ABREVIATIONS

NAFLD: non-alcoholic fatty liver disease

NASH: non-alcoholic steatohepatitis

LSM: liver stiffness measurement

NFS: NAFLD Fibrosis Score

PLAN

ABSTRACT

Objective: NAFLD is highly prevalent but only few patients develop advanced liver fibrosis with impaired liver-related prognosis. We aimed to compare liver stiffness measurement by Fibroscan (LSM) and blood fibrosis tests for the diagnosis of liver fibrosis and the prognostic assessment in NAFLD.

Design: Diagnostic accuracy was evaluated in a cross-sectional study including 588 NAFLD patients with liver biopsy (NASH-CRN fibrosis stage), LSM, and 8 blood fibrosis tests (BARD, NAFLD Fibrosis Score, FibroMeter^{S}, APRI, FIB4, Fibrotest, Hepascore, FibroMeter^{V2G}). Prognostic accuracy was evaluated in a longitudinal study including 626 NAFLD patients.

Results: LSM and FibroMeter^{V2G} were the two best-performing tests in the cross-sectional study. AUROC for advanced F3/4 fibrosis was, respectively: 0.831±0.019 and 0.817±0.020 ($p \leq 0.041$ vs other tests), rate of patients with ≥90% negative/positive predictive values for F3/4: 56.4% and 46.7% ($p < 0.001$ vs other tests), and Obuchowski index: 0.834±0.014 and 0.798±0.016 ($p \leq 0.036$ vs other tests). Two fibrosis classifications were developed to precisely estimate the histological fibrosis stage from LSM or FibroMeter^{V2G} result without any liver biopsy (diagnostic accuracy, respectively: 80.8% vs 77.4%, $p=0.190$). The longitudinal study showed that LSM and FibroMeter^{V2G} fibrosis classifications categorize NAFLD patients in several subgroups with significant different prognosis ($p < 0.001$). Overall survival and survival without liver-related death progressively decreased as a function of these subgroups.

Conclusion: Among 9 fibrosis tests evaluated, LSM and FibroMeter^{V2G} are the most accurate for the non-invasive diagnosis of liver fibrosis in NAFLD. LSM and FibroMeter^{V2G} fibrosis classifications help physicians to estimate both the fibrosis stage and the patient prognosis in clinical practice.

INTRODUCTION

Non-alcoholic fatty liver disease (NAFLD) is the liver manifestation of the metabolic syndrome. With the worldwide burden of obesity, the median prevalence of NAFLD in the general population has reached 20-30%, placing it as the most prevalent cause of chronic liver disease worldwide [1, 2]. NAFLD is a heterogeneous entity that covers a wide spectrum of liver lesions ranging from bland steatosis to non-alcoholic steatohepatitis (NASH), fibrosis, and finally cirrhosis with its life-threatening complications. Several pathological lesions are used to describe NAFLD severity [3, 4, 5], but recent longitudinal studies agreed that liver fibrosis amount is the main determinant of patient outcome [6, 7, 8]. Consequently, as in other causes of chronic liver disease, liver fibrosis in NAFLD patients must be accurately evaluated in clinical practice.

Liver biopsy currently remains the reference method to evaluate liver lesions and liver fibrosis in NAFLD [9]. However, this invasive procedure with potentially severe or fatal complications [10] appears unsuitable for evaluating prognosis in NAFLD, as it would induce a large number of biopsies completely disproportionate to the low rate of patients who develop advanced fibrosis [11]. The NAFLD Fibrosis Score (NFS) is a blood test specifically dedicated for the non-invasive diagnosis of advanced fibrosis in NAFLD [12]. Numerous studies have validated its diagnostic accuracy [13], leading the American Association for the Study of Liver Diseases to recently recommend its use as “a clinically useful tool for identifying NAFLD patients with higher likelihood of having bridging fibrosis and/or cirrhosis” [9].

However, the NFS has two main limitations. First, it is used with two diagnostic cut-offs, the one for the exclusion and the other for the affirmation of advanced fibrosis, leaving thus one third of the patients in the ‘grey zone’ where liver biopsy remains still required [14]. In this setting, we have previously developed methods to suppress this grey zone through establishment of fibrosis classifications that give an accurate estimation of the histological fibrosis stage from the non-invasive test result without any liver biopsy requirement [15, 16, 17]. Second, the NFS includes only indirect markers of liver fibrosis. In chronic hepatitis C, it has been shown that blood tests including both direct and indirect markers of liver fibrosis are more accurate than tests including

only indirect markers [18, 19]. Liver stiffness measurement (LSM) by transient elastography (Fibroscan) is another accurate method for the non-invasive diagnosis of liver fibrosis in NAFLD [14]. However, studies that investigated LSM specifically in NAFLD remain scarce and included small samples of patients [20]. Overall, the latest guidelines of the European Association for the Study of the Liver conclude that LSM and blood fibrosis tests in NAFLD require further validation and direct comparison [20].

The aims of the present study were to evaluate and directly compare the accuracy of 8 blood tests and LSM for the non-invasive diagnosis of liver fibrosis in a large cohort of NAFLD patients, to develop fibrosis classifications for the most accurate of them, and to validate the clinical significance of these classifications in a longitudinal prognostic cohort.

PATIENTS AND METHODS

The study protocol of the present study conformed to the ethical guidelines of the current Declaration of Helsinki. All patients included in both cross-sectional and longitudinal cohorts gave informed written consent to participate.

Cross-sectional cohort

The purpose of this cohort was to evaluate and compare the diagnostic accuracy of the non-invasive fibrosis tests, and to develop the fibrosis classifications.

Patients

Patients with biopsy-proven NAFLD were consecutively included from January 2004 to June 2014 at Angers University Hospital and from October 2003 to April 2014 at Bordeaux University Hospital. NAFLD was defined as liver steatosis on liver biopsy after exclusion of concomitant steatosis-inducing drugs, excessive alcohol consumption (>210 g/week in men or >140 g/week 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 (ascites, variceal bleeding, systemic infection, or hepatocellular carcinoma).

Liver biopsy

In each centre, pathological examination was performed by a senior expert specialized in hepatology and blinded for patient data. Liver fibrosis was evaluated according to the NASH CRN scoring system [3]: F0 = no fibrosis; F1 = perisinusoidal or portal/periportal fibrosis, F2 = perisinusoidal and portal/periportal fibrosis, F3 = bridging fibrosis, and F4 = cirrhosis. Significant fibrosis was defined as F≥2, advanced fibrosis as F≥3, and cirrhosis as F4. Because previous longitudinal studies have demonstrated that liver-related prognosis is impaired when advanced fibrosis occurs [6, 7, 8, 13], and as recommended by the latest EASL guidelines [20], we chose advanced F≥3 fibrosis as our primary diagnostic target.

Blood fibrosis tests

Fasting blood samples were taken the day of or within the week preceding liver biopsy. Eight blood fibrosis tests were calculated according to published or patented formulas: NFS [12], BARD [21], FibroMeterS [22], APRI [23], FIB4 [24], Fibrotest [25], Hepascore [26], and

FibroMeterV2G [27]. BARD, NFS and FibroMeterS were specifically developed for liver fibrosis assessment in NAFLD, whereas the 5 other tests were developed in patients with chronic viral hepatitis. FibroMeterV2G and Hepascore include both direct and indirect markers of liver fibrosis, whereas the 6 other blood tests include only indirect markers. NFS was interpreted according to published cut-offs [12]: patients with NFS results >0.676 are considered as having advanced fibrosis, those with NFS <-1.455 as having F0-2 stages, and those between the 2 cut-offs as having an indeterminate diagnosis (grey zone). All blood assays were performed in the laboratories of the Angers or Bordeaux centres. We have previously demonstrated the excellent inter-laboratory reproducibility of blood fibrosis tests [28].

Liver stiffness measurement

In each centre, LSM with Fibroscan was performed using the standard M probe by an experienced observer blinded for patient data. LSM was performed the day of liver biopsy or no more than 3 months around. Examination conditions were those recommended by the manufacturer [29]. LSM was stopped when 10 valid measurements were recorded and the result (kilo Pascals: kPa) was expressed as the median of these valid measurements.

Longitudinal cohort

The purpose of this prognostic cohort was to validate the clinical significance of the fibrosis classifications previously developed in the cross-sectional cohort. All NAFLD patients seen between January 2005 and December 2009 in the Hepatology Department of the Angers University Hospital for a non-invasive evaluation of liver fibrosis were retrospectively included. The follow-up started the day of the non-invasive testing and ended November 15th, 2014. The date and cause of death were obtained from the computerized National Registry of Individuals (CepiDC-Inserm, France). For some patients with unsuccessful individual matching with the national registry, mortality data were obtained from the hospital database, or from the concerned general practitioner.

Statistical analysis

In the cross-sectional cohort, diagnostic accuracy of fibrosis tests was evaluated using the classical indexes for binary diagnostic targets: AUROC, rate of well-classified patients according to the highest Youden index that maximizes sensitivity and specificity, and the rate of patients included in the intervals of ≥90% negative or positive predictive values (for the latter, see Supplementary Material for precise definitions). The diagnostic accuracy of fibrosis tests was also evaluated using the Obuchowski index [30]. The Obuchowski index is a multinomial version of the area under the receiver operating characteristics (AUROC) adapted to ordinal references such as pathological fibrosis staging. With N (=5: F0 to F4) categories of the gold standard outcome and AUROCst, it estimates the AUROC of diagnostic tests differentiating between categories s and t. The Obuchowski measure is a weighted average of the N(N-1)/2 (=10) different AUROCst corresponding to all the pair-wise comparisons between two of the N categories. In addition, the Obuchowski measure was assessed using a penalty function proportional to the difference in fibrosis stages, i.e., a penalty of 1 when the difference between stages was 1, 2 when the difference was 2, 3 when the difference was 3, and 4 when the difference was 4. Finally, the result can be interpreted as the probability that the non-invasive test will correctly rank two randomly chosen patients with different fibrosis stages.

In the longitudinal cohort, prognostic accuracy of fibrosis test was evaluated using the C-index of Harrell, as previously described [31]. Briefly, the Harrell C-index is an extension of the AUROC for time-to-event (survival) data and evaluates the concordance between the predicted risk of event and the observed survival time. Its results varies from 0 to 1: 1 shows a perfect concordance (discriminative power of the risk score), 0.5 shows random prediction, and a value less than 0.5 indicates discrimination in the opposite direction to that expected. Survival curves were determined using the Kaplan-Meier method and compared with the log rank test.

