APPENDICES

Supplementary Methods

Measurement of serum cytokeratin-18

CK-18 M30 was measured by commercial ELISA kits provided by Herui Biomed Company Limited, Suzhou, China. Coefficients of variation for intra-assay and inter-assay were <15%.

Other established non-invasive prediction models for NASH

Briefly, ION,[1] HAIR,[2] and NICE model[3] were calculated according to published formulas.

Body composition analysis

Body composition was evaluated by impedance analysis (InBody 720, Biospace, Seoul, Korea), which was a multifrequency impedance plethysmograph body composition analyzer and takes readings from the body using an eight-point tactile electrode method, measuring resistance at five specific frequencies (1 kHz, 50 kHz, 250 kHz, 500 kHz, and 1MHz) and reactance at three specific frequencies (5 kHz, 50 kHz, and 250 kHz).[4] Skeletal muscle mass, muscle mass, fat free mass and fat mass were measured after emptying the bladder and in light underwear using a calibrated InBody 720 bio-impedance device. All participants received similar instructions prior to the assessment of body composition and were required to be in a fasted state.[5]

Supplementary Table 1 – External validation cohorts description

		Derivation cohort	French cohort	Turkish cohort	Malaysian cohort	Egyptian cohort	Spanish cohort
Study description	**Funding**	Training funding by the High-level creative Talents from Department of public health in Zhejiang province, China	No funding		Research grant from the University of Malaya, Malaysia		
	Enrolment dates (first and last inclusion)	From 2017/12 to present	From 2004/04 to 2019/02	From to 2017 to present	First cohort: from 2012/11 to 2015/10 Second cohort: from 2016/09 to 2018/03	Between 2015/01 to 2019/10	From 2015 to present
	Study design	Prospective cross-sectional single center study	Prospective cross-sectional single center study	Prospective cross-sectional single center study	Prospective cross-sectional single center study	Prospective cross-sectional single center study	Prospective cross-sectional single center study
	PMID if data were used for publication	PMID: 31677195 PMID: 31625959 PMID: 31519069 PMID: 31195161 PMID: 31786360	PMID: 29577364		First cohort: PMID 28419855 Second cohort: PMID 31310032		PMID 31195161 PMID 30810330 PMID 30353552
	Center description	Hepatology tertiary care	Hepatology tertiary care	Hepatology tertiary care	Hepatology tertiary care	Hepatology tertiary care	Hepatology tertiary care
	Eligibility criteria	Inclusion: age 18-75 years; BMI < 35 kg/m^2; US, CT or MRI imaging showing fatty liver disease; abnormal ALT but below 5 ULN; no alcohol drinking history or daily alcohol intake < 20 g for male and 10 g for female	Inclusion: LB scheduled of the evaluation of NAFLD	Inclusion: 1) evidence of hepatic steatosis on ultrasound and/or fibrosis on transient elastography; 2) hepatomegaly or elevated aminotransferase levels, and 3) absence of secondary causes of hepatic fat accumulation (e.g., significant alcohol consumption [>21 units of alcohol per week in men and >14 units of alcohol per week in women] and previous history of steatogenic drugs use). Exclusion: patients with viral hepatitis, DILI, autoimmune hepatitis,	Inclusion: NAFLD patients diagnosed on US following exclusion of other cause of CLD including alcohol	Inclusion: FLD patients: ≥18 years old, with elevated ALT or significant fibrosis (≥ F2) by VCTE or FIB-4 with exclusion of other possible causes of CLD	Inclusion: age 18-75 years; BMI < 35 kg/m^2; US, CT or MRI imaging showing fatty liver disease; abnormal ALT but below 5 ULN; no alcohol drinking history or daily alcohol intake < 20 g for male and 10 g for female

