BMJ Open

Can routine blood tests be modelled to detect advanced liver disease in the community: model derivation and validation using UK primary and secondary care data

Theresa Hydes ,[1] Michael Moore ,[1] Beth Stuart,[1] Miranda Kim,[1] Fangzhong Su,[1] Colin Newell,[2,3] David Cable,[4] Alan Hales,[1,5] Nick Sheron[6]

To cite: Hydes T, Moore M, Stuart B, *et al*. Can routine blood tests be modelled to detect advanced liver disease in the community: model derivation and validation using UK primary and secondary care data. *BMJ Open* 2021;**11**:e044952. doi:10.1136/bmjopen-2020-044952

► Prepublication history and additional material for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2020-044952).

Received 17 September 2020
Revised 27 October 2020
Accepted 06 January 2021

For numbered affiliations see end of article.

Correspondence to
Professor Nick Sheron;
nicksheronwork@outlook.com

ABSTRACT

Objectives Most patients are unaware they have liver cirrhosis until they present with a decompensating event. We therefore aimed to develop and validate an algorithm to predict advanced liver disease (AdvLD) using data widely available in primary care.

Design, setting and participants Logistic regression was performed on routinely collected blood result data from the University Hospital Southampton (UHS) information systems for 16 967 individuals who underwent an upper gastrointestinal endoscopy (2005–2016). Data were used to create a model aimed at detecting AdvLD: 'CIRRhosis Using Standard tests' (CIRRUS). Prediction of a first serious liver event (SLE) was then validated in two cohorts of 394 253 (UHS: primary and secondary care) and 183 045 individuals (Care and Health Information Exchange (CHIE): primary care).

Primary outcome measures Model creation dataset: cirrhosis or portal hypertension. Validation datasets: SLE (gastro-oesophageal varices, liver-related ascites or cirrhosis).

Results In the model creation dataset, 931 SLEs were recorded (5.5%). CIRRUS detected cirrhosis or portal hypertension with an area under the curve (AUC) of 0.90 (95% CI 0.88 to 0.92). Overall, 3044 (0.8%) and 1170 (0.6%) SLEs were recorded in the UHS and CHIE validation cohorts, respectively. In the UHS cohort, CIRRUS predicted a first SLE within 5 years with an AUC of 0.90 (0.89 to 0.91) continuous, 0.88 (0.87 to 0.89) categorised (crimson, red, amber, green grades); and AUC 0.84 (0.82 to 0.86) and 0.83 (0.81 to 0.85) for the CHIE cohort. In patients with a specified liver risk factor (alcohol, diabetes, viral hepatitis), a crimson/red cut-off predicted a first SLE with a sensitivity of 72%/59%, specificity 87%/93%, positive predictive value 26%/18% and negative predictive value 98%/99% for the UHS/CHIE validation cohorts, respectively.

Conclusion Identification of individuals at risk of AdvLD within primary care using routinely available data may provide an opportunity for earlier intervention and prevention of liver-related morbidity and mortality.

Strengths and limitations of this study

- The 'CIRRhosis Using Standard tests' (CIRRUS) algorithm was designed using blood tests commonly requested routinely in general practice (albumin, creatinine, bilirubin, mean corpuscular volume, sodium, platelet count, protein) to be able to identify individuals at risk of advanced liver disease (AdvLD) in the community without the need for specialist tests.
- Unlike most diagnostic liver algorithms, the CIRRUS model was created using a dataset comprised of patients from both primary and secondary care without the primary intent of preselecting for liver disease, rendering it better suited for detecting liver disease outside a secondary care hepatology environment.
- However, bias was inevitably introduced in the model creation cohort as the primary inclusion criteria for this dataset was referral for an upper gastrointestinal endoscopy, and subsequently, this population had a higher prevalence of AdvLD compared with the validation datasets.
- Not all individuals within the model creation dataset had undergone a full liver assessment so we could not completely exclude cirrhosis in the control group.
- The primary care dataset used for validation of the CIRRUS score only contained participants who are alive, missing patients who may have died prior to data extraction, even if they had a serious liver event; this would have led to a likely reduction of the positive predictive value for the performance of CIRRUS in this setting.

INTRODUCTION

Mortality rates for most diseases have reduced substantially in recent decades with the exception of liver disease, where mortality rates are four times higher than in the 1970s. Around three-quarters of fatal liver disease is detected when a patient presents as an emergency with complications of advanced liver disease (AdvLD).[1] This late presentation may be a

factor in the poor survival of liver admissions, currently around 40%.[1]

Cirrhosis, portal hypertension and liver failure are the result of progressive liver fibrosis which has no signs or symptoms. Historically, general practitioners have used serum liver enzymes to detect or rule out significant liver disease, most commonly an elevated alanine transaminase (ALT). The problem being that 90% of liver mortality occurs as a result of alcohol or obesity, where liver enzymes are frequently normal and are not an accurate guide of disease progression.[2]

Algorithms such as the Fibrosis-4 index, AST to platelet ratio, FibroTest and Enhanced Liver Fibrosis score can help stage liver fibrosis but require at least one non-standard test, for example, aspartate transaminase (AST), gamma glutamyl transferase (GGT), α2-macroglobulin, haptoglobin, apolipoprotein A1, hyaluronic acid, collagen p3 peptide or tissue inhibitor of metalloproteinase 1.[3] The recent UK NICE Guideline Development Group concluded that there are currently no validated tools to predict which individuals are at the highest risk of developing cirrhosis and recommend this as a subject for future research.[2]

The prime recommendation of the UK Lancet Liver Commission was to 'strengthen detection of liver disease at early stages and its treatment in Primary Care'.[1] A risk evaluation tool designed to predict a serious liver event (SLE) that uses routine blood results (which are often already available) and is not dependent on prior suspicion of liver disease could thus be significantly clinically useful.

METHODS

A retrospective analysis (5 April 2002–7 January 2018) of routinely collected anonymised National Health Service data from the University Hospital Southampton (UHS) and Care and Health Information Exchange (CHIE) was performed in order to create and then validate a risk score for AdvLD. Three datasets were used as described in figure 1. (1) The model creation dataset: data from 16 967 individuals extracted from UHS information systems (primary and secondary care) who sequentially underwent an oesophagogastroduodenoscopy (OGD). (2) Validation dataset 1: pseudoanonymised data from 394 253 patients identified by the UHS pathology system (primary and secondary care) cross-referenced with admission data from within the hospital domain. (3) Validation dataset 2: fully anonymised blood test, diagnosis and admissions data from 183 045 subjects in participating general practices from across Hampshire from the CHIE dataset (primary care). We did not have access to cause of death from death certificates for any of these datasets. Ethical approval was obtained as part of the LOCATE study[4] (South Central – Southampton A, 13/SC/0012), and data were examined in compliance with the UHS Data Governance Group and CHIE Information Governance Group.

