Screening for therapeutic trials and treatment indication in clinical practice : MACK-3, a new blood test for the diagnosis of fibrotic NASH

Received: 9 October 2017 | First decision: 23 October 2017 | Accepted: 1 March 2018

DOI: 10.1111/apt.14621

WILEY AP&T Alimentary Pharmacology & Therapeutics

Screening for therapeutic trials and treatment indication in clinical practice: MACK-3, a new blood test for the diagnosis of fibrotic NASH

J. Boursier[1] | R. Anty[2] | L. Vonghia[3] | V. Moal[1] | T. Vanwolleghem[3] | C. M. Canivet[2] | S. Michalak[1] | S. Bonnafous[2] | P. Michielsen[3] | F. Oberti[1] | A. Iannelli[2] | L. Van Gaal[3] | S. Patouraux[2] | O. Blanchet[1] | A. Verrijken[3] | P. Gual[2] | M.-C. Rousselet[1] | A. Driessen[3] | G. Hunault[1] | S. Bertrais[1] | A. Tran[2] | P. Calès[1] | S. Francque[3]

[1]Angers, France

[2]Nice, France

[3]Antwerp, Belgium

Correspondence
Prof. J Boursier, Service d'Hépato-Gastroentérologie, Angers, France.
Email: JeBoursier@chu-angers.fr

Funding information
Nice University; Agence Nationale de la Recherche; Angers University Hospital; Institut National de la Santé et de la Recherche Médicale; National Research Agency

Summary

Background: The composite histological endpoint comprising nonalcoholic steatohepatitis (NASH) and NAFLD activity score ≥4 and advanced fibrosis ($F \geq 2$) ("fibrotic NASH") is becoming an important diagnostic target in NAFLD: it is currently used to select patients for inclusion in phase III therapeutic trials and will ultimately be used to indicate treatment in clinical practice once the new drugs are approved.

Aim: To develop a new blood test specifically dedicated for this new diagnostic target of interest.

Methods: Eight Hundred and forty-six biopsy-proven NAFLD patients from three centres (Angers, Nice, Antwerp) were randomised into derivation and validation sets.

Results: The blood fibrosis tests BARD, NFS and FIB4 had poor accuracy for fibrotic NASH with respective AUROC: 0.566 ± 0.023, 0.654 ± 0.023, 0.732 ± 0.021. In the derivation set, fibrotic NASH was independently predicted by AST, HOMA and CK18; all three were combined in the new blood test MACK-3 (hoMa, Ast, CK18) for which 90% sensitivity and 95% specificity cut-offs were calculated. In the validation set, MACK-3 had a significantly higher AUROC (0.847 ± 0.030, $P \leq 0.002$) than blood fibrosis tests. Using liver biopsy in the grey zone between the two cut-offs (36.0% of the patients), MACK-3 provided excellent accuracy for the diagnosis of fibrotic NASH with 93.3% well-classified patients, sensitivity: 90.0%, specificity: 94.2%, positive predictive value: 81.8% and negative predictive value: 97.0%.

Conclusion: The new blood test MACK-3 accurately diagnoses fibrotic NASH. This new test will facilitate patient screening and inclusion in NAFLD therapeutic trials and will enable the identification of patients who will benefit from the treatments once approved.

The Handling Editor for this article was Professor Stephen Harrison, and it was accepted for publication after full peer-review.
The authors' complete affiliations are listed in Appendix 1.

1 | INTRODUCTION

Nonalcoholic fatty liver disease (NAFLD), the liver manifestation of the metabolic syndrome (MetS), encompasses a wide spectrum of liver lesions.[1] While isolated steatosis is generally considered benign, nonalcoholic steatohepatitis (NASH) represents the aggressive form of the disease responsible for accelerated liver fibrosis progression, higher rates of cirrhosis and its associated complications, and higher rates of hepatocellular carcinoma.[2] Now that the burden of NAFLD among the general population has reached 25%,[3] NASH is expected to become the leading indication for liver transplantation in the United States within 10-20 years.[4]

Knowledge of the pathophysiological mechanisms driving NAFLD has advanced considerably and resultantly research in drug development has become a highly active field within which numerous therapeutic trials are underway.[5] A recent workshop conducted by the American Association for the Study of Liver Diseases concluded that therapeutic trials in NAFLD should enrol patients having an active disease defined by the presence of steatohepatitis and a NAFLD activity score (NAS) ≥ 4.[6] The first phase III trials in NASH were recently launched with the aim of demonstrating the clinical benefits of obeticholic acid (REGENERATE study, NCT02548351) and Elafibranor (RESOLVE-IT study, NCT02704403). These two trials are using clinical outcomes as study endpoints and, to be included, patients must have biopsy-proven NASH, a NAS ≥ 4 and significant/advanced F2-3 fibrosis. This enrichment of the study populations with fibrotic patients is intended to ensure that the number of required clinical events will occur during patient follow-up and within a reasonable time frame. Because it is a condition required for inclusion in phase III trials, the composite criterion of "NASH and NAS $\geq$4 and fibrosis $F \geq 2$" is therefore becoming an important new diagnostic target in NAFLD. More importantly, it will become an indication for treatment in clinical practice when and if these drugs receive their market authorisations. This second aspect particularly reinforces the interest of non-invasive tests able to accurately diagnose this composite histological endpoint. Several blood fibrosis tests are available but they have been usually developed for the diagnosis of advanced fibrosis and they are less accurate for earlier stages of fibrosis.[7] Serum level of apoptotic caspase-3 generated cytokeratin-18 fragments (CK18) has been shown as a candidate biomarker for NASH[8] but a large study has recently challenged its diagnostic accuracy.[9]

In the present work, we aimed to evaluate the accuracy of blood fibrosis tests and CK18 for the diagnosis of "NASH and NAS $\geq$4 and fibrosis $F \geq 2$" in a large multicentre series of NAFLD patients, and to improve the accuracy by developing an accurate blood test specifically designed for this new diagnostic target of interest.