				metabolic/genetic liver disease or low platelets count (< 100 x 10^9/L)			
Histological information	**Reason to send a patient to LB**	Persistent elevated transaminase or elevated LSM by VCTE or CAP (especially LSM)	Abnormal liver function tests, hyperferritinaemia, metabolic syndrome, abnormal non-invasive tests of liver fibrosis (Fib4, NFS, FibroMeter, LSM by VCTE)	Evidence of hepatic steatosis on US, abnormal liver enzymes or hepatomegaly, absence of secondary causes of hepatic fat accumulation (e.g. significant alcohol consumption and previous use of steatogenic drugs), LSM by VCTE >6 kPa or rarely patients with LSM by VCTE <6 kPa to exclude other CLD	Persistent ALT or AST ≥ 40, or reasons for NASH to be suspected (e.g. significant liver fibrosis based on liver stiffness measurement, obese patient with metabolic syndrome)	Elevated ALT or significant fibrosis (≥ F2) by VCTE or FIB-4	Persistent ALT or AST ≥ 40, or reasons for NASH to be suspected (e.g. significant liver fibrosis based on liver stiffness measurement, obese patient with metabolic syndrome, etc)
	LB reading	Central reading by a single expert pathologist	Prospective protocolized reading by a single expert pathologist	Central reading by a single expert pathologist	Central reading by a single expert pathologist	Reading by two independent expert pathologists.	Central reading by a single expert pathologist

BMI: body mass index, US: ultrasound, CT: computed tomography, MRI: magnetic resonance imaging, ALT: alanine aminotransferase; ULN: upper limit of normal, LB: liver biopsy, NAFLD: nonalcoholic fatty liver disease, DILI: drug induced liver injury, CLD: chronic liver disease, FLD: fatty liver disease, VCTE: vibration controlled transient elastography, FIB-4: fibrosis-4 socre, LSM: liver stiffness measurement, CAP: controlled attenuation parameter, NFS: NAFLD fibrosis score, NASH: nonalcoholic steatohepatitis.

Supplementary Table 2 – Potential risk of bias in derivation and external validation cohorts

		Derivation cohort	French cohort	Turkish cohort	Malaysian cohort	Egyptian cohort	Spanish cohort
Patients selection	**Potential bias due to patients selected for LB based on SCr or AST results**	✓ (LB in patients with FLD on US or FibroScan or CT or MRI)	✓ (LB in patients with abnormal liver function tests, hyperferritinaemia, metabolic syndrome, abnormal non-invasive tests of liver fibrosis by FibroMeter or LSM by VCTE)	✓ (LB in patients with FLD on US or fibrosis on FibroScan or hepatomegaly or elevated ALT levels)	× (LB in patients with persistent ALT or AST ≥ 40, or with suspected NASH)	✓ (LB in patients with FLD and elevated ALT or significant fibrosis (≥ F2) by VCTE or FIB-4)	✓ (LB in patients with FLD on US or fibrosis on FibroScan or hepatomegaly or elevated ALT levels)
LB quality	**Potential bias in LB quality**	✓ (90% have a LB length ≥ 15mm)	✓ (93% have a LB length ≥ 15mm)	✓ (97% have a LB length ≥ 15mm)	× (45% have a LB length ≥ 15mm)	✓ (90% have a LB length ≥ 15mm)	✓ (90% have a LB length ≥ 15mm)
	Potential bias in LB reading	✓ (double-blind central reading but single pathologist)	✓ (Prospective protocolized reading by a single pathologist)	✓ (double-blind central reading but single pathologist)	✓ (double-blind central reading but single pathologist)	× (routine reading)	✓ (double-blind central reading but single pathologist)
Timing	**Potential bias due to time interval between SCr evaluation and LB**	✓ (same day for all)	✓ (same day for all)	! (maximum 3 months' time interval between evaluation and LB)	✓ (same day for all)	! (no more than 3 months' time interval between evaluation and LB)	✓ (same day for all)
	Potential bias due to time interval between AST evaluation and LB	✓ (same day for all)	✓ (same day for all)	! (maximum 3 months' time interval between evaluation and LB)	✓ (same day for all)	! (no more than 3 months' time interval between evaluation and LB)	✓ (same day for all)

✓: low risk, ×: high risk, !: unclear

LB: liver biopsy, SCr: serum creatinine, AST: aspartate aminotransferase, US: ultrasound, CT: computed tomography, MRI: magnetic resonance imaging, LSM: liver stiffness measurement, VCTE: vibration controlled transient elastography, FLD: fatty liver disease, ALT: alanine aminotransferase; NASH: nonalcoholic steatohepatitis.