Statistical analyses were performed using SPSS version 18.0 software (IBM, Armonk, NY, USA) and SAS 9.1 (SAS Institute Inc., Cary, NC, USA). This study was reported in accordance with the recently published LiverFibroSTARD statements [32].

RESULTS

Cross-sectional cohort: diagnostic accuracy of blood fibrosis tests and LSM

Patients

The flow chart of the cross-sectional study is depicted in the **Figure 1**. A total of 588 patients were included, 243 in the Angers centre and 345 in the Bordeaux centre. Patient characteristics at inclusion are detailed in **Table 1**. Mean biopsy length was 26±12 mm. Failure of LSM with no valid measurement occurred in 83 patients (14.1%). Median LSM result in the 505 remaining patients was 9.1 kPa (1^{st} quartile: 6.4 kPa; 3^{rd} quartile: 13.9 kPa). Finally, LSM and all 8 blood tests were available in 452 patients (core group). The prevalence of histological fibrosis stages was not significantly different between the core group and the 136 other patients.

Comparison of the diagnostic accuracy of the non-invasive fibrosis tests

AUROC – Blood tests and LSM were directly compared in the core group where all tests were available. For the diagnosis of advanced F≥3 fibrosis, the primary diagnostic target of the study, FibroMeter^{V2G} had a significantly higher AUROC (0.817±0.020) than the 7 other blood tests (p≤0.025; **Table 2,** see **Supplementary Table s1** for detailed pairwise comparisons). LSM has a significantly higher AUROC for advanced fibrosis (0.831±0.019, p≤0.041) than blood tests, except when compared to FibroMeter^{V2G} (p=0.559).

Binary diagnosis of advanced fibrosis – The best diagnostic cut-off for advanced fibrosis was calculated for each fibrosis test according to the highest Youden index that maximizes sensitivity and specificity. LSM, FibroMeter^{V2G} and Hepascore provided the highest rate of well-classified patients using this cut-off (around 73%, **Table 3**). The negative predictive values of the fibrosis tests were quite good, ranging from 76% to 90%, but the positive predictive values were insufficient with no more than 63% for the best test.

Intervals of reliable diagnosis – Due to the insufficient diagnostic accuracy obtained with a single diagnostic cut-off, we evaluated whether the fibrosis tests are able to give an accurate diagnosis in the largest rate of patients. In this setting, 2 diagnostic cut-offs have been published for NFS, the one for the exclusion (-1.455) and the other for the affirmation (0.676) of advanced fibrosis.

32.8% of patients had NFS <-1.455 and 18.2% had NFS >0.676, thus leaving the remaining 49.0% in the grey zone between these 2 thresholds. 83.8% of patients with NFS <-1.455 had F0-2 stages at liver biopsy (negative predictive value) and 72.6% of patients with NFS >0.676 had advanced fibrosis (positive predictive value). Finally, 41.4% of patients in the grey zone had advanced fibrosis.

To optimize both negative and positive predictive values, we calculated the thresholds of ≥90% positive or negative predictive values for each fibrosis test. In the core group, the rates of patients included in the grey zone between the 2 calculated thresholds, i.e., those for whom both negative or positive predictive values were <90%, were: BARD: 87.6%, APRI: 87.2%, Fibrotest: 81.9%, NFS: 78.5%, FibroMeter^{S}: 70.1%, Hepascore: 64.8%, FIB4: 64.2%, FibroMeter^{V2G}: 53.3%, LSM: 43.6% (**Figure 2**). Thus, LSM provided the lowest rate of patients in the grey zone ($p \leq 0.001$ vs blood tests). Among blood tests, the rate of patients included in the grey zone was the lowest using FibroMeter^{V2G} ($p < 0.001$ vs the 7 other blood tests). Detailed results and 90% predictive value thresholds are presented in **Table s2**.

Obuchowski index – Beyond the binary diagnosis of advanced fibrosis, we used the Obuchowski index to evaluate the ability of fibrosis tests to discriminate individual fibrosis stages. Among blood tests, FibroMeter^{V2G} had the highest Obuchowski index (0.798±0.016) with a significant difference compared to the 7 other tests ($p \leq 0.036$, **Table 2**; see **Table s1** for detailed pairwise comparisons). LSM had a significantly higher Obuchowski index than blood tests (0.834±0.014, $p \leq 0.001$), except when compared to FibroMeter^{V2G} that showed borderline significance ($p = 0.063$).

Fibrosis classifications

Fibrosis classifications that give an estimation of the histological fibrosis stage from the non-invasive fibrosis test results have already been developed in chronic hepatitis C [15]. Such classifications are very useful for the correct interpretation of fibrosis tests results in clinical practice. However, all published fibrosis classifications are based on the Metavir fibrosis staging and none has been specifically developed for NAFLD using the NASH-CRN scoring system. We thus developed fibrosis classifications for Fibroscan and FibroMeter^{V2G} in NAFLD. We chose

these 2 tests because the previous results showed they were the most accurate, especially the Obuchowski index analysis that suggested they were the best to discriminate individual fibrosis stages.

Details of the methodology used to develop the fibrosis classifications are presented in **Supplementary Material**. The LSM fibrosis classification included 7 classes (F0/1, F1±1, F1/2, F2/3, F3±1, F3/4, F4) and the FibroMeter^{V2G} one included 6 classes (F1±1, F1/2, F2/3, F3±1, F3/4, F4; **Figure 3**). The rate of well-classified patients by the LSM and the FibroMeter^{V2G} fibrosis classifications was, respectively, 80.8% vs 77.4% (p=0.190). Discrepancy between the histological fibrosis stage and the fibrosis classification was ≥2 stages in only 2.8% of patients with LSM and 4.0% with FibroMeter^{V2G} (p=0.362).

Intention-to-diagnose analysis

As stated above, the diagnostic accuracy of LSM and FibroMeter^{V2G} fibrosis classifications was not significantly different in the per-protocol analysis performed in the 505 patients having both LSM and FibroMeter^{V2G} available. We then conducted an intention-to-diagnose analysis by taking into account LSM failure in the statistical analysis. LSM failure occurred in 83 of the 588 patients included and, among the 505 remaining patients, LSM well-classified 408 patients (**Figure s1**). On the other hand, no measurement failure occurred with FibroMeter^{V2G} whom fibrosis classification well-classified 460 of the 588 included patients. Finally, in an intention-to-diagnose basis, the FibroMeter^{V2G} fibrosis classification well-classified significantly more patients than the LSM fibrosis classification (460/588 vs 408/588, p=0.001).

The Fibroscan M probe is limited by its high rate of measurement failure in obese patients (14.1% in the present study). To circumvent this limitation, the Fibroscan manufacturer has recently developed the XL probe specifically dedicated for LSM in obese patients. This new probe provides similar diagnostic accuracy than the M probe, but failure of measurement occurs in only 2-5% of patients [33, 34, 35]. Unfortunately, LSM with the XL probe was not available for the present study. We thus simulated its results with the two following hypotheses: same diagnostic accuracy than the M probe and a 5% rate of measurement failure. Under these conditions, LSM fibrosis classification using the XL probe would have well-classified 452 of the

588 included patients, which was not significantly different from the 460/588 patients with the FibroMeter^{V2G} fibrosis classification (p=0.161, **Figure s1**).

Longitudinal cohort: prognostic accuracy of blood fibrosis tests and LSM

626 NAFLD patients had a non-invasive evaluation of liver fibrosis in the Hepatology Department of the Angers University Hospital from January 2005 to December 2009. Of them, 567 had LSM using the Fibroscan M probe, and 429 had blood sampling allowing for the retrospective calculation of FibroMeter^{V2G}, Hepascore, APRI, and FIB4 (**Figure s2**). All the 4 blood fibrosis tests and LSM were available in 370 patients. The characteristics of these 370 patients are detailed in the **Table s3**. Mean age was 59.7±14.4 years and 66.2% were male. 93 patients died during the median follow-up of 6.3 years (interquartile range: 5.0 - 7.8 years; 3842 person-years). The cause of death was unknown in 10 patients and liver-related in 27 patients.

Prognostic accuracy of fibrosis tests

Harrell C-indexes calculated in the 370 patients showed that APRI was the fibrosis test with the lowest prognostic accuracy (**Table 4**). FibroMeter^{V2G} had the best discriminative ability for the prediction of all-cause mortality with a significantly higher C-index compared to the 4 other fibrosis tests. For the prediction of liver-related mortality, best C-indexes were obtained with LSM, FibroMeter^{V2G} and Hepascore.

Prognostic accuracy of fibrosis classifications

Figure 4 shows that overall survival progressively decreased with increasing LSM or FibroMeter^{V2G} result, and that LSM and FibroMeter^{V2G} fibrosis classifications categorized patients in several subgroups with significant different prognosis. The same pattern was observed when survival without liver-related death was considered (**Figure s3**). Overall survival and survival without liver-related death as a function of subgroups defined by the previously published cut-offs for APRI, FIB4, and Hepascore are depicted in **Figures s4** to **s6**.

DISCUSSION

By evaluating 8 blood tests and LSM in 588 patients, the present work is the largest cross-sectional study about non-invasive fibrosis tests in NAFLD. Among the 9 tests evaluated, our results show that LSM and FibroMeter^{V2G} are the most accurate for the non-invasive diagnosis of liver fibrosis in NAFLD. In addition, the longitudinal study demonstrates that non-invasive fibrosis tests, especially the LSM and FibroMeter^{V2G} fibrosis classifications, are prognostic markers able to categorize NAFLD patients in several subgroups with significant different prognosis. The strengths of our work are: 1/ the large sample size of the study, 2/ the evaluation of a large panel of 9 non-invasive fibrosis tests including blood tests specifically developed for NAFLD (BARD, NFS, FibroMeter^{S}) and their comparison to popular blood tests initially developed for chronic hepatitis C (APRI, FIB4, Fibrotest, Hepascore, FibroMeter^{V2G}), 3/ the evaluation of fibrosis tests by using global indexes of diagnostic accuracy (Obuchowski index, AUROC) and other indexes highly relevant to the use and interpretation of fibrosis test results in clinical practice (diagnostic cut-offs, intervals of ≥90% predictive value), 4/ for the first time in NAFLD, the development of fibrosis classifications which help physicians to interpret LSM and FibroMeter^{V2G} results in clinical practice, and 5/ the validation of the prognostic significance of these fibrosis classifications in a longitudinal cohort. In this setting, the present study is the first to evaluate and compare the prognostic accuracy of FibroMeter^{V2G}, Hepascore, and LSM in NAFLD.