Patient and public involvement

Patients or the public were not involved in the design, or conduct, or reporting or dissemination plans of our research.

Model creation dataset

To create a risk evaluation tool for AdvLD, we extracted blood result data from UHS information systems for 40 427 unselected mixed primary and secondary care subjects who sequentially underwent an OGD between 1 January 2005 and 1 November 2016. This study group was selected on the basis that it comprised a large mixed primary and secondary care population, not principally selected according to the presence of liver risk factors, yet with objective evidence of one aspect of AdvLD, namely, portal hypertension on a defined date. Of these subjects, 20 222 were within the age range of 18–70 and had at least one full set of blood test results (online supplemental table 1). The reference standards for AdvLD in this cohort were evidence of portal hypertension on OGD, or cirrhosis determined by liver biopsy, fibroscan >15 kPa, or liver fibrosis markers (hyaluronic acid and collagen P3 peptide) within the cirrhotic range.[5] Relevant matching data were extracted from linked information systems including hospital admission and pathology records. Subjects with liver disease, but undetermined cirrhosis or portal hypertension status, were not included, leaving 16 967 participants (online supplemental table 2). The fully linked dataset was anonymised before being received by the analysts.

The model creation cohort was split into two interchangeable creation and validation cohorts (cohort 0 and cohort 1) using SPSS uniform random number generation and entered into a logistic regression. Portal hypertension at endoscopy and the presence of cirrhosis were dependent variables, and mean routine blood test results (7–365 days prior to diagnosis) were independent variables. Independent variables were chosen on the basis that they are included in every routine biochemistry (sodium, potassium, creatinine, urea, ALT, alkaline phosphatase, bilirubin, total protein) or full blood count profile (haemoglobin, white cell count, platelet count, mean corpuscular volume). Urea and alkaline phosphatase were excluded because they did not form part of the routine primary care workup. Models for each of the two random cohorts were forward fitted, retaining variables which were significant at p=0.05 and which significantly increased the value of the scaled deviance. This approach is in line with that procedure recommended by Collett.[6] The model was named CIRRUS, corresponding to 'CIRRhosis Using Standard tests'.

Validation datasets

The performance of the CIRRUS model as a predictor of clinical risk was examined in two large validation datasets. The reference standard outcomes for the validation cohorts were the presence of liver disease and/or an SLE. These were determined by examining

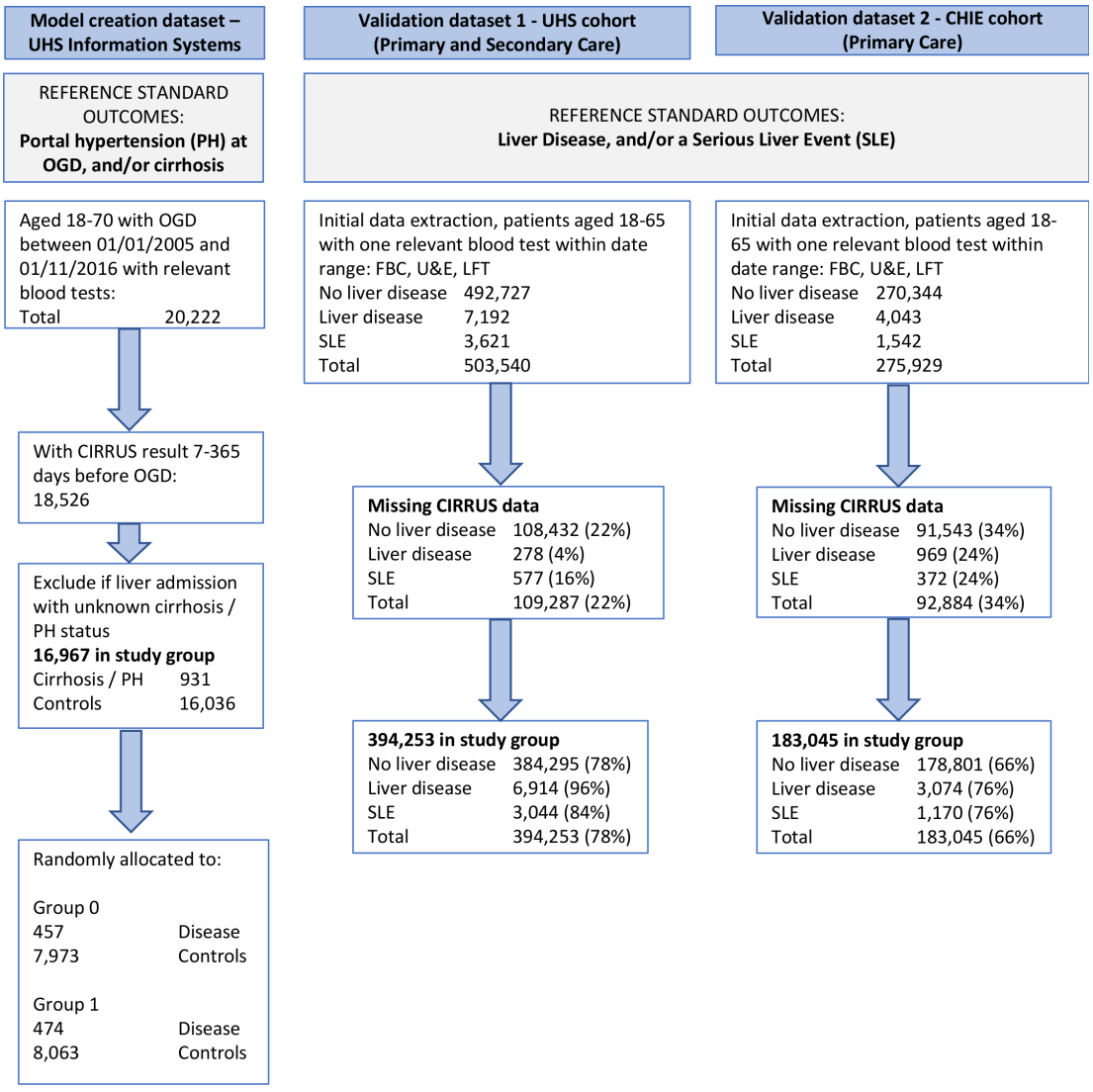


Figure 1 Flow chart showing participants and exclusions for all three datasets. Validation data were calculated for each date on which a full set of tests were available. Definition of cirrhosis (model creation dataset): cirrhosis diagnosed according to a liver biopsy, fibroscan result >15 kPa or serum fibrosis markers (hyaluronic acid and collagen P3 peptide) within the cirrhotic range. Serious liver event (SLE): a hospital admission with an ICD code for varices, liver-related ascites or cirrhosis. CHIE, Care and Health Information Exchange; CIRRUS, CIRRhosis Using Standard tests; OGD, oesophagogastroduodenoscopy; UHS, University Hospital Southampton.

