Supplementary Information to: Prognostic value of multiparametric MRI, transient elastography and serum fibrosis markers in patients with chronic liver disease

Authors: Arjun N. A. Jayaswal, Christina Levick, Emmanuel A. Selvaraj, Andrea Dennis, Jonathan C. Booth, Jane Collier, Jeremy Cobbold, Elizabeth M. Tunnicliffe, Matthew Kelly, Eleanor Barnes, Stefan Neubauer, Rajarshi Banerjee, Michael Pavlides*.

TABLE OF CONTENTS:

Supplementary figures

SUPPLEMENTARY METHODS

Patient selection

Patients were recruited as part of the Rapid Imaging Assessment of the Liver (RIAL) and Non-Invasive and COmprehensive Liver Imaging Assessment (NICOLA) studies (ethics refs: 11/H0504/2 and 13/SC/0243, respectively). The breakdown of these studies' patient groups is described in Table S1. Prognostic results from some of the patients from these cohort studies have been reported previously[1].

Table S1: Description of patient cohort

Study 1: Rapid Imaging Assessment of the Liver (RIAL)	Study 2: Non-Invasive and COmprehensive Liver Imaging Assessment (NICOLA)
Patients with suspected liver disease of any aetiology attending liver clinic and scheduled for a clinically indicated liver biopsy having paired MR assessment (n = 121)	Patients undergoing weight loss surgery (n = 19) Patients with suspected NAFLD/NASH scheduled to undergo clinically indicated biopsy (n = 16; 1 did not undergo biopsy). Patients with a diagnosis of cirrhosis based on histology or clinical or radiological findings (n = 20). Patients scheduled for either clinically indicated transjugular biopsy and hepatic venous pressure gradient measurements, or clinically indicated percutaneous biopsy for suspected cirrhosis, who agreed to undergo transjugular biopsy instead to obtain HVPG measurements (n = 21).

Biopsies were performed a median (IQR) of 11 (3-29) days from MRI. They contained a median of 10 (7-14) portal tracts and total core length of median (IQR) 18 (14-27) mm.

Image analysis

T_1 and T_2* maps were analysed using Liver*MultiScan*™ Discover – a software program developed by Perspectum Diagnostics Ltd for the purpose of generating cT_1 maps.

1) Three regions of interest (ROIs) are manually placed on T_2* colour maps to obtain a representative iron concentration within the liver.

2) This iron concentration is fed into an algorithm[2] that transforms the measured T_1 map into a cT_1 map in such a way that corrects for any bias caused by excess iron.

3) 3 ROIs are placed on the cT_1 map over parts of the liver that are considered representative of the liver parenchyma. This means avoiding blood vessels, organ boundaries and potential cysts/ascites that are present. A single ROI measurement calculates the mean of the cT_1 pixel values within it; the mean of three ROIs is taken as the final cT_1 value.

The r square parameter is a measure of how well the T_1 and T_2* measurements fit the model curves. R square values must be > 0.95 for T_2* and > 0.99 for T_1. Scans can also be rejected for poor scan quality criteria including mistriggering of the ECG-gated sequence and banding artefacts. Example liver cT_1 images are shown in Figure S8.

Selection of cut-off values

The cut off for T_2* from our previous study[3] was used to define two severity categories for iron (high: T_2* ≤12.5 ms, low: T_2* >12.5 ms) which represents the best threshold for identifying those with any grade of histological iron deposition and corresponds to a dry weight of iron of 1.7 mg/g[4]. The cut-off for fat assessed was 5.6%[5]. A cut-off of 8 kPa for TE was used, which is reported to correspond to presence of marked fibrosis[6]. TE was also split into 4 groups (<8kPa, 8-20 kPa, 20-39 kPa and ≥40kPa) for a more precise stratification of risk, similarly to previous studies[7,8]. For the APRI score, a cut-off of 1.00 was used, which has shown sensitivity of 76% and specificity of 72% for predicting cirrhosis[9]. For the FIB-4 score, a cut-off of 1.45 was used, which has been shown to

have a negative predictive value of 90% for advanced fibrosis[10]. For AST/ALT score, a cut-off of >1 has shown to predict cirrhosis[11].

Statistical analysis

The influence of failure/unreliability of liver cT_1 and TE on prediction of liver events was assessed by the analysis of 3 x 2 tables. In a traditional 2 x 2 table, true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN) would be extracted from all successful and reliable measurements. 3 x 2 tables are calculated by taking into account *all* attempts at the test in question including failures and unreliable readings. Unreliable readings are simply included as if they were to be used for a clinical diagnosis and thus do not require any adjustment. Failures (termed non-evaluable (positive) if associated with a clinical event, and non-evaluable (negative) if associated with survival) are included in a 3rd row in the table. Non-evaluable "positives" are counted as an incorrect negative test result and add to the false negative rate. Non-evaluable "negatives" are counted as an incorrect positive test result and add to the false positive rate (see tables S15-18 for liver cT_1, T_1 and TE 3 x 2 tables).

Tables S2: 3 x 2 example table.

	Clinical event	No clinical event
Index test positive	True positive	**False positive**
Index test non-evaluable	**Non-evaluable (positive)** ↓	**Non-evaluable (negative)** ↑
Index test negative	**False negative**	True negative

Arrows indicate how non-evaluable results are classified, depending on patient outcome.