LSM and FibroMeter^{V2G} provided the highest AUROCs for advanced fibrosis (F≥3), a relevant diagnostic target in clinical practice as previously stated [6, 7, 8, 13, 20]. Because diagnostic accuracy appeared insufficient using a single diagnostic cut-off, we determined the thresholds of 90% negative and positive predictive value for advanced fibrosis. As expected from AUROC results, LSM and FibroMeterV2G provided the lowest rates of patients in the grey zone between the intervals of ≥90% predictive value. To note, used with its recommended -1.455 and 0.676 cut-offs, NFS included half of the patients in the grey zone and provided suboptimal negative and positive predictive values, respectively 84% and 73%. The new Obuchowski index has been

developed to evaluate diagnostic tests against ordinal references, such as fibrosis staging on liver biopsy, and is thus particularly relevant for non-invasive tests of liver fibrosis [30]. In this setting, the Obuchowski indexes of LSM and FibroMeter^{V2G} were significantly higher than those of the 7 other blood tests. Taken together, all these results demonstrate that LSM and FibroMeter^{V2G} are the most accurate tests for the non-invasive diagnosis of liver fibrosis in NAFLD.

It may appear quite surprising that FibroMeterV2G, a blood fibrosis test initially developed for chronic hepatitis C, shows better results than tests targeting NAFLD specifically, such as NFS or FibroMeterS. In fact, NFS and FibroMeterS include indirect markers of liver fibrosis that are, individually, only moderately correlated with fibrosis stages. On the other hand, FibroMeterV2G also includes direct markers of fibrosis (alpha2-macroglobulin and hyaluronate) known to be well correlated with fibrosis stages in NAFLD and individually more accurate than indirect markers [36, 37]. The fibrogenic process in the liver uses pathways that are similar among the various causes of chronic liver diseases. Consequently, our results suggest the choice of the markers included in the test, especially direct markers of liver fibrosis, is more important than the concept of tests specifically dedicated to the cause of the chronic liver disease.

As in chronic hepatitis C [18, 19], the results of the present study show that non-invasive fibrosis tests used with a single diagnostic cut-off have insufficient diagnostic accuracy: the rate of well classified patients for advanced fibrosis did not exceed 74% and positive predictive value was very poor (<64%). In this setting, some tests such as APRI, FIB4 or NFS are used with 2 thresholds to optimize both negative and predictive values [12, 23, 24]. However, this method leaves an intermediate grey zone with undetermined diagnosis which further requires the use of liver biopsy in a significant proportion of patients. Longitudinal studies have circumvented this limitation by showing that the 2 diagnostic thresholds finally individualize 3 prognostic categories with low, intermediate, or high-risk of liver-related complications or mortality [38, 39]. At the light of these concepts, our fibrosis classifications provide several advantages. Compared to a simple binary diagnosis of advanced fibrosis (F0-2 vs F3/4), they allow for a more

precise estimation of the liver fibrosis stage (≥6 diagnostic classes) with a good diagnostic accuracy (around 80%) and without any liver biopsy requirement. In addition, the prognostic study shows that the fibrosis classifications categorize patients in several subgroups with significant different prognosis. In this setting, the pattern of the survival curves obtained with our fibrosis classifications is very similar that the one obtained with the 5 histological fibrosis stages [6]. Finally, the fibrosis classifications we developed are clinically relevant in NAFLD because they help the physician to precisely estimate both the liver fibrosis stage and the patient prognosis without any liver biopsy.

Our study has 2 main limitations. LSM with the XL probe of Fibroscan was not available because it is been commercialised only since a few years. As it increases in parallel with body mass index [40], failure of measurement with the M probe represents a significant limitation of LSM in NAFLD patients. In this setting, our result show that the diagnostic accuracy of the LSM fibrosis classification lost 9% when evaluated in an intention-to-diagnosis manner. The recent XL probe specifically dedicated to LSM in obese patients provides similar accuracy with better feasibility (≥95% success rate) than the M probe [33, 34, 35]. By using these previously published results, we simulated the results that would have been obtained with a fibrosis classification for LSM with the XL probe. In an intention-to-diagnose basis, the XL probe provided similar diagnostic accuracy that the best-performing blood fibrosis test. The second limitation of our work is the lack of NFS in the longitudinal prognostic study. It has been recently suggested that NFS performs better than APRI or NFS for the prognostic assessment in NAFLD [38, 39]. Further works will have to compare the prognostic accuracy of NFS, FibroMeterV2G, and LSM in NAFLD.

In conclusion, by comparing 9 fibrosis tests in a large population, we identified FibroMeter^{V2G} and Fibroscan as the best tools for the non-invasive diagnosis of liver fibrosis in NAFLD. Beyond the classical binary diagnosis of advanced fibrosis, new fibrosis classifications developed for these 2 fibrosis tests allowed for a precise estimation of the histological fibrosis stage without any liver biopsy requirement. In addition to the diagnostic evaluation, these classifications also provide prognostic information by categorizing patients in several subgroups with a significant increase in mortality risk.

REFERENCES

1 Younossi ZM, Stepanova M, Afendy M, Fang Y, Younossi Y, Mir H*, et al.* Changes in the prevalence of the most common causes of chronic liver diseases in the United States from 1988 to 2008. Clin Gastroenterol Hepatol 2011;**9**:524-30.

2 Das K, Mukherjee PS, Ghosh A, Ghosh S, Mridha AR, Dhibar T*, et al.* Nonobese population in a developing country has a high prevalence of nonalcoholic fatty liver and significant liver disease. Hepatology 2010;**51**:1593-602.

3 Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW*, et al.* Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 2005;**41**:1313-21.

4 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**:810-20.

5 Brunt EM, Kleiner DE, Wilson LA, Unalp A, Behling CE, Lavine JE*, et al.* Portal chronic inflammation in nonalcoholic fatty liver disease (NAFLD): a histologic marker of advanced NAFLD-Clinicopathologic correlations from the nonalcoholic steatohepatitis clinical research network. Hepatology 2009;**49**:809-20.

6 Angulo P, Kleiner DE, Dam-Larsen S, Adams LA, Bjornsson ES, Charatcharoenwitthaya P*, et al.* Liver Fibrosis, but no Other Histologic Features, Associates with Long-term Outcomes of Patients With Nonalcoholic Fatty Liver Disease. Gastroenterology 2015; in press.

7 Ekstedt M, Hagstrom H, Nasr P, Fredrikson M, Stal P, Kechagias S*, et al.* Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. Hepatology 2015;**61**:1547-54.

8 Younossi ZM, Stepanova M, Rafiq N, Makhlouf H, Younoszai Z, Agrawal R*, et al.* Pathologic criteria for nonalcoholic steatohepatitis: interprotocol agreement and ability to predict liver-related mortality. Hepatology 2011;**53**:1874-82.

9 Chalasani N, Younossi Z, Lavine JE, Diehl AM, Brunt EM, Cusi K*, et al.* The diagnosis and management of non-alcoholic fatty liver disease: practice guideline by the American Gastroenterological Association, American Association for the Study of Liver Diseases, and American College of Gastroenterology. Gastroenterology 2012;**142**:1592-609.

10 Bravo AA, Sheth SG, Chopra S. Liver biopsy. N Engl J Med 2001;**344**:495-500.

11 Singh S, Allen AM, Wang Z, Prokop LJ, Murad MH, Loomba R. Fibrosis progression in nonalcoholic fatty liver vs nonalcoholic steatohepatitis: a systematic review and meta-analysis of paired-biopsy studies. Clin Gastroenterol Hepatol 2015;**13**:643-54 e1-9; quiz e39-40.

12 Angulo P, Hui JM, Marchesini G, Bugianesi E, George J, Farrell GC*, et al.* The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. Hepatology 2007;**45**:846-54.

13 Musso G, Gambino R, Cassader M, Pagano G. Meta-analysis: natural history of non-alcoholic fatty liver disease (NAFLD) and diagnostic accuracy of non-invasive tests for liver disease severity. Ann Med 2011;**43**:617-49.

14 Boursier J, Rousselet MC, Aube C, Cales P. Liver fibrosis in patients with non-alcoholic fatty liver disease: diagnostic options in clinical practice. Expert Opin Med Diagn 2012;**6**:381-94.