2 | PATIENTS AND METHODS

2.1 | Patients

The study population was obtained by pooling the data from three published studies performed in Angers (France),[7] Antwerp (Belgium)[10] and Nice (France).[11] Inclusion criteria were biopsy-proven NAFLD as defined by the presence of liver steatosis after exclusion of concomitant steatosis-inducing drugs (such as corticosteroids, tamoxifen, amiodarone, or methotrexate), excessive alcohol consumption (>30 g/d in men or >20 g/d in women), chronic hepatitis B or C infection, and histological evidence of other concomitant chronic liver disease. Patients were excluded if they had cirrhosis complications (hepatic encephalopathy, ascites, variceal bleeding, systemic infection or hepatocellular carcinoma). Patient recruitment differed across the three centres: Angers centre included out-patients from its hepatology department; Antwerp centre included obese patients without pre-existing diabetes who presented at their obesity clinic; and Nice centre included morbidly obese patients referred for bariatric surgery. Metabolic syndrome was defined as the presence of at least three of the following parameters:[12] elevated waist circumference ($\geq$94 cm in men or $\geq$80 cm in women), elevated blood pressure (systolic blood pressure $\geq$130 mm Hg or diastolic blood pressure $\geq$85 mm Hg or antihypertensive drug), elevated glycemia ($\geq$100 mg/dL or anti-diabetic drug), elevated triglycerides ($\geq$150 mg/dl or lipid-lowering drug), low HDL cholesterol (<40 mg/dL in men or <50 mg/dL in women or lipid-lowering drug).

2.2 | Histology

Liver biopsy was performed percutaneously (16-gauge Menghini in Angers and Antwerp) or during surgery (14-gauge Tru-Cut in Antwerp or hepatic wedges in Nice). In each centre, pathological examination of liver biopsies was performed by a senior expert specialised in hepatology and blinded for patient data. Steatosis, lobular inflammation and ballooning were semi-quantitatively graded using the NASH-CRN scoring system.[13] The NAS, ranging from 0 to 8, corresponded to the sum of the scores for steatosis, lobular inflammation and ballooning.[14] NASH was defined according to the latest recommendations[6,15] as the presence of a score $\geq$1 for each of the three components of the NAS. "Active NASH" was defined as the presence of NASH with a NAS score $\geq$4. Liver fibrosis was staged from F0 to F4 according to the NASH-CRN scoring system.[13] Significant fibrosis was defined as $F \geq 2$ and advanced fibrosis as $F \geq 3$. Finally, the presence of an active NASH with significant fibrosis, ie NASH and NAS$\geq$4 and fibrosis $F \geq 2$, was termed "fibrotic NASH". This latter was the primary diagnostic target of the study.

A subgroup of 30 patients covering the histological spectrum of NAFLD lesions was selected in the Angers population to study the inter-centre agreement on histopathological evaluation of liver lesions. Liver biopsies were digitised using an Aperio digital slide scanner (Scanscope CS System, Aperio Technologies, Vista CA 92081, USA) image processor that provided high quality 30 000 $\times$ 30 000 pixel images at a resolution of 0.5 μm/pixel (magnification $\times$20). Two slides were digitised for each patient: one stained with haematoxylin and eosin and one with picrosirius red. The digitised slides were sent to the expert pathologist from each of the three investigating centres (SM, SP, AD) to score steatosis,

lobular inflammation, ballooning and fibrosis according to the NASH-CRN scoring system.

2.3 | Blood parameters

All blood samples from the three original studies were taken in fasting conditions.[7,10,11] The blood results available from these works enabled the calculation of the following blood fibrosis tests according to published formulas: BARD,[16] FIB4,[17] and NAFLD fibrosis score (NFS).[18] Serum level of insulin were measured on fresh samples using Chimiluminescence I2000 (Abbott) in Angers, Cobas Instrument (Roche) in Antwerp, and Centaur XP (Siemens) in Nice. HOMA was calculated as follow: fasting glucose (mmol/L) * fasting insulin (μU/mL)/22.5. To limit the influence of supra-physiological results of insulinemia induced by insulin therapy, the HOMA was capped at 10. Serum levels of CK18 were measured on frozen (−80°C) samples in each centre using the M30-Apoptosense enzyme-linked immunosorbent assay kit (PEVIVA, Bromma, Sweden).

2.4 | Statistics

Quantitative variables were expressed as mean ± standard deviation. Inter-observer agreement of the histological scoring was determined using the Fleiss' Kappa score.[19] Correlations between quantitative variables were determined using the Spearman correlation coefficient (Rs). The accuracy of non-invasive tests was evaluated using the area under the receiver operating characteristics (AUROC) and compared between tests using the Delong test.

2.4.1 | Development of the new blood test for the diagnosis of fibrotic NASH

To facilitate future automatic calculation by laboratory systems, only biological parameters were considered for the development of the new blood test dedicated to the diagnosis of fibrotic NASH. Variables with skewed distribution were ln transformed. The study population was randomly divided into derivation and validation sets. In the derivation set, the statistically significant variables identified by the univariate analysis ($P < 0.05$) were introduced in a multivariate analysis (stepwise forward binary logistic regression) whose regression formula was used to develop the new blood test. The diagnostic accuracy of the new test was then evaluated in the validation set.

2.4.2 | Sample size calculation

With α risk: 5%, β risk: 10%, fibrotic NASH prevalence: 25%, expected AUROC for fibrotic NASH: 0.70 for the best blood fibrosis test and 0.85 for the new test, AUROC correlation: 0.30 and bilateral testing, the required sample size for the validation set was 264 patients. The whole study population (n = 846) was therefore 2:1 randomly divided so that the validation set included 282 patients.

Statistical analyses were performed using SPSS version 18.0 software (IBM, Armonk, NY, USA) and SAS 9.1 (SAS Institute Inc., Cary, NC, USA). All authors had access to the study data and reviewed and approved the final manuscript.

3 | RESULTS

3.1 | Patient characteristics

The characteristics of the 846 patients included in the study are detailed in Table 1. A total of 62.1% were female and mean age was 47.5 ± 13.6 years. MetS was present in 67.7% of the patients and diabetes in 27.2% of them. In Angers and Antwerp, mean liver biopsy length was 24 ± 12 mm, 91.0% of biopsies were ≥10 mm, 73.9% were ≥15 mm and 60.8% were ≥20 mm. NASH was diagnosed in 54.0% of patients, NAS≥4 in 46.9% and $F \geq 2$ in 51.4%. Finally, 23.3% had fibrotic NASH as defined by the presence of all the three criteria: NASH, NAS≥4 and $F \geq 2$.

3.2 | Inter-observer agreement on histopathological evaluation of liver lesions

The liver biopsies of the 30 patients were all read by the expert pathologist from each investigating centre, for a total of 90 observations. The distribution of the recorded scores is shown in Table S1. Fleiss' Kappa score was 0.77 for steatosis, 0.62 for lobular inflammation, 0.64 for ballooning, 0.58 for NASH, 0.81 for the fibrosis stage, and 0.60 for fibrotic NASH. These results were similar to those obtained in a previous inter-observer reproducibility study performed between nine experts of the NASH-CRN (Table S2).