International Classification of Disease (ICD)–10 codes of hospital admissions (online supplemental table 3). An SLE was defined as a hospital admission with either varices, liver-related ascites or cirrhosis. Assessors of the reference standards were blinded from the test results. For each dataset, patients aged 18–65 on the day of their blood test were selected on the basis that they had at least one full blood count, electrolyte or liver blood test result on the pathology record system. Blood results were excluded if they were taken after the diagnosis of the first SLE, but those taken on the day of diagnosis were included. An upper age limit of 65 was chosen for both cohorts to maximise the positive predictive value (PPV) and better power the study. This followed a sensitivity analysis which identified that the PPV of CIRRUS dropped to 0.12 above age 65 as a result of a reduced incidence of SLE in older age groups (online supplemental table 4).

Validation dataset 1 (UHS)

Pseudoanonymised blood test data were obtained for unselected patients from the UHS pathology database between 5 April 2002 and 17 January 2018. This comprised blood tests requested from both primary and secondary care. UHS provides services to around 1.3 million subjects in the South of England. Data were aggregated by the test day and merged with a dataset comprising information about admissions to UHS. The analysis dataset comprised 503 540 primary and secondary care patients aged 18–65 with UHS blood test results within the study period. Overall, 109 287 of 503 540 (22%) patients were excluded because they were missing CIRRUS data, of whom 577

out of 3621 (16%) went on to have a first SLE. The final study population comprised 394 253 patients.

Validation dataset 2 (CHIE)

Data were extracted from 379 279 patients from an estimated total population of 1 094 456 active on the CHIE database between 1 January 2013 and 1 July 2015,[7] who had either a liver blood test result or evidence of liver disease on Read coding. Retrospective blood test data from this cohort were available from 1 January 2007 onwards, and a further tranche of updated outcome data was collected in September 2016. While the CHIE dataset has the advantage of exclusively representing individuals seen in primary care, it is significantly limited by the fact it only contained data on individuals who are alive, that is, patients who died between 2007 and 2013 would have been missed from the analysis even if they had been admitted with an SLE. Overall, 92 884 of 275 929 (34%) patients in the initial dataset were excluded due to missing CIRRUS data, of whom 372 out of 1542 (24%) went on to have a first SLE. The final study population comprised 183 045 patients.

Categorisation of the CIRRUS score into grades according to severity

Blood result data used to derive the CIRRUS score were aggregated (mean) by test day. Following calculation of a CIRRUS score, test results were converted into four categories: crimson (C), red (R), amber (A), green (G) to aid interpretation corresponding to the level of risk. Outputs were converted to a specificity to aid clinical interpretation of where the cut-offs should be set. High levels of specificity were chosen given the low prevalence of cirrhosis in primary care. Multiple CIRRUS results were obtained for each subject dependent on the number of blood tests in the system. The index test date was the first date on which the most severe result was obtained within the dataset. If a subject progressed from green to amber to red to crimson, the first index test was crimson, and the index test date was the date of the first crimson result.

Study subgroups according to liver risk factors

To create study subgroups with a higher prevalence of SLEs, and thus potentially more clinically useful PPVs, we categorised patients according to known information about their risk factors for liver disease. For validation dataset 1, liver risk factors included alcohol (previous admission with an alcohol ICD code), diabetes (elevated HbA1C or previous admission with type 2 diabetes) and viral hepatitis (hepatitis B surface antigen or DNA, or hepatitis C virus RNA). For validation dataset 2, liver risk factors included alcohol (previous admission with an alcohol ICD code, harmful dependent drinking or an alcohol comorbidity in the primary care record) and diabetes (type 2 diabetes in primary or secondary care records). It was not possible to extract viral serology data from the CHIE dataset.

Box 1 CIRRUS model

Predicted risk (PR)=exp (alb*−0.210+cr*−0.010+tb*0.015+mcv *0.083+na*−0.067+plt*−0.018+tp*0.145+0.886) / (1+exp (alb*−0.210+cr*−0.010+tb*0.015+mcv*0.083+na*−0.067+plt*−0.018 +tp*0.145+0.886))

Albumin (alb), g/L; CIRRUS, CIRRhosis Using Standard tests; creatinine (cr), µmol/L; total bilirubin (tb), µmol/L; mean corpuscular volume (mcv), fL; sodium (na), mmol/L; platelet count (plt), 109/L; total protein, g/L.

Statistical analysis

The predicted risk using the CIRRUS algorithm was calculated for each subject who had the required data and its performance evaluated using sensitivity, specificity, predictive values and likelihood ratios, area under the curve (AUC) analysis and Kaplan-Meier plots to represent time from index test to admission with an SLE. Data were censored on the date of death or date of data extraction.

RESULTS

Creation of the CIRRUS model

Population characteristics for the dataset used to create the CIRRUS model are given in online supplemental table 1. Patients were randomly allocated into two cohorts 0 and 1, and a binary logistic regression model calculated in each and validated in the same and opposing cohorts. The AUC for the model 0 applied to cohort 1 was 0.90 (95% CI 0.88 to 0.92), and for model 1 applied to cohort 0 was also 0.90 (95% CI 0.88 to 0.92). As there was no overlap between subjects in cohorts 0 and 1, the two cohorts were recombined to create the final CIRRUS model (box 1).

Using the receiver operator curve (ROC) analysis, CIRRUS risk cut-offs were set to correspond to the following levels of specificity: crimson ≥99%, red 98%, amber 94%–97% and green <94%. This resulted in the following levels of predicted probabilities: crimson/red 0.284, red/amber 0.174, amber/green 0.077. For these different cut-offs, sensitivities ranged from 55% (crimson 'vs' red/amber/green) to 77% (crimson/red/amber 'vs' green), PPV 0.71–0.43, negative predictive value (NPV) 0.97–0.99 and PLR 43–13, respectively (table 1).

The median intervals between the index test and reference standard were between 201 and 236 days in the control group, and 244–281 days in the cirrhosis/portal hypertension group (online supplemental table 5).

