Liver Fibrosis: A Meta-Analysis. Plos One. 2014;9(4).
238. Xia BQ, Wang FY, Friedrich-Rust M, Zhou F, Zhu JY, Yang H, et al. Feasibility and Efficacy of Transient Elastography using the XL probe to diagnose liver fibrosis and cirrhosis A meta-analysis. Medicine. 2018;97(39).
239. Tsochatzis EA, Gurusamy KS, Ntaoula S, Cholongitas E, Davidson BR, Burroughs AK. Elastography for the diagnosis of severity of fibrosis in chronic liver disease: A meta-analysis of diagnostic accuracy. J Hepatol. 2011;54(4):650-9.
240. Stevenson M, Lloyd-Jones M, Morgan MY, Wong R. Non-invasive diagnostic assessment tools for the detection of liver fibrosis in patients with suspected alcohol-related liver disease: A systematic review and economic evaluation. Health Technology Assessment. 2012;16(4).
241. Stebbing J, Farouk L, Panos G, Anderson M, Jiao LR, Mandalia S, et al. A Meta-analysis of transient elastography for the detection of hepatic fibrosis. Journal of Clinical Gastroenterology. 2010;44(3):214-9.
242. Poynard T, Morra R, Halfon P, Castera L, Ratziu V, Imbert-Bismut F, et al. Meta-analyses of FibroTest diagnostic value in chronic liver disease. BMC Gastroenterology. 2007;7 (no pagination).
243. Nguyen-Khac E, Thiele M, Voican C, Nahon P, Moreno C, Boursier J, et al. Non-invasive diagnosis of liver fibrosis in patients with alcohol-related liver disease by transient elastography: an individual patient data meta-analysis. The Lancet Gastroenterology and Hepatology. 2018;3(9):614-25.
244. Lv TT, Liu LX. A meta-analysis of fibroscan for the staging of liver fibrosis. Journal of Digestive Diseases. 2016;17 (Supplement 1):28-9.
245. Geng XX, Huang RG, Lin JM, Jiang N, Yang XX. Transient Elastography in Clinical Detection of Liver Cirrhosis: A Systematic Review and Meta-analysis. Saudi Journal of Gastroenterology. 2016;22(4):294-303.
246. Friedrich-Rust M, Ong MF, Martens S, Sarrazin C, Bojunga J, Zeuzem S, et al. Performance of transient elastography for the staging of liver fibrosis: A meta-analysis. Gastroenterology. 2008;134(4):960-74.
247. Bota S, Herkner H, Sporea I, Salzl P, Sirli R, Neghina AM, et al. Meta-analysis: ARFI elastography versus transient elastography for the evaluation of liver fibrosis. Liver International. 2013;33(8):1138-47.
248. Parkes J, Guha IN, Harris S, Rosenberg WM, Roderick PJ. Systematic review of the diagnostic performance of serum markers of liver fibrosis in alcoholic liver disease. Comp Hepatol. 2012;11(1):5.
249. Papatheodoridi M, Hiriart JB, Lupsor-Platon M, Bronte F, Boursier J, Elshaarawy O, et al. Refining the Baveno VI elastography criteria for the definition of compensated advanced chronic liver disease. J Hepatol. 2020.

250. Reiberger T, Ferlitsch A, Payer BA, Pinter M, Schwabl P, Stift J, et al. Noninvasive screening for liver fibrosis and portal hypertension by transient elastography - A large single center experience. Wiener Klinische Wochenschrift. 2012;124(11-12):395-402.
251. Janssens F, de Suray N, Piessevaux H, Horsmans Y, de Timary P, Starkel P. Can Transient Elastography Replace Liver Histology for Determination of Advanced Fibrosis in Alcoholic Patients A Real-life Study. Journal of Clinical Gastroenterology. 2010;44(8):575-82.
252. Voican CS, Louvet A, Trabut JB, Njike-Nakseu M, Dharancy S, Sanchez A, et al. Transient elastography alone and in combination with FibroTest for the diagnosis of hepatic fibrosis in alcoholic liver disease. Liver International. 2017;37(11):1697-705.
253. Naveau S, Essoh BM, Ghinoiu M, Marthey L, Njike-Nakseu M, Balian A, et al. Comparison of Fibrotest and PGAA for the diagnosis of fibrosis stage in patients with alcoholic liver disease. European Journal of Gastroenterology & Hepatology. 2014;26(4):404-11.
254. Nahon P, Kettaneh A, Tengher-Barna I, Ziol M, de Ledinghen V, Douvin C, et al. Assessment of liver fibrosis using transient elastography in patients with alcoholic liver disease. J Hepatol. 2008;49(6):1062-8.
255. Kim SG, Kim YS, Jung SW, Kim HK, Jang JY, Moon JH, et al. [The usefulness of transient elastography to diagnose cirrhosis in patients with alcoholic liver disease]. Korean Journal of Hepatology. 2009;15(1):42-51.
256. Fernandez M, Trepo E, Degre D, Gustot T, Verset L, Demetter P, et al. Transient elastography using Fibroscan is the most reliable noninvasive method for the diagnosis of advanced fibrosis and cirrhosis in alcoholic liver disease. European Journal of Gastroenterology & Hepatology. 2015;27(9):1074-9.
257. Cho Y, Choi YI, Oh S, Han J, Joo SK, Lee DH, et al. Point shear wave elastography predicts fibrosis severity and steatohepatitis in alcohol-related liver disease. Hepatol Int. 2020;14(2):270-80.
258. Mueller S, Millonig G, Sarovska L, Friedrich S, Reimann FM, Pritsch M, et al. Increased liver stiffness in alcoholic liver disease: differentiating fibrosis from steatohepatitis. World Journal of Gastroenterology. 2010;16(8):966-72.