3.3 | Accuracy of blood fibrosis tests and CK18 for the diagnosis of fibrotic NASH

As expected, blood fibrosis tests performed better in diagnosing advanced fibrosis and cirrhosis (Table 2). In contrast, their accuracy was poor for the diagnosis of NASH or NAS ≥ 4 with AUROCs ≤ 0.65. For the diagnosis of fibrotic NASH, FIB4 had a higher AUROC (0.732 ± 0.021) than NFS (0.654 ± 0.023, $P < 0.001$) or BARD (0.566 ± 0.023, $P < 0.001$). Two diagnostic cut-offs are published for FIB4 (1.30 and 2.67) and for NFS (−1.455 and 0.676) in NAFLD.[18, 20] Their low cut-offs provided only poor to moderate sensitivity for fibrotic NASH: 49.0% of the patients with fibrotic NASH were included in the lower interval of the FIB4 (<1.30) and 30.4% in the lower interval of the NFS (<−1.455; Figure S1A). In addition, their high cut-offs identified very few patients with fibrotic NASH (FIB4: 10.8%, NFS: 18.0%; Figure S1A) and had poor positive predictive value (FIB4: 55.3%, NFS: 38.0%; Figure S1B). Taken together, these results suggest that BARD, FIB4, and NFS do not perform optimally for the detection of fibrotic NASH in NAFLD patients.

CK18 correlated significantly with steatosis, lobular inflammation, ballooning and fibrosis with significant differences between most of the adjacent grades/stages (Table S3). Furthermore, CK18 was significantly higher in patients with NASH compared to those without

TABLE 1 Patient characteristics at inclusion

	Angers (n = 309)	Antwerp (n = 270)	Nice (n = 267)	P[a]
Age (y)	56.1 ± 12.4	44.9 ± 12.3	40.3 ± 10.7	<0.001
Male sex (%)	60.8	37.8	11.6	<0.001
BMI (kg/m^2)	32.5 ± 6.0	40.0 ± 6.6	43.8 ± 4.8	<0.001
Elevated waist circumference (%)[b]	95.8	100.0	100.0	<0.001
Elevated blood pressure (%)[c]	79.5	69.7	51.3	<0.001
Elevated glycemia (%)[d]	75.7	19.3	44.4	<0.001
Elevated triglycerides (%)[e]	60.4	48.7	37.5	<0.001
Reduced HDL cholesterol (%)[f]	68.4	57.8	36.1	<0.001
Metabolic syndrome (%)[g]	83.1	62.8	54.4	<0.001
Diabetes (%)[h]	53.4	3.3	21.1	<0.001
Biopsy length (mm)	30 ± 12	16 ± 8	—[i]	<0.001
Steatosis (0/1/2/3, %)	5.8/46.3/29.1/18.8	0.0/44.1/31.5/24.4	6.4/36.3/26.2/31.1	<0.001
Lobular inflammation (0/1/2/3, %)	17.2/69.9/12.9/0.0	18.9/44.4/25.6/11.1	83.5/16.1/0.4/0.0	<0.001
Ballooning (0/1/2, %)	23.0/50.2/26.9	17.8/44.1/38.1	82.4/16.9/0.7	<0.001
NAS	3.6 ± 1.5	4.3 ± 1.8	2.2 ± 1.4	<0.001
NASH (%)	70.2	72.6	16.5	<0.001
Fibrosis stage (0/1/2/3/4, %)	8.4/26.2/29.8/29.4/6.1	60.4/18.1/12.6/8.1/0.7	1.5/33.0/60.3/5.2/0.0	<0.001
Fibrotic NASH (%)[j]	42.1	15.9	9.0	<0.001
AST (IU/l)	45 ± 24	33 ± 18	27 ± 16	<0.001
ALT (IU/l)	64 ± 38	48 ± 28	35 ± 28	<0.001
Gamma GT (IU/L)	118 ± 141	44 ± 33	43 ± 40	<0.001
Bilirubin (μmol/L)	11 ± 6	10 ± 4	8 ± 4	<0.001
Prothrombin time (%)	98 ± 12	94 ± 10	92 ± 9	<0.001
Platelets (G/L)	224 ± 67	294 ± 66	278 ± 66	<0.001
CK18 (IU/L)	289 ± 321	264 ± 243	260 ± 172	0.082
HbA1c (%)	6.6 ± 1.3	5.7 ± 0.6	6.0 ± 1.2	<0.001

BMI, body mass index.

[a]By Kruskal–Wallis test; [b]waist circumference ≥94 cm in men or ≥80 cm in women; [c]systolic blood pressure ≥130 mm Hg or diastolic blood pressure ≥85 mm Hg or antihypertensive drug; [d]glycemia ≥100 mg/dL or antidiabetic drug; [e]triglycerides ≥150 mg/dL or lipid-lowering drug; [f]HDL cholesterol <40 mg/dL in men or <50 mg/dL in women or lipid-lowering drug; [g]defined as ≥3 of the following parameters: elevated waist circumference, elevated blood pressure, elevated glycemia, elevated triglycerides, low HDL cholesterol; [h]glycemia ≥126 mg/dL or antidiabetic drug; [i]surgical samples; [j]defined as the presence of NASH and NAS ≥4 and fibrosis $F \geq 2$.

(respectively: 319 ± 316 vs 217 ± 142 IU/l, $P < 0.001$). Despite this, CK18 showed low accuracy for the diagnosis of NASH, with an AUROC at 0.595 ± 0.019 (Table 2). Thereafter, CK18 was evaluated as a function of NASH, NAS ≥ 4 and $F \geq 2$. In the NAS < 4 and NAS ≥ 4 subgroups, CK18 levels did not differ significantly between patients with or without NASH (Figure S2A). Interestingly, in the four subgroups defined by NASH and NAS ≥ 4, CK18 levels were significantly higher in $F \geq 2$ patients compared to F0-1 patients (Figure S2B). Consequently, CK18 performed better in diagnosing fibrotic NASH (Table 2) and appeared to be a candidate biomarker for this diagnostic target.

3.4 | New blood test for the non-invasive diagnosis of fibrotic NASH

The 846 patients were 2:1 randomly divided into derivation and validation sets, between which patients characteristics were not significantly different (Table S4). To facilitate a future automatic calculation of the new test by laboratory systems, only biological parameters were tested in the derivation set. Multivariate analysis identified AST, HOMA and CK18 as independent predictors of fibrotic NASH (Table S5). A new blood test combining these three parameters was derived using the regression formula of the multivariate analysis. This new test, called MACK-3 (hoMa, Ast, CK18), provides a result ranging from 0 to 1. In the derivation set, the MACK-3 thresholds corresponding to 90% sensitivity and 95% specificity for fibrotic NASH were, respectively, 0.134 and 0.550.