Supplementary Table 3 – Univariable and multivariable regression analyses of variables with definite NASH in the derivation cohort of Chinese patients with NAFLD.

Variables	univariable analysis		multivariable analysis	
	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Demographics				
Age (years)	0·960 (0·943-0·977)	<0·001	0·974 (0·954-0·995)	0·003
Sex (male)		0·049		
Body measurements				
Height (cm)		0·416		
Weight (kg)		0·053		
BMI (kg/m^2)	1·104 (1·037-1·175)	0·001		0·073
Waist circumference (cm)		0·056		
WHR		0·612		
Laboratory parameters				
AST (U/L)	1·051 (1·037-1·065)	<0·001	1·050 (1·035-1·066)	<0·001
ALT (U/L)	1·020 (1·014-1·026)	<0·001		0·261
AST/ALT ratio	0·346 (0·167-0·719)	0·004		
GGT (U/L)		0·120		
Alkaline phosphatase (U/L)		0·063		
Albumin (g/L)		0·087		
Platelet count (x10^9/L)		0·112		
Hemoglobin (g/L)		0·414		
Fasting glucose (mmol/L)		0·228		
HbA1c (%)		0·414		
HOMA-IR		0·058		0.122
Creatinine (μmol/L)	0·972 (0·957-0·986)	<0·001	0·964 (0·947-0·982)	<0·001

eGFR	1·025 (1·014-1·037)	<0·001	0·197
INR		0·966	
Total bilirubin (μmol/L)		0·844	
Total cholesterol (mmol/L)	1·353 (1·131-1·619)	0·006	0·128
Triglyceride (mmol/L)		0·221	
Uric acid (μmol/L)	1·003 (1·001-1·005)	0·001	0·229
Alpha-fetal protein (ng/ml)		0·601	
Hyaluronic acid (ng/ml)		0·564	
P3NP (ng/ml)	1·086 (1·048-1·125)	<0·001	
Ⅳ-C (ng/ml)	1·094 (1·056-1·134)	<0·001	
Laminin (ng/ml)		0·117	
Novel biomarkers related to NASH			
CK-18 M30 (U/L)	1·004 (1·002-1·005)	<0·001	
Concomitant diseases			
Hypertension (%)		0·218	
Type 2 diabetes (%)		0·517	

Note: The variables with no linear relationship between ALT/AST ratio, P3NP, Ⅳ-C and logit p (a probability of NASH occurrence), and still not related to logit p after conversion, were not included in multivariate analysis. CK-18 M30 was not included in multivariate analysis due to small sample of patients with available data (n =349).

Abbreviations: Ⅳ-C = type IV collagen

Supplementary Table 4 – Pairwise comparison of ROC curves between acNASH and the HAIR, ION, NICE model in the derivation cohort of Chinese patients with NAFLD.

Variable	AUROC	95% CI
acNASH	0·818	0·777-0·860
HAIR ***	0·621	0·570-0·669
ION ***	0·720	0·673-0·765
NICE model *	0·776	0·731-0·817

The HAIR score for each patient (0–3) was calculated by adding hypertension =1, ALT > 40 IU = 1, and HOMA-IR index > 5·0 = 1.

The index of NASH (ION) was calculated according to the following equation: 1·33 waist-to-hip ratio +0·03 × triglycerides (mg/dl) + 0·18 × ALT (U/L) +8·53 × HOMA-IR – 13·93 in men; 0·02 × triglycerides (mg/dl) + 0·24 × ALT (U/L) + 9·61 × HOMA-IR – 13·99 in women.