15 Boursier J, Bertrais S, Oberti F, Gallois Y, Fouchard-Hubert I, Rousselet MC*, et al.* Comparison of accuracy of fibrosis degree classifications by liver biospy and non invasive tests in chronic hepatitis C. BMC Gastroenterol 2011;**30**:132.
16 Boursier J, de Ledinghen V, Zarski JP, Fouchard-Hubert I, Gallois Y, Oberti F*, et al.* Comparison of 8 diagnostic algorithms for liver fibrosis in hepatitis C: new algorithms are more precise and entirely non-invasive. Hepatology 2012;**55**:58-67.
17 Cales P, Boursier J, Ducancelle A, Oberti F, Hubert I, Hunault G*, et al.* Improved fibrosis staging by elastometry and blood test in chronic hepatitis C. Liver Int 2014;**34**:907-17.
18 Degos F, Perez P, Roche B, Mahmoudi A, Asselineau J, Voitot H*, et al.* Diagnostic accuracy of FibroScan and comparison to liver fibrosis biomarkers in chronic viral hepatitis: a multicenter prospective study (the FIBROSTIC study). J Hepatol 2010;**53**:1013-21.
19 Zarski JP, Sturm N, Guechot J, Paris A, Zafrani ES, Asselah T*, et al.* Comparison of nine blood tests and transient elastography for liver fibrosis in chronic hepatitis C: The ANRS HCEP-23 study. J Hepatol 2012;**56**:55-62.
20 European Association for the Study of the Liver. Electronic address, easloffice@easloffice.eu; Asociacion Latinoamericana para el Estudio del Higado. EASL-ALEH Clinical Practice Guidelines: Non-invasive tests for evaluation of liver disease severity and prognosis. J Hepatol 2015;**63**:237-64.
21 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**:1441-7.
22 Cales P, Laine F, Boursier J, Deugnier Y, Moal V, Oberti F*, et al.* Comparison of blood tests for liver fibrosis specific or not to NAFLD. J Hepatol 2009;**50**:165-73.
23 Wai CT, Greenson JK, Fontana RJ, Kalbfleisch JD, Marrero JA, Conjeevaram HS*, et al.* A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. Hepatology 2003;**38**:518-26.
24 Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, Montaner J*, et al.* Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology 2006;**43**:1317-25.
25 Imbert-Bismut F, Ratziu V, Pieroni L, Charlotte F, Benhamou Y, Poynard T. Biochemical markers of liver fibrosis in patients with hepatitis C virus infection: a prospective study. Lancet 2001;**357**:1069-75.
26 Adams LA, Bulsara M, Rossi E, DeBoer B, Speers D, George J*, et al.* Hepascore: an accurate validated predictor of liver fibrosis in chronic hepatitis C infection. Clin Chem 2005;**51**:1867-73.
27 Calès P, Oberti F, Michalak S, Hubert-Fouchard I, Rousselet MC, Konate A*, et al.* A novel panel of blood markers to assess the degree of liver fibrosis. Hepatology 2005;**42**:1373-81.
28 Cales P, Veillon P, Konate A, Mathieu E, Ternisien C, Chevailler A*, et al.* Reproducibility of blood tests of liver fibrosis in clinical practice. Clin Biochem 2008;**41**:10-8.
29 Castera L, Forns X, Alberti A. Non-invasive evaluation of liver fibrosis using transient elastography. J Hepatol 2008;**48**:835-47.
30 Lambert J, Halfon P, Penaranda G, Bedossa P, Cacoub P, Carrat F. How to measure the diagnostic accuracy of noninvasive liver fibrosis indices: the area under the ROC curve revisited. Clin Chem 2008;**54**:1372-8.
31 Boursier J, Brochard C, Bertrais S, Michalak S, Gallois Y, Fouchard-Hubert I*, 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**:178-88.
32 Boursier J, de Ledinghen V, Poynard T, Guechot J, Carrat F, Leroy V*, et al.* An extension of STARD statements for reporting diagnostic accuracy studies on liver fibrosis tests: The Liver-FibroSTARD standards. J Hepatol 2014.

33 de Ledinghen V, Wong VW, Vergniol J, Wong GL, Foucher J, Chu SH*, et al.* Diagnosis of liver fibrosis and cirrhosis using liver stiffness measurement: comparison between M and XL probe of FibroScan(R). J Hepatol 2012;**56**:833-9.
34 Myers RP, Pomier-Layrargues G, Kirsch R, Pollett A, Duarte-Rojo A, Wong D*, et al.* Feasibility and diagnostic performance of the Fibroscan XL probe for liver stiffness measurement in overweight and obese patients. Hepatology 2012;**2012**:199-208.
35 Wong VW, Vergniol J, Wong GL, Foucher J, Chan AW, Chermak F*, et al.* Liver stiffness measurement using XL probe in patients with nonalcoholic fatty liver disease. Am J Gastroenterol 2012;**107**:1862-71.
36 Afdhal NH, Nunes D. Evaluation of liver fibrosis: a concise review. Am J Gastroenterol 2004;**99**:1160-74.
37 Guha IN, Parkes J, Roderick PR, Harris S, Rosenberg WM. Non-invasive markers associated with liver fibrosis in non-alcoholic fatty liver disease. Gut 2006;**55**:1650-60.
38 Angulo P, Bugianesi E, Bjornsson ES, Charatcharoenwitthaya P, Mills PR, Barrera F*, et al.* Simple Noninvasive Systems Predict Long-term Outcomes of Patients With Nonalcoholic Fatty Liver Disease. Gastroenterology 2013;**145**:782-9 e4.
39 Kim D, Kim WR, Kim HJ, Therneau TM. Association between noninvasive fibrosis markers and mortality among adults with nonalcoholic fatty liver disease in the United States. Hepatology 2013;**57**:1357-65.
40 Castera L, Foucher J, Bernard PH, Carvalho F, Allaix D, Merrouche W*, et al.* Pitfalls of liver stiffness measurement: a 5-year prospective study of 13,369 examinations. Hepatology 2010;**51**:828-35.

TABLE

Table 1: Patient characteristics at inclusion in the cross-sectional study.

Table 2: AUROCs and Obuchowski indexes of the non-invasive fibrosis tests in the core group.

Table 3: Accuracy of non-invasive fibrosis tests for the binary diagnosis of advanced F≥3 fibrosis in the core group.

Table 4: Prognostic accuracy of non-invasive fibrosis tests evaluated by the Harrell C-index (95% CI into brackets).

Table 1: Patient characteristics at inclusion in the cross-sectional study.

	Group			p
	All (n=588)	**Core [a] (n=452)**	**Others (n=136)**	
Age (years)	55.9 ± 12.0	55.9 ± 12.0	56.1 ± 11.9	0.695
Male sex (%)	57.3	60.0	48.5	0.023
Diabetes (%) [b]	48.0	46.7	52.2	0.282
BMI (kg/m^2)	31.7 ± 5.8	31.1 ± 5.2	33.8 ± 7.2	<0.001
Biopsy length (mm)	26 ± 12	27 ± 11	23 ± 11	<0.001
Fibrosis stage (%):				0.518
- 0	9.0	8.6	10.3	
- 1	25.9	27.2	21.3	
- 2	26.5	26.1	27.9	
- 3	24.8	25.2	23.5	
- 4	13.8	12.8	16.9	
AST (IU/l)	48 ± 30	48 ± 29	47 ± 31	0.133
ALT (IU/l)	69 ± 49	68 ± 39	71 ± 72	0.010
GammaGT (IU/l)	142 ± 191	142 ± 187	142 ± 205	0.710
Total bilirubin (µmol/l)	12 ± 8	12 ± 9	12 ± 8	0.850
Prothrombin time (%)	94 ± 15	95 ± 14	93 ± 17	0.041
Platelets (G/l)	217 ± 70	216 ± 65	220 ± 86	0.855
LSM median (kPa) [c]	12.7 ± 11.5	12.7 ± 11.2	12.3 ± 13.5	0.094

BMI: body mass index; NAFLD: non-alcoholic fatty liver disease; AST: aspartate aminotransferase; ALT: alanine aminotransferase; LSM: liver stiffness measurement by Fibroscan

[a] all 8 studied blood fibrosis tests and LSM available for each patient

[b] either anti-diabetic treatment or fasting glycemia ≥126 mg/dl

[c] in the 505 patients with available results for liver stiffness

Table 2: AUROCs and Obuchowski indexes of the non-invasive fibrosis tests in the core group. Detailed pairwise comparisons are presented in Table s1 in Supplementary Material.

Fibrosis test	**AUROC**			**Obuchowski**
	F ≥2	**F ≥3**	**F4**	**index**
BARD	0.698 ± 0.025	0.695 ± 0.024	0.694 ± 0.031	0.698 ± 0.019
NFS	0.717 ± 0.024	0.732 ± 0.024	0.766 ± 0.032	0.730 ± 0.019
FibroMeter^{S}	0.764 ± 0.023	0.759 ± 0.023	0.779 ± 0.029	0.763 ± 0.017
APRI	0.719 ± 0.025	0.754 ± 0.023	0.767 ± 0.034	0.735 ± 0.019
FIB4	0.721 ± 0.024	0.780 ± 0.022	0.777 ± 0.033	0.748 ± 0.019
Fibrotest	0.716 ± 0.025	0.736 ± 0.024	0.761 ± 0.034	0.722 ± 0.019
Hepascore	0.753 ± 0.023	0.778 ± 0.022	0.807 ± 0.034	0.765 ± 0.018
FibroMeter^{V2G}	0.786 ± 0.022	0.817 ± 0.020	0.824 ± 0.029	0.798 ± 0.016
LSM	0.842 ± 0.019	0.831 ± 0.019	0.864 ± 0.024	0.834 ± 0.014

NFS: NAFLD Fibrosis Score; LSM: liver stiffness measurement by Fibroscan

Table 3: Accuracy of non-invasive fibrosis tests for the binary diagnosis of advanced F≥3 fibrosis in the core group. Diagnostic cut-offs were calculated according to the highest Youden index that maximizes sensitivity and specificity.

Fibrosis test	**Cut-off**	**DA**	**Se**	**Spe**	**NPV**	**PPV**	**-LR**	**+LR**	**OR**
BARD	2	61.5 [a]	79.1	50.7	79.8	49.6	0.41	1.60	3.9
NFS	-1.036	66.4	76.7	60.0	80.8	54.1	0.39	1.92	5.0
FibroMeter^{S}	0.311	68.6	79.7	61.8	83.2	56.1	0.33	2.08	6.3
APRI	0.559	70.6	61.0	76.4	76.2	61.4	0.51	2.59	5.1
FIB4	1.515	70.4	75.6	67.1	81.7	58.6	0.36	2.30	6.3
Fibrotest	0.316	66.2	81.4	56.8	83.2	53.6	0.33	1.88	5.7
Hepascore	0.322	72.8 [b]	67.4	76.1	79.2	63.4	0.43	2.82	6.6
FibroMeter^{V2G}	0.453	73.7 [c]	76.7	71.8	83.4	62.6	0.32	2.72	8.4
LSM	8.7	72.6 [b]	88.4	62.9	89.8	59.4	0.18	2.38	12.9

DA: diagnostic accuracy; Se: sensitivity; Spe: specificity; NPV negative predictive value; PPV: positive predictive value; -LR: negative likelihood ratio; +LR: positive likelihood ratio; OR: odd ratio; NFS: NAFLD Fibrosis Score; LSM: liver stiffness measurement by Fibroscan

[a] p≤0.059 vs other fibrosis tests (except vs Fibrotest: p=0.121)

[b] p≤0.025 vs BARD, NFS, and Fibrotest

[c] p≤0.031 vs BARD, NFS, FibroMeter^{S}, and Fibrotest

Table 4: Prognostic accuracy of non-invasive fibrosis tests evaluated by the Harrell C-index (95% CI into brackets).