In the validation set, MACK-3 had a significantly higher AUROC for the diagnosis of fibrotic NASH (0.847 ± 0.030) than BARD, NFS and FIB4 ($P \leq 0.002$, Table 3). 48.3% of the patients had MACK-3 ≤ 0.134 (90% sensitivity threshold), 36.0% were in the intermediate grey zone (0.135-0.549) and 15.7% had MACK-3 ≥ 0.550 (95% specificity threshold). The prevalence of fibrotic NASH in these three intervals was, respectively, 4.7%, 25.0% and 71.4% ($P < 0.001$,

TABLE 2 AUROCs of blood fibrosis tests and CK18 for the diagnosis of NAFLD lesions

Diagnostic target	BARD	NFS	FIB4	CK18
Mild fibrosis ($F \geq 1$)	0.626 ± 0.022	0.721 ± 0.020	0.653 ± 0.021	0.607 ± 0.022
Significant fibrosis ($F \geq 2$)	0.641 ± 0.019	0.686 ± 0.018	0.648 ± 0.019	0.634 ± 0.019
Advanced fibrosis ($F \geq 3$)	0.633 ± 0.025	0.720 ± 0.023	0.808 ± 0.021	0.659 ± 0.027
Cirrhosis (F4)	0.648 ± 0.065	0.865 ± 0.036	0.874 ± 0.040	0.572 ± 0.061
NASH	0.466 ± 0.020	0.521 ± 0.020	0.651 ± 0.019	0.595 ± 0.019
NAS ≥4	0.433 ± 0.020	0.490 ± 0.020	0.604 ± 0.019	0.643 ± 0.019
NASH and NAS ≥4	0.449 ± 0.020	0.493 ± 0.020	0.597 ± 0.020	0.638 ± 0.019
Fibrotic NASH[a]	0.566 ± 0.023	0.654 ± 0.023	0.732 ± 0.021	0.715 ± 0.022

NAS, NAFLD activity score.
[a]Defined as the presence of NASH and NAS ≥4 and $F \geq 2$.

TABLE 3 AUROCs of MACK-3 and blood fibrosis tests for the diagnosis of fibrotic NASH

Test	All	Derivation set	Validation set	*P*
BARD	0.566 ± 0.023	0.584 ± 0.028	0.528 ± 0.042	0.267
NFS	0.654 ± 0.023	0.671 ± 0.027	0.618 ± 0.040	0.272
FIB4	0.732 ± 0.021	0.738 ± 0.025	0.721 ± 0.037	0.703
MACK-3	0.846 ± 0.016	0.845 ± 0.019	0.847 ± 0.030	0.955
Comparison (*P*)				
BARD vs NFS	<0.001	0.003	0.015	—
BARD vs FIB4	<0.001	<0.001	<0.001	—
BARD vs MACK-3	<0.001	<0.001	<0.001	—
NFS vs FIB4	<0.001	0.007	<0.001	—
NFS vs MACK-3	<0.001	<0.001	<0.001	—
FIB4 vs MACK-3	<0.001	<0.001	0.002	—

Figure 1). MACK-3 ≤ 0.134 provided 90.0% sensitivity for fibrotic NASH, and MACK-3 ≥ 0.550 had 94.2% specificity. Among the patients in the highest interval (MACK-3 ≥ 0.550), 85.7% had fibrotic NASH or advanced F3-4 fibrosis (Figure 1). Using liver biopsy in the grey zone between the 0.134 and 0.550 thresholds, MACK-3 showed an excellent diagnostic accuracy for fibrotic NASH with 93.3% well-classified patients (Table 4).

3.5 | Sensitivity analysis

The accuracy of MACK-3 for the diagnosis of fibrotic NASH was different across the three centres (Table S6), likely reflecting the different populations they included. To evaluate whether there was a centre effect or this was due to the different characteristics of the three populations, we tested the influence of centre, age, sex, BMI, MetS, oral antidiabetic treatment, insulin treatment, biopsy length ≥20 mm, NAFLD/bariatric patients and derivation/validation set. By multivariate analysis including all these parameters, only MetS was independently associated with the rate of well-classified patients by MACK-3 (Table S7). Diagnostic accuracy of MACK-3 as a function of MetS presence/absence is presented in Table S6.

3.6 | Practical algorithm for clinical practice

NAFLD represents a very large population in which using MACK-3 to identify the subgroup of patients who need to be treated would require extensive measurement of CK18 and HOMA-IR. To limit unnecessary measurements of these biomarkers, we evaluated whether simple parameters available in routine practice (age, sex, BMI, MetS, AST, ALT, GGT) can be used in a first step to identify a subgroup with no need for MACK-3 calculation because of a very low prevalence of fibrotic NASH. In the derivation set, multivariate analysis including the seven simple parameters identified AST and MetS as independent predictors of fibrotic NASH. 25.1% of the patients had neither MetS nor elevated AST ≥35 UI/L with very low prevalence of fibrotic NASH in this subgroup (0.7%). This result was confirmed in the validation set where the prevalence of fibrotic NASH was 1.5% in patients having neither MetS nor AST ≥35 UI/L. This very low prevalence was also consistent across the three centres: 0.0% in Angers, 1.3% in Antwerp and 1.0% in Nice ($P = 0.876$).

Therefore, we derived the MACK-3 algorithm where we propose to use MACK-3 only in patients having MetS and/or AST ≥35 UI/L (Figure 2). Accuracy of the MACK-3 algorithm was not significantly different between the derivation and the validation sets (Table S8).

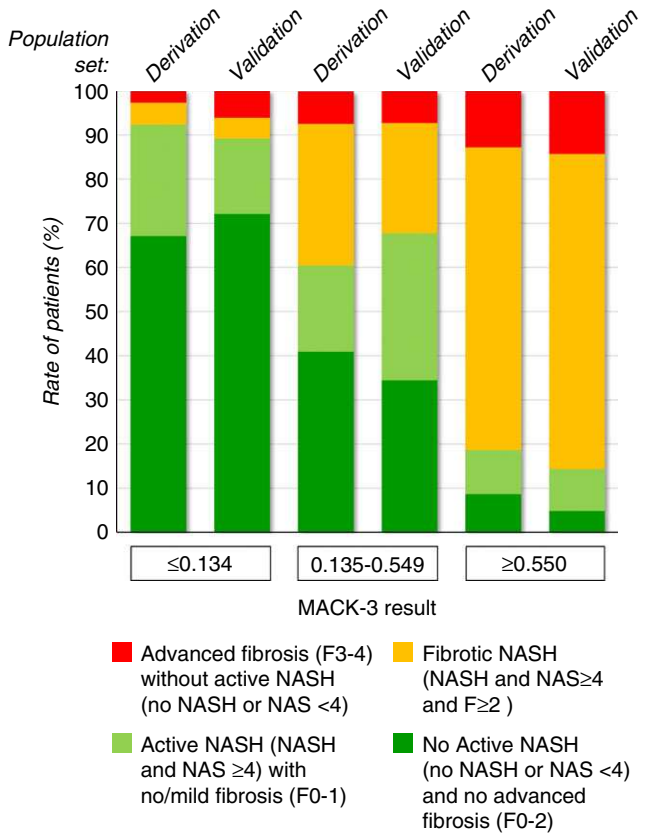

FIGURE 1 Prevalence of NAFLD lesions as a function of the 3 intervals of MACK-3 results and derivation/validation sets