SUPPLEMENTARY RESULTS

Table S3: Characteristics of individuals who experienced clinical events.

Liver cT_1 (ms)	Diagnosis	Index event (months)	Patient mortality (months)	Ishak fibrosis stage	TE (kPa)
840	NAFLD	Encephalopathy (15)	Yes (23)	6	44
1125	NAFLD	Ascites (21)	Yes (64)	6	39.1
896	Viral hepatitis	HCC (10)	Yes (18)	6	35.3
766	PSC	Encephalopathy (47)	Yes (64)	2	29.9
994	NAFLD	None	Yes – non liver-related death (43)	6	12.8
945	ArLD	Ascites (14)	No	6	5.2
871	NAFLD	Ascites (22)	Yes (34)	6	15.1
900	Viral hepatitis	HCC/ascites (17)	Yes – non liver-related death (39)	6	27.7
861	ArLD	HCC (7)	No	nm	69.1
1186	ArLD	Variceal bleed (19)	Yes (20)	6	na
1163	Viral hepatitis	Ascites (6)	No	6	na
na	Haemochromatosis	None	Yes – non liver-related death (56)	1	na
900	NAFLD	None	Yes – non liver-related (65)	6	nm

The only disagreements between reviewers were in 3 cases ascertaining whether cause of death was liver-related or not. Data shown include cT_1 value (na if failed), Ishak score (if available, nm if not clinically indicated) and TE (nm if not measured, na if failure/unreliable reading).

Event-free survival analysis

Table S4: Univariate Cox regression analysis and ROC curve analysis of continuous variables for the development of a new clinical event in the full cohort.

Variable	HR (% per unit)	95% CI	P value	AUC	95% CI	Events / observations (%)
cT_1 (ms)	1.007	1.003-1.011	0.0014	0.74	0.61-0.88	13/182 (7.1%)
T_1 (ms)	1.006	1.002-1.010	0.0061	0.68	0.49-0.87	13/182 (7.1%)
Liver fat (%)	0.954	0.879-1.034	0.252	0.56	0.43-0.70	13/179 (7.3%)
Liver iron (T_2*)	1.024	0.926-1.133	0.454	0.56	0.39-0.72	13/182 (7.1%)
TE (kPa)	1.062	1.032-1.092	<0.0001	0.83	0.67-0.99	10/121 (7.4%)
Ishak stage	1.994	1.323-3.008	0.0010	0.81	0.68-0.95	14/178 (7.9%)
FIB-4	1.241	1.147-1.344	<0.0001	0.83	0.68-0.98	13/185 (7.0%)
APRI	1.008	0.949-1.070	0.794	0.80	0.69-0.91	13/186 (7.0%)
AST/ALT	2.605	1.612-4.211	<0.0001	0.75	0.58-0.92	13/187 (7.0%)

Table S5: Cox regression analysis for binary cut-offs of variables for the development of a new clinical event and performance in identifying patients that develop new clinical events in the full cohort.

Test cut-off (n)	HR	95% CI	P value (log-rank)	Sens (%)	Spec (%)	PPV (%)	NPV (%)
cT_1 >825 ms	9.91	1.287-76.24	0.007	92.3	47.3	11.9	98.8
cT_1 >840 ms	12.11	1.573-93.2	0.0013	92.3	52.7	13.0	98.9
T_1 >825 ms	3.021	0.829-11.01	0.078	76.9	48.5	10.3	96.5
T_1 >868 ms	5.47	1.502-19.95	0.0038	76.9	63.9	14.1	97.3
Liver fat >5.6%	0.539	0.166-1.752	0.30	30.8	52.4	4.8	90.6
T_2* <12.5 ms	0.804	0.178-3.63	0.78	15.4	80.5	5.7	92.5
T_2* <8 ms	1.728	0.225-13.29	0.59	7.7	95.9	12.5	93.1
TE >8 kPa	7.79	0.974-62.3	0.022	88.9	51.8	12.9	98.3
Ishak >F4	12.64	2.80-57.06	0.0010	84.6	73.9	20.4	98.4
FIB-4 >1.45	4.110	0.910-18.56	0.029	84.6	47.7	10.9	97.6
APRI >1	2.645	0.886-7.900	0.07	46.2	79.2	14.3	95.1

AST/ALT >1	6.093	1.673-22.19	0.0017	76.9	65.5	14.3	97.4

Table S6: Cox regression analysis for the association of liver cT_1 grouped cut-offs derived from 90% sensitivity (840 ms) and 90% specificity (990 ms) with event-free survival in the full cohort.

	Hazard ratio (HR)	95% CI	P value
cT_1 < 840 ms (n=91)	Reference	-	-
840 – 990 ms (n=70)	9.33	1.15 – 75.9	0.037
cT_1 > 990 ms (n=21)	22.0	2.56 – 189	0.005

Table S7: Cox regression analysis for the association of TE grouped cut-offs (based on examples from literature) with event-free survival in the full cohort.