259. Altman D. Practical statistics for medical research. . London: Chapman and Hall; 1991. 611 p.
260. Zhou Q, Lin S, Zhu Y. Comparison of transient elastography and MRI in diagnosis of liver cirrhosis. Hepatology International. 2017;11 (1 Supplement 1):S572.
261. Foucher J, Chanteloup E, Vergniol J, Castera L, Le Bail B, Adhoute X, et al. Diagnosis of cirrhosis by transient elastography (FibroScan): a prospective study. Gut. 2006;55(3):403-8.
262. Halfon P, Munteanu M, Poynard T. FibroTest-ActiTest as a non-invasive marker of liver fibrosis. Gastroenterol Clin Biol. 2008;32(6 Suppl 1):22-39.
263. Poynard T, Morra R, Halfon P, Castera L, Ratziu V, Imbert-Bismut F, et al. Meta-analyses of FibroTest diagnostic value in chronic liver disease. BMC Gastroenterol. 2007;7:40.

264. Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, Montaner J, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology. 2006;43(6):1317-25.
265. 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):1265-9.
266. Papatheodoridi M, Hiriart JB, Lupsor-Platon M, Bronte F, Boursier J, Elshaarawy O, et al. Refining the Baveno VI elastography criteria for the definition of compensated advanced chronic liver disease. J Hepatol. 2021;74(5):1109-16.
267. Bardou-Jacquet E, Legros L, Soro D, Latournerie M, Guillygomarc'h A, Le Lan C, et al. Effect of alcohol consumption on liver stiffness measured by transient elastography. World Journal of Gastroenterology. 2013;19(4):516-22.
268. de Ledinghen V, Wong VWS, Vergniol J, Wong GLH, Foucher J, Chu SHT, et al. Diagnosis of liver fibrosis and cirrhosis using liver stiffness measurement: Comparison between M and XL probe of FibroScan (R). J Hepatol. 2012;56(4):833-9.
269. Lannerstedt H, Konopski Z, Sandvik L, Haaland T, Loberg EM, Haukeland JW. Combining transient elastography with FIB4 enhances sensitivity in detecting advanced fibrosis of the liver. Scand J Gastroenterol. 2013;48(1):93-100.
270. Dolman GE, Nieboer D, Steyerberg EW, Harris S, Ferguson A, Zaitoun AM, et al. The performance of transient elastography compared to clinical acumen and routine tests - what is the incremental diagnostic value? Liver International. 2013;33(2):172-9.
271. Lieber CS, Weiss DG, Morgan TR, Paronetto F. Aspartate aminotransferase to platelet ratio index in patients with alcoholic liver fibrosis. American Journal of Gastroenterology. 2006;101(7):1500-8.
272. Angulo P, Bugianesi E, Bjornsson ES, Charatcharoenwitthaya P, Mills PR, Barrera F, et al. Simple noninvasive systems predict long-term outcomes of patients with nonalcoholic fatty liver disease. Gastroenterology. 2013;145(4):782-9 e4.
273. Pang JXQ, Pradhan F, Zimmer S, Niu S, Crotty P, Tracey J, et al. The feasibility and reliability of transient elastography using Fibroscan (R) : A practice audit of 2335 examinations. Canadian Journal of Gastroenterology and Hepatology. 2014;28(3):143-9.
274. Rhodes F, Cococcia S, Panovska-Griffiths J, Tanwar S, Westbrook RH, Rodger A, et al. Uncovering unsuspected advanced liver fibrosis in patients referred to alcohol nurse specialists using the ELF test. BMC Gastroenterol. 2021;21(1):143.
275. Genesca J, Abraldes JG, Bosch J. Do we need to re-define the baveno VI elastography criteria for compensataed advanced chronic liver disease (cACLD)? Letter to the editor in relation to "refining the baveno VI elastography criteria for the definition of compensated advanced chronic liver disease" by papatheodoridi et al. J Hepatol. 2021.
276. Wong VW, Irles M, Wong GL, Shili S, Chan AW, Merrouche W, et al. Unified interpretation of liver stiffness measurement by M and XL probes in non-alcoholic fatty liver disease. Gut. 2019;68(11):2057-64.
277. Miele L, De Michele T, Marrone G, Antonietta Isgro M, Basile U, Cefalo C, et al. Enhanced liver fibrosis test as a reliable tool for assessing fibrosis in nonalcoholic fatty liver disease in a clinical setting. Int J Biol Markers. 2017;32(4):e397-e402.

278. Glyn-Owen K, Bohning D, Parkes J, Roderick P, Buchanan R. The combined effect of alcohol and body mass index on risk of chronic liver disease: A systematic review and meta-analysis of cohort studies. Liver Int. 2020.
279. Rhodes FA, Cococcia S, Patel P, Panovska-Griffiths J, Tanwar S, Westbrook RH, et al. Is there scope to improve the selection of patients with alcohol-related liver disease for referral to secondary care? A retrospective analysis of primary care referrals to a UK liver centre, incorporating simple blood tests. BMJ Open. 2021;11(6):e047786.
280. 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):274-85.
281. Hussain A, Patel PJ, Rhodes F, Srivastava A, Patch D, Rosenberg W. Decompensated cirrhosis is the commonest presentation for NAFLD patients undergoing liver transplant assessment. Clin Med (Lond). 2020;20(3):313-8.
282. Ekstedt M, Franzen L, M H, Bendtsen P, Mathiesen UL, Bodemar G, et al. Alcohol consumption is associated with progression of hepatic fibrosis in non-alcoholic fatty liver disease. Scand J Gastroenterol. 2009;44(3):366-74.
283. Bataller R, Gao B. Liver fibrosis in alcoholic liver disease. Semin Liver Dis. 2015;35(2):146-56.
284. Becker U, Deis A, Sorensen TI, Gronbaek M, Borch-Johnsen K, Muller CF, et al. Prediction of risk of liver disease by alcohol intake, sex, and age: a prospective population study. Hepatology. 1996;23(5):1025-9.