	All-cause mortality	**Liver-related mortality**
APRI	0.572 [0.469-0.638] [a]	0.726 [0.602-0.842] [a]
FIB4	0.714 [0.657-0.763]	0.811 [0.731-0.887]
Hepascore	0.748 [0.688-0.803]	0.862 [0.788-0.927]
FibroMeter^{V2G}	0.801 [0.754-0.842] [b]	0.872 [0.808-0.928] [c]
LSM	0.741 [0.681-0.797]	0.877 [0.814-0.934]

LSM: liver stiffness measurement by Fibroscan
[a] p ≤0.021 vs other fibrosis tests
[b] p ≤0.010 vs other fibrosis tests
[c] p=0.047 vs FIB4

FIGURE LEGENDS

Figure 1: Flow chart of the cross-sectional study. LSM: liver stiffness measurement by Fibroscan.

Figure 2: Rate of patients included in the intervals of ≥90% negative (NPV) or positive (PPV) predictive values, as a function of blood fibrosis tests or liver stiffness measurement by Fibroscan (LSM). The patients in the 'grey zone' are those for whom NPV and PPV are <90%. Results presented are those obtained in the core group that allowed for direct comparison between fibrosis tests. NFS: NAFLD Fibrosis Score, FM^{S}: FibroMeter^{S}, FM^{V2G}: FibroMeter^{V2G}.

Figure 3: Fibrosis classifications of Fibroscan and FibroMeterV2G which give an estimation of the histological fibrosis stage from the fibrosis test result. LSM: liver stiffness measurement by Fibroscan.

Figure 4: Overall survival as a function of the subgroups defined by the fibrosis classification developed for liver stiffness measurement by Fibroscan (LSM, ***panel 4a***) or the fibrosis classification developed for FibroMeter^{V2G} (***panel 4b***).

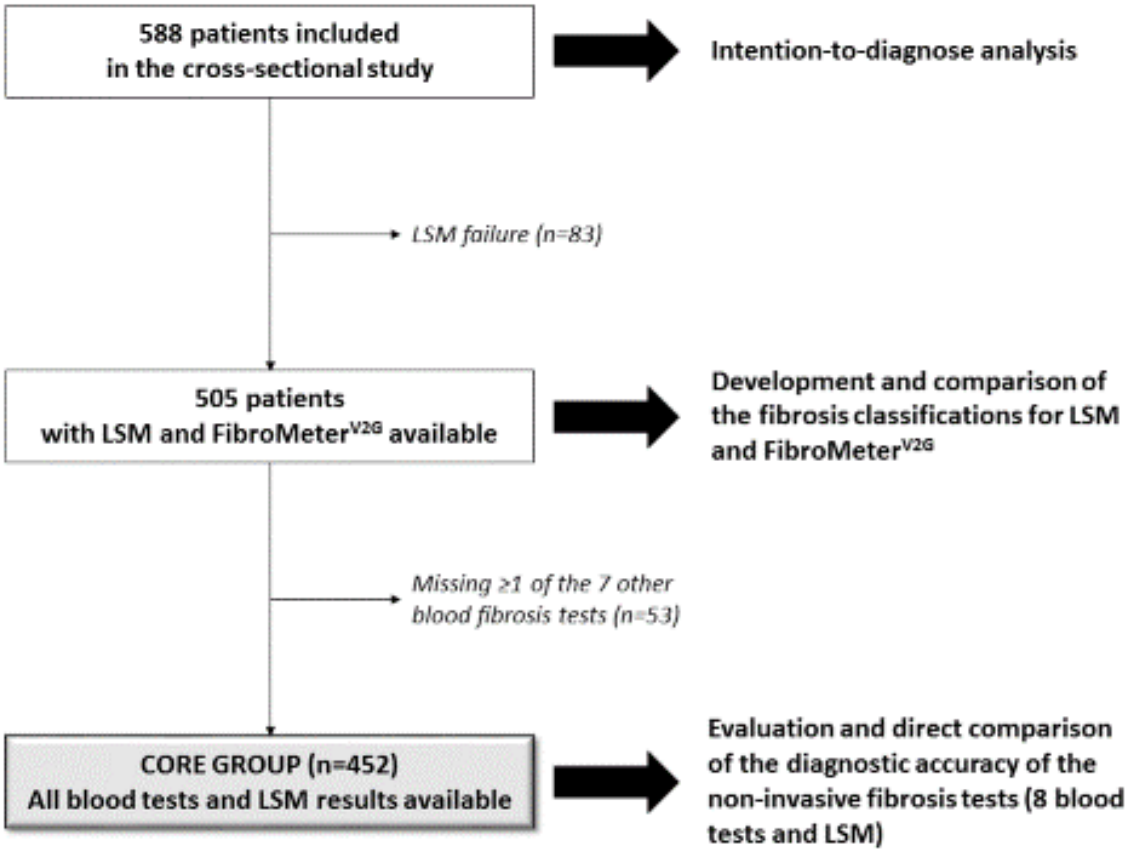

Figure 1: Flow chart of the cross-sectional study. LSM: liver stiffness measurement by Fibroscan.

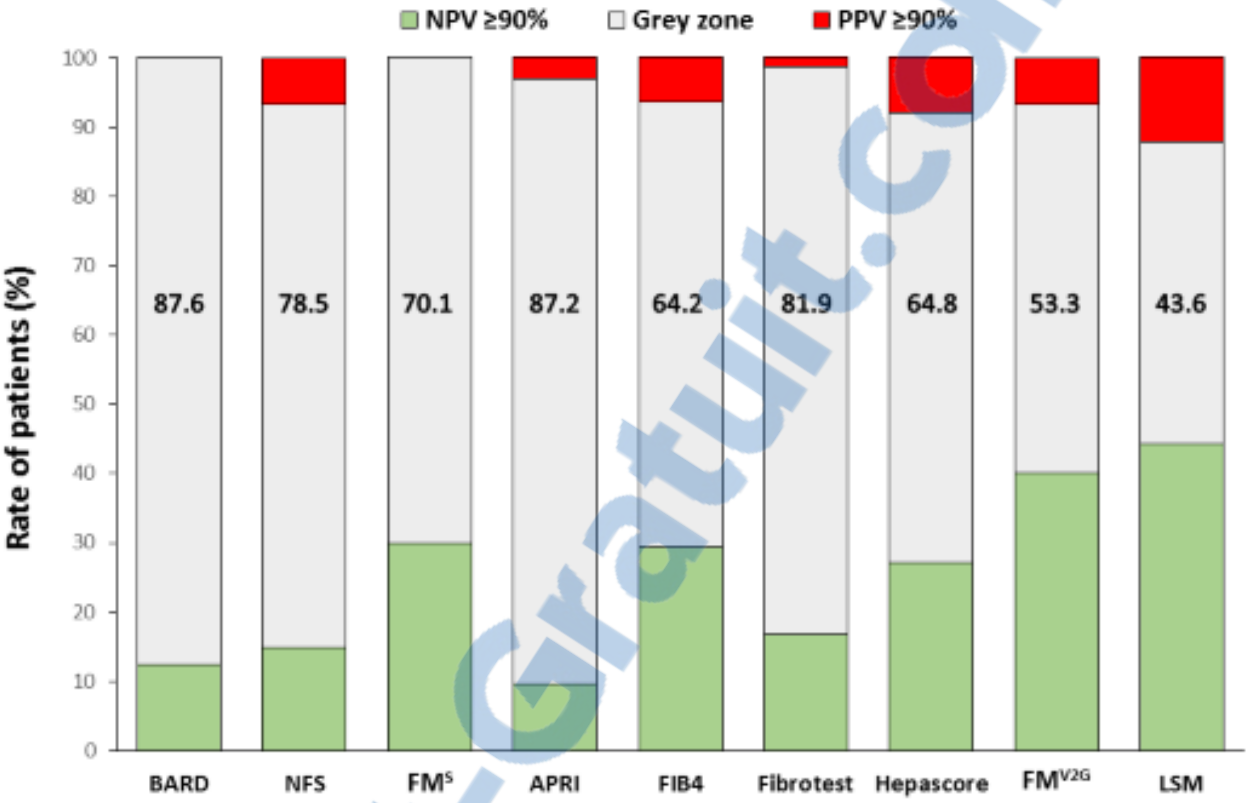

Figure 2: Rate of patients included in the intervals of ≥90% negative (NPV) or positive (PPV) predictive values, as a function of blood fibrosis tests or liver stiffness measurement by Fibroscan (LSM). The patients in the 'grey zone' are those for whom NPV and PPV are <90%. Results presented are those obtained in the core group that allowed for direct comparison between fibrosis tests. NFS: NAFLD Fibrosis Score, FM^{S}: FibroMeter^{S}, FM^{V2G}: FibroMeter^{V2G}.

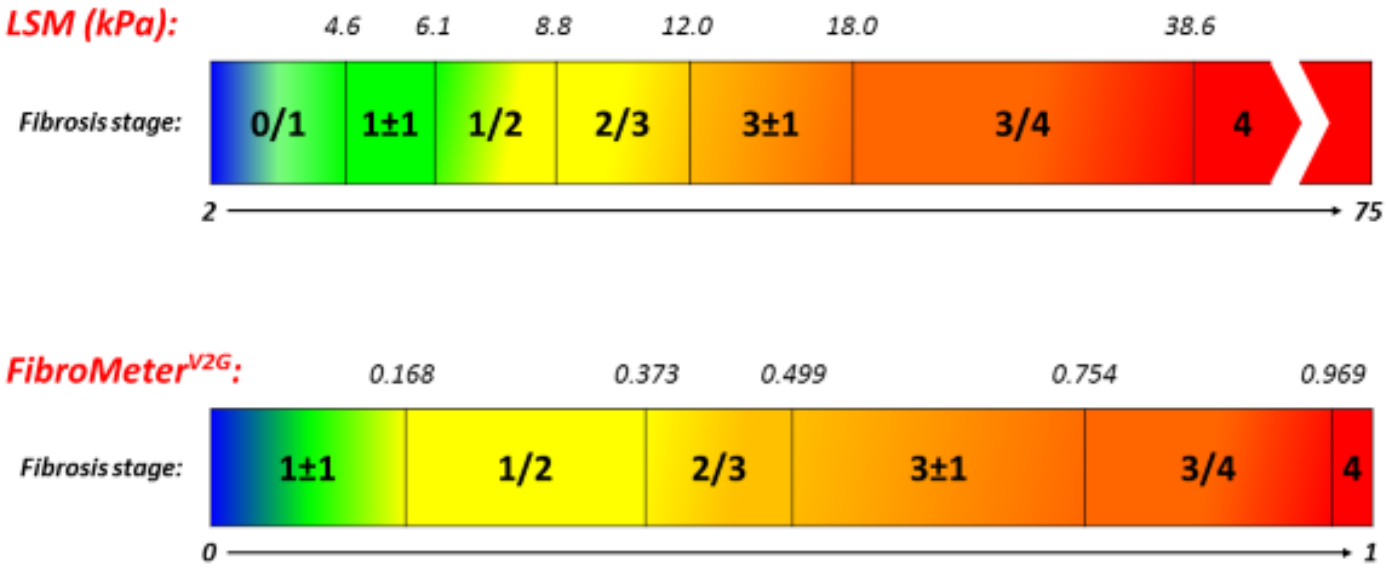

Figure 3: Fibrosis classifications of Fibroscan and FibroMeterV2G which give an estimation of the histological fibrosis stage from the fibrosis test result. LSM: liver stiffness measurement by Fibroscan.