	Hazard ratio (HR)	95% CI	P value
TE < 8 kPa (n=59)	Reference	-	-
8 – 20 kPa (n=39)	2.9	0.26 - 32.1	0.384
20 – 39 kPa (n=18)	14.0	1.56 - 124	0.018
TE ≥ 40 kPa (n=5)	44.3	3.79 - 518	0.003

Table S8: Cox regression analysis for the association of Ishak fibrosis stage stratified by mild (F0-2), moderate (F3-4), severe (F5-6) with event-free survival in the full cohort.

	Hazard ratio (HR)	95% CI	P value
Mild vs Moderate vs Severe	3.71 (per increasing group) *	1.585 – 8.684	0.0025

*one group (F3-4) had zero events in it, so only global HR and p value could be calculated as direct comparisons cannot generate HRs.

Table S9: Univariate Cox regression analysis and ROC analysis of continuous variables for the development of a new clinical event in patients with NAFLD/ArLD/viral hepatitis (VH) only

Test	HR	95% CI	P value	AUC	95% CI	Events / observations (%)
cT_1 (ms)	1.007	1.002-1.011	0.005	0.75	0.63-0.87	11/148 (7.4%)

T_1 (ms)	1.005	1.001-1.010	0.020	0.68	0.48-0.88	11/148 (7.4%)
Liver fat (%)	0.906	0.802-1.023	0.111	0.65	0.51-0.78	11/146 (7.5%)
T_2* (ms)	1.014	0.911-1.128	0.802	0.54	0.37-0.72	11/148 (7.4%)
TE (kPa)	1.083	1.039-1.129	0.0002	0.80	0.60-1.00	7/97 (7.2%)
Ishak stage**	-	-	-	-	-	10/142 (7.0%)
FIB-4	1.522	1.293-1.793	<0.0001	0.85	0.69-1.00	10/148 (6.8%)
APRI	1.929	1.353-2.749	0.0003	0.82	0.70-0.94	10/149 (6.7%)
AST/ALT	3.343	1.966-5.684	<0.0001	0.81	0.62-0.99	10/150 (6.7%)

** No patients with Ishak stages 0-5 had any clinical events, only with stage 6.

Table S10: Cox regression analysis for binary cut-offs of variables for the development of a new clinical event and performance in identifying patients that develop new clinical events in patients with NAFLD/ArLD/VH only.

Test	HR*	95% CI	P value (log-rank)	Sens (%)	Spec (%)	PPV (%)	NPV (%)
cT_1 >825 ms	-	-	0.006	100.0	43.1	12.4	100.0
cT_1 >840 ms	-	-	0.002	100.0	48.9	13.6	100.0
T_1 >825 ms	3.508	0.754-16.33	0.088	81.8	44.5	10.6	96.8
T_1 >868ms	6.174	1.329-28.68	0.008	81.8	59.9	14.1	97.6
Liver fat >5.6%	0.331	0.0878-1.249	0.086	27.3	45.2	3.9	88.4
T_2* <12.5 ms	0.815	0.1761-3.772	0.79	18.2	77.4	6.1	92.2
T_2* <8 ms	1.888	0.241-14.78	0.54	9.1	95.6	14.3	92.9
TE >8kPa	5.875	0.7068-48.8	0.063	85.7	52.2	12.2	97.9
Ishak ≥F5	29.56	3.881-225.2	<0.0001	100.0	75.0	23.3	100.0
FIB-4 >1.45	7.083	0.896-55.96	0.03	90.0	48.6	11.3	98.5
APRI >1	4.373	1.23-15.56	0.012	60.0	78.4	16.7	96.5
AST/ALT >1	18.33	2.321-145	<0.0001	90.0	68.6	17.0	99.0

* Hazard ratio incalculable for binary cut-offs in which one group had zero events

Table S11: Cox regression analysis for the association of liver cT_1 grouped cut-offs derived from 90% sensitivity (840 ms) and 90% specificity (990 ms) with event-free survival in patients with NAFLD/ArLD/VH only.

	Hazard ratio (HR)	95% CI	P value
cT_1 < 840 ms (n=73) vs 840 – 990 ms (n=44) vs cT_1 > 990 ms (n=47)	3.67 (per increasing group)*	1.554-8.644	0.001

*one group (<840 ms) had zero events in it, so global HR and p value only could be calculated as direct comparisons cannot generate HRs.

Table S12: Cox regression analysis for the association of TE grouped cut-offs (based on examples from literature) with event-free survival in patients with NAFLD/ArLD/VH only.

	Hazard ratio (HR)	95% CI	P value
TE < 8 kPa (n=48)	Reference	-	-
8 – 20 kPa (n=31)	2.9	0.26-31.4	0.392
20 – 39 kPa (n=14)	7.70	0.70-85.2	0.096
TE ≥ 40 kPa (n=4)	39.9	3.43-466	0.003

Cox regression analysis for the association of Ishak fibrosis stage stratified by mild (F0-2), moderate (F3-4), severe (F5-6) with event-free survival in patients with NAFLD/ArLD/VH only could not be carried out as only one group (F5-6) had any events in it, so no HR could be calculated as direct comparisons cannot generate HRs.

Table S13: Intention to treat analysis of the association of MRI and TE results with event-free survival (n = 197).

