File, Supplemental-Digital-Content-1

The FLIP partners' consortium

Vlad Ratziu, Thierry Poynard, Assistance Publique Hôpitaux de Paris, Groupe Hospitalier Pitie Salpetriere, Universite Pierre et Marie Curie

Biopredictive SAS, Jean-Marie Castille, jean-marie.castille@biopredictive.com, Yen Ngo

Alma Consulting Group SAS, Tania Langon, tlangon@almacg.com

University of Newcastle, Chris Day, c.p.day@newcastle.ac.uk

University of Bristol, Debbie Lawlor, d.a.lawlor@bristol.ac.uk

Universita di Bologna, Giulio Marchesini, giulio.marchesini@unibo.it

University di Firenze, Fabio Marra, f.marra@dmi.unifi.it

Universita degli studi di torino , Elisabetta Bugianesi, ebugianesi@yahoo.it

Universita di Modena e Reggio Emilia, Stefano Bellentani, bellentanistefano@gmail.com

Université de Berne, Jean-François Dufour, jf.dufour@ikp.unibe.ch

Servicio Andaluz de Salud, Manuel Romero Gomez, manuel.romero.sspa@juntadeandalucia.es

Bispebjerg hospital, Region Hovedstaden, Thorkild Sørensen, TIAS@ipm.regionh.dk

Fondazione Italiana Fegato-Fonlus, Claudio Tribelli, ctliver@csf.units.it

Universita di Ancona, Samuele De Minicis, s.deminicis@yahoo.it

Medizinischen Universitaet Wien, Michael Trauner, michael.trauner@meduniwien.ac.at

The FLIP Pathology Consortium:

Pierre Bedossa, Assistance Publique-Hopitaux de Paris, hopital Beaujon, University Paris-Diderot, Paris, France. Alastair D. Burt, School of Medicine, University of Adelaide, Adelaide, Australia. Annette S.H. Gouw, Dept. of Pathology & Medical Biology, University Medical Center Groningen, Groningen, The Netherlands. Carolin Lackner, Institute of Pathology, Medical University of Graz, Graz, Austria. Peter Schirmacher, Institut Fur Pathologie, Universitatsklinikum Heidelberg, Heidelberg, Germany. Luigi Terracciano, Institute of Pathology, University Hospital, Basel, Switzerland. D. Tiniakos, Medical School, National and Kapodistrian University of Athens, Greece and Medical School, Newcastle University, Newcastle-upon-Tyne, United Kingdom. J. Brain, Medical School, Newcastle University, Newcastle-upon-Tyne, United Kingdom. Yvonne Bury, Royal Victoria Infirmary, Hospitals Foundation Trust, Newcastle upon Tyne, United Kingdom. Daniela Cabibi, University of Palermo, Palermo, Italy. Frederic Charlotte, Assistance Publique Hopitaux de Paris, Groupe Hospitalier Pitie-Salpetriere, Universite Pierre et Marie Curie, Paris, France. Ezio David, University of Torino, Ospedale Molinette, Torino, Italy. Luisa Losi, Modena University Hospital, Modena, Italy. Matteo Montani, Institute of Pathology, University of Bern, CH-3010 Bern, Switzerland. Marıa Jesus Pareja, Hospital Universitario de Valme, Sevilla, Spain. Dominique Wendum, Assistance Publique-Hopitaux de Paris, Hopital St Antoine Universite Pierre et Marie Curie. Fritz Wrba, Medical University of Vienna, Vienna, Austria. Marianne Ziol, Assistance Publique-Hopitaux de Paris, Hopital Jean Verdier, Universit_e Paris 13, Bobigny, France. Vlad Ratziu, Assistance Publique Hopitaux de Paris, Hopital Pitie Salpetriere, Université Pierre et Marie Curie, Paris, France.