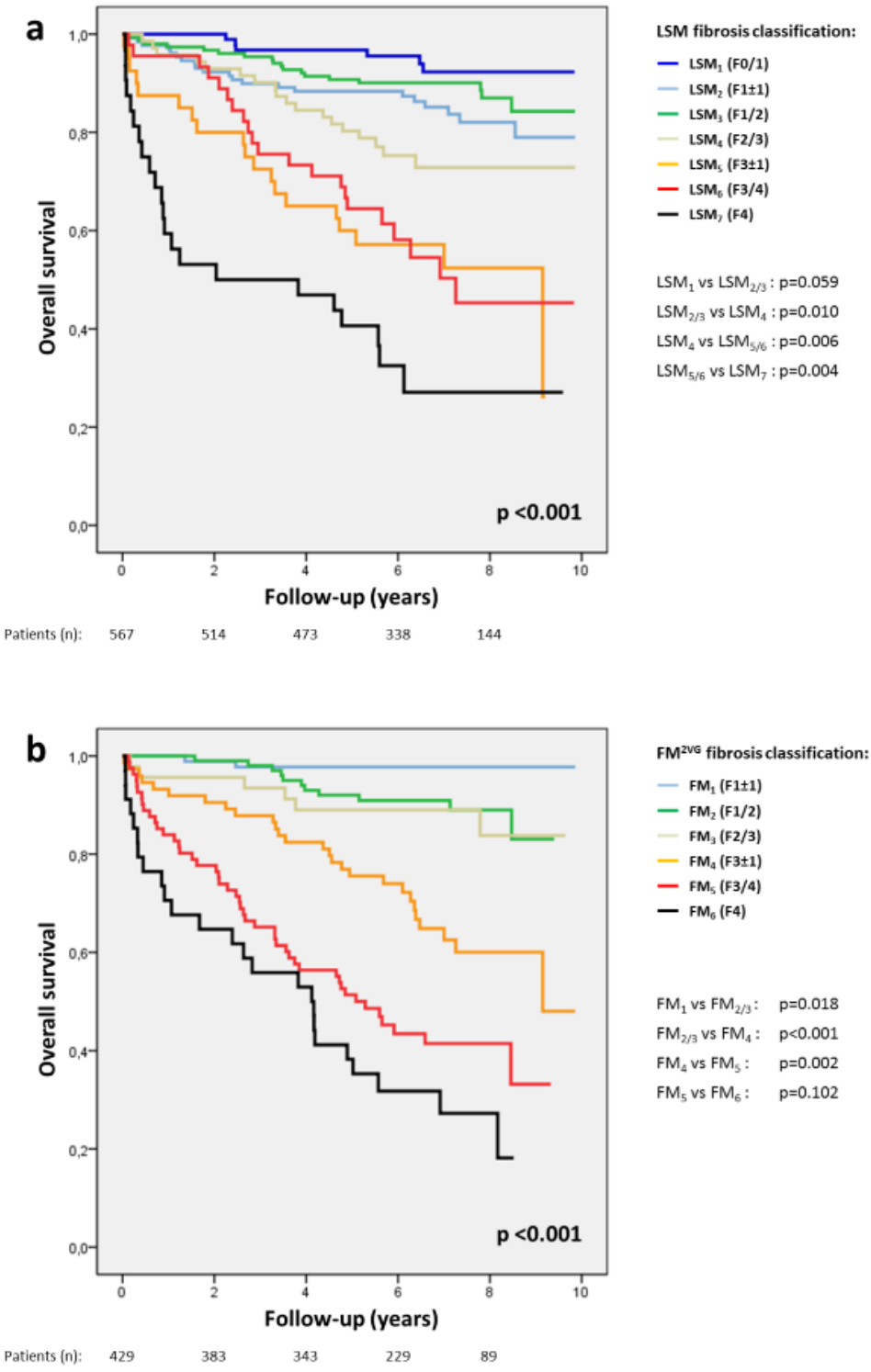

Figure 4: Overall survival as a function of the subgroups defined by the fibrosis classification developed for liver stiffness measurement by Fibroscan (LSM, ***panel 4a***) or the fibrosis classification developed for FibroMeter^{V2G} (***panel 4b***).

SUPPLEMENTARY MATERIAL

Intervals of ≥90% predictive values

Intervals of ≥90% predictive values correspond to the intervals of fibrosis test values where the accuracy for a diagnostic target is considered sufficiently reliable for clinical practice. The thresholds of 90% predictive values for the diagnostic target are calculated (**Figure A**), and they define two intervals of blood tests values:

- A lower interval, defined by a blood test value ≤ the 90% negative predictive value threshold, where patients have a ≥90% chance of not having the diagnostic target;
- And a higher interval, defined by a blood test value ≥ the 90% positive predictive value threshold, where patients have a ≥90% risk of having the diagnostic target.

Between these two thresholds, in the "grey zone", the diagnostic accuracy is insufficient (i.e., negative and positive predictive values <90%) and liver biopsy is theoretically required.

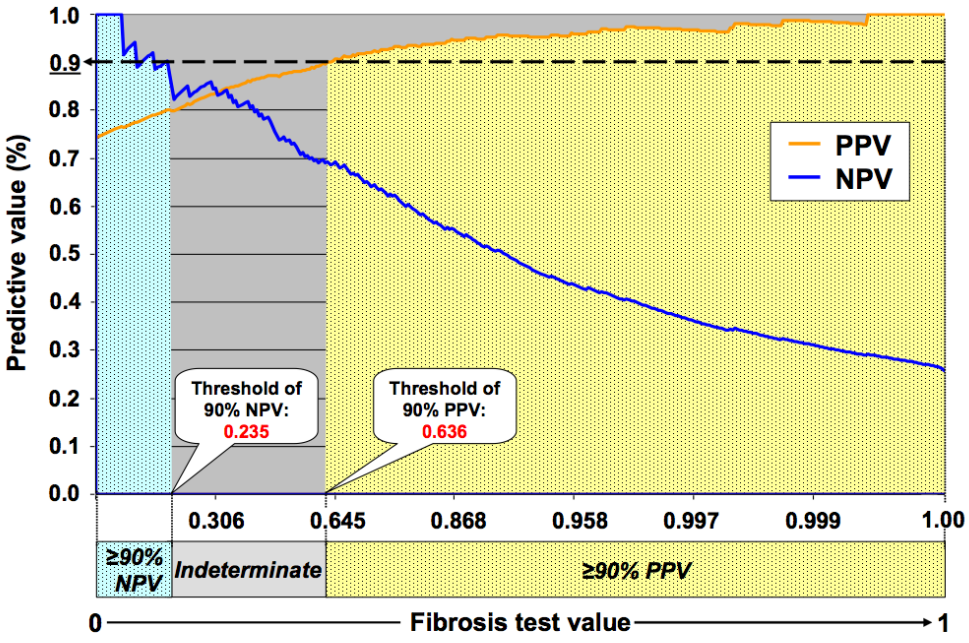

Figure A: *Intervals of ≥90% negative (NPV) and positive (PPV) predictive values for the diagnosis of advanced fibrosis (F≥3). NPV for significant fibrosis is ≥90% in patients with a score ≤0.235; PPV is ≥90% in patients with a score ≥0.636 (this figure is an example)*

Development of a fibrosis classification for LSM in NAFLD

Development of a fibrosis classification for LSM in NAFLD

The patients were ranked according to increasing LSM results. The prevalence of fibrosis stages was then calculated in each n to n+20 patient subgroup: patients 1 to 21, 2 to 22, 3 to 23, … , 485 to 505. Results were reported in **Figure B** that precisely depicts the prevalence of fibrosis stages as a function of LSM result. We then determined the intervals of LSM results where the same fibrosis stages have ≥80% prevalence (delimited by dashed lines in **Figure B**). These intervals represent the new fibrosis classification for LSM in NAFLD.

The same methodology was used to derive a fibrosis classification for FibroMeter^{V2G} (**Figure C**).