Test	HR	95% CI	P value	Sens (%)	Spec (%)	PPV (%)	NPV (%)
cT_1 >825 ms	4.47	1.00-20.0	0.03	81.3	44.2	11.4	96.4
cT_1 >840 ms	5.36	1.20-24.0	0.01	81.3	49.2	12.4	96.7
T_1 >825 ms	2.12	0.660-6.79	0.21	68.8	45.3	10.0	94.3
T_1 >868 ms	3.64	1.134-11.7	0.02	62.5	59.1	11.9	94.7
TE >8 kPa	1.72	0.526-5.60	0.4	64.3	43.2	9.8	92.6

Table S14: 3 x 2 table for event-free survival as assessed by liver cT_1 >825 ms.

Liver cT_1	Clinical event	No clinical event
Liver cT_1 >825 ms	12	**89 → 103**
Non-evaluable	**1 ↓**	**14 ↑**
Liver cT_1 ≤825 ms	**1 → 2**	80

Arrows indicate how non-evaluable results are classified, depending on patient outcome.

Table S15: 3 x 2 table for event-free survival as assessed by liver cT_1 >840 ms.

Liver cT_1	Clinical event	No clinical event
Liver cT_1 >840 ms	12	**80 → 94**
Non-evaluable	**1 ↓**	**14↑**
Liver cT_1 ≤840 ms	**1 → 2**	89

Arrows indicate how non-evaluable results are classified, depending on patient outcome.

Table S16: 3 x 2 table for event-free survival as assessed by liver T_1 >825 ms.

Liver T_1	Clinical event	No clinical event
Liver T_1 >825 ms	10	**87 → 101**
Non-evaluable	**1 ↓**	**14 ↑**
Liver T_1 ≤825 ms	**3 → 4**	82

Arrows indicate how non-evaluable results are classified, depending on patient outcome.

Table S17: 3 x 2 table for event-free survival as assessed by liver T_1 >868 ms.

Liver T_1	Clinical event	No clinical event
Liver T_1 >868 ms	10	**61 → 75**

Non-evaluable	**1** ↓	**14** ↑
Liver T_1 ≤868 ms	**3** → **4**	108

Arrows indicate how non-evaluable results are classified, depending on patient outcome.

Table S18: 3 x 2 table for event-free survival as assessed by TE >8kPa.

TE	Clinical event	No clinical event
TE >8 kPa	9	**66** → **83**
Non-evaluable	**2**↓	**17**↑
TE ≤8 kPa	**2** → **4**	64

Arrows indicate how 19 non-evaluable results (failures) are classified, depending on patient outcome in addition to the unreliable results, which have simply been treated as valid measurements.

2 x 2 tables for all clinical endpoints

Table S19: 2 x 2 tables for binary cut-offs of non-invasive tests used in intention to diagnose analysis (n = 193)

	Event-free survival				All-cause mortality			
	TP	FP	TN	FN	TP	FP	TN	FN
cT_1 > 825	12	89	80	1	9	92	80	1
cT_1 > 840	12	80	89	1	9	83	89	1
T_1 >825 ms	10	87	82	3	8	89	83	2
T_1 >868 ms	10	61	108	3	8	63	109	2
liver fat > 5.6%	4	79	87	9	3	86	83	7
T_2* < 12.5 ms	2	33	136	11	1	34	138	9
T_2* < 8 ms	1	7	162	12	0	8	164	10
TE > 8 kPa	8	54	58	1	7	55	59	0
Ishak ≥ F5	11	43	122	2	9	45	122	2
FIB-4 >1.45	11	90	82	2	9	92	82	2

APRI >1	6	36	137	7	5	37	138	6
AST/ALT >1	10	60	114	3	8	62	114	3

Multivariate Cox regression analysis

Table S20: Multivariate cox regression analysis to assess independence of liver cT_1 and age for prediction of event-free survival.

	Hazard ratio	95 % CI	P value	Number of observations and events
cT_1	1.0067	1.002-1.011	0.002	182 observations, 13 events
Age	1.0574	1.001-1.117	0.047	
Cox model	-	-	0.0004	