Performance of CIRRUS in validation dataset 1 (UHS cohort)

Validation dataset 1 comprised 394 253 patients, of whom 3044 had an SLE, 6914 had a liver admission not fitting SLE criteria (liver disease) and 384 295 had no suggestion of liver disease (controls) (figure 1 and online supplemental table 6). CIRRUS predicted a first admission for an SLE within 5 years with an AUC of 0.90 (95% CI 0.89 to 0.91) continuous, or 0.88 (95% CI 0.87 to 0.89) when categorised into grades of severity (green to crimson)

Table 1 Sensitivity, specificity, positive and negative predictive values, and likelihood ratios of CIRRUS in the model creation dataset

Cut-off	Event/Total	Prevalence of cirrhosis/portal hypertension	Sensitivity	Specificity	PPV	NPV	PLR	NLR
CvRAG	931/16967	0.05	0.55	0.99	0.71	0.97	43.02	0.45
CRvAG	931/16967	0.05	0.64	0.98	0.63	0.98	29.15	0.37
CRAvG	931/16967	0.05	0.77	0.94	0.43	0.99	12.81	0.24

CRAvG, crimson, red or amber test 'vs' green test result; ; CRvAG, crimson or red test 'vs' amber or green test result; CvRAG, crimson test 'vs' red, amber or green test result; NLR, negative likelihood ratio.; NPV, negative predictive value ; PLR, positive likelihood ratio; PPV, positive predictive value .

(figure 2). For patients preselected according to known liver risk factors, CIRRUS predicted a subsequent SLE with a sensitivity of 72%, specificity 87%, PPV 26%, NPV 98% for a crimson/red 'vs' amber/green cut-off (table 2 and figure 3). A sensitivity analysis of CIRRUS calculated using mean blood test data aggregated over 4-week period was additionally performed in this cohort. This evened out day-to-day fluctuations reducing the sensitivities slightly, but increasing the PPVs (online supplemental table 7).

Performance of CIRRUS in validation dataset 2 (CHIE cohort)

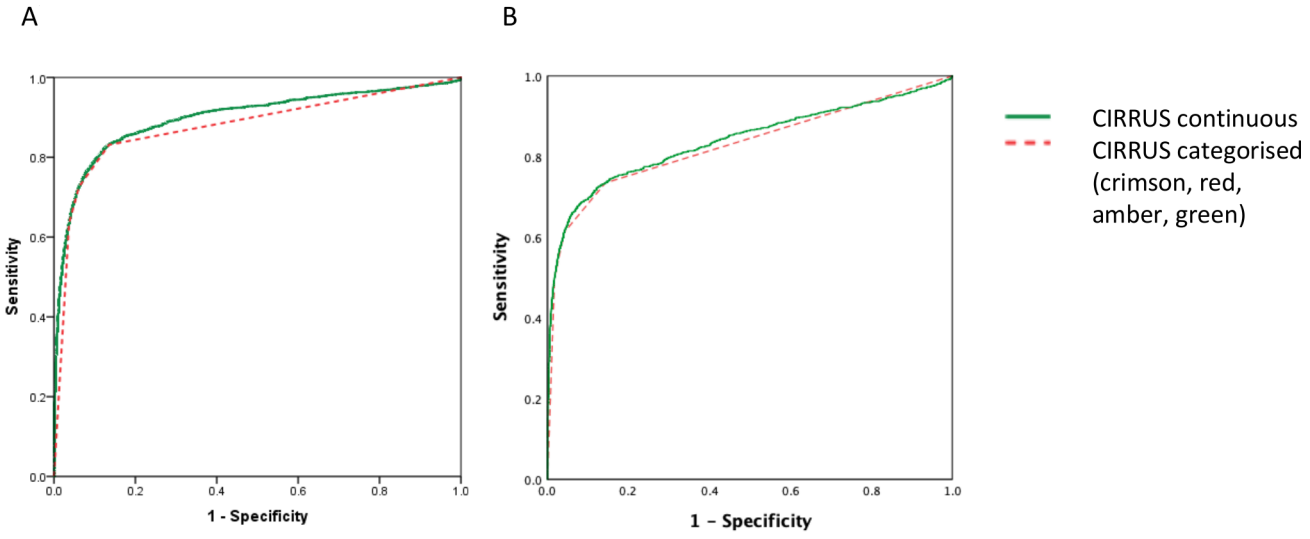
Validation dataset 2 comprised 183045 patients, of whom 1170 had an SLE, 3074 had liver disease not fitting SLE criteria and 178801 had no evidence of liver disease (figure 1 and online supplemental table 6).

For a first admission with an SLE within 5 years of the index test, AUCs were 0.84 (95% CI 0.82 to 0.86) continuous or 0.83 (95% CI 0.81 to 0.85) when categorised into severity grades (figure 2). In patients with a specified risk factor for liver disease, CIRRUS predicted a subsequent SLE with a sensitivity of 59%, specificity 93%, PPV 18%, NPV 99% for a crimson/red 'vs' amber/green cut-off (table 2 and figure 3).

Of note, the majority of individual blood tests were normal in both developmental datasets, even in patients with a crimson CIRRUS test (online supplemental table 8).

Time from index test to SLE in the validation cohorts

The median time (days) from the index test to the first SLE for validation dataset 1 (UHS) were as follows: green 2257 (95% CI 1978 to 2553), amber 187 (95% CI 63 to



Test result variable (s)	CIRRUS continuous	CIRRUS categorised
SLE within 5 yrs	n=2,289	
No liver disease	n=385,050	
AUC	0.90	0.88
95% CI	0.89-0.91	0.87-0.89

Test result variable (s)	CIRRUS continuous	CIRRUS categorised
SLE within 5 yrs	n=920	
No liver disease	n=178,801	
AUC	0.84	0.83
95% CI	0.82-0.86	0.81-0.85

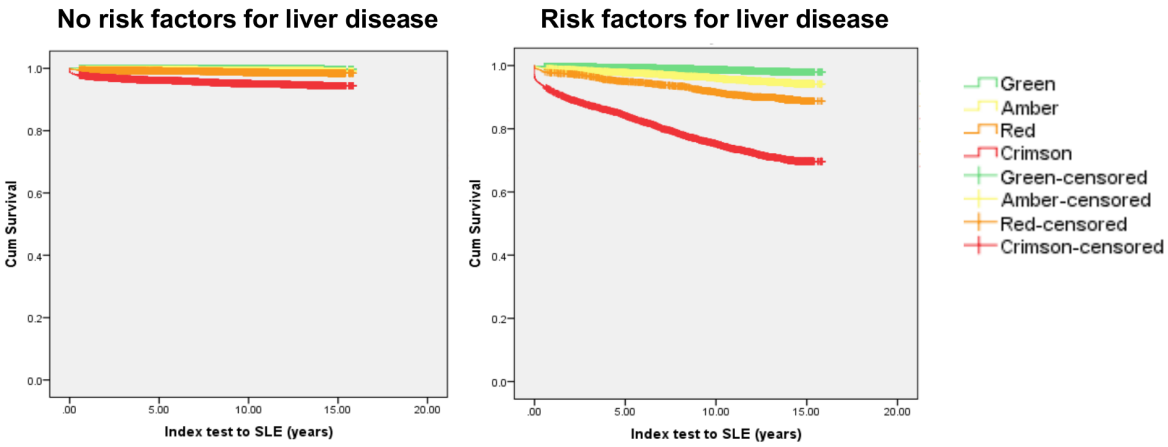
Figure 2 Area under curve (AUC) analysis for (A) validation dataset 1 (University Hospital Southampton) and (B) validation dataset 2 (Care and Health Information Exchange) for the prediction of a serious liver event (SLE) within 5 years of the index test.