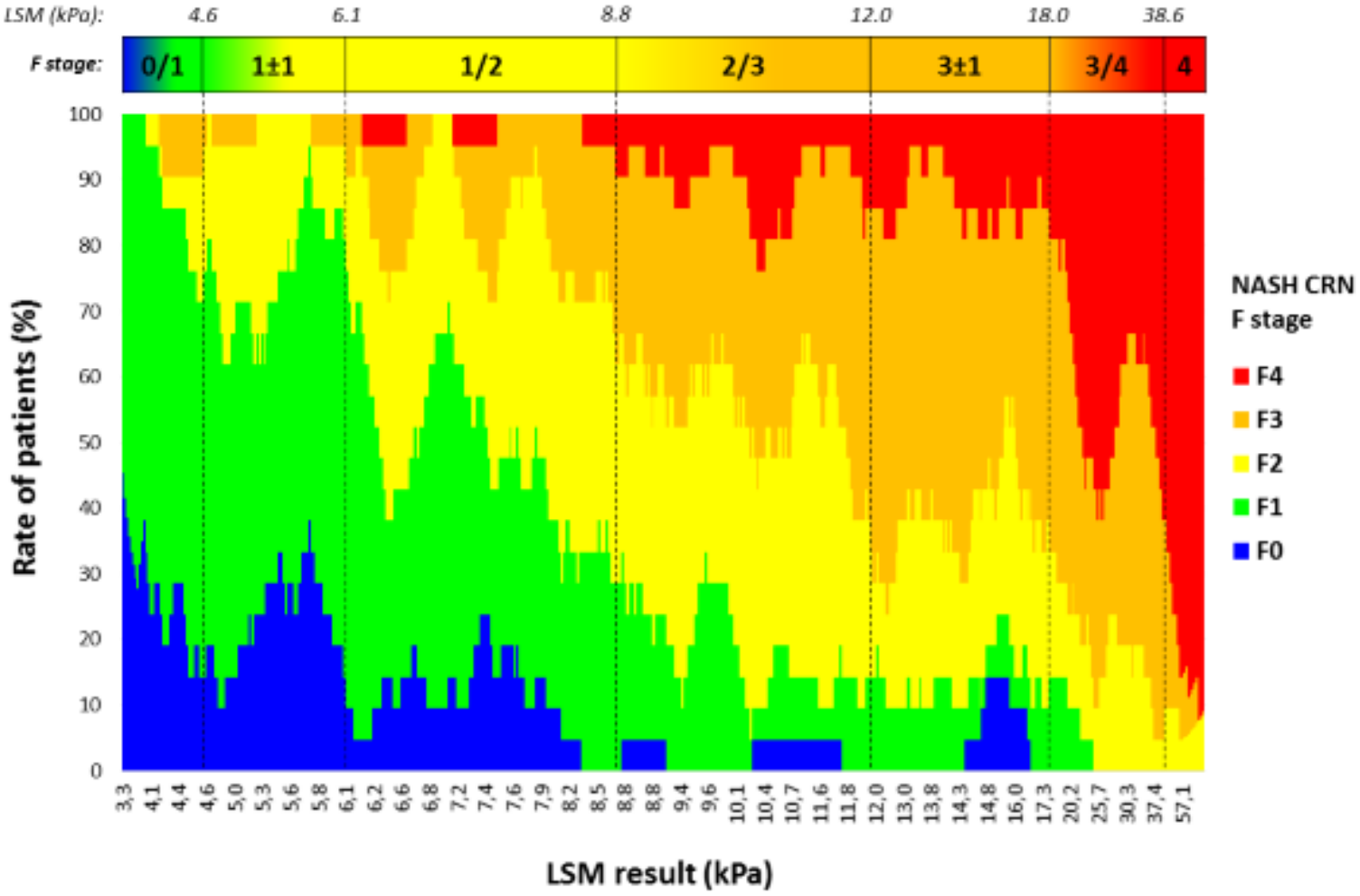

Figure B: *Development of a fibrosis classification for LSM*

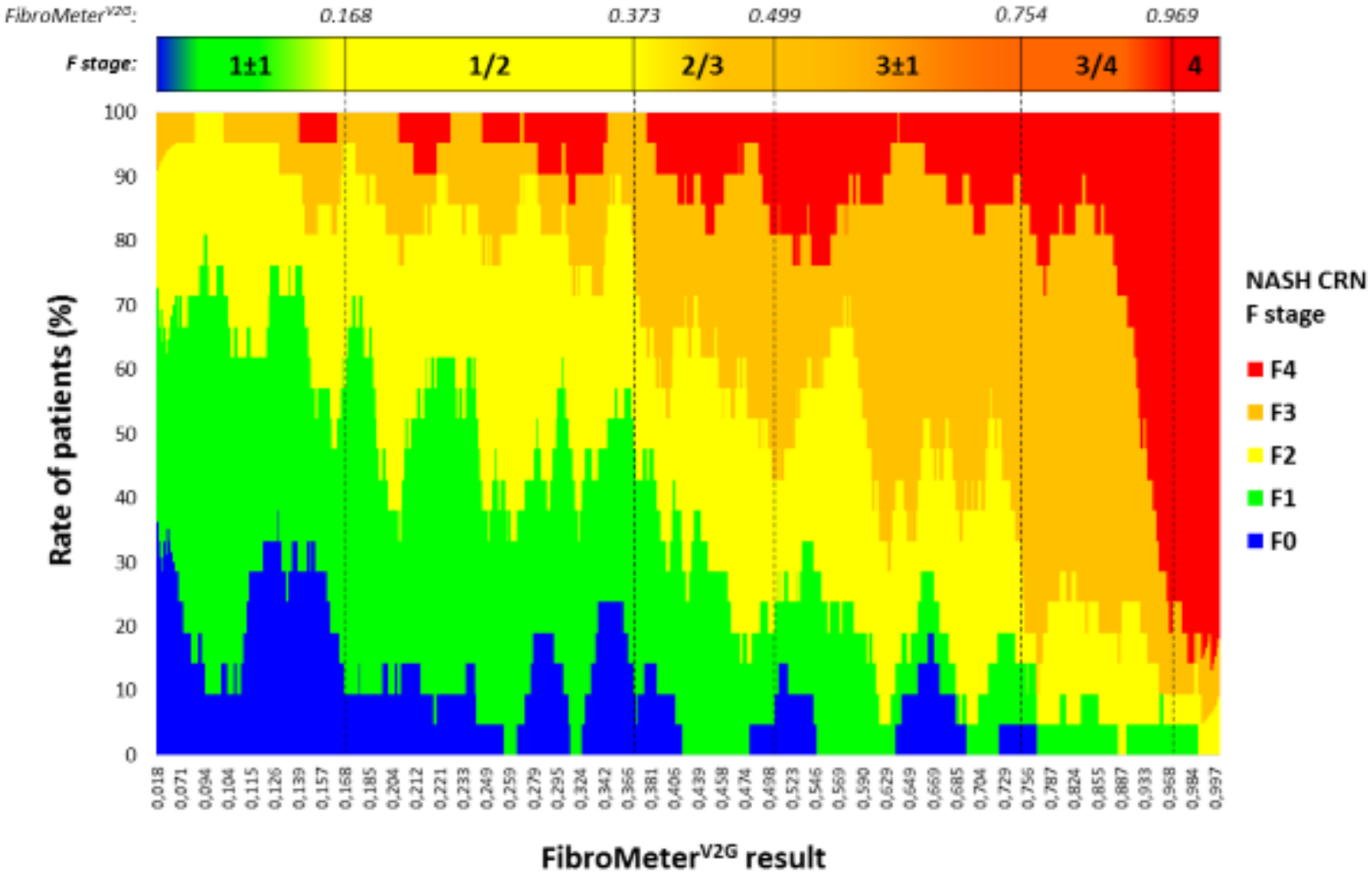

Figure C: *Development of a fibrosis classification for FibroMeter^{V2G}*

Table s1: Direct comparison of AUROCs and Obuchowski indexes of non-invasive fibrosis tests in the core group.

Fibrosis test	**AUROC**			**Obuchowski**
	F ≥2	**F ≥3**	**F4**	**index**
BARD	0.698 ± 0.025	0.695 ± 0.024	0.694 ± 0.031	0.698 ± 0.019
NFS	0.717 ± 0.024	0.732 ± 0.024	0.766 ± 0.032	0.730 ± 0.019
FibroMeter^{S}	0.764 ± 0.023	0.759 ± 0.023	0.779 ± 0.029	0.763 ± 0.017
APRI	0.719 ± 0.025	0.754 ± 0.023	0.767 ± 0.034	0.735 ± 0.019
FIB4	0.721 ± 0.024	0.780 ± 0.022	0.777 ± 0.033	0.748 ± 0.019
Fibrotest	0.716 ± 0.025	0.736 ± 0.024	0.761 ± 0.034	0.722 ± 0.019
Hepascore	0.753 ± 0.023	0.778 ± 0.022	0.807 ± 0.034	0.765 ± 0.018
FibroMeter^{V2G}	0.786 ± 0.022	0.817 ± 0.020	0.824 ± 0.029	0.798 ± 0.016
LSM	0.842 ± 0.019	0.831 ± 0.019	0.864 ± 0.024	0.834 ± 0.014
Comparison (p)				
BARD vs NFS	0.452	0.128	0.049	0.128
BARD vs FibroMeter^{S}	0.010	0.010	0.012	0.002
BARD vs APRI	0.556	0.068	0.077	0.161
BARD vs FIB4	0.419	0.002	0.026	0.027
BARD vs Fibrotest	0.593	0.200	0.108	0.356
BARD vs Hepascore	0.083	0.007	0.005	0.009
BARD vs FibroMeter^{V2G}	0.004	<0.001	<0.001	<0.001
BARD vs LSM	<0.001	<0.001	<0.001	<0.001
NFS vs FibroMeter^{S}	0.002	0.123	0.643	0.014
NFS vs APRI	0.950	0.427	0.974	0.807
NFS vs FIB4	0.833	0.015	0.678	0.270
NFS vs Fibrotest	0.977	0.899	0.911	0.737
NFS vs Hepascore	0.193	0.097	0.194	0.119
NFS vs FibroMeter^{V2G}	0.003	<0.001	0.043	<0.001

NFS vs LSM	<0.001	<0.001	0.006	<0.001
FibroMeter^{S} vs APRI	0.071	0.823	0.704	0.144
FibroMeter^{S} vs FIB4	0.006	0.222	0.956	0.263
FibroMeter^{S} vs Fibrotest	0.061	0.350	0.617	0.051
FibroMeter^{S} S vs Hepascore	0.688	0.474	0.466	0.954
FibroMeter^{S} vs FibroMeter^{V2G}	0.259	0.003	0.114	0.025
FibroMeter^{S} S vs LSM	0.003	0.003	0.010	<0.001
APRI vs FIB4	0.915	0.181	0.690	0.453
APRI vs Fibrotest	0.923	0.511	0.892	0.556
APRI vs Hepascore	0.247	0.417	0.362	0.235
APRI vs FibroMeter^{V2G}	0.004	0.006	0.090	0.001
APRI vs LSM	<0.001	0.004	0.013	<0.001
FIB4 vs Fibrotest	0.836	0.061	0.623	0.198
FIB4 vs Hepascore	0.221	0.937	0.438	0.446
FIB4 vs FibroMeter^{V2G}	<0.001	0.025	0.046	<0.001
FIB4 vs LSM	<0.001	0.041	0.021	<0.001
Fibrotest vs Hepascore	0.083	0.026	0.122	0.012
Fibrotest vs FibroMeter^{V2G}	<0.001	<0.001	0.003	<0.001
Fibrotest vs LSM	<0.001	<0.001	0.010	<0.001
Hepascore vs FibroMeter^{V2G}	0.072	0.014	0.521	0.036
Hepascore vs LSM	0.001	0.040	0.110	0.001
FibroMeter^{V2G} vs LSM	0.033	0.559	0.219	0.063

NFS: NAFLD Fibrosis Score, LSM: liver stiffness measurement by Fibroscan

Table s2: Thresholds of 90% negative (NPV) or positive (PPV) predictive value for advanced F≥3 fibrosis, and rates of patients included in the intermediate grey zone between these 2 thresholds (i.e., with both NPV and PPV <90%) in the core group.