The NICE model was calculated as follows: -5·654 + 3·780E-02 × ALT (IU/L) + 2·215E-03 × CK18 fragment (IU/L) + 1·825 × (presence of metabolic syndrome = 1).

Note: Pairwise comparisons with acNASH, ***P value <0·001, *P value <0·05

Supplementary Table 5 – Pairwise comparison of ROC curves between acNASH and the HAIR, ION, NICE model in patients with established T2DM of the derivation cohort of Chinese patients with NAFLD.

Variable	AUROC	95% CI
acNASH	0·857	0·790-0·924
HAIR ***	0·635	0·532-0·738
ION ***	0·717	0·621-0·814
NICE model ***	0·769	0·680-0·857

The HAIR score for each patient (0–3) was calculated by adding Hypertension =1, ALT > 40 IU = 1, and HOMA-IR index > 5·0 = 1.

The index of NASH (ION) was calculated according to the following equation: 1·33 waist to hip ratio +0·03 * triglycerides (mg/dl) + 0·18 * ALT (U/L) +8·53 * HOMA-IR – 13·93 in men; 0·02 * triglycerides (mg/dl) + 0·24 * ALT (U/L) + 9·61 * HOMA-IR – 13·99 in women.

The NICE model was calculated as follows: -5·654 + 3·780E-02 * ALT (IU/L) + 2·215E-03 * CK18 fragment (IU/L) + 1·825 * (presence of metabolic syndrome = 1).

Note: Pairwise comparisons with acNASH, ***P value <0·001.

Supplementary Table 6 – Performance of the acNASH for the diagnosis of definite NASH on liver histology in patients with established T2DM of the derivation cohort and external validation cohorts.

Cohorts	AUROC (95% CI)	N	Prevalence of definite NASH	Diagnostic performance using dual cut-offs (cut-offs from derivation cohort)		
				rule-out zone	grey zone	rule-in zone
Derivation cohort	0·857 (0·790-0·924)	111	50 (45·0%)	**acNASH<4·15**	**acNASH: 4·15-7·73**	**acNASH >7·73**
				n=32 (29%)	n= 44 (40%)	n=35 (32%)
				Se= 0·96		Sp= 0·92
				Sp=0·49		Se= 0·60
				NPV= 0·94		PPV=0·86
French cohort	0·816 (0·764-0·869)	231	94 (40·7%)	**acNASH<4·15**	**acNASH: 4·15-7·73**	**acNASH >7·73**
				n=60 (26%)	n=118 (51%)	n=53 (23%)
				Se=1·00		Sp=0·91
				Sp=0·44		Se=0·43
				NPV=1·00		PPV=0·75
Turkish cohort	0·865(0·785-0·944)	93	67 (72·0%)	**acNASH<4·15**	**acNASH: 4·15-7·73**	**acNASH >7·73**
				n=14 (15%)	n= 49(53%)	n=30 (32%)
				Se= 0·97		Sp=0·96
				Sp= 0·46		Se=0·43
				NPV=0·86		PPV=0·97
Malaysian cohort	0·841 (0·787-0·895)	209	102 (48·8%)	**acNASH<4·15**	**acNASH: 4·15-7·73**	**acNASH >7·73**
				n=55 (26%)	n=88 (42%)	n=66 (32%)

				Se=0·93		Sp=0·90
				Sp=0·51		Se=0·54
				NPV=0·89		PPV=0·83
Spanish cohort	0·796 (0·676-0·916)	58	19 (29·7%)	**acNASH<4·15**	**acNASH: 4·15-7·73**	**acNASH >7·73**
				n=9 (16%)	n=19 (33%)	n=30 (52%)
				Se=1·00		Sp=0·72
				Sp=0·23		Se=0·58
				NPV=1·00		PPV=0·43
Pooled external cohort	0·814 (0·781-0·848)	591	281 (47·5%)	**acNASH<4·15**	**acNASH: 4·15-7·73**	**acNASH >7·73**
				n=145 (25%)	n=267 (45%)	n=179 (30%)
				Se=0·97		Sp=0·86
				Sp=0·44		Se=0·49
				NPV=0·94		PPV=0·77