The FibroFrance-CPAM group:

Thierry Poynard, Pascal Lebray, Patrick Ingiliz, Anne Varaut, Yen Ngo1, Pascal Norha, Julien Massard, and Vlad Ratziu, APHP-UPMC-Liver Center, Assistance Publique Hôpitaux de Paris, Université Pierre et Marie Curie, Paris, France

Brigitte Varsat, Jean Pierre Carrau, and Jean Pierre Giordanella, Caisse Primaire Assurance Maladie, Paris, France,

Mona Munteanu and Fabienne Drane, Biopredictive, Paris, France

Djamila Messous, and Françoise Imbert Bismut, and Biochemistry Unit, Assistance Publique Hôpitaux de Paris, Université Pierre et Marie Curie, Paris, France

The FibroFrance-Obese group:

Thierry Poynard, Hugo Perazzo, Frederic Charlotte and Vlad Ratziu, APHP UPMC Liver Center, Paris, France

Guillaume Lassailly, Robert Caıazzo, Francois Pattou and Philippe Mathurin, Hepatology Department, Centre Hospitalier Lille, Lille, France,

Emmanuel Diaz, Robert Callafe, Bernard Demol, Hepatology Department, Centre Hospitalier Bethune, Bethune

Karine Clement and JoanTordjman, Institute of Cardiometabolisme and Nutrition, APHP Pitie-Salpetriere Hospital, Paris, France, 5 INSERM, U872, Nutriomique, Paris, France, Universite Pierre et Marie Curie-Paris 6, Centre de Recherche des Cordeliers, Paris, France,

Mona Munteanu, Biopredictive, Paris, France,

Pierre Bedossa, Assistance Publique-Hopitaux de Paris, Hopital Beaujon, Departement de Pathologie Clichy, Clichy, France,

Table, Supplemental-Digital-Content-2. Comparison between histological scoring systems and pre-determined cutoffs according to histological reference scoring systems

SAF scores	SAF categories	Previous METAVIR and sensitive steatosis category	Standard Cutoffs	Extended Cutoffs
Fibrosis			**FibroTest**	**FibroTest**
F0	None	None	0.00	0.00
F1	Perisinusoidal or portal	Portal fibrosis	0.27	0.27
F2	Perisinusoidal and portal without bridging	Few septa	0.48	0.48
F3	Bridging	Many septa	0.58	0.58
F4	Cirrhosis	Cirrhosis	0.74	0.74
Activity			**ActiTest**	**ActiTest**
A0	Ballooning + lobular inflammation =0	None	0.00	0.00
A1	Ballooning + lobular inflammation =1	Minimal periportal necrosis or inflammation	0.29	0.17 A0A1
A2	Ballooning + lobular inflammation=2	Moderate periportal necrosis-inflammation	0.52	0.29 A1
A3	Ballooning + lobular inflammation =3	Severe periportal necrosis-inflammation	0.52	0.52 A2A3
A4	Ballooning + lobular inflammation =4	Severe periportal necrosis-inflammation	0.52	0.52 A2A3
Steatosis			**SteatoTest**	**SteatoTest**
S0	<5%	0% and >0-5%	0.00	0.00
S1	5%-33%	>5-33%	0.57	0.57
S2	>33-66%	>33%	0.69	0.69
S3	>66%	>33%	0.69	0.81

In ActiTest, the possible choices of predetermined cutoffs were those previously [19] and based on 3 levels, as follows: Level 1= 0.52 METAVIR A2, the Standard for chronic hepatitis C and B; Level 2= 0.29 METAVIR A1; and Level 3= 0.17 METAVIR A0-A1. For SteatoTest, the four possible choices were predetermined cutoffs for SAF-S1, previously published [11,12], as follows: Level 1= 0.57 S1 ≥5%; Level 2= 0.48 S1 >0%; Level 3= 0.38 S0-S1 ≥5%; and Level 4 =0.30 S0-S1 >0%.

Imbert-Bismut F, Messous D, Thibault V, et al. Intra-laboratory analytical variability of biochemical markers of fibrosis (Fibrotest) and activity (Actitest) and reference ranges in healthy blood donors. Clin Chem Lab Med. 2004;42:323-33.

Table, Supplemental-Digital-Content-3.