Fibrosis test	**Thresholds**		**Rate of patients (%)**		
	90% NPV	**90% PPV**	**NPV ≥90%**	**Intermediate zone**	**PPV ≥90%**
BARD	0	-	12.4	87.6	0.0
NFS	-2.697	1.348	14.8	78.5	6.6
FibroMeter^{S}	0.132	-	29.9	70.1	0.0
APRI	0.244	1.740	9.7	87.2	3.1
FIB4	1.063	3.618	29.4	64.2	6.4
Fibrotest	0.167	0.952	16.8	81.9	1.3
Hepascore	0.120	0.955	27.2	64.8	8.0
FibroMeter^{V2G}	0.325	0.925	40.0 [a]	53.3 [a]	6.6
LSM	8.7	20.5	44.2 [b]	43.6 [c]	12.2 [c]

NFS: NAFLD Fibrosis Score; LSM: liver stiffness measurement by Fibroscan
[a] p<0.001 vs other blood fibrosis tests
[b] p<0.001 vs blood fibrosis tests excepted FibroMeter^{V2G} (p=0.121)
[c] p≤0.004 vs blood fibrosis tests

Table s3: Baseline characteristics of the 626 NAFLD patients included in the longitudinal cohort

Age (year)	59.7 ± 14.4
Male sex (%)	66.2
AST (IU/l)	41 ± 34
ALT (IU/l)	51 ± 45
GammaGT (IU/l)	152 ± 203
Total bilirubin (µmol/l)	12 ± 15
Prothrombin time (%)	91 ± 20
Platelets (G/l)	231 ± 87
LSM median (kPa)	13.4 ± 16.1
Follow-up duration (year):	6.3 (5.0 - 7.8) [a]
All-cause death during follow-up (n)	93
Liver-related death during follow-up (n)	27

[a] Result is expressed as median with 1st and 3rd quartile into brackets

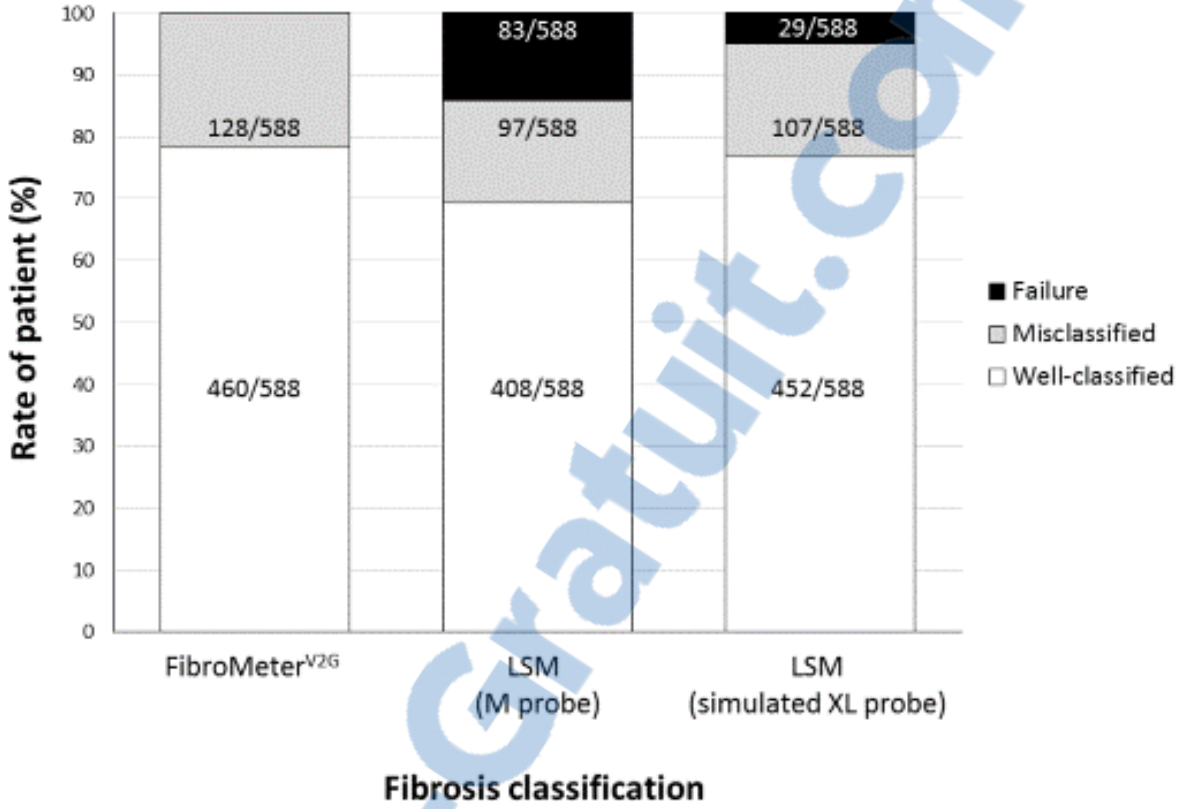

Figure s1: Diagnostic accuracy of FibroMeter^{V2G} and Fibroscan (M and XL probe) fibrosis classifications by taking into account measurement failure in an intention-to-diagnose basis in the 588 included patients. The results for XL probe were simulated from those of the M probe considering the following hypotheses: same diagnostic accuracy than the M probe and a 5% rate of measurement failure.

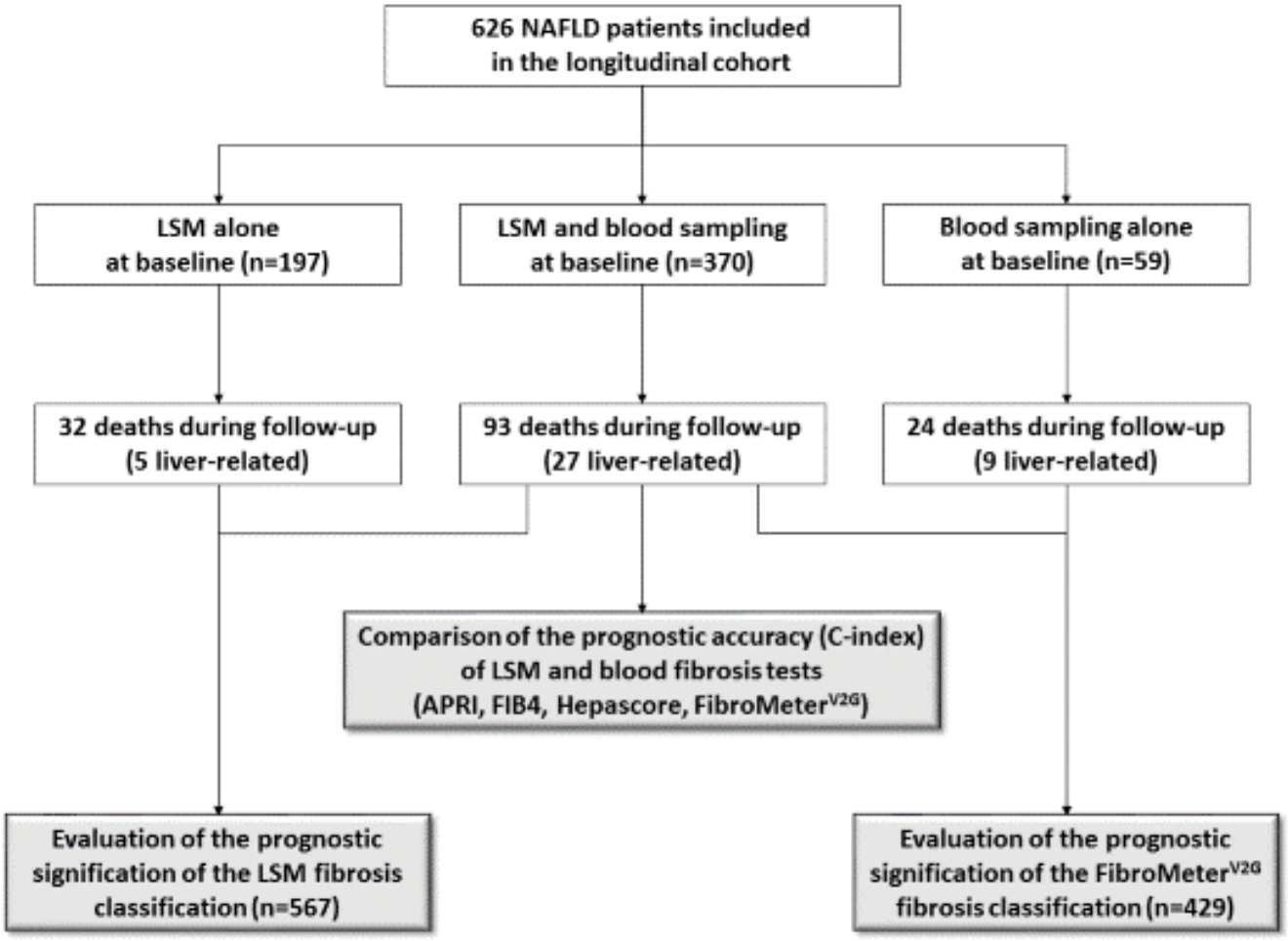

Figure s2: Flow chart of the longitudinal study. LSM: liver stiffness measurement by Fibroscan.