Performance associated with a dual cut-off approach is evaluated using the acNASH index when the cut-offs are calculated in the derivation cohort and applied in several external validation cohorts. The lower cut-off constitutes a rule-out cut-off and is based on a sensitivity≥0·91 in the derivation cohort. The higher cut-off constitutes a rule-in cut-off and is based on a specificity≥0·91 in the derivation cohort. Individuals with an acNASH score between the rule-out and rule-in cut-offs are in the grey zone. In the rule-out group, the sensitivity is provided together with the specificity and negative predictive value to appraise the rule-out performance of the score. In the rule-in group, the specificity is provided together with the sensitivity and positive predictive value to appraise the rule-in performance of the score.

NB: The Egyptian cohort was excluded from this analysis because of the small sample of patients with established diabetes (n=15).

Abbreviations: AUROC: area under the receiver operating curve, NASH: non-alcoholic fatty liver disease, NPV: negative predictive value, PPV: positive predictive value, Se: sensitivity, Sp: specificity.

Supplementary Table 7 – Pairwise comparison of ROC curves between acNASH and the HAIR, ION, NICE model in patients with normal ALT of the derivation cohort of Chinese patients with NAFLD.

Variable	AUROC	95% CI
acNASH	0·821	0·743-0·884
HAIR ***	0·523	0·432-0·613
ION ***	0·642	0·552-0·726
NICE model ***	0·631	0·540-0·715

The HAIR score for each patient (0–3) was calculated by adding Hypertension =1, ALT > 40 IU = 1, and HOMA-IR index > 5·0 = 1.

The index of NASH (ION) was calculated according to the following equation: 1·33 waist to hip ratio +0·03 * triglycerides (mg/dl) + 0·18 * ALT (U/L) +8·53 * HOMA-IR – 13·93 in men; 0·02 * triglycerides (mg/dl) + 0·24 * ALT (U/L) + 9·61 * HOMA-IR – 13·99 in women.

The NICE model was calculated as follows: -5·654 + 3·780E-02 * ALT (IU/L) + 2·215E-03 * CK18 fragment (IU/L) + 1·825 * (presence of metabolic syndrome = 1).

Note: Pairwise comparison with acNASH, ***P value <0·001.

Supplementary Table 8 – Performance of the acNASH for the diagnosis of definite NASH on liver histology in patients with normal ALT levels of the derivation cohort and external validation cohorts.

Cohorts	AUROC (95% CI)	N	Prevalence of definite NASH	Diagnostic performance using dual cut-offs (cut-offs from derivation cohort)		
				rule-out zone	grey zone	rule-in zone
Derivation cohort	0·829 (0·744-0·914)	129	32 (24·8%)	**acNASH<4·15**	**acNASH: 4·15-7·73**	**acNASH >7·73**
				n=81 (63%)	n=40 (31%)	n=8 (6%)
				Se=0·78		Sp=1·00
				Sp=0·76		Se=0·25
				NPV=0·91		PPV=1·00
French cohort	0·835(0·768-0·902)	150	27 (18·0%)	**acNASH<4·15**	**acNASH: 4·15-7·73**	**acNASH >7·73**
				n=86 (57%)	n=62 (41%)	n=2 (2%)
				Se=0·89		Sp=1·00
				Sp=0·67		Se=0·07
				NPV=0·97		PPV=1·00
Turkish cohort	0·826(0·719-0·933)	54	23 (42·6%)	**acNASH<4·15**	**acNASH: 4·15-7·73**	**acNASH >7·73**
				n=21 (39%)	n=30 (56%)	n=3 (5%)
				Se=0·87		Sp=1·00
				Sp=0·58		Se=0·13
				NPV=0·86		PPV=1·00
Malaysian cohort	0·876(0·782-0·971)	87	12 (13·7%)	**acNASH<4·15**	**acNASH: 4·15-7·73**	**acNASH >7·73**
				n=51 (59%)	n=35 (40%)	n=1 (1%)