285. Sanchez-Jimenez BA, Brizuela-Alcantara DC, Ramos-Ostos MH, Alva-lopez F, Uribe-Esquivel M, Chavez-Tapia N. Both alcoholic and non-alcoholic steatohepatitis association with cardiovascular risk and liver fibrosis. Alcohol. 2018;69:63-7.
286. Glyn-Owen K, Bohning D, Parkes J, Roderick P, Buchanan R. The combined effect of alcohol and obesity on risk of liver disease: a systematic review and meta-analysis. . Hepatology. 2019;70(S1):753A-A.
287. Chan WC, Treeprasertsuk S, Goh BBG, al. e. Optimising use of nonalcoholic fatty liver disease fibrosis score, Fibrosis-4 score, and liver stiffness measurement to identify patients with advanced fibrosis. Clin Gastroenterol Hepatol. 2019;17:2570-80.
288. Alberti KG, Eckel RH, Grundy SM, Zimmet PZ, Cleeman JI, Donato KA, 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. Circulation. 2009;120(16):1640-5.
289. Moreno C, Mueller S, Szabo G. Non-invasive diagnosis and biomarkers in alcohol-related liver disease. J Hepatol. 2019;70(2):273-83.
290. Lock CA, Kaner EF. Implementation of brief alcohol interventions by nurses in primary care: do non-clinical factors influence practice? Fam Pract. 2004;21(3):270-5.
291. Eggleston J, Gallagher J, Gallagher M, Hares T, Murray E, Naroz N, et al. Who should give lifestyle advice in general practice and what factors influence attendance at health promotion clinics? Survey of patients' views. Br J Gen Pract. 1995;45(401):669-71.

292. Trembling PM, Apostolidou S, Gentry-Maharaj A, Parkes J, Ryan A, Tanwar S, et al. Risk of chronic liver disease in post-menopausal women due to body mass index, alcohol and their interaction: a prospective nested cohort study within the United Kingdom Collaborative Trial of Ovarian Cancer Screening (UKCTOCS). BMC Public Health. 2017;17(1):603.
293. Gill MG, Majumdar A. Metabolic associated fatty liver disease: Addressing a new era in liver transplantation. World J Hepatol. 2020;12(12):1168-81.
294. Hagstrom H, Nasr P, Ekstedt M, Kechagias S, Onnerhag K, Nilsson E, et al. Low to moderate lifetime alcohol consumption is associated with less advanced stages of fibrosis in non-alcoholic fatty liver disease. Scandinavian Journal of Gastroenterology. 2017;52(2):159-65.
295. Corrao G, Bagnardi V, Zambon A, Arico S. Exploring the dose-response relationship between alcohol consumption and the risk of several alcohol-related conditions: a meta-analysis. Addiction. 1999;94(10):1551-73.
296. Patel PJ, Smith D, Connor JP, Horsfall LU, Hayward KL, Hossain F, et al. Alcohol Consumption in Diabetic Patients with Nonalcoholic Fatty Liver Disease. Can J Gastroenterol Hepatol. 2017;2017:7927685.
297. Srivastava A, Jong S, Gola A, Gailer R, Morgan S, Sennett K, et al. Cost-comparison analysis of FIB-4, ELF and fibroscan in community pathways for non-alcoholic fatty liver disease. BMC Gastroenterol. 2019;19(1):122.
298. Becker U, Gronbaek M, Johansen D, Sorensen TI. Lower risk for alcohol-induced cirrhosis in wine drinkers. Hepatology. 2002;35(4):868-75.
299. Blackwelder WC, Yano K, Rhoads GG, Kagan A, Gordon T, Palesch Y. Alcohol and mortality: the Honolulu Heart Study. Am J Med. 1980;68(2):164-9.
300. Fuchs CS, Stampfer MJ, Colditz GA, Giovannucci EL, Manson JE, Kawachi I, et al. Alcohol consumption and mortality among women. N Engl J Med. 1995;332(19):1245-50.
301. Schult A, Eriksson H, Wallerstedt S, Kaczynski J. Overweight and hypertriglyceridemia are risk factors for liver cirrhosis in middle-aged Swedish men. Scand J Gastroenterol. 2011;46(6):738-44.
302. NIHR-Involve. UK standards for Public Involvement: Better public involvement for better health and social care research 2019 [Available from: https://www.invo.org.uk/wp-content/uploads/2019/11/UK-standards-for-public-involvement-v6.pdf].
303. Public-Health-England. Alcohol Use Screening Tests 2017 [Available from: https://www.gov.uk/government/publications/alcohol-use-screening-tests].
304. Select-Stastistical-Services. Comparing Two Proportions – Sample Size 2021 [Available from: https://select-statistics.co.uk/calculators/sample-size-calculator-two-proportions/.
305. Wang H, Chow SC. Sample size calculation for comparing proportions. Wiley Encyclopedia of Clinical Trials 2007.
306. NHS-Digital. Hospital Episode Statistics for Admitted Patient Care and Outpatient Data 2021 [Available from: https://digital.nhs.uk/data-and-information/publications/statistical/hospital-episode-statistics-for-admitted-patient-care-outpatient-and-accident-and-emergency-data.
307. The-Health-Foundation. Non-COVID-19 NHS care during the pandemic: Activity trends for key NHS services in England 2021 [Available from:

https://www.health.org.uk/news-and-comment/charts-and-infographics/non-covid-19-nhs-care-during-the-pandemic.]
308. Tan JH, Tong J, Ho HH. Delayed presentation of acute coronary syndrome with mechanical complication during COVID-19 pandemic: a case report. Eur Heart J Case Rep. 2021;5(2):ytaa506.
309. British-Liver-Trust. Alcohol Related Liver Disease: Facts about Alcohol Related Liver Disease 2019 [Available from: https://britishlivertrust.org.uk/information-and-support/living-with-a-liver-condition/liver-conditions/alcohol/].