LIVER-FIBROSTARD CHECKLIST

The Liver-FibroSTARD checklist summarizes the important information that must be present in the manuscripts of diagnostic studies on non-invasive tools for liver fibrosis evaluation. Compared to STARD, the Liver-FibroSTARD checklist includes 2 additional items (#12 and #26) and 44 sub-items. The sub-items correspond to those proposals that clearly depicted, within the items, each of the particular features of diagnostic studies on liver fibrosis tests. Finally, Liver-FibroSTARD presents as a complementary module of the STARD checklist.

Some items or sub/items include several criteria; major criteria are indicated by an asterisk (*). Example: item #3: "The study population: The inclusion and exclusion criteria, setting, and locations* where data were collected". If a major item is missing, the corresponding criterion has to be rated absent. Some items/sub-items (#12.1 and #23.1, #13.10 and #22.2) are redundant since they can be found in different locations of the article.*

FibroSTARD check list adapted for NASH tests:Only items 14, 20, 21, 23, 24, 25, 26 were not addressed or discussed in the manuscript or in supplementary files. For population 2 details are given in Reference 17..

TITLE/ABSTRACT /KEYWORDS Page 1 and page 3 Feature is NASH not Fibrosis	**1. Identify the article as a study of diagnostic accuracy (recommend MeSH heading "sensitivity and specificity").**	
	1.1. Identify the article, especially in the title, as a study of the diagnostic performance of liver fibrosis/cirrhosis biomarker(s)/test(s).	✓
	1.2. Recommended key words (choose the most appropriate): "liver fibrosis", "cirrhosis", "diagnosis", "biomarker", "diagnostic test", "noninvasive diagnosis".	✓
INTRODUCTION Page 5,6, NASH and not fibrosis	**2. State the research questions or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups.**	
	In study aims, specify:	
	2.1. If the aim is to identify new marker(s)/develop new test(s), or to evaluate published marker(s)/test(s).	✓
	2.2. Whether the study is performed in a single or multiple cause(s) of chronic liver disease.	✓
	2.3. The reference used for fibrosis diagnosis in the study.	✓
	2.4. The diagnostic target used as the primary aim of the study and, if appropriate, other diagnostic targets used as secondary aims.	✓
METHODS	*Describe:*	
Participants Figure 1 Page 6 Here we focus on NASH test	**3. The study population: The inclusion and exclusion criteria*, setting, and locations* where data were collected.**	✓
	4. Participant recruitment: Was recruitment based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?	
	4.1. State if healthy subjects without chronic liver disease are included or not in the study.	✓
	4.2. State if patients were selected by one abnormal or several discordant fibrosis test(s).	✓
	4.3. State if patients were selected according to the availability of reference or index test(s) result(s).	✓
	5. Participant sampling: Was the study population a consecutive series of participants defined by the selection criteria in item 3 and 4? If not, specify how participants were further selected.	✓
	6. Data collection: Was data collection planned before the index test and reference standard were performed (prospective study) or after (retrospective study)?	
	6.1. The chronology between patient inclusion, data collection (reference/index tests)*, and data analysis is well described.*	✓
	6.2. Has the study population been previously used/published for the evaluation of the studied fibrosis test(s)?	✓