SUPPLEMENTARY FIGURES

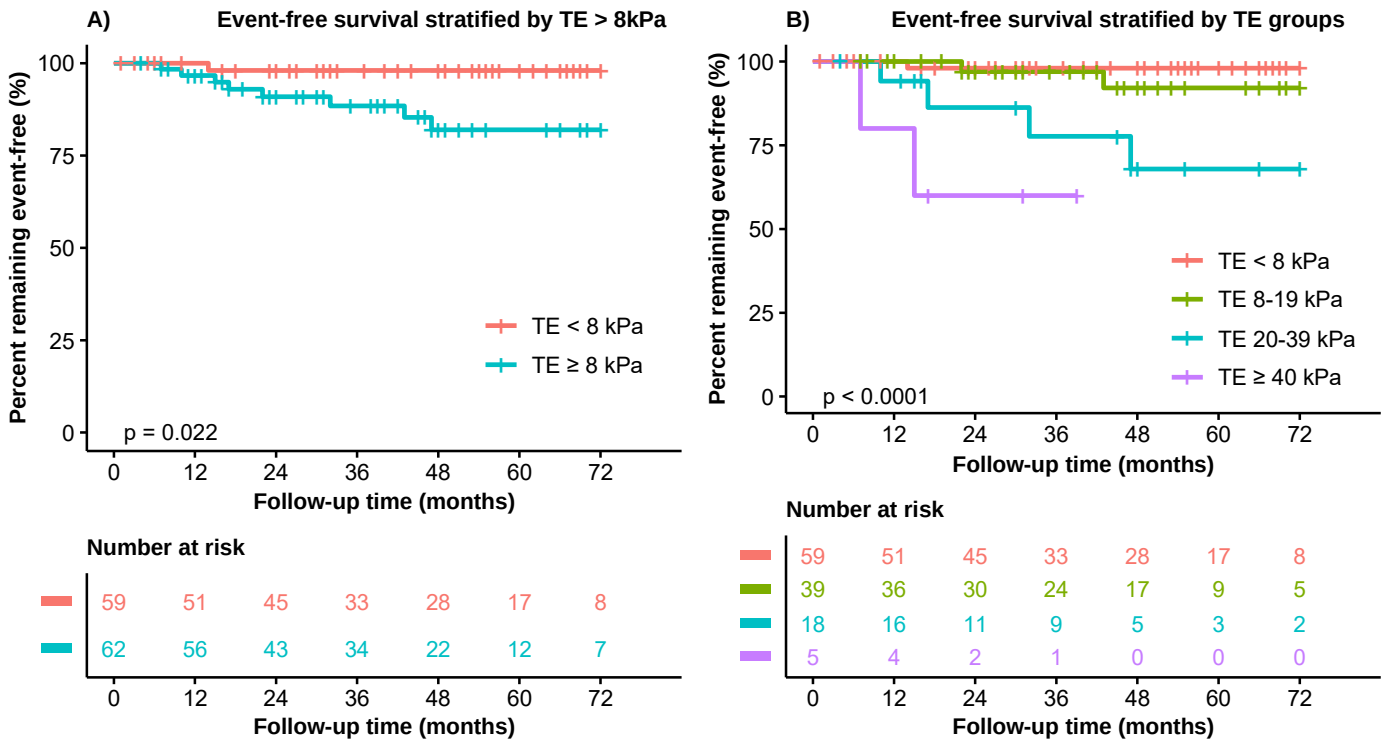


Figure S1: Kaplan-Meier curves for event-free survival stratified by A) TE ≥ 8kPa and B) TE < 8kPa, 8-19 kPa, 20-39 kPa, ≥ 40 kPa. P values represent log-rank test.