TABLE 4 Accuracy of MACK-3 for the diagnosis of fibrotic NASH. MACK-3 was used with its 2 thresholds (≤0.134: no fibrotic NASH; ≥0.550: fibrotic NASH). Liver biopsy was performed in the grey zone

	All	Derivation set	Validation set	*P* value
DA	93.5	93.6	93.3	0.881
Se	90.0	90.0	90.0	1.000
Spe	94.5	94.7	94.2	0.852
NPV	96.9	96.8	97.0	—
PPV	83.4	84.2	81.8	—
−LR	0.11	0.11	0.11	—
+LR	16.4	16.9	15.5	—
OR	155.1	160.0	146.3	—
LB	38.4	39.6	36.0	0.356

DA, overall diagnostic accuracy (ie, rate of well-classified patients, %); Se, sensitivity (%); Spe, specificity (%); NPV, negative predictive value (%); PPV, positive predictive value; −LR, negative likelihood ratio; +LR, positive likelihood ratio; OR, diagnostic odds ratio; LB, liver biopsy requirement (%).

In the whole population, the rates of patients included in the four subgroups of the MACK-3 algorithm were, respectively, 25.6%, 26.1%, 34.4% and 13.9%. The prevalence of fibrotic NASH in the 4 subgroups of the MACK-3 algorithm was, respectively, 1.0%, 8.7%, 33.5% and 69.1% (Figure 3). Among the patients included in the last subgroup (MetS and/or AST ≥35 UI/L with MACK-3 ≥ 0.550), 82.7% had fibrotic NASH or advanced F3-4 fibrosis (Figure 3). Only 10.7% of the patients with fibrotic NASH and 15.3% of the patients with advanced F3-4 fibrosis were missed by the MACK-3 algorithm, which respectively represented 2.5% and 2.8% of the whole population. Finally, the MACK-3 algorithm provided an excellent accuracy for the diagnosis of fibrotic NASH with 93.2% well-classified patients, 89.3% sensitivity, 94.4% specificity, 83.1% positive predictive value and 96.6% negative predictive value (Table S8).

4 | DISCUSSION

NAFLD is closely linked to obesity and insulin resistance, and consequently its treatment is primarily based on lifestyle modifications.[1] Weight loss of ≥10% is associated with a 90% resolution in NASH but less than 10% of patients are able to achieve and maintain that objective.[21] In this context, the development of new drugs to reduce liver inflammation and stop fibrosis progression in NAFLD is receiving a great amount of attention.[22-26] Fibrotic NASH is now used in the stratification of NAFLD severity by the latest international EASL–EASD–EASO Clinical Practice Guidelines.[1] Two recently launched phase III therapeutic trials use the presence of fibrotic NASH as an inclusion criterion. As it is very difficult to correctly identify this subset of patients within the very large population of NAFLD patients, we performed a large multicentre study that resulted in the development of MACK-3. This new blood test showed a very good 0.846 ± 0.016 AUROC for the diagnosis of fibrotic NASH. We finally proposed the new three-step MACK-3 algorithm with an excellent 93.2% diagnostic accuracy. To facilitate MACK-3 calculation for a patient in clinical practice or for an entire dataset in a clinical study, we have developed an online calculator available at the following weblink: http://forge.info.univ-angers.fr/~gh/wstat/mack3-calculator.php

Because fibrosis is part of the definition of fibrotic NASH, we first evaluated known blood fibrosis tests for this diagnostic target. NFS and FIB4 showed only moderate diagnostic accuracy, with respective AUROCs at 0.65 and 0.73. We also evaluated CK18 and our results (AUROC: 0.60) confirmed the poor diagnostic accuracy of CK18 for the diagnosis of NASH recently reported by Cusi et al[9] Interestingly, however, CK18 was significantly correlated with each of the individual liver lesions observed in NAFLD (Table S3) and was highest in patients with fibrotic NASH (Figure S2B). Finally, rather than a biomarker for the single diagnosis of NASH, CK18 was more globally associated with the severity of NAFLD and showed promise as a potential candidate for the non-invasive diagnosis of fibrotic NASH.

To improve the non-invasive diagnosis of fibrotic NASH, we identified relevant biomarkers by multivariate analysis in the derivation set and combined them in a new blood test. The new blood test MACK-3 includes biomarkers associated with liver inflammation (AST), insulin resistance (HOMA), and apoptosis (CK18). All these

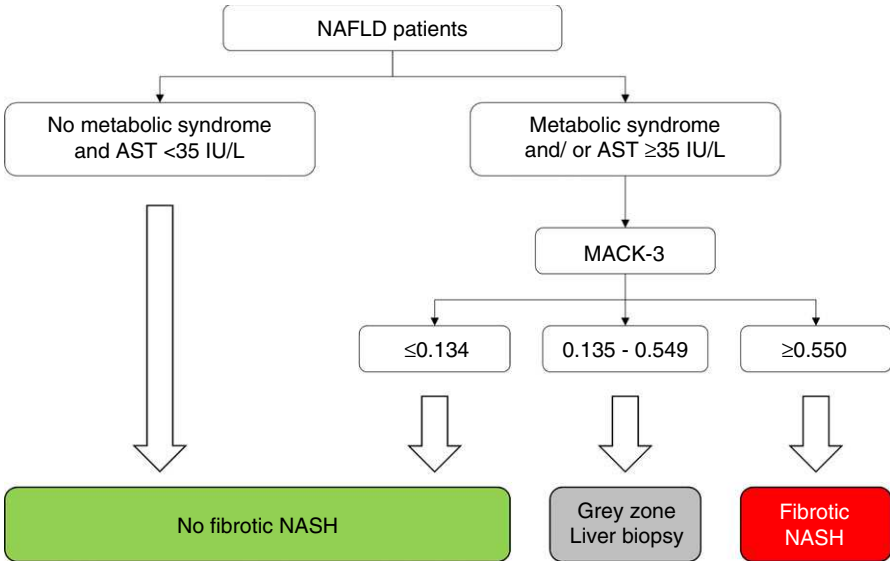

FIGURE 2 MACK-3 algorithm for the diagnosis of fibrotic NASH in clinical practice