Test methods	**7. The reference standard and its rationale.**	✓
Described for population 1 in previous publication Munteanu 2016 Ref 17 Detailed page 7 for histology Supplementary-Table-S3 FLIP pathologists References 4-6, 17 FLIP and CRN	**8. Technical specifications of material and methods involved including how and when measurements were taken, and/or cite references for index tests and reference standard.** *For the reference and index test(s), specify characteristics with sufficient detail to permit exact reoperation, when appropriate:* *8.1. Center: standardization of procedures across centers.* *8.2. Patient: fasting conditions*, time, posture, etc. (give information about the influence of conditions on the intra-individual variability).* *8.3. Delay: time interval between reference and index test(s).* *8.4. Material: technical specifications (name, generation, manufacturer, instrument), method of measurement, applicability (failure/reliability criteria)*. Specifically for liver biopsy, indicate material used per center, i.e. percutaneous/transjugular/other, needle diameter.* *8.5. Biological samples: description of method of collection, transport, storage*.* *8.6. Specify how the index tests were calculated.* *8.7. Specify how the risk for false negative/positive results was taken into account.* *Specifically for liver biopsy:* *8.8. How sample bias was limited: minimal biopsy size (length)*, number of portal tracts required, number of fragments.* *8.9. Methods for histological assessment: human/automated reading*, local/central reading*, number and expertise of pathologists*, single/double reading*, consensus methods.* *8.10. Scoring system used (Metavir, Ishak, Scheuer, etc.).*	✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓
	9. Definition of and rationale for the units, cut-offs*, and/or categories of the results of the index tests and the reference standard.	✓
	10. The number*, training and expertise* of the persons executing and reading the index tests and the reference standard.	✓
	11. Whether or not the readers of the index tests and reference standard were blind (masked) to the results of the other test and describe any other clinical information available to the readers.	✓
Statistical methods A companion article is also submitted, which is a methodological analysis of pitfalls in assessing accuracy	**12. State if the study is conducted on an intention-to-diagnose basis or if the analysis is per-protocol (i.e. with exclusion of failed/unreliable fibrosis test(s)/reference measurements).** *12.1. If intention-to-diagnose analysis, specify how failure and unreliable test(s)/reference are taken into account in the analysis.* [a]	✓ ✓
	13. Methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g. 95% confidence intervals). *Specify:* *13.1. Detailed sample size calculation.* *13.2. Statistical methods used to quantify uncertainty (e.g. 95% confidence intervals).* *13.3. Control of multiple comparisons that increases type I error: multiple comparisons of tests (e.g. Bonferroni correction, etc.), multiple diagnostic targets.* *13.4. Method for calculation of fibrosis test(s) diagnostic cut-offs.* *13.5. Method for validation of new test(s) or new calculated diagnostic cut-off(s) (e.g. external validation set, internal validation by bootstrapping, etc.).* *13.6. Method for control of center/operator effect.* *13.7. Method for control of spectrum effect if unrepresentative prevalence of fibrosis stages (e.g. Obuchowski index, DANA, etc.).* *13.8. Method for control of misclassification errors by the reference test.* *13.9. Use of a reference without gold standard.* *13.10. Analysis of discordances between reference/index test(s).* [b]	✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓

	14. Methods for calculating test reproducibility.	○
RESULTS	*Report:*	
Participants	**15. When study was performed, including beginning and end dates of recruitment.**	✓
Supplementary-Table-S3	**16. Clinical and demographic characteristics of the study population (e.g. age*, sex*, spectrum of presenting symptoms, comorbidity, current treatments, recruitment centers).**	✓
	16.1. For liver biopsy: size (length), number of portal tracts, number of fragments.*	✓
	16.2. For index test(s): confounding factors that potentially influence the test(s) results (flare-up, inflammation, other liver lesions, intrinsic characteristics, etc.).	✓
Supplementary-Table-S14	**17. The number of participants satisfying the criteria for inclusion who did or did not undergo the index tests and/or the reference standard*; describe why participants failed to undergo either test (a flow diagram is strongly recommended).**	✓
	17.1. If per-protocol analysis, report comparisons between patients excluded due to failed/unreliable test(s)/reference and patients with reliable fibrosis test(s)/reference.	✓
Test results	**18. Time-interval* between the index tests and the reference standard, and any treatment administered between.**	✓
Supplementary-Table-S3	**19. Distribution of severity of disease (define criteria) in those with the target condition*; other diagnoses in participants without the target condition.**	✓
	19.1. Specify the prevalence of the diagnostic condition (spectrum effect).*	✓
Page 12	**20. A cross tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard.**	
	20.1. Presentation of contingency tables, box/scatter plots.	○
	21. Any adverse events from performing the index tests or the reference standard.	○
Estimates	**22. Estimates of diagnostic accuracy* and measures of statistical uncertainty (e.g. 95% confidence intervals).**	✓
Page 12 Table S10 Table S11 Table S12 Table S13	*22.1. Specify sensitivity* and specificity* with 95% confidence intervals; ROC analysis.*	✓
	22.2. Analyzing discordances between fibrosis tests(s)/reference. [b]	✓
	23. How indeterminate results, missing data and outliers of the index tests were handled.	
	23.1. How missing/failure/unreliable results of index test(s)/reference were handled (intention-to-diagnose/per-protocol analysis). [a]	○
	23.2. How outliers of the index tests were handled.	○
	24. Estimates of variability of diagnostic accuracy between subgroups of participants, readers or centers, if done.	○
	25. Estimates of test reproducibility, if done.	○
	26. Estimates of cost-benefit.	○
DISCUSSION	**27. Discuss the clinical applicability of the study findings.**	
Page 14 to 17	*27.1. Discuss the representativeness of the study sample and recruiting centers (i.e. spectrum effect, etc.).*	✓
	27.2. Discuss the interpretation of fibrosis test(s) results in clinical practice.	✓
	27.3. Discuss the clinical relevance of the study results.	✓