Table 2 Prediction of a first serious liver event (SLE), liver admission with or without SLE (liver disease or SLE), or mortality in validation dataset 1 (UHS cohort) and prediction of a first SLE or liver admission with or without SLE in validation dataset 2 (CHIE cohort), within 5 years of the CIRRUS index test

	CIRRUS	Risk factor	Event (SLE)/ total	Prevalence	Sensitivity	Specificity	PPV	NPV	PLR	NLR
Validation dataset 1 (UHS cohort)										
SLE	**C**vRAG	Liver risk factor (alc/ DM2/VH)	1733/31 939	5.93	0.63	0.92	0.33	0.98	7.71	0.40
	CRvAG				0.72	0.87	0.26	0.98	5.65	0.32
	CvRAG	No liver risk factors	1311/362 314	0.37	0.49	0.97	0.05	1.00	14.85	0.53
	CRvAG				0.57	0.94	0.04	1.00	10.18	0.46
	CvRAG	Alcohol	1053/10 642	10.89	0.76	0.87	0.42	0.97	5.91	0.28
	CRvAG				0.83	0.82	0.35	0.97	4.48	0.21
	CvRAG	Type 2 diabetes	649/20 499	3.35	0.48	0.94	0.21	0.98	7.53	0.56
	CRvAG				0.58	0.89	0.16	0.98	5.50	0.47
	CvRAG	Viral hepatitis	288/1890	30.00	0.59	0.88	0.68	0.83	5.06	0.47
	CRvAG				0.72	0.83	0.65	0.87	4.26	0.33
Liver disease or SLE	**C**vRAG	Liver risk factor (alc/ DM2/VH)	4467/31 939	13.99	0.44	0.92	0.47	0.91	5.36	0.61
	CRvAG				0.54	0.87	0.41	0.92	4.26	0.52
	CvRAG	No liver risk factors	5491/362 314	1.52	0.28	0.97	0.12	0.99	8.46	0.75
	CRvAG				0.35	0.94	0.09	0.99	6.33	0.69
Dead	**C**vRAG	Liver risk factor (alc/ DM2/VH)	4638/27 734	14.5	0.36	0.91	0.40	0.89	3.9	0.70
	CRvAG				0.45	0.86	0.35	0.90	3.2	0.64
	CvRAG	No liver risk factors	4612/362 314	5.3	0.24	0.97	0.35	0.96	9.6	0.78
	CRvAG				0.31	0.95	0.28	0.96	6.8	0.72
Validation dataset 2 (CHIE cohort)										
SLE	**C**vRAG	Liver risk factor (alc/ DM2)	876/35 832	2.54	0.49	0.97	0.29	0.99	15.56	0.53
	CRvAG				0.59	0.93	0.18	0.99	8.62	0.44
	CvRAG	No liver risk factors	294/147 213	0.20	0.30	0.99	0.04	1.00	21.20	0.71
	CRvAG				0.39	0.97	0.02	1.00	11.36	0.63
	CvRAG	Alcohol	636/14 666	4.55	0.56	0.95	0.36	0.98	11.57	0.47
	CRvAG				0.65	0.90	0.25	0.98	6.82	0.38
	CvRAG	Type 2 diabetes	441/23 896	1.92	0.41	0.97	0.24	0.99	16.35	0.60
	CRvAG				0.52	0.94	0.15	0.99	8.91	0.51
Liver disease or SLE	**C**vRAG	Liver risk factor (alc/ DM2)	2259/35 832	6.30	0.28	0.97	0.37	0.95	8.88	0.74
	CRvAG				0.37	0.93	0.27	0.96	5.40	0.68
	CvRAG	No liver risk factors	1985/147 213	1.35	0.10	0.99	0.09	0.99	7.36	0.91
	CRvAG				0.15	0.97	0.06	0.99	4.41	0.88

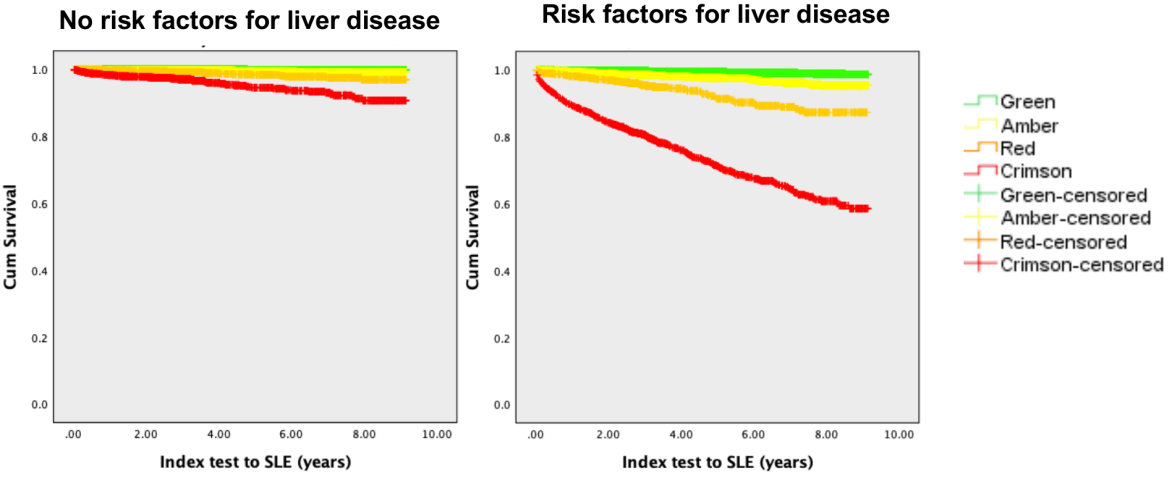
Validation dataset 1 (UHS cohort) liver risk factors: any of the following preidentified risk factors for liver disease: alcohol (previous admission with an alcohol ICD code), diabetes (elevated HbA1C or a previous admission with type 2 diabetes), viral hepatitis (HBsAg, HBV DNA or HCV RNA).
Validation dataset 2 (CHIE cohort) liver risk factors: any of the following preidentified risk factors for liver disease: alcohol (previous admission with an alcohol ICD code, harmful dependent drinking or an alcohol comorbidity in the primary care record) or diabetes (type 2 diabetes in primary or secondary care records).
In each dataset, the control population were patients with no liver admissions.
An SLE was defined as a hospital admission with an ICD code for either varices, liver-related ascites or cirrhosis.
CHIE, Care and Health Information Exchange; CIRRUS, CIRRhosis Using Standard tests; CRvAG, crimson or red test 'vs' amber or green result; CvRAG, crimson test 'vs' red, amber or green result; DM2, type 2 diabetes mellitus; NLR, negative likelihood ratio; NPV, negative predictive value; PLR, positive likelihood ratio; PPV, positive predictive value; SLE, serious liver event; UHS, University Hospital Southampton.