				Se=0·92 Sp=0·67 NPV=0·98		Sp=1·00 Se=0·08 PPV=1·00
Pooled external patients' cohort	0·849 (0·803-0·894)	291	62 (21·3%)	**acNASH<4·15**	**acNASH: 4·15-7·73**	**acNASH >7·73**
				n=158 (54%) Se=0·89 Sp=0·66 NPV=0·96	n=127 (44%)	n=6 (2%) Sp=1·00 Se=0·10 PPV=1·00

Performance associated with a dual cut-off approach is evaluated using the acNASH index when the cut-offs are calculated in the derivation cohort and applied in several external validation cohorts. The lower cut-off constitutes a rule-out cut-off and is based on a sensitivity ≥0·91 in the derivation cohort. The higher cut-off constitutes a rule-in cut-off and is based on a specificity ≥0·91 in the derivation cohort. Individuals with an acNASH score between the rule-out and rule-in cut-offs are in the grey zone. In the rule-out group, the sensitivity is provided together with the specificity and negative predictive value to appraise the rule-out performance of the score. In the rule-in group, the specificity is provided together with the sensitivity and positive predictive value to appraise the rule-in performance of the score. AUROC: area under the receiver operating curve, NASH: non-alcoholic fatty liver disease, NPV: negative predictive value, PPV: positive predictive value, Se: sensitivity, Sp: specificity.

NB: The Egyptian (n=25) and Spain cohorts (n=25) were excluded from this analysis because of their small sample of patients with normal ALT levels.

Supplementary Table 9 – Performance of acNASH for the diagnosis of definite NASH in women from the derivation cohort and external validation cohorts.

Cohorts	AUROC (95% CI)	N	Prevalence of definite NASH	Diagnostic performance using dual cut-offs (cut-offs from derivation cohort)		
				rule-out zone	grey zone	rule-in zone
Derivation cohort	0·789 (0·702-0·876)	106	62 (58%)	**acNASH<4·15**	**acNASH: 4·15-7·73**	**acNASH >7·73**
				n=16 (15%)	n=49 (46%)	n=41 (39%)
				Se=0·95		Sp=0·89
				Sp=0·30		Se=0·58
				NPV=0·81		PPV=0·88
French cohort	0·819(0·757-0·880)	170	77 (45%)	**acNASH<4·15**	**acNASH: 4·15-7·73**	**acNASH >7·73**
				n=35 (21%)	n=78 (46%)	n=57 (34%)
				Se=0·99		Sp=0·82
				Sp=0·37		Se=0·55
				NPV=0·97		PPV=0·74
Turkish cohort	0·858(0·773-0·943)	81	56 (69%)	**acNASH<4·15**	**acNASH: 4·15-7·73**	**acNASH >7·73**
				n=8 (10%)	n=45 (56%)	n=31 (38%)
				Se=1·00		Sp=0·92
				Sp=0·32		Se=0·52
				NPV=1·00		PPV=0·94
Malaysian cohort	0·898(0·845-0·951)	132	53 (40%)	**acNASH<4·15**	**acNASH: 4·15-7·73**	**acNASH >7·73**
				n=40 (30%)	n=55 (42%)	n=37 (28%)

				Se=0·98		Sp=0·91
				Sp=0·49		Se=0·57
				NPV=0·98		PPV=0·81
Egyptian cohort	0·740(0·587-0·893)	40	18 (45%)	**acNASH<4·15**	**acNASH:4·15-7·53**	**acNASH>7·73**
				n=5 (13%)	n=26 (65%)	n=9 (23%)
				Se=1·00		Sp=0·86
				Sp=0·23		Se=0·33
				NPV=1·00		PPV=0·67
Spain cohort	0·795(0·696-0·893)	84	18 (21%)	**acNASH<4·15**	**acNASH:4·15-7·53**	**acNASH>7·73**
				n=25 (30%)	n=41 (49%)	n=18 (21%)
				Se=1·00		Sp=0·85
				Sp=0·38		Se=0·44
				NPV=1·00		PPV=0·44
Pooled external patients cohort	0·836 (0·803-0·870)	507	222 (44%)	**acNASH<4·15**	**acNASH: 4·15-7·73**	**acNASH >7·73**
				n=113 (22%)	n=242 (48%)	n=152 (30%)
				Se=0·99		Sp=0·87
				Sp=0·39		Se=0·52
				NPV=0·98		PPV=0·76