[a] *Items 12.1 and 23.1 are redundant but retained since they can be located in different paragraphs within an article*

[b] *Items 13.10 and 22.2 are redundant but retained since they can be located in different paragraphs within an article*

Explanations: see glossary

Authors: ARDENT group (see details in glossary) and AFEF (French Association for the Study of the Liver)

Version: February 2015

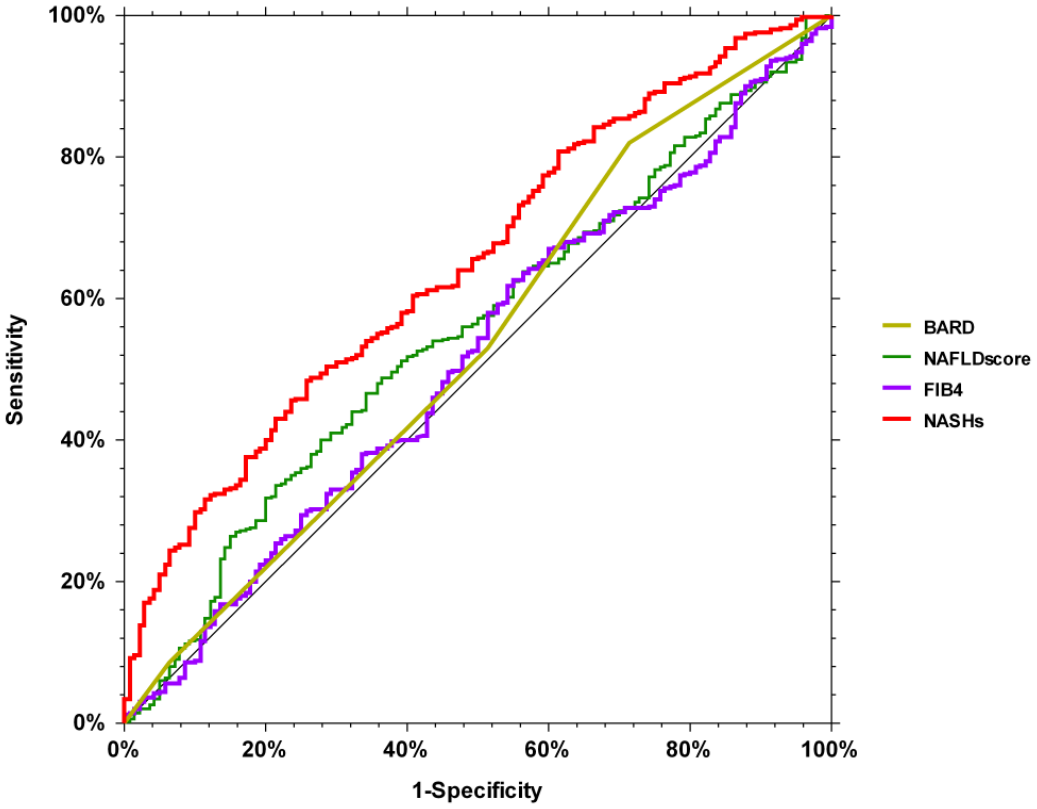
Table, Supplemental-Digital-Content-4. Prevalence of histological NASH according to the 27 possible definitions of steatosis, ballooning and lobular inflammation, in Population-1 (n=1081)