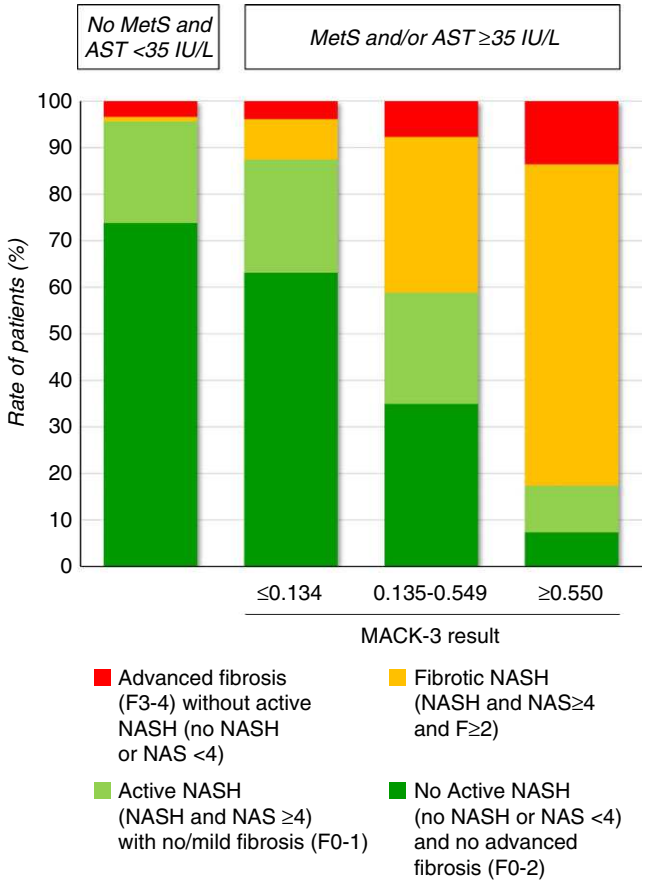

FIGURE 3 Prevalence of NAFLD lesions as a function of the 4 subgroups defined by the MACK-3 algorithm

conditions are closely linked with NAFLD severity, providing thus pathophysiological relevance for this new diagnostic test. Inclusions in phase III therapeutic trials in the setting of NASH are currently challenged by high rates of screening failure. Since the MACK-3 ≥ 0.550 threshold identifies a subset of NAFLD patients with a high rate of fibrotic NASH, the use of this new test should considerably reduce the rate of screening failure.

NAFLD now concerns 25% of the general population.[3] Consequently, with the upcoming arrival of new treatments in the clinic, many patients will need to be tested for the presence of fibrotic NASH. With the aim of limiting extensive measurements of specialised biomarkers and therefore costs related to the use of MACK-3 in clinical practice, we sought to identify simple parameters able to exclude fibrotic NASH. Such approach has been previously developed for the non-invasive diagnosis of advanced fibrosis in chronic liver diseases.[27] In the derivation set, we found that the absence of MetS combined with AST <35 UI/L defined a subgroup of patients within which the prevalence of fibrotic NASH was only 0.7%. This result was confirmed in the validation set and was consistent across the three centres. Thus, in the MACK-3 algorithm, we propose to perform MACK-3 calculation only for patients who have either MetS or AST ≥35 UI/L, and to use liver biopsy in a third step if MACK-3 results fall in the grey zone. The MACK-3 algorithm provided excellent accuracy for the diagnosis of fibrotic NASH: the rate of correctly classified patients was 93.2%, sensitivity was 89.3%, and positive predictive value was 83.1%.

The populations included by the three centres in our study were significantly different, due to different recruitment sources. That difference does however reflect the heterogeneity of NAFLD patients in clinical practice. We decided to pool our three populations because large multicentre samples are more suitable for the development of diagnostic tests than single centre populations, the latter being closely linked to intrinsic characteristics of the local population, as well as centre procedures and expertise. Finally, we estimate that pooling our three NAFLD populations was an advantage, allowing us to cover a large spectrum of NAFLD patients and the wide heterogeneity of the dysmetabolic disease. Our results are thus probably more robust than they would have been if we had limited our study population to a single centre.

The reading of liver biopsies by local pathologists as the different assays used to measure insulinemia could have introduced inter-centre variability in our work.[13,28] To test this, we performed a multivariate analysis including the "center" variable and the parameters that differed significantly between the three populations. The only parameter associated with the rate of correctly classified patients by MACK-3 was the metabolic syndrome, with no effect of the "center." This suggested that the difference in correctly classified patients we observed between the three investigating sites was linked to the different prevalence of the metabolic syndrome and not to local expertise or procedures. The reproducibility study between the three pathologists involved in the study showed an inter-observer agreement that was very close to what has been previously observed between nine expert pathologists of the NASH-CRN.[13] These results validate the expertise of our pathologists and therefore the quality of the liver biopsy examinations performed in our work. Nonetheless, we acknowledge that further studies are required to independently validate the diagnostic accuracy of MACK-3 in larger, "real-life" NAFLD patient sets, and to evaluate inter-laboratory reproducibility.

Fibrotic NASH, defined by the presence of NASH + NAS≥4 + $F \geq 2$, has the limitation to exclude F3-4 patients without NASH. Several accurate methods exist for the non-invasive evaluation of liver fibrosis in NAFLD.[29] In a recent large study where we directly compared Fibroscan and eight blood fibrosis tests, FibroMeter^{V} and Fibroscan showed the best accuracies for the non-invasive diagnosis of advanced F3-4 fibrosis in NAFLD.[7] Further works will have to evaluate whether the combination of MACK-3 with FibroMeter^{V} and/or Fibroscan improves the stratification of NAFLD patients.

Despite it is well established that fibrosis stage is the main predictor of NAFLD prognosis,[30] it remains to demonstrate that fibrotic NASH is predictive of patient outcome. Liver biopsy is an imperfect reference for liver lesions because of sample bias and suboptimal inter-observer reproducibility.[31,32] This is well illustrated by our results: despite our three pathologists involved were all senior experts specialised in hepatology, inter-observer reproducibility for fibrotic NASH was modest with kappa index at 0.60. Due to this lack of gold standard, it is not possible to assess the true diagnostic accuracy of MACK-3. Longitudinal studies circumvent this limitation by evaluating the prognostic significance of non-invasive tests.[33] Such longitudinal works are now required to validate if a blood test targeted for fibrotic NASH, like MACK-3, accurately identifies the NAFLD patients who have impaired prognosis and therefore require treatment with the upcoming new drugs.