310. NHS-Digital. Statistics on Alcohol, England 2020 [Available from: https://digital.nhs.uk/data-and-information/publications/statistical/statistics-on-alcohol/2020].
311. Terg R, Fassio E, Guevara M, Cartier M, Longo C, Lucero R, et al. Ciprofloxacin in primary prophylaxis of spontaneous bacterial peritonitis: a randomized, placebo-controlled study. J Hepatol. 2008;48(5):774-9.
312. Harrison P, Hogan BJ, Floros L, Davies E, Guideline-Development-Group. Assessment and management of cirrhosis in people older than 16 years: summary of NICE guidance. BMJ. 2016;354:i2850.
313. Fagan KJ, Pretorius CJ, Horsfall LU, Irvine KM, Wilgen U, Choi K, et al. ELF score >= 9.8 indicates advanced hepatic fibrosis and is influenced by age, steatosis and histological activity. Liver International. 2015;35(6):1673-81.
314. Lichtinghagen R, Pietsch D, Bantel H, Manns MP, Brand K, Bahr MJ. The Enhanced Liver Fibrosis (ELF) score: Normal values, influence factors and proposed cut-off values. J Hepatol. 2013;59(2):236-42.
315. Parkes J, Guha IN, Roderick P, Harris S, Cross R, Manos MM, et al. Enhanced Liver Fibrosis (ELF) test accurately identifies liver fibrosis in patients with chronic hepatitis C. J Viral Hepat. 2011;18(1):23-31.
316. Poynard T, Ratziu V, Charlotte F, Goodman Z, McHutchison J, Albrecht J. Rates and risk factors of liver fibrosis progression in patients with chronic hepatitis c. J Hepatol. 2001;34(5):730-9.
317. McPherson S, Hardy T, Dufour JF, Petta S, Romero-Gomez M, Allison M, et al. Age as a Confounding Factor for the Accurate Non-Invasive Diagnosis of Advanced NAFLD Fibrosis. American Journal of Gastroenterology. 2017;112(5):740-51.
318. Wahl K, Rosenberg W, Vaske B, Manns MP, Schulze-Osthoff K, Bahr MJ, et al. Biopsy-Controlled Liver Fibrosis Staging Using the Enhanced Liver Fibrosis (ELF) Score Compared to Transient Elastography. Plos One. 2012;7(12).
319. Mueller S, Millonig G, Sarovska L, Friedrich S, Reimann FM, Pritsch M, et al. Increased liver stiffness in alcoholic liver disease: Differentiating fibrosis from steatohepatitis. World Journal of Gastroenterology. 2010;16(8):966-72.
320. Fernandez M, Trepo E, Degre D, Gustot T, Verset L, Demetter P, et al. Transient elastography using Fibroscan is the most reliable noninvasive method for the diagnosis of advanced fibrosis and cirrhosis in alcoholic liver disease. Eur J Gastroenterol Hepatol. 2015;27(9):1074-9.
321. Nguyen-Khac E, Thiele M, Voican C, Nahon P, Moreno C, Boursier J, et al. Non-invasive diagnosis of liver fibrosis in patients with alcohol-related liver disease

by transient elastography: an individual patient data meta-analysis. Lancet Gastroenterol Hepatol. 2018;3(9):614-25.
322. Fernandez M, Trepo E, Degre D, Gustot T, Verset L, Demetter P, et al. Transient elastography using Fibroscan is the most reliable noninvasive method for the diagnosis of advanced fibrosis and cirrhosis in alcoholic liver disease. European Journal of Gastroenterology & Hepatology. 2015;27(9):1074-9.
323. Trabut JB, Thepot V, Nalpas B, Lavielle B, Cosconea S, Corouge M, et al. Rapid decline of liver stiffness following alcohol withdrawal in heavy drinkers. Alcoholism: Clinical & Experimental Research. 2012;36(8):1407-11.
324. Ponomarenko Y, Leo MA, Kroll W, Lieber CS. Effects of alcohol consumption on eight circulating markers of liver fibrosis. Alcohol Alcohol. 2002;37(3):252-5.
325. Chapman R, Middleton J. The NHS long term plan and public health. BMJ. 2019;364:l218.
326. Robyn Burton CH, Don Lavoie et al. The Public Health Burden of Alcohol and the Effectiveness and Cost-Effectiveness of Alcohol Control Policies: An evidence review 2016 [Available from: https://www.gov.uk/government/publications/the-public-health-burden-of-alcohol-evidence-review].
327. Lackner C, Spindelboeck W, Haybaeck J, Douschan P, Rainer F, Terracciano L, et al. Histological parameters and alcohol abstinence determine long-term prognosis in patients with alcoholic liver disease. J Hepatol. 2017;66(3):610-8.
328. Sharma S, Khalili K, Nguyen GC. Non-invasive diagnosis of advanced fibrosis and cirrhosis. World Journal of Gastroenterology. 2014;20(45):16820-30.
329. Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. BMJ. 2009;338:b605.
330. Parkes J, Roderick P, Harris S, Day C, Mutimer D, Collier J, et al. Enhanced liver fibrosis test can predict clinical outcomes in patients with chronic liver disease. Gut. 2010;59(9):1245-51.
331. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane Handbook for Systematic Reviews of Interventions version 6.0 2019 [Available from: www.training.cochrane.org/handbook].
332. Hayden JA, van der Windt DA, Cartwright JL, Cote P, Bombardier C. Assessing bias in studies of prognostic factors. Ann Intern Med. 2013;158(4):280-6.
333. Shen Y, Wu SD, Wu L, Wang SQ, Chen Y, Liu LL, et al. The prognostic role of liver stiffness in patients with chronic liver disease: a systematic review and dose-response meta-analysis. Hepatology International. 2019.