27 combinations of SAF scoring system features			NASH definitions according to 27 combinations		Prevalence of each combination		
Steatosis grade	**Activity grade**		**FLIP-CRN**	**Simplified H-NASHs**			
	Ballooning (B)	**Lobular Inflammation (L)**					
3 levels	**3 levels**	**3 levels**	**0=absence**	**1=presence**	**n**	**%**	**95%CI**
0%	0	0	0	0	38	3.5	2.5-4.8
0%	0	1	0	0	5	0.5	0.1-1.1
0%	0	2	0	1	0	0.0	0.0-0.3
0%	1	0	0	0	3	0.3	0.1-0.8
0%	1	1	0	1	3*	0.3	0.1-0.8
0%	1	2	0	1	0	0.0	0.0-0.3
0%	2	0	0	1	1*	0.1	0.0-0.5
0%	2	1	0	1	0	0.0	0.0-0.3
0%	2	2	0	1	1*	0.1	0.0-0.5
Subtotal	**Steatosis 0%**				**51**	**4.7**	**3.5-6.2**
1-4%	0	0	0	0	34	3.1	2.2-4.4
1-4%	0	1	0	0	1	0.1	0.0-0.5
1-4%	0	2	0	1	0	0.0	0.0-0.3
1-4%	1	0	0	0	3	0.3	0.1-0.8
1-4%	1	1	0	1	1*	0.1	0.0-0.5
1-4%	1	2	0	1	0	0.0	0.0-0.3
1-4%	2	0	0	1	0	0.0	0.0-0.3
1-4%	2	1	0	1	0	0.0	0.0-0.3
1-4%	2	2	0	1	0	0.0	0.0-0.3
Subtotal	**Steatosis 1-4%**				**39***	**3.6**	**2.6-4.9**
≥5%	0	0	0	0	241	22.3	19.8-24.9
≥5%	0	1	0	0	90	8.3	6.7-10.1
≥5%	0	2	0	1	12*	1.1	0.6-1.9
≥5%	1	0	0	0	78	7.2	5.7-8.9
≥5%	1	1	1	1	229	21.2	18.8-23.7
≥5%	1	2	1	1	47	4.3	3.2-5.7
≥5%	2	0	0	1	21*	1.9	1.2-3.0
≥5%	2	1	1	1	158	14.6	12.6-16.9
≥5%	2	2	1	1	115	10.6	8.9-12.6
Subtotal	**Steatosis ≥5%**				**991**	**91.7**	**89.9-93.3**
Total					**1081**	**100**	**99.7-1.00**

The prevalence of NASH using standard definition was 50.8% (47.8-53.8) (549/1081), and using simplified definition, 54.4% (51.4-57.4) (588/1081).
*These 39 NASH cases defined by ballooning+ lobular inflammation stages ≥2, (3.6%;2.6-4.9) that were missed by the FLIP-algorithm included 15 cases with significant fibrosis (6 F2, 5 F3 and 4 cirrhosis).