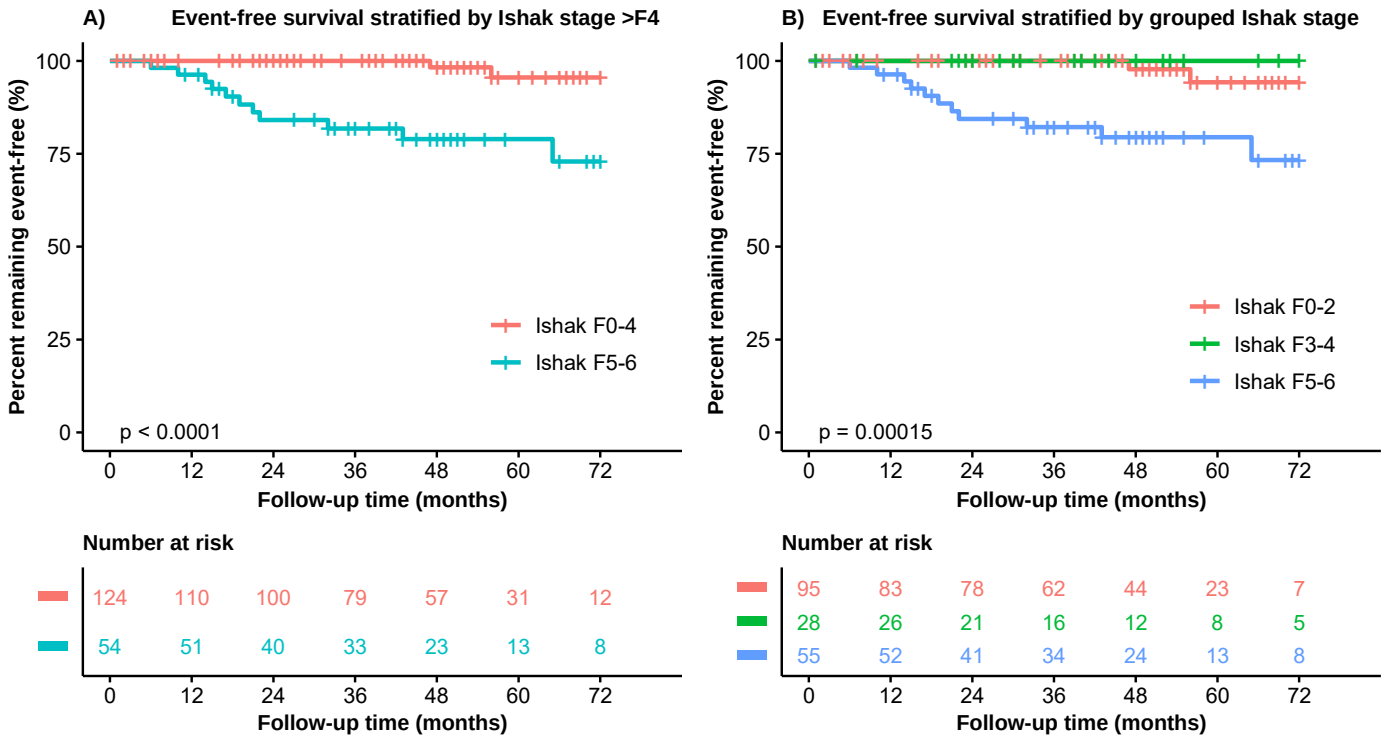


Figure S2: Kaplan-Meier curves for event-free survival stratified by A) Ishak stage ≥ F5 and B) Mild, moderate and severe fibrosis. P values represent log-rank test.

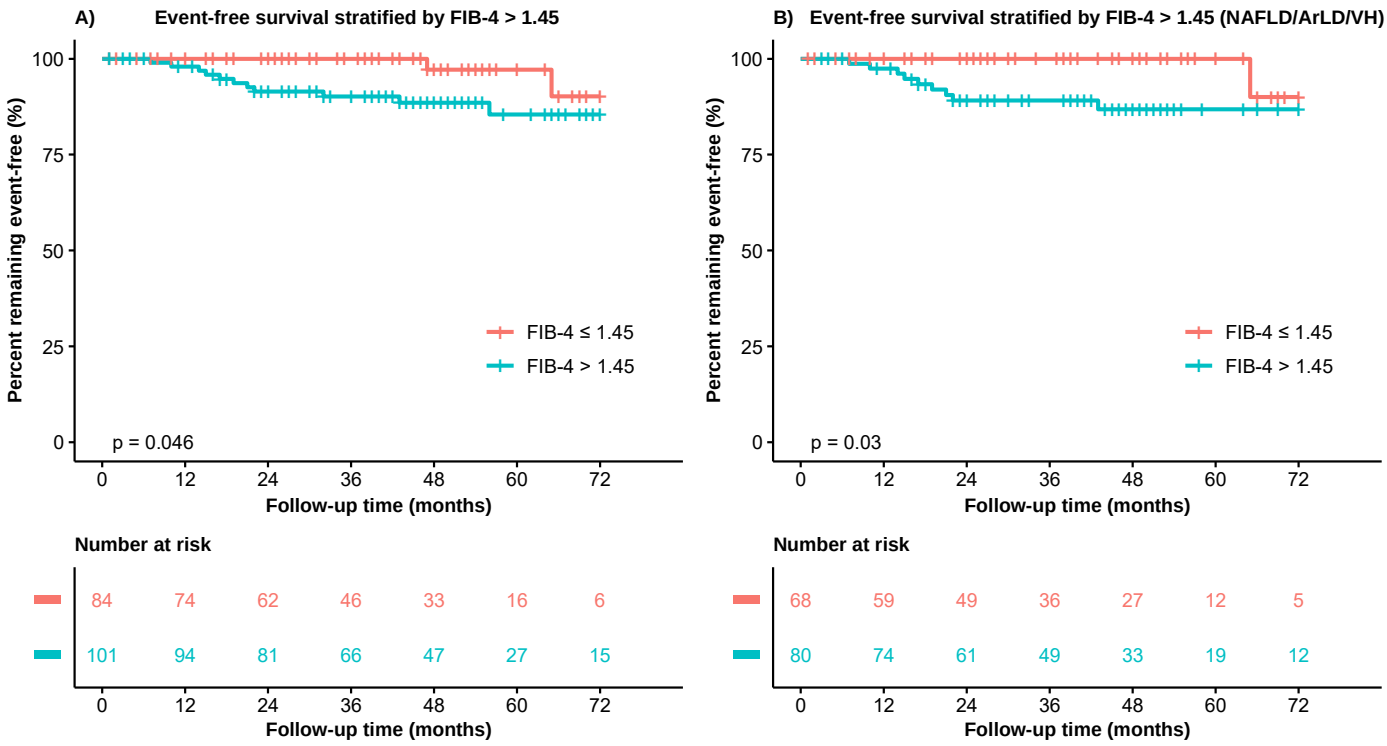


Figure S3: Kaplan-Meier curves for event-free survival stratified by FIB-4 >1.45 in A) all patients and B) NAFLD/ArLD/VH subgroup. P values represent log-rank test.