334. NICE. The guidelines manual: Process and methods [PMG6] : appendices B-I, Appendix I: Methodology checklist: prognostic studies 2012 [Available from: https://www.nice.org.uk/process/pmg6/resources/the-guidelines-manual-appendices-bi-2549703709/chapter/appendix-i-methodology-checklist-prognostic-studies#notes-on-use-of-methodology-checklist-prognostic-studies].
335. Parkes J, Roderick P, Harris S, Day C, Mutimer D, Collier J, et al. Enhanced liver fibrosis test can predict clinical outcomes in patients with chronic liver disease. Gut. 2010;59(9):1245-51.
336. Hashim A, Haddadin Y, Macken L, Bremner S, Keller M, File A, et al. Can transient elastrography predict development of portal hypertension and or hepatic decompensation in individuals with cirrhosis? Gut. 2017;66 (Supplement 2):A49.

337. Cho E, Lee JG, Sohn JH, Jeong JY, Kim TY, Kim SM, et al. Combined Effect of hepatic venous pressure gradient and liver stiffness on long-term mortality in patients with cirrhosis. Journal of Gastroenterology and Hepatology. 2018;33:401-.
338. Hashim A, Parnell B, Haddadin Y, Macken L, Bremner S, Keller ME, et al. Does liver stiffness measurement and controlled attenuation parameter predict portal hypertension, liver related events and overall mortality in individuals with cirrhosis? Hepatology. 2017;66 (Supplement 1):263A-4A.
339. Day JW, Rosenberg WM, Parkes J. ELF test thresholds for disease stratification and prognosis in chronic liver disease. Hepatology. 2015;62:596A.
340. Hogan BJ, O'Beirne J, Patch DW, Yu D, Parisi I, Dhillon AP, et al. ELFTM predicts clinical outcomes in liver disease. Hepatology. 2015;62:604A-5A.
341. Stauber RE, Spindelboeck W, Putz-Bankuti C, Pock H, Stojakovic T, Obermayer-Pietsch B. Enhanced Liver Fibrosis (ELF) score predicts mortality in cirrhosis. Hepatology International. 2013;7:S492-S3.
342. Parkes J, Roderick P, Harris S, Gough C, Wheatley M, Alexander GJ, et al. European liver fibrosis (ELF) panel of serum markers can predict clinical outcome in a cohort of patients from England with mixed aetiology chronic liver disease. Hepatology. 2007;46(4):832A-A.
343. Adler M, Larocca L, Trovato FM, Marcinkowski H, Pasha Y, Taylor-Robinson SD. Evaluating the risk of hepatocellular carcinoma in patients with prominently elevated liver stiffness measurements by FibroScan: a multicentre study. HPB. 2016;18(8):678-83.
344. Singh S, Fujii LL, Murad MH, Wang Z, Asrani S, Ehman R, et al. Liver stiffness measurement predicts risk of decompensation, hepatocellular cancer and mortality in patients with chronic liver diseases: A systematic review and meta-analysis. Hepatology. 2013;58(4):956A.
345. Munteanu M, Rudler M, Lebray P, Perazzo H, Massard J, Pais R, et al. Long term prognostic value of the FibroTest in patients with non-alcoholic-fatty-liver disease (NAFLD), compared to chronic hepatitis C (CHC), B (CHB), and alcoholic liver disease (ALD). J Hepatol. 2018;68 (Supplement 1):S99-S100.
346. Boursier J, Bertrais S, Marsault P, Rapilly P, Fouchard-Hubert I, Oberti F, et al. Prediction of mortality by fibroscan, Fibrometer and their combination in a cohort of 3,623 patients with chronic liver disease. Hepatology. 2011;54:1227A.
347. Ngo Y, Perazzo H, Munteanu M, Lebray P, Moussalli J, Thabut D, et al. Meta-Analysis of the prognostic value of liver fibrosis biomarkers. Hepatology. 2011;54:1225A.
348. Rouyer M, Royer L, Perazzo H, Munteanu M, Ngo Y, Luckina E, et al. Non-invasive prediction of upper gastrointestinal bleeding by rupture of esophageal varices (ROV) using liver and spleen stiffness (Aixplorer), liver stiffness by fibroscan and fibrotest. J Hepatol. 2013;58:S96.
349. Irvine KM, Wockner LF, Shanker M, Fagan KJ, Horsfall LU, Fletcher LM, et al. Predicting clinical outcomes in chronic liver disease: The ELF test is superior to histology and simple scores. Journal of Gastroenterology and Hepatology (Australia). 2015;30:104-5.
350. Kim JH, Choe WH, Kwon SY. Predictive value of MELD, APRI, and FIB-4 for mortality and development of HCC in cirrhotic patients. Hepatology International. 2016;10(1):S356.

351. Buechter M, Kahraman A, Manka P, Gerken G, Jochum C, Canbay A, et al. Spleen and Liver Stiffness Is Positively Correlated with the Risk of Esophageal Variceal Bleeding. Digestion. 2016;94(3):138-44.
352. Festi D, Colecchia A, Schiumerini R, Marzi L, Mandolesi D, Reggiani MLB. Spleen and liver stiffness measurement can predict clinical complications in compensated cirrhotic patients: A prospective study. Hepatology. 2012;56:926A.
353. Kumar A, Shastri SM, Kumar M, Sharma P, Garg H, Sarin SK. Transient elastography correlates with portal pressure and identifies variceal bleeders. Indian Journal of Gastroenterology. 2010;29(1):A96-A7.
354. Sporea I, Ratiu I, Sirli R, Popescu A, Bota S. Value of transient elastography for the prediction of variceal bleeding. World Journal of Gastroenterology. 2011;17(17):2206-10.
355. Klibansky DA, Mehta SH, Curry M, Nasser I, Challies T, Afdhal NH. Transient elastography for predicting clinical outcomes in patients with chronic liver disease. Journal of Viral Hepatitis. 2012;19(2):e184-e93.