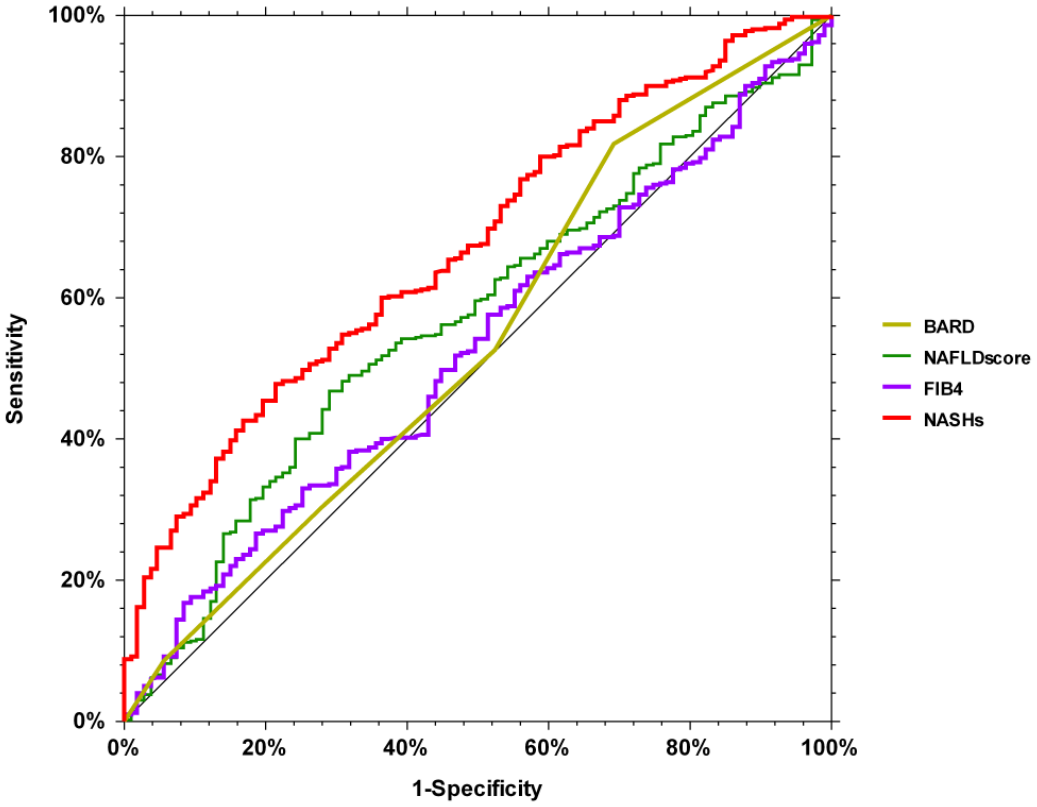
Figure, Supplemental-Digital-Content-5
Performance of NIT-NASHs versus non-patented tests.

NIT-NASHs had a significantly higher AUROC (0.671;0.614-0.721), than NAFLD-score AUROC (0.570;0.510-0.626;P=0.006), FIB4 (0.528;0.467-0.584;P=0.0003) and BARD index (0.541;0.476-0.599;P=0.003).



Figure, Supplemental-Digital-Content-6:
Performance of NIT-A2orF2 versus non-patented tests.

NIT-A2orF2 had a significantly higher AUROC (0.671;0.613-0.721) than the NAFLD-score (0.570;0.510-0.626; P=0.006), FIB4 (0.528;0.467-0.584;P=0.0003) and BARD (0.541;0.476-0.599;P=0.003)



Table, Supplemental-Digital-Content-7: Characteristics of patients included in the construction populations, with missing non-patented NITs (NAFLD-score, FIB4 and BARD), in comparison with cases with both patented and not-patented NITs.

Characteristics	Patented NITs missing	Patented NITs assessed	P-value
n	507 (100%)	574 (100%)	1
Presumed NAFLD	507 (100%)	574 (100%)	1
Gender male	128 (25.2%)	359 (62.5%)	<0.0001
Diabetes treated or glucose6.1mmol/L	145 (28.6%)	209 (36.4%)	0.006
Age (year)	43.2 (41.3-44.7)	53.0 (51.1-54.0)	<0.0001
BMI (weight/heigth^2)>= 30	491 (96.8%)	268 (46.7%)	<0.0001
Biopsy number	**507 (100%)**	**574 (100%)**	
Biopsy length (mm)	12 (11-12)	25 (22-25)	<0.0001
Biopsy-test days	0.0 (0)	0.0 (0.0-0.1)	1
Stage of fibrosis (SAF F biopsy)			<0.0001
F0 no fibrosis	237 (46.7%)	117 (20.4%)	
F1 perisinusoidal or portal	214 (42.2%)	173 (30.1%	
F2 sinusoidal or periportal without bridging	33 (6.5%)	135 (23.5%)	
F3 bridging fibrosis	15 (3.0%)	118 (20.6%)	
F4 cirrhosis	8 (1.6%)	31 (5.4%)	
Ballooning			<0.0001
Grade 0	313 (61.7%)	108 (18.8%)	
Grade 1	124 (24.5%)	240 (41.8%)	
Grade 2	70 (13.8%)	228 (39.4%)	
Lobular inflammation			<0.0001
Grade 0	307 (60.6%)	112 (19.5%)	
Grade 1	170 (33.5%)	317 (55.2%)	
Grade 2	30 (5.9%)	145 (25.3%)	
Grade of activity (SAF A biopsy)			<0.0001
A0 no activity	252 (49.7%)	61 (10.6%)	
A1 mild	101 (19.9%)	79 (13.8%)	
A2 moderate	86 (17.0%)	181 (31.5%)	
A3 severe or A4 very severe	68 (13.4%)	253 (44.1%)	
Grade of steatosis (sensitive)			<0.0001
S0 no steatosis 0%	32 (6.3%)	19 (3.3%)	
S0 1-4%	18 (7.5%)	1 (0.2%)	
S1 mild 5%-100%	437 (86.2%)	554 (96.5%)	
FLIP-algo Steatosis ≥5%			<0.0001
No-steatosis ("No-NAFLD")	70 (13.8%)	20 (3.5%)	
Steatosis only	299 (59.0%)	143 (24.9%)	
NASH	138 (27.2%)	411 (71.6%)	