A



Liver risk	CIRRUS category	Events (SLE), N	Total, N
No liver risk	Green	407	314,720
	Amber	160	25,760
	Red	107	8,636
	Crimson	637	13,198
	Overall	1,311	362,314
Liver risk (alcohol, T2 diabetes, viral hepatitis)	Green	304	22,018
	Amber	180	3,986
	Red	156	1,730
	Crimson	1,093	4,205
	Overall	1,733	31,939
Overall	Overall	3,044	394,253

B



Liver risk	CIRRUS category	Events (SLE), N	Total, N
No liver risk	Green	136	127,813
	Amber	42	14,053
	Red	29	3,116
	Crimson	87	2,231
	Overall	294	147,213
Liver risk (alcohol, T2 diabetes)	Green	253	28,029
	Amber	103	4,653
	Red	90	1,458
	Crimson	430	1,692
	Overall	876	35,832
Overall	Overall	1,170	183,045

Figure 3 Kaplan-Meier curves for the time from the CIRRUS index test to the first serious liver event (SLE) 'vs' no liver disease controls in years, categorised according to liver risk factors and severity grades for (A) validation dataset 1 (University Hospital Southampton) and (B) validation dataset 2 (Care and Health Information Exchange). CIRRUS, CIRRhosis Using Standard tests.

Box 2 Validation dataset 1 (UHS): AUC (95% CI) for CIRRUS, APRI, FIB-4, AST/ALT ratio (n=6105)

CIRRUS categorised 0.83 (0.80–0.86)
APRI categorised 0.79 (0.76–0.83)
FIB-4 categorised 0.79 (0.75–0.83)
AST/ALT ratio categorised 0.58 (0.54–0.62)

465), red 120 (95% CI 36 to 380) and crimson 106 (95% CI 81 to 147). For validation dataset 2 (CHIE), these were as follows: green 1513 (95% CI 1368 to 1637), amber 770 (95% CI 637 to 1058), red 784 (95% CI 602 to 938) and crimson 550 (95% CI 443 to 637). As some of the time periods between the index date and SLE were short for validation dataset 1, we performed a sensitivity analysis excluding subjects where the time between the index CIRRUS score and SLE was less than 100 days (online supplemental tables 9A and B). In this case, the AUC was 0.91 (95% CI 0.90 to 0.92) continuous and 0.88 (95% CI 0.86 to 0.89) when categorised by grades.

CIRRUS compared with AST to platelet ratio index (APRI), fibrosis-4 (FIB-4) and AST/ALT ratio

In a subset of 6105 patients derived from validation dataset 1 (UHS), we were able to compare the performance of the CIRRUS test with FIB-4, APRI and AST/ALT ratio as a predictor of a first SLE within 5 years of the index test, taking the maximum value of each test in each case. Within this subgroup, 186 individuals had a first SLE. A positive result was defined as follows: CIRRUS (red/crimson grade), FIB-4 3.25, APRI≥1.5, AST/ALT ratio>1. Data were aggregated (mean) by month of test prior to calculating the algorithms. Our results show that the performance of CIRRUS is comparable to all three of these well-established tests (box 2 and online supplemental figure 1).

We were able to make the same comparisons in a subset of 4297 patients derived from validation dataset 2 (CHIE). The CHIE data were first aggregated by quarters (3-month period) to enable more results to be calculated (box 3).

An overall comparison of the availability of risk score data was assessed for validation dataset 1 (UHS). In total, 3044 first SLE patients had a CIRRUS result (1993 (65%) had a crimson or red result). In comparison FIB-4, APRI and AST/ALT ratio results were available in 312, 319 and 337 patients, respectively, of whom 132 (42%), 120 (38%) and 243 (72%) patients had a positive result. As a result of greater test availability within this cohort, the CIRRUS model could detect at least eight times as many patients prior to the first SLE than either FIB-4, APRI or AST/ALT ratio.

Box 3 Validation dataset 2 (CHIE): AUC (95% CI) for CIRRUS, APRI, FIB-4, AST/ALT ratio (n=4297)

CIRRUS categorised 0.88 (0.83–0.93)
APRI categorised 0.86 (0.81–0.91)
FIB-4 categorised 0.87 (0.82–0.92)
AST/ALT categorised 0.68 (0.62–0.74)

DISCUSSION

Principal findings

We have shown that routinely available blood tests can be used to identify individuals at high risk of a first admission with AdvLD. In the model creation population, CIRRUS detected AdvLD with an AUC of 0.90 (95% CI 0.88 to 0.92). In the UHS (mixed primary and secondary care) and CHIE (primary care) validation cohorts, CIRRUS predicted an admission for an SLE within 5 years with an AUC of 0.90 (95% CI 0.89 to 0.91) and 0.84 (0.82–0.86), respectively. In patients with known liver risk factors and a crimson or red CIRRUS test, the PPV to predict a first SLE was 26% and 18% for the UHS and CHIE datasets (alcohol 35%/25%, type 2 diabetes 16%/15%). If a crimson cut-off point was employed, the PPVs increased to 33% and 29%, respectively. In individuals with no identified risk factors for liver disease, SLE prevalence was low (0.4%/0.2%), and PPVs reduced to 4% for the UHS and 2% for the CHIE dataset. Where data were aggregated by month, in the case of the UHS cohort, the PPVs were 34% for individuals with a crimson/red result and 43% for people with a crimson result.