Performance associated with a dual cut-off approach is evaluated using the acNASH index when the cut-offs are calculated in the derivation cohort and applied in several external validation cohorts. The lower cut-off constitutes a rule-out cut-off and is based on a sensitivity ≥0·91 in the derivation cohort. The higher cut-off constitutes a rule-in cut-off and is based on a specificity ≥0·91 in the derivation cohort. Individuals with an acNASH score between the rule-out and rule-in cut-offs are in the grey zone. In the rule-out group, the sensitivity is provided together with the specificity and negative predictive value to appraise the rule-out performance of the score. In the rule-in group, the specificity is provided together with the sensitivity and positive predictive value to appraise the rule-in performance of the score. AUROC: area under the receiver operating curve, NASH: non-alcoholic fatty liver disease, NPV: negative predictive value, PPV: positive predictive value, Se: sensitivity, Sp: specificity

Supplementary Table 10– Performance of acNASH for the diagnosis of definite NASH in men from the derivation cohort and external validation cohorts.

Cohorts	AUROC (95% CI)	N	Prevalence of definite NASH	Diagnostic performance using dual cut-offs (cut-offs from derivation cohort)		
				rule-out zone	grey zone	rule-in zone
Derivation cohort	0·822 (0·774-0·871)	284	128 (45%)	**acNASH<4·15**	**acNASH: 4·15-7·73**	**acNASH >7·73**
				n=98 (65%)	n=108 (38%)	n=78 (27%)
				Se=0·89		Sp=0·92
				Sp=0·54		Se=0·51
				NPV=0·86		PPV=0·83
French cohort	0·782(0·725-0·839)	278	71 (26%)	**acNASH<4·15**	**acNASH: 4·15-7·73**	**acNASH >7·73**
				n=91 (67%)	n=146 (53%)	n=41 (15%)
				Se=0·94		Sp=0·93
				Sp=0·42		Se=0·37
				NPV=0·96		PPV=0·63
Turkish cohort	0·787(0·667-0·907)	91	68 (75%)	**acNASH<4·15**	**acNASH: 4·15-7·73**	**acNASH >7·73**
				n=20 (78%)	n=50 (55%)	n=21 (23%)
				Se=0·90		Sp=0·91
				Sp=0·57		Se=0·28
				NPV=0·65		PPV=0·90
Malaysian cohort	0·801(0·723-0·880)	138	89 (64%)	**acNASH<4·15**	**acNASH: 4·15-7·73**	**acNASH >7·73**

				n=49 (36%)	n=55 (40%)	n=34 (25%)
				Se=0·87		Sp=0·92
				Sp=0·49		Se=0·51
				NPV=0·86		PPV=0·79
Spain cohort	0·795 (0·661-0·928)	54	17(31%)	**acNASH<4·15**	**acNASH:4·15-7·53**	**acNASH>7·73**
				n=6 (89%)	n=19 (35%)	n=29 (54%)
				Se=1·00		Sp=0·59
				Sp=0·16		Se=0·82
				NPV=1·00		PPV=0·48
Pooled external patients cohort	0·775 (0·737-0·812)	582	225 (39%)	**acNASH<4·15**	**acNASH: 4·15-7·73**	**acNASH >7·73**
				n=170 (71%)	n=279 (48%)	n=133 (23%)
				Se=0·92		Sp=0·89
				Sp=0·43		Se=0·42
				NPV=0·89		PPV=0·71