[1] Cases with histological steatosis (5%) or activity (A>0) were excluded. [2] One case had steatosis 2% and Ballooning and Lobular inflammation grade 1 and therefore classified NASH with FLIP algorithm using 0% cutoff and "no steatosis" using 5% cutoff

Table, Supplemental-Digital-Content-8. Large studies (>500 cases) in adults with histological steatosis grading

In order to describe the construction of NITs, and the impact of definitions on their accuracy, we review the literature to clarify the main definitions of the population of interest (the appropriate context of use was defined as carriers of metabolic risk factor), the definition of the disease of interest (metabolic liver diseases included steatosis, activity and fibrosis [SAF], in the absence of other known liver disease). We screened PUBMED with the following tags: "NAFLD metabolic liver disease biopsy human" (January 5th 2017). The criteria of inclusion were studies in adults, with 500 or more biopsies and giving the definition of histological steatosis.

Context of use	Author, year	Number	Prevalence S0 0%-4%	0%	Prevalence A0	Prevalence A0S0 S0<5%*	Prevalence S0 S0<5%	Cirrhosis F4
NAFLD	Kleiner, 2005	576	58 (10%)	NA	NA (14%)*	13 (2.2%)	NA	35 (6%)
NAFLD	Brunt, 2011	934	37 (4%)	NA	3 (0.3%)	3 (0.3%)	3 (0.4%)	0 (0%)
NAFLD Obese	Bedossa, 2012	679	158 (23%)	NA	248 (36.5%)	147 (21.6%)	147 (21.6%)	6 (0.9%)
NAFLD	Kessoku, 2014	1,048	0 (0%)	NA	NA	0 (0%)	NA	38 (3.6%)
NAFLD	Angulo, 2015	619	0 (0%)	NA	157 (25.4%)	0 (0%)	NA	18 (2.9%)
Total		3856	253 (6.6%)	NA	NA	163 (4.2%)	NA	NA

NA=not available.

* Lobular inflammation taken if no details overall activity as most sensitive than ballooning for grade 1. 13 out of 575 (2.2%)cases only were A0S0 (see Figure 4 in the Kleiner article).

Only 163 out of 3856 (4.2%) were A0S0, which should be the appropriate controls for assessing NITs performance for NASH prediction.

No study detailed the full spectrum of steatosis, including cases without any steatosis (0%). One study included presumed NAFLD cases with steatosis 1-4% without excluding any cases, but did not specify the prevalence of 0% versus 1-4%. One study included presumed NAFLD, only with severe obesity, with steatosis 1-4% without excluding any cases, but did not specify the prevalence of 0% steatosis versus 1-4%. One study included presumed NAFLD cases with steatosis 1-4%, did not specify the prevalence of 0% versus 1-4%, although cases with cirrhosis were excluded. The two last studies excluded cases with steatosis 1-4%. Finally, no large study estimated the prevalence of minimal steatosis (1-4%) among presumed NAFLD and therefore the estimated prevalence of absence of any steatosis.

Biopsy length's medians were not given in these 5 studies, but mentioned in Angulo study as a confounding factor analyzed in the prognostic analysis.