In conclusion, MACK-3 is a new blood test combining AST, HOMA and CK18 for an accurate diagnosis of fibrotic NASH. This test will improve the screening of patients for therapeutic trials in NASH and, in the future, the identification of patients who will benefit from treatment with the new drugs currently under evaluation.

ACKNOWLEDGEMENTS

Data Collection and data management: Sandra Girre, Pascal Veillon, Audrey Morisset, Marc de Saint Loup.

Declaration of personal interests: None.

Declaration of funding interests: This work was supported by grants from the University Hospital of Angers (France), INSERM (France), the University of Nice and charities (Association Française pour l'Etude du Foie (AFEF)/LFB and Société Francophone du Diabète/Roche Pharma to PG). This work was also funded by the French Government (National Research Agency, ANR) through #ANR-15-CE14-0016-01.

AUTHORSHIP

Guarantor of the article: Pr Jerome Boursier.

Author contributions: Jerome Boursier, study design, data acquisition, analysis, and draft writing/critical revision; Rodolphe Anty, study design, data acquisition, and draft writing/critical revision; Luisa Vonghia, Valérie Moal, Thomas Vanwolleghem, Clemence Canivet, Sophie Michalak, Stephanie Bonnafous, Peter Michielsen, Frédéric Oberti, Antonio Iannelli, Luc Van Gaal, Stéphanie Patouraux, Odile Blanchet, An Verrijken, Philippe Gual, Marie-Christine Rousselet, Ann Driessen, data acquisition; Gilles Hunault, Sandrine Bertrais, analysis; Albert Tran, data acquisition; Paul Cales, data acquisition, analysis, and draft writing/critical revision; Sven Francque, study design, data acquisition, and draft writing/critical revision.

ORCID

J. Boursier http://orcid.org/0000-0002-7282-1436
P. Calès http://orcid.org/0000-0003-4866-5274

REFERENCES

1. EASL. EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. *J Hepatol*. 2016;64:1388-1402.
2. Marengo A, Jouness RI, Bugianesi E. Progression and natural history of nonalcoholic fatty liver disease in adults. *Clin Liver Dis*. 2016;20:313-324.
3. Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. *Hepatology*. 2016;64:73-84.
4. Charlton MR, Burns JM, Pedersen RA, Watt KD, Heimbach JK, Dierkhising RA. Frequency and outcomes of liver transplantation for nonalcoholic steatohepatitis in the United States. *Gastroenterology*. 2011;141:1249-1253.
5. Ratziu V, Goodman Z, Sanyal A. Current efforts and trends in the treatment of NASH. *J Hepatol*. 2015;62:S65-S75.
6. Sanyal AJ, Brunt EM, Kleiner DE, et al. Endpoints and clinical trial design for nonalcoholic steatohepatitis. *Hepatology*. 2011;54:344-353.
7. Boursier J, Vergniol J, Guillet A, et al. Diagnostic accuracy and prognostic significance of blood fibrosis tests and liver stiffness

measurement by FibroScan in non-alcoholic fatty liver disease. *J Hepatol*. 2016;65:570-578.
8. Kwok R, Tse YK, Wong GL, et al. Systematic review with meta-analysis: non-invasive assessment of non-alcoholic fatty liver disease–the role of transient elastography and plasma cytokeratin-18 fragments. *Aliment Pharmacol Ther*. 2014;39:254-269.
9. Cusi K, Chang Z, Harrison S, et al. Limited value of plasma cytokeratin-18 as a biomarker for NASH and fibrosis in patients with non-alcoholic fatty liver disease. *J Hepatol*. 2014;60:167-174.
10. Francque SM, Verrijken A, Mertens I, et al. Noninvasive assessment of nonalcoholic fatty liver disease in obese or overweight patients. *Clin Gastroenterol Hepatol*. 2012;10:1162-1168.
11. Anty R, Iannelli A, Patouraux S, 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. *Aliment Pharmacol Ther*. 2010;32:1315-1322.
12. Alberti KG, Eckel RH, Grundy SM, et al. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation*. 2009;120:1640-1645.
13. Kleiner DE, Brunt EM, van Natta M, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology*. 2005;41:1313-1321.
14. Brunt EM, Kleiner DE, Wilson LA, Belt P, Neuschwander-Tetri BA, Network NCR. Nonalcoholic fatty liver disease (NAFLD) activity score and the histopathologic diagnosis in NAFLD: distinct clinicopathologic meanings. *Hepatology*. 2011;53:810-820.
15. Sanyal AJ, Friedman SL, McCullough AJ, et al. American Association for the Study of Liver D. Challenges and opportunities in drug and biomarker development for nonalcoholic steatohepatitis: findings and recommendations from an American Association for the Study of Liver Diseases-U.S. Food and Drug Administration Joint Workshop. *Hepatology*. 2015;61:1392-1405.
16. Harrison SA, Oliver D, Arnold HL, Gogia S, Neuschwander-Tetri BA. Development and validation of a simple NAFLD clinical scoring system for identifying patients without advanced disease. *Gut*. 2008;57:1441-1447.
17. 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;43:1317-1325.
18. Angulo P, Hui JM, Marchesini G, et al. The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. *Hepatology*. 2007;45:846-854.
19. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86:420-428.
20. Shah AG, Lydecker A, Murray K, Tetri BN, Contos MJ, Sanyal AJ. Comparison of noninvasive markers of fibrosis in patients with nonalcoholic fatty liver disease. *Clin Gastroenterol Hepatol*. 2009;7:1104-1112.
21. Vilar-Gomez E, Martinez-Perez Y, Calzadilla-Bertot L, et al. Weight loss through lifestyle modification significantly reduces features of nonalcoholic steatohepatitis. *Gastroenterology*. 2015;149:367-378. e365.
22. Friedman SL, Ratziu V, Harrison SA, et al. A randomized, placebo-controlled trial of cenicriviroc for treatment of nonalcoholic steatohepatitis with fibrosis. *Hepatology*. 2017; doi: 10.1002/hep.29477.
23. Loomba R, Lawitz E, Mantry PS, et al. The ASK1 Inhibitor Selonsertib in Patients with Nonalcoholic Steatohepatitis: a Randomized, Phase 2 Trial. *Hepatology*. 2017; in press.
24. Neuschwander-Tetri BA, Loomba R, Sanyal AJ, et al. Farnesoid X nuclear receptor ligand obeticholic acid for non-cirrhotic, non-alcoholic steatohepatitis (FLINT): a multicentre, randomised, placebo-controlled trial. *Lancet*. 2015;385:956-965.
25. Ratziu V, Harrison SA, Francque S, et al. Elafibranor, an Agonist of the Peroxisome Proliferator-Activated Receptor-alpha and -delta, Induces Resolution of Nonalcoholic Steatohepatitis Without Fibrosis Worsening. *Gastroenterology*. 2016;150:e1145.
26. Townsend SA, Newsome PN. Review article: new treatments in non-alcoholic fatty liver disease. *Aliment Pharmacol Ther*. 2017;46:494-507.
27. Boursier J, de Ledinghen V, Leroy V, et al. A stepwise algorithm using an at-a-glance first-line test for the non-invasive diagnosis of advanced liver fibrosis and cirrhosis. *J Hepatol*. 2017;66:1158-1165.
28. Tohidi M, Arbab P, Ghasemi A. Assay-dependent variability of serum insulin concentrations: a comparison of eight assays. *Scand J Clin Lab Invest*. 2017;77:122-129.
29. Festi D, Schiumerini R, Marzi L, et al. Review article: the diagnosis of non-alcoholic fatty liver disease – availability and accuracy of non-invasive methods. *Aliment Pharmacol Ther*. 2013;37:392-400.
30. Dulai PS, Singh S, Patel J, et al. Increased risk of mortality by fibrosis stage in nonalcoholic fatty liver disease: systematic review and meta-analysis. *Hepatology*. 2017;65:1557-1565.
31. Bedossa P, Dargere D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology*. 2003;38:1449-1457.
32. Rousselet MC, Michalak S, Dupre F, et al. Sources of variability in histological scoring of chronic viral hepatitis. *Hepatology*. 2005;41:257-264.
33. Boursier J, Brochard C, Bertrais S, et al. Combination of blood tests for significant fibrosis and cirrhosis improves the assessment of liver-prognosis in chronic hepatitis C. *Aliment Pharmacol Ther*. 2014;40:178-188.