356. Robic MA, Procopet B, Metivier S, Peron JM, Selves J, Vinel JP, et al. Liver stiffness accurately predicts portal hypertension related complications in patients with chronic liver disease: a prospective study. J Hepatol. 2011;55(5):1017-24.
357. Pang JXQ, Zimmer S, Niu S, Crotty P, Tracey J, Pradhan F, et al. Liver Stiffness by Transient Elastography Predicts Liver-Related Complications and Mortality in Patients with Chronic Liver Disease. Plos One. 2014;9(4).
358. Paternostro R, Ferlitsch M, Etschmaier A, Schwarzer R, Reiberger T, Mandorfer M, et al. Transient elastography related cut off parameters in patients with advanced chronic liver disease independently predict decompensation and mortality but lack prediction of varices. Hepatology. 2015;62:925A-6A.
359. Shen Y, Wu SD, Wu L, Wang SQ, Chen Y, Liu LL, et al. The prognostic role of liver stiffness in patients with chronic liver disease: a systematic review and dose-response meta-analysis. Hepatol Int. 2019;13(5):560-72.
360. Pons M, Simon-Talero M, Millan L, Ventura-Cots M, Santos B, Augustin S, et al. Basal values and changes of liver stiffness predict the risk of disease progression in compensated advanced chronic liver disease. Dig Liver Dis. 2016;48(10):1214-9.
361. Kim M, Lim Y, Baik SK, al. e. MELD score and liver stiffness are predictive for the development of acute decompensation that induce acute-on chronic liver failure. . Hepatology. 2015;62:1222A.
362. Kitson MT, Roberts SK, Colman JC, Paul E, Button P, Kemp W. Liver stiffness and the prediction of clinically significant portal hypertension and portal hypertensive complications. Scandinavian Journal of Gastroenterology. 2015;50(4):462-9.
363. Vergniol J, Foucher J, Terrebonne E, al. e. Prognostic value of transient elastography and non-invasive markers of fibrosis in patients with chronic liver disease: A prospective analysis of 4,935 person-years. . Hepatology. 2009;50.
364. Chang Y, Cho YK, Cho J, Jung HS, Yun KE, Ahn J, et al. Alcoholic and Nonalcoholic Fatty Liver Disease and Liver-Related Mortality: A Cohort Study. American Journal of Gastroenterology. 2019;114(4):620-9.
365. Chaudhari SR, Nigam N, Sharma CB, Rastogi A. Comparison of non-invasive fibrosis estimation scores to predict the outcome in decompensated alcoholic cirrhosis patients. Hepatology International. 2017;11 (1 Supplement 1):S204.

366. Raker JM, Sivananthan A, Heald E, Chhaya V, Shalabi A, Forton DM. Defining liver stiffness measurement cut-offs to predict mortality and complications in patients with cirrhosis of mixed etiology. Hepatology. 2016;64 (1 Supplement 1):326A.
367. Bertrais S, Couffon C, Rondeau L, Favre M, Oberti F, Fouchard-Hubert I, et al. PERFORMANCE OF NON-INVASIVE LIVER FIBROSIS TESTS IN PREDICTING DEATH RATE IN PATIENTS WITH ALCOHOLIC CIRRHOSIS. J Hepatol. 2012;56:S527-S8.
368. Mueller J, Rausch V, Silva I, Peccerella T, Piecha F, Dietrich C, et al. Survival in a 10 year prospective cohort of heavy drinkers: Liver stiffness is the best long-term prognostic parameter. J Hepatol. 2019;70 (1):e107.
369. Gomez A, Redondo B, Fraile M, Castano-Garcia A, Torner M, Dieguez MLG, et al. Usefulness of liver stiffness measurement by transient elastography for predicting complications in patients with alcoholic liver disease. J Hepatol.68:S638.
370. Hyun Kim J, Lee M, Woo Park S, Kang M, Kim M, Hoon Lee S, et al. Validation of modified fibrosis-4 index for predicting hepatocellular carcinoma in patients with compensated alcoholic liver cirrhosis. Medicine (United States). 2018;97(48).
371. Kothari HG, Gupta SJ, Gaikwad NR, Sankalecha TH, Samarth AR. Role of non-invasive markers in prediction of esophageal varices and variceal bleeding in patients of alcoholic liver cirrhosis from central India. The Turkish journal of gastroenterology : the official journal of Turkish Society of Gastroenterology. 2019;30(12):1036-43.
372. Kothari HG, Gupta SJ, Gaikwad NR, Sankalecha TH, Samarth AR. Role of non-invasive markers in prediction of esophageal varices and variceal bleeding in patients of alcoholic liver cirrhosis from central India. Turkish Journal of Gastroenterology. 2019;30(12):1036-43.
373. Lee JG, Sohn JH, Jeong JY, Kim TY, Kim SM, Cho YS, et al. Combined Effect of hepatic venous pressure gradient and liver stiffness on long-term mortality in patients with cirrhosis. Journal of Gastroenterology and Hepatology. 2018;33:401-.
374. Crossan C, Tsochatzis EA, Longworth L, Gurusamy K, Papastergiou V, Thalassinos E, et al. Cost-effectiveness of noninvasive liver fibrosis tests for treatment decisions in patients with chronic hepatitis B in the UK: systematic review and economic evaluation. J Viral Hepat. 2016;23(2):139-49.
375. Poynard T, Ngo Y, Perazzo H, Munteanu M, Lebray P, Moussalli J, et al. Prognostic value of liver fibrosis biomarkers: a meta-analysis. Gastroenterol Hepatol (N Y). 2011;7(7):445-54.
376. Cochrane-MethodS-Group. Cochrane Methods:Prognosis. 'What is prognosis research?' [Available from: https://methods.cochrane.org/prognosis/about-us].
377. Debray TPA, de Jong VMT, Moons KGM, Riley RD. Evidence synthesis in prognosis research. Diagn Progn Res. 2019;3:13.
378. Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. PLoS Med. 2013;10(2):e1001381.