Strengths and weaknesses in relation to other studies

Existing diagnostic liver algorithms were largely created using data from patients referred with suspected liver disease.[8 9] A review of 6000 papers on liver blood tests found that practically all were based on hospital practice.[10] A selection of algorithms using liver blood tests to stage liver fibrosis or cirrhosis were evaluated by the National Institute for Health and Care Excellence (NICE) in 2016.[2] The AST/ALT ratio was developed in 57 patients with chronic active hepatitis and 177 biopsied liver patients.[11 12] Fibrotest was developed in 339 biopsied patients with hepatitis C,[13] APRI in 270 biopsied patients with hepatitis C[14] and FIB-4 in 832 biopsied patients coinfected with hepatitis C and HIV.[15] None of the algorithms evaluated by NICE were designed to detect AdvLD in a primary care setting, with the exception of the ALFI (Algorithm for Liver Function Investigations) study which analysed data from 95 977 primary care subjects. The ALFI model used some liver blood tests alongside methadone and alcohol use to predict a liver diagnosis with a C-stat of 0.79 (95% CI 0.72 to 0.84). Multiple imputations were used to compensate for missing GGT data, which was available in only 11% of subjects.[16 17] Data required for the CIRRUS score were available in 84% of UHS patients who had an SLE, and as a result, CIRRUS picked up over eight times as many first SLE patients than FIB-4, APRI or the AST/ALT ratio. Therefore, CIRRUS could potentially be used as a first-line screening tool to flag up individuals at risk of significant liver disease in a

standard mixed primary and secondary care population in a way that more specialist scores could not.

Study limitations

The model creation population was chosen with the hope of selecting a cohort with a low prevalence of liver disease. However, the fact that all individuals had been referred for an OGD inevitably introduced bias, given that alcohol and obesity are risk factors for peptic ulcer disease and oesophagitis, and in a minority of cases, the indication for endoscopy would have been liver disease with the intent of screening for varices. Furthermore, not all patients within this cohort had undergone a full liver assessment, and we were unable to completely exclude underlying cirrhosis in the controls. Despite these compromises, our algorithm performed as well as the current alternatives and was applicable in many more patients.

The value of the CHIE dataset is unfortunately significantly limited by the fact that it did not include deceased patients, rendering it subject to survival bias. This would have reduced the PPV of an SLE slightly compared with the UHS cohort. It also prevented us from analysing mortality in this group. We had hoped to follow this cohort of patients prospectively, but it was not possible to obtain clinical information due to national information governance changes. We have included the CHIE analysis despite these shortcomings as it was sourced entirely from primary care and hence represents performance in a lower prevalence setting where the test is intended to be used.

The timescales between a positive CIRRUS result and subsequent SLE were short, particularly for the UHS validation cohort; however, a sensitivity analysis confirmed that CIRRUS performed equally well when patients with a result less than 100 days before an SLE were excluded. There will also have been some overlap between the cohorts. As both the UHS and CHIE datasets were fully anonymised, it was not possible to perform a sensitivity analysis; however, given the identical AUC results in the two mutually exclusive creation cohorts, this is not likely to have influenced the results. All tests and investigations were from a single UK region, although there is no reason to suspect this region is atypical we recommend further validation in a completely independent cohort. Furthermore, outcomes were reliant on ICD coding of liver admissions, for which there are known issues.[18] However, this enabled us to study a large number of individuals, allowing greater precision of estimates and reducing the confidence intervals of our findings. It would also have been beneficial to have included obesity as a risk factor for liver disease; however, unfortunately, body mass index data were not available to us. Finally, this study lacked public and patient involvement. This would have been advantageous given the large number of anonymised patient records required to create and validate the CIRRUS score and to understand how this test may be received.

Implications for practice and future work

We show that a model combining routine data is more accurate than individual tests, and has the potential to identify a majority of patients with evidence of alcohol misuse and a substantial proportion of patients with type 2 diabetes before their first admission with a serious, and frequently fatal, liver event. Alcohol and the metabolic syndrome are responsible for around 90% of UK liver deaths,[1] and we estimate around 75% of patients are not recognised as having AdvLD until they present with a liver-related admission. The PPVs in these selected groups are clinically useful and could improve levels of detection of AdvLD in the community. In patients without evidence of alcohol misuse, type 2 diabetes or positive viral serology, the low prevalence of liver disease reduced the PPV, and with it the applicability of the test in practice. Further validation in well-characterised research cohorts could identify further preselection variables, increasing the applicability of the test.

While the CIRRUS score benefits from high levels of specificity and NPVs, sensitivity values are lower, with NLRs ranging from 0.32 to 0.40/0.44–0.53 (liver risk factors) and 0.46–0.53/0.63–0.71 (no liver risk factors) for a crimson to red cut-off for validation cohorts 1 and 2, respectively. In other words, this cut-off results in a higher frequency of false-negative results than is ideal for a community-based screening test. However, we felt a red cut-off was appropriate to balance acceptable levels of sensitivity against the work load that would be generated in terms of managing a greater frequency of patients with a raised CIRRUS score but normal fibrosis markers or fibroscan. For those without a risk factor for liver disease, there was a low prevalence of SLEs (0.2%–0.4%) in our validation cohorts. The PLRs were 15–21 for a crimson test and 10–11 for a red/crimson test, showing the test does still discriminate, but the high rate of false positives in this group may reduce clinical utility. The metabolic derangement associated with other severe medical conditions, for example, severe sepsis, malignancy (particularly haematological) and other organ failures, is likely to be responsible for false-positive results . In a clinical use scenario, many of these patients are likely to have signs or symptoms suggestive of alternative diagnose. Furthermore, for the UHS validation dataset,both ‘any risk’ and ‘no risk’ groups had high PPVs for death (28%–40%), so a positive CIRRUS test remains an indication of potentially serious health issues.

The CIRRUS test could be employed in two ways. Retrospective screening of existing results in hospital or primary care can indicate patients at increased risk of a future liver event. These individuals could be contacted to allow confirmation of any liver disease using non-invasive methods and supported with an intervention to help reduce their liver risk factor burden. We aim to test this strategy in the multicentre POLeMMIC (Prevention of Liver Mortality and Morbidity in the Community) study. AdvLD creates a metabolic fingerprint which is relatively stable, whereas severe metabolic derangements that cause

false-positive test results may improve over days. Using mean data over a 1-month period increased the PPV at the price of a small reduction in sensitivity, and this method may be best for retrospective case finding or clinical trials. An alternative approach would be to analyse the data in real-time in a primary care information system. It would be relatively straightforward to modify existing clinical results systems to calculate a CIRRUS grade and flag up patients at risk, with a reasonable level of expectation that they have a significant underlying health problem of some sort. An appointment with a practice nurse could stratify patients based on their underlying liver risk factors. Where indicated, liver disease could be confirmed using a non-invasive methodology.[2] Subsequently, community or hospital-based fibroscanning can stage cirrhosis and identify patients fitting the Baveno criteria for portal hypertension.[19]

The lifetime national health service costs for a patient admitted with alcohol-related liver cirrhosis are between £50 and £1 20 000 and it would not be necessary to prevent many liver admissions to render this approach cost effective.[20] We estimate around 40%–50% of patients with alcohol-related liver disease will revert to safe drinking following an admission with liver disease[21] and there is some evidence that a liver diagnosis has a similar impact in primary care.[22 23] CIRRUS also indicates an impending SLE in patients with viral hepatitis, but as all these patients are now referred for antiviral therapy, the clinical utility of this test is probably low. The information needed to calculate the CIRRUS test is probably already available to most UK primary care doctors for a majority of patients in the community who are likely to suffer their first SLE over the next 5 years. This first event is often fatal; if our study is validated in different cohorts, it is possible many of these patients could be warned of an impending disaster that could potentially be averted.