SUPPORTING INFORMATION

Additional Supporting Information will be found online in the supporting information tab for this article.

How to cite this article: Boursier J, Anty R, Vonghia L, et al. Screening for therapeutic trials and treatment indication in clinical practice: MACK-3, a new blood test for the diagnosis of fibrotic NASH. *Aliment Pharmacol Ther*. 2018;00:1–10. https://doi.org/10.1111/apt.14621

APPENDIX 1

THE AUTHORS' COMPLETE AFFILATIONS

Jerome Boursier, Hepatology Department, Angers University Hospital and HIFIH Laboratory, UPRES 3859, SFR 4208, Bretagne Loire University, Angers, France; Rodolphe Anty, CHU of Nice, Digestive Center and INSERM, U1065, C3M, Team 8 "Hepatic complications in obesity" and Université Côte d'Azur, Nice, France; Luisa Vonghia, Department of Gastroenterology and Hepatology, Antwerp University Hospital & Laboratory of Experimental Medicine and Pediatrics, University of Antwerp, Antwerp, Belgium; Valérie Moal, HIFIH

Laboratory, UPRES 3859, SFR 4208, Bretagne Loire University and Biochemistry Department, Angers University Hospital, Angers, France; Thomas Vanwolleghem, Department of Gastroenterology and Hepatology, Antwerp University Hospital & Laboratory of Experimental Medicine and Pediatrics, University of Antwerp, Antwerp, Belgium; Clemence M Canivet, CHU of Nice, Digestive Center and INSERM, U1065, C3M, Team 8 "Hepatic complications in obesity" and Université Côte d'Azur, Nice, France; Sophie Michalak, Pathology Department, Angers University Hospital, Angers, France; Stephanie Bonnafous, CHU of Nice, Digestive Center and INSERM, U1065, C3M, Team 8 "Hepatic complications in obesity" and Université Côte d'Azur, Nice, France; Peter Michielsen, Department of Gastroenterology and Hepatology, Antwerp University Hospital & Laboratory of Experimental Medicine and Pediatrics, University of Antwerp, Antwerp, Belgium; Frédéric Oberti, Hepatology Department, Angers University Hospital and HIFIH Laboratory, UPRES 3859, SFR 4208, Bretagne Loire University, Angers, France; Antonio Iannelli, CHU of Nice, Digestive Center and INSERM, U1065, C3M, Team 8 "Hepatic complications in obesity" and Université Côte d'Azur, Nice, France; Luc Van Gaal, Department of Gastroenterology and Hepatology, Antwerp University Hospital & Laboratory of Experimental Medicine and Pediatrics, University of Antwerp, Antwerp, Belgium; Stéphanie Patouraux, INSERM, U1065, C3M, Team 8 "Hepatic complications in obesity" and Université Côte d'Azur and CHU of Nice, Biological Center, Pasteur Hôpital, Nice, France; Odile Blanchet, CHU of Nice, Biological Center, Pasteur Hôpital, Nice, France; An Verrijken, Department of Gastroenterology and Hepatology, Antwerp University Hospital & Laboratory of Experimental Medicine and Pediatrics, University of Antwerp, Antwerp, Belgium; Philippe Gual, CHU of Nice, Digestive Center and INSERM, U1065, C3M, Team 8 "Hepatic complications in obesity" and Université Côte d'Azur, Nice, France; Marie-Christine Rousselet, HIFIH Laboratory, UPRES 3859, SFR 4208, Bretagne Loire University and Pathology Department, Angers University Hospital, Angers, France; Ann Driessen, Department of Pathology, Antwerp University Hospital, Antwerp, Belgium; Gilles Hunault, HIFIH Laboratory, UPRES 3859, SFR 4208, Bretagne Loire University, Angers, France; Sandrine Bertrais, HIFIH Laboratory, UPRES 3859, SFR 4208, Bretagne Loire University, Angers, France; Albert Tran, CHU of Nice, Digestive Center and INSERM, U1065, C3M, Team 8 "Hepatic complications in obesity" and Université Côte d'Azur, Nice, France; Paul Calès, Hepatology Department, Angers University Hospital and HIFIH Laboratory, UPRES 3859, SFR 4208, Bretagne Loire University, Angers, France; Sven Francque, Department of Gastroenterology and Hepatology, Antwerp University Hospital & Laboratory of Experimental Medicine and Pediatrics, University of Antwerp, Antwerp, Belgium.