379. Hosseini N, Shor J, Szabo G. Alcoholic Hepatitis: A Review. Alcohol Alcohol. 2019;54(4):408-16.
380. Stickel F, Osterreicher CH. The role of genetic polymorphisms in alcoholic liver disease. Alcohol Alcohol. 2006;41(3):209-24.

381. Sarin SK, Pande A, Schnabl B. Microbiome as a therapeutic target in alcohol-related liver disease. J Hepatol. 2019;70(2):260-72.
382. Mendenhall CL, Moritz TE, Roselle GA, Morgan TR, Nemchausky BA, Tamburro CH, et al. Protein energy malnutrition in severe alcoholic hepatitis: diagnosis and response to treatment. The VA Cooperative Study Group #275. JPEN J Parenter Enteral Nutr. 1995;19(4):258-65.
383. Bellentani S, Saccoccio G, Costa G, Tiribelli C, Manenti F, Sodde M, et al. Drinking habits as cofactors of risk for alcohol induced liver damage. The Dionysos Study Group. Gut. 1997;41(6):845-50.
384. Lu XL, Luo JY, Tao M, Gen Y, Zhao P, Zhao HL, et al. Risk factors for alcoholic liver disease in China. World J Gastroenterol. 2004;10(16):2423-6.
385. Rotily M, Durbec JP, Berthezene P, Sarles H. Diet and alcohol in liver cirrhosis: a case-control study. Eur J Clin Nutr. 1990;44(8):595-603.
386. Thursz M, Forrest E, Roderick P, Day C, Austin A, O'Grady J, et al. The clinical effectiveness and cost-effectiveness of STeroids Or Pentoxifylline for Alcoholic Hepatitis (STOPAH): a 2 x 2 factorial randomised controlled trial. Health Technol Assess. 2015;19(102):1-104.
387. Im GY, Cameron AM, Lucey MR. Liver transplantation for alcoholic hepatitis. J Hepatol. 2019;70(2):328-34.
388. Shipley LC, Singal AK. Liver transplantation for alcoholic hepatitis. Transl Gastroenterol Hepatol. 2020;5:26.
389. Trebicka J, Sundaram V, Moreau R, Jalan R, Arroyo V. Liver Transplantation for Acute-on-Chronic Liver Failure: Science or Fiction? Liver Transpl. 2020;26(7):906-15.
390. Lee BP, Mehta N, Platt L, Gurakar A, Rice JP, Lucey MR, et al. Outcomes of Early Liver Transplantation for Patients With Severe Alcoholic Hepatitis. Gastroenterology. 2018;155(2):422-30 e1.
391. Lee BP, Samur S, Dalgic OO, Bethea ED, Lucey MR, Weinberg E, et al. Model to Calculate Harms and Benefits of Early vs Delayed Liver Transplantation for Patients With Alcohol-Associated Hepatitis. Gastroenterology. 2019;157(2):472-80 e5.
392. Baganate F, Beal EW, Tumin D, Azoulay D, Mumtaz K, Black SM, et al. Early mortality after liver transplantation: Defining the course and the cause. Surgery. 2018;164(4):694-704.
393. Altamirano J, Miquel R, Katoonizadeh A, Abraldes JG, Duarte-Rojo A, Louvet A, et al. A histologic scoring system for prognosis of patients with alcoholic hepatitis. Gastroenterology. 2014;146(5):1231-9 e1-6.
394. Arena U, Vizzutti F, Corti G, Ambu S, Stasi C, Bresci S, et al. Acute viral hepatitis increases liver stiffness values measured by transient elastography. Hepatology. 2008;47(2):380-4.
395. Coco B, Oliveri F, Maina AM, Ciccorossi P, Sacco R, Colombatto P, et al. Transient elastography: a new surrogate marker of liver fibrosis influenced by major changes of transaminases. J Viral Hepat. 2007;14(5):360-9.
396. Sagir A, Erhardt A, Schmitt M. Transient elastography is unreliable for detection of cirrhosis in patients with acute liver damage. . Hepatology. 2007;47:592-5.

397. Millonig G, Reimann FM, Friedrich S, Fonouni H, Mehrabi A, Buchler MW, et al. Extrahepatic cholestasis increases liver stiffness (FibroScan) irrespective of fibrosis. Hepatology. 2008;48(5):1718-23.
398. Wahl K, Rosenberg W, Vaske B, Manns MP, Schulze-Osthoff K, Bahr MJ, et al. Biopsy-controlled liver fibrosis staging using the enhanced liver fibrosis (ELF) score compared to transient elastography. PLoS One. 2012;7(12):e51906.
399. Forrest EH, Storey N, Sinha R, Atkinson SR, Vergis N, Richardson P, et al. Baseline neutrophil-to-lymphocyte ratio predicts response to corticosteroids and is associated with infection and renal dysfunction in alcoholic hepatitis. Aliment Pharmacol Ther. 2019;50(4):442-53.
400. Rahimi E, Pan JJ. Prognostic models for alcoholic hepatitis. Biomark Res. 2015;3:20.
401. Louvet A, Naveau S, Abdelnour M, Ramond MJ, Diaz E, Fartoux L, et al. The Lille model: a new tool for therapeutic strategy in patients with severe alcoholic hepatitis treated with steroids. Hepatology. 2007;45(6):1348-54.
402. Forrest EH, Atkinson SR, Richardson P, Masson S, Ryder S, Thursz MR, et al. Reply to: "The long-term prognosis of alcoholic hepatitis is poor and independent of disease severity for patients surviving an acute episode". J Hepatol. 2018;68(6):1332.
403. Dominguez M, Rincon D, Abraldes JG, Miquel R, Colmenero J, Bellot P, et al. A new scoring system for prognostic stratification of patients with alcoholic hepatitis. Am J Gastroenterol. 2008;103(11):2747-56.