Author affiliations
[1]School of Primary Care and Population Sciences, University of Southampton, Southampton, UK
[2]Human Development and Health, University of Southampton Faculty of Medicine, Southampton, UK
[3]Southampton Biomedical Research Centre, Southampton, UK
[4]Informatics, University Hospital Southampton NHS Foundation Trust, Southampton, UK
[5]AH IT Solutions, Southampton, Hampshire, UK
[6]The Institute of Hepatology, Foundation for Liver Research, London, UK

Acknowledgements The authors thank NHS South, Central and West Commissioning Support (SCWCSU) Unit and the Care and Health Information Exchange Information Governance Group (CHIEIGG) for their support and for provision of access to CHIE data. We would also like to thank Hugh Sanderson for the initial CHIE SQL queries, Mandy Lu for the help in digitising text-based pathology reports, and Hazel Inskip and Paul Roderick for advice on analysis. This research was funded by the British Liver Trust and Southampton Biomedical Research Centre.

Contributors TH was responsible for data analysis and interpretation and writing this manuscript. NS conceived and designed the study, performed initial analyses, interpreted the data and drafted the paper. MM commented on various aspects of the study design, analysis and drafting. BS and MK repeated all analysis from the original raw data independently, and commented on the manuscript. DC and AH facilitated the extraction and data governance of UHS data, this involved linking a variety of UHS datasets that had not previously been linked. FS constructed and ran CHIE data queries. CN constructed and ran CHIE data queries, constructed the main SQL database, entered index test data and commented on the manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests NS undertook paid consultancy work and received travelling expenses from pharmaceutical companies Norgine (2014) and Kyowa Kirin Limited (2014), as well as under taking medicolegal work in hepatitis C and alcohol-related liver disease. None of the remaining authors have any competing interests.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. Data were obtained from routinely collected anonymised National Health Service data from the University Hospital Southampton and from the Care and Health Information Exchange (CHIE).

ORCID iDs
Theresa Hydes http://orcid.org/0000-0002-7768-6886
Michael Moore http://orcid.org/0000-0002-5127-4509

REFERENCES

1 Williams R, Aspinall R, Bellis M. Lancet Commission. addressing the crisis of the liver disease in the UK: a blueprint for attaining excellence in healthcare for liver disease and reducing premature mortality from the major lifestyle issues of excess alcohol consumption, obesity and viral. *Lancet* 2014;384:1953–97.
2 National Institute for Health and Care Excellence. *Cirrhosis in over 16S: assessment and management NG50*, 2016.
3 Harris R, Harman DJ, Card TR, *et al*. Prevalence of clinically significant liver disease within the general population, as defined by non-invasive markers of liver fibrosis: a systematic review. *Lancet Gastroenterol Hepatol* 2017;2:288–97.
4 ISRCTN R. *LOCATE - LOcal Care And Treatment Evaluation*, 2018.
5 Sheron N, Moore M, Ansett S, *et al*. Developing a 'traffic light' test with potential for rational early diagnosis of liver fibrosis and cirrhosis in the community. *Br J Gen Pract* 2012;62:e616–24.
6 Collett D. *Modelling binary data*. 2nd edn. Chapman and Hall, 1991: 93.
7 Exchange Cal. Care and information exchange, 2018. Available: https://www.careinformationexchange-nwl.nhs.uk/
8 Kunutsor SK, Apekey TA, Seddoh D, *et al*. Liver enzymes and risk of all-cause mortality in general populations: a systematic review and meta-analysis. *Int J Epidemiol* 2014;43:187–201.
9 Radcke S, Dillon JF, Murray AL. A systematic review of the prevalence of mildly abnormal liver function tests and associated health outcomes. *Eur J Gastroenterol Hepatol* 2015;27:1–7.
10 Badrick T, Turner P. Review and recommendations for the component tests in the liver function test profile. *Indian J Clin Biochem* 2016;31:21–9.
11 Adelasco L, Monarca A, Dantes M, *et al*. Features of chronic hepatitis in alcoholics. A survey in Milan. *Liver* 2008;7:283–9.
12 Williams AL, Hoofnagle JH. Ratio of serum aspartate to alanine aminotransferase in chronic hepatitis. Relationship to cirrhosis. *Gastroenterology* 1988;95:734–9.
13 Imbert-Bismut F, Ratziu V, Pieroni L, *et al*. Biochemical markers of liver fibrosis in patients with hepatitis C virus infection: a prospective study. *Lancet* 2001;357:1069–75.

14 Wai C-T, Greenson JK, Fontana RJ, *et al*. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. *Hepatology* 2003;38:518–26.
15 Sterling RK, Lissen E, Clumeck N, *et al*. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology* 2006;43:1317–25.
16 McLernon DJ, Dillon JF, Sullivan FM, *et al*. The utility of liver function tests for mortality prediction within one year in primary care using the algorithm for liver function investigations (ALFI). *PLoS One* 2012;7:e50965.
17 McLernon DJ, Donnan PT, Sullivan FM, *et al*. Prediction of liver disease in patients whose liver function tests have been checked in primary care: model development and validation using population-based observational cohorts. *BMJ Open* 2014;4:e004837.
18 O'Malley KJ, Cook KF, Price MD, *et al*. Measuring diagnoses: ICD code accuracy. *Health Serv Res* 2005;40:1620–39.
19 Bae J, Sinn DH, Kang W, *et al*. Validation of the Baveno VI and the expanded Baveno VI criteria to identify patients who could avoid screening endoscopy. *Liver Int* 2018;38:1442–8.
20 Bouttell J, Lewsey J, Geue C, *et al*. The Scottish alcoholic liver disease evaluation: a population-level matched cohort study of hospital-based costs, 1991-2011. *PLoS One* 2016;11:e0162980.
21 Verrill C, Markham H, Templeton A, *et al*. Alcohol-related cirrhosis--early abstinence is a key factor in prognosis, even in the most severe cases. *Addiction* 2009;104:768–74.
22 Sheron N, Moore M, O'Brien W, *et al*. 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:e698–705.
23 Eyles C, Moore M, Sheron N, *et al*. Acceptability of screening for early detection of liver disease in hazardous/harmful drinkers in primary care. *Br J Gen Pract* 2013;63:e516–22.