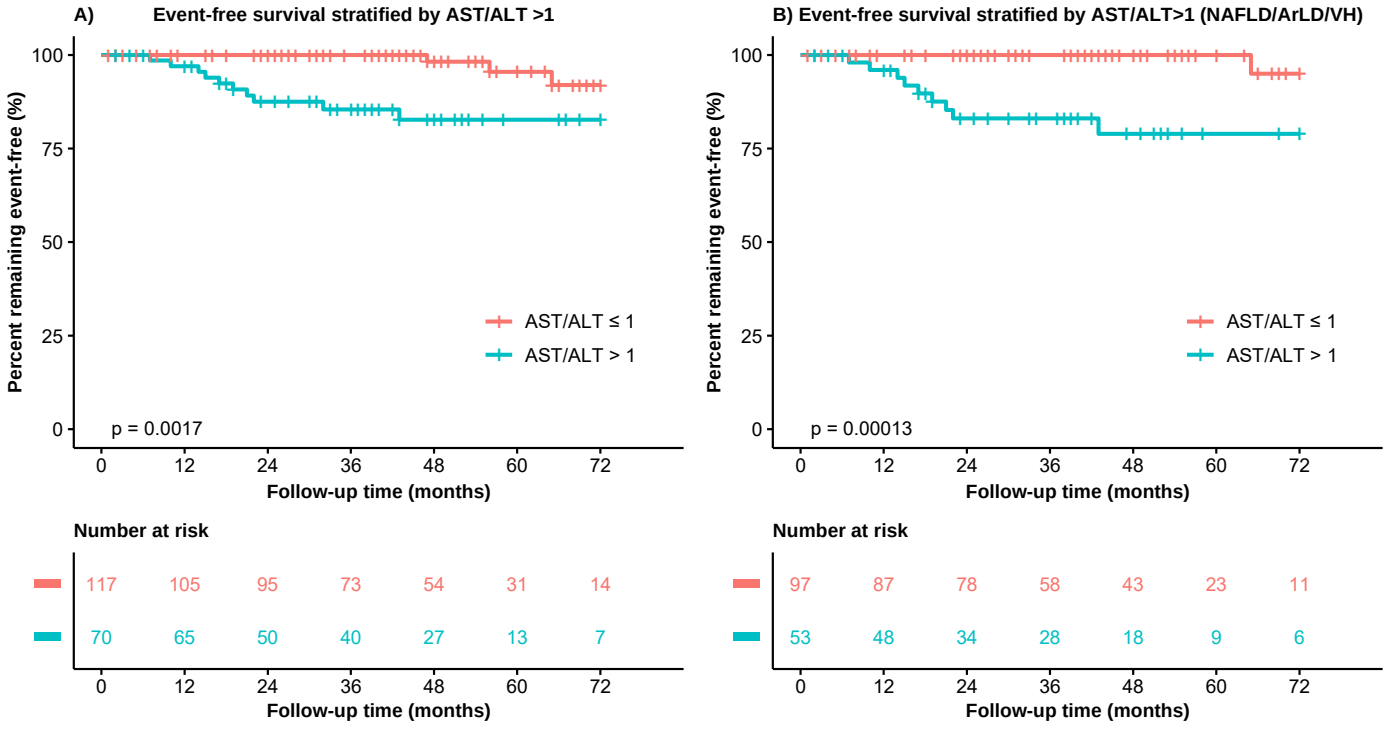
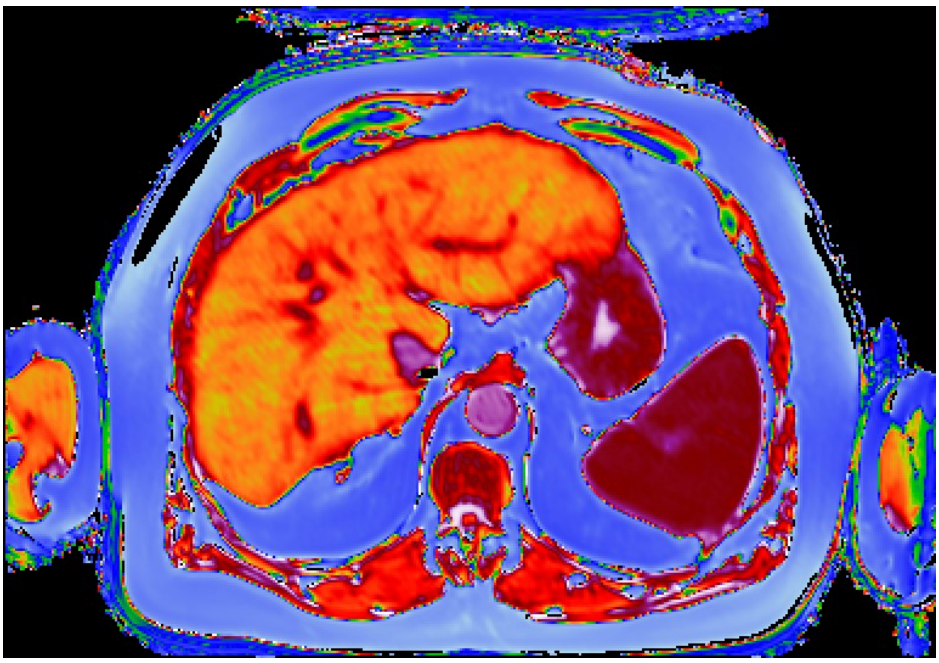
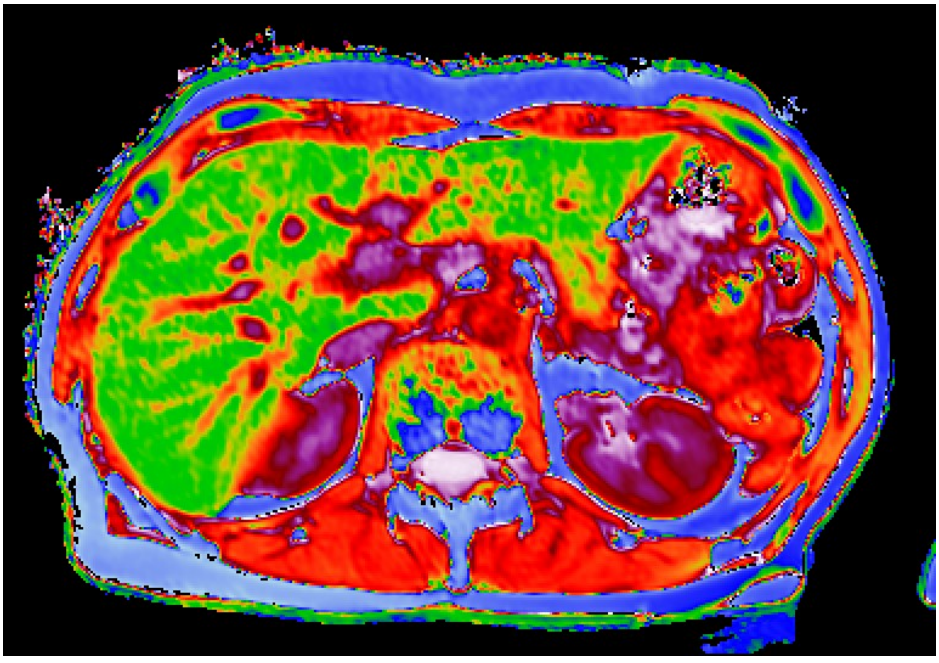