404. Vergis N, Atkinson SR, Knapp S, Maurice J, Allison M, Austin A, et al. In Patients With Severe Alcoholic Hepatitis, Prednisolone Increases Susceptibility to Infection and Infection-Related Mortality, and Is Associated With High Circulating Levels of Bacterial DNA. Gastroenterology. 2017;152(5):1068-77 e4.
405. Bajaj JS, O'Leary JG, Reddy KR, Wong F, Olson JC, Subramanian RM, et al. Second infections independently increase mortality in hospitalized patients with cirrhosis: the North American consortium for the study of end-stage liver disease (NACSELD) experience. Hepatology. 2012;56(6):2328-35.
406. Tanwar S, Trembling PM, Guha IN, Parkes J, Kaye P, Burt AD, et al. Validation of terminal peptide of procollagen III for the detection and assessment of nonalcoholic steatohepatitis in patients with nonalcoholic fatty liver disease. Hepatology. 2013;57(1):103-11.
407. Pares A, Deulofeu R, Gimenez A, Caballeria L, Bruguera M, Caballeria J, et al. Serum hyaluronate reflects hepatic fibrogenesis in alcoholic liver disease and is useful as a marker of fibrosis. Hepatology. 1996;24(6):1399-403.
408. Murawaki Y, Ikuta Y, Idobe Y, Kitamura Y, Kawasaki H. Tissue inhibitor of metalloproteinase-1 in the liver of patients with chronic liver disease. J Hepatol. 1997;26(6):1213-9.
409. Tenny S, Hoffman MR. Prevalence. StatPearls. Treasure Island (FL)2021.
410. ONS. Quarterly alcohol specific deaths in England and Wales: 2001 to 2019 registrations and Quarter 1 (Jan to Mar) to Quarter 3 (July to Sept) provisional registrations. 2021 [Available from: www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/quarterlyalcoholspecificdeathsinenglandandwales/2001to2019registrationsandquarter1jantomartoquarter3julytosept2020provisionalregistrations.].

411. Winstock AR, al. e. GDS COVID-19 Special edition: Key findings report. Global drug survey. 2020 [Available from: https://www.globaldrugsurvey.com/gds-covid-19-special-edition-key-findings-report/].
412. Limb M. Deaths from alcohol hit record high during 2020, show figures. BMJ. 2021;372:n317.
413. Public-Health-England. Monitoring alcohol consumption and harm during the COVID-19 pandemic: summary 2021 [Available from: https://www.gov.uk/government/publications/alcohol-consumption-and-harm-during-the-covid-19-pandemic/monitoring-alcohol-consumption-and-harm-during-the-covid-19-pandemic-summary].
414. Propper C, Stockton I, Stoye G. COVID-19 and disruptions to the health and social care of older people in England 2020 [Available from: https://ifs.org.uk/publications/15160#].
415. Finlay I, Gilmore I. Covid-19 and alcohol-a dangerous cocktail. BMJ. 2020;369:m1987.
416. Mahase E. Covid-19: EU states report 60% rise in emergency calls about domestic violence. BMJ. 2020;369:m1872.
417. Garnett C, Jackson S, Oldham M, Brown J, Steptoe A, Fancourt D. Factors associated with drinking behaviour during COVID-19 social distancing and lockdown among adults in the UK. Drug Alcohol Depend. 2021;219:108461.
418. Stewart C. Liver disease in the United Kingdom (UK) - Statistics & Facts 2021 [Available from: https://www.statista.com/topics/5619/liver-disease-in-the-uk/].
419. Santi V, Trevisani F, Gramenzi A, Grignaschi A, Mirici-Cappa F, Del Poggio P, et al. Semiannual surveillance is superior to annual surveillance for the detection of early hepatocellular carcinoma and patient survival. J Hepatol. 2010;53(2):291-7.
420. Stevenson M, Lloyd-Jones M, Morgan MY, Wong R. Non-invasive diagnostic assessment tools for the detection of liver fibrosis in patients with suspected alcohol-related liver disease: a systematic review and economic evaluation. Health Technol Assess. 2012;16(4):1-174.
421. Nahon P, Kettaneh A, Tengher-Barna I, Ziol M, de Ledinghen V, Douvin C, et al. Assessment of liver fibrosis using transient elastography in patients with alcoholic liver disease. J Hepatol. 2008;49(6):1062-8.
422. Trabut JB, Thepot V, Nalpas B, Lavielle B, Cosconea S, Corouge M, et al. Rapid decline of liver stiffness following alcohol withdrawal in heavy drinkers. Alcohol Clin Exp Res. 2012;36(8):1407-11.
423. Bardou-Jacquet E, Legros L, Soro D, Latournerie M, Guillygomarc'h A, Le Lan C, et al. Effect of alcohol consumption on liver stiffness measured by transient elastography. World J Gastroenterol. 2013;19(4):516-22.
424. Kothari HG, Gupta SJ, Gaikwad NR, Sankalecha TH, Samarth AR. Role of non-invasive markers in prediction of esophageal varices and variceal bleeding in patients of alcoholic liver cirrhosis from central India. Turk J Gastroenterol. 2019;30(12):1036-43.
425. Daswani R, Kumar A, Anikhindi SA, Sharma P, Singla V, Bansal N, et al. Predictors of 90-day mortality in patients with severe alcoholic hepatitis: Experience with 183 patients at a tertiary care center from India. Indian J Gastroenterol. 2018;37(2):141-52.

426. Soultati AS, Dourakis SP, Alexopoulou A, Deutsch M, Vasilieva L, Archimandritis AJ. Predicting utility of a model for end stage liver disease in alcoholic liver disease. World J Gastroenterol. 2006;12(25):4020-5.
427. Stockwell T, Zhao J, Greenfield T, Li J, Livingston M, Meng Y. Estimating under- and over-reporting of drinking in national surveys of alcohol consumption: identification of consistent biases across four English-speaking countries. Addiction. 2016;111(7):1203-13.