Performance associated with a dual cut-off approach is evaluated using the acNASH index when the cut-offs are calculated in the derivation cohort and applied in several external validation cohorts. The lower cut-off constitutes a rule-out cut-off and is based on a sensitivity ≥0·91 in the derivation cohort. The higher cut-off constitutes a rule-in cut-off and is based on a specificity ≥0·91 in the derivation cohort. Individuals with an acNASH score between the rule-out and rule-in cut-offs are in the grey zone. In the rule-out group, the sensitivity is provided together with the specificity and negative predictive value to appraise the rule-out performance of the score. In the rule-in group, the specificity is provided together with the sensitivity and positive predictive value to appraise the rule-in performance of the score. AUROC: area under the receiver operating curve, NASH: non-alcoholic fatty liver disease, NPV: negative predictive value, PPV: positive predictive value, Se: sensitivity, Sp: specificity.

NB: The Egyptian (n=21) was excluded from this analysis because of their small sample of patients of male.

Supplementary Table 11 – Regression models with different combination of variables for predicting NASH in both the derivation cohort and the pooled external validation cohorts.

Variables in the Models	Derivation cohort AUC (95% CI)	Pooled external cohorts AUC (95% CI)
age, AST, SCr	0·823 (0·782-0·863)	0·765 (0·723-0·808)
AST, SCr	0·818 (0·777-0·859)	0·805 (0·780-0·830)
AST, e-$GFR_{CKD-EPI}$	0·756 (0·708-0·804)	0·717 (0·687-0·748)

Abbreviations: SCr, serum creatinine; AST, aspartate aminotransferase; e-GFR, estimated glomerular filtration rate (using the CKD-EPI study equation)

Supplementary Table 12 – Comparisons of ROC curves between acNASH, AST, ALT and AST/ALT ratio of subgroups in the derivation cohort of Chinese patients with NAFLD.

Subgroups	acNASH AUROC (95% CI)	AST AUROC (95% CI)	ALT AUROC (95% CI)	AST/ALT ratio AUROC (95% CI)
Normal ALT	0·829 (0·744-0·914)	0·704 (0·586-0·821)	0·594 (0·487-0·702)	0·638 (0·526-0·749)
Hypertension	0·824 (0·753-0·894)	0·786 (0·708-0·864)	0·753 (0·669-0·837)	0·387 (0·288-0·486)
Type 2 diabetes	0·825 (0·750-0·900)	0·800 (0·718-0·882)	0·750 (0·660-0·839)	0·454 (0·345-0·562)

References

1. Otgonsuren M, Estep MJ, Hossain N, Younossi E, Frost S, Henry L, et al. Single non-invasive model to diagnose non-alcoholic fatty liver disease (NAFLD) and non-alcoholic steatohepatitis (NASH). *Journal of gastroenterology and hepatology* 2014; **29**(12): 2006-13.
2. Dixon JB, Bhathal PS, O'Brien PE. Nonalcoholic fatty liver disease: predictors of nonalcoholic steatohepatitis and liver fibrosis in the severely obese. *Gastroenterology* 2001; **121**(1): 91-100.
3. Anty R, Iannelli A, Patouraux S, Bonnafous S, Lavallard VJ, Senni-Buratti M, et al. A new composite model including metabolic syndrome, alanine aminotransferase and cytokeratin-18 for the diagnosis of non-alcoholic steatohepatitis in morbidly obese patients. *Alimentary pharmacology & therapeutics* 2010; **32**(11-12): 1315-22.
4. Volgyi E, Tylavsky FA, Lyytikainen A, Suominen H, Alen M, Cheng S. Assessing body composition with DXA and bioimpedance: effects of obesity, physical activity, and age. *Obesity (Silver Spring, Md)* 2008; **16**(3): 700-5.
5. Ogawa H, Fujitani K, Tsujinaka T, Imanishi K, Shirakata H, Kantani A, et al. InBody 720 as a new method of evaluating visceral obesity. *Hepatogastroenterology* 2011; **58**(105): 42-4.