Figure S4: Kaplan-Meier curves for event-free survival stratified by AST/ALT >1 in A) all patients and B) NAFLD/ArLD/VH subgroup. P values represent log-rank test.





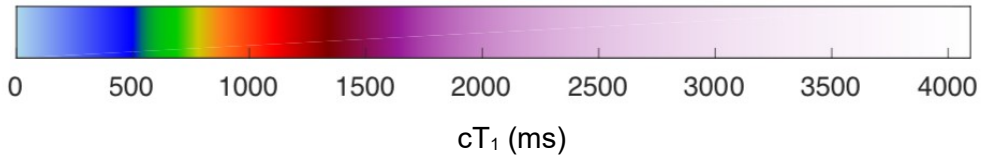


Figure S5*:* Example of abdominal iron-corrected T_1 (cT_1) colour maps of (top) a patient with compensated cirrhosis (Ishak F6) at time of scan who went on to develop hepatocellular carcinoma, ascites and death; (bottom) a patient with no histological or clinical evidence of cirrhosis at time of scan who did not develop any clinical event.

References

1. Pavlides, M. *et al.* Multiparametric magnetic resonance imaging predicts clinical outcomes in patients with chronic liver disease. *J. Hepatol.* **64**, 308–15 (2016).
2. Tunnicliffe, E. M., Banerjee, R., Pavlides, M., Neubauer, S. & Robson, M. D. A model for hepatic fibrosis: the competing effects of cell loss and iron on shortened modified Look-Locker inversion recovery T1(shMOLLI-T1) in the liver. *J. Magn. Reson. Imaging* **45**, 450–462 (2017).
3. Banerjee, R. *et al.* Multiparametric magnetic resonance for the non-invasive diagnosis of liver disease. *J. Hepatol.* **60**, 69–77 (2014).
4. Nuttall, K. L., Palaty, J. & Lockitch, G. Reference limits for copper and iron in liver biopsies. *Ann. Clin. Lab. Sci.* **33**, 443–50 (2003).
5. Szczepaniak, L. S. *et al.* Magnetic resonance spectroscopy to measure hepatic triglyceride content: prevalence of hepatic steatosis in the general population. *Am. J. Physiol. Metab.* **288**, E462–E468 (2005).
6. Fraquelli, M. *et al.* Reproducibility of transient elastography in the evaluation of liver fibrosis in patients with chronic liver disease. *Gut* **56**, 968–973 (2007).
7. Vergniol, J. *et al.* Noninvasive Tests for Fibrosis and Liver Stiffness Predict 5-Year Outcomes of Patients With Chronic Hepatitis C. *Gastroenterology* **140**, 1970-1979.e3 (2011).
8. Pang, J. X. Q. *et al.* Liver Stiffness by Transient Elastography Predicts Liver-Related Complications and Mortality in Patients with Chronic Liver Disease. *PLoS One* **9**, e95776 (2014).
9. Lin, Z.-H. *et al.* Performance of the aspartate aminotransferase-to-platelet ratio index for the staging of hepatitis C-related fibrosis: An updated meta-analysis. *Hepatology* **53**, 726–736 (2011).
10. Sterling, R. K. *et al.* Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology* **43**, 1317–1325 (2006).
11. Sheth, S. G., Flamm, S. L., Gordon, F. D. & Chopra, S. AST/ALT Ratio Predicts Cirrhosis in Patients With Chronic Hepatitis C Virus Infection. *Am. J. Gastroenterol.* **93**, 44–48 (1998).
12. Parkes, J. *et al.* Enhanced liver fibrosis test can predict clinical outcomes in patients with chronic liver disease. *Gut* **59**, 1245–1251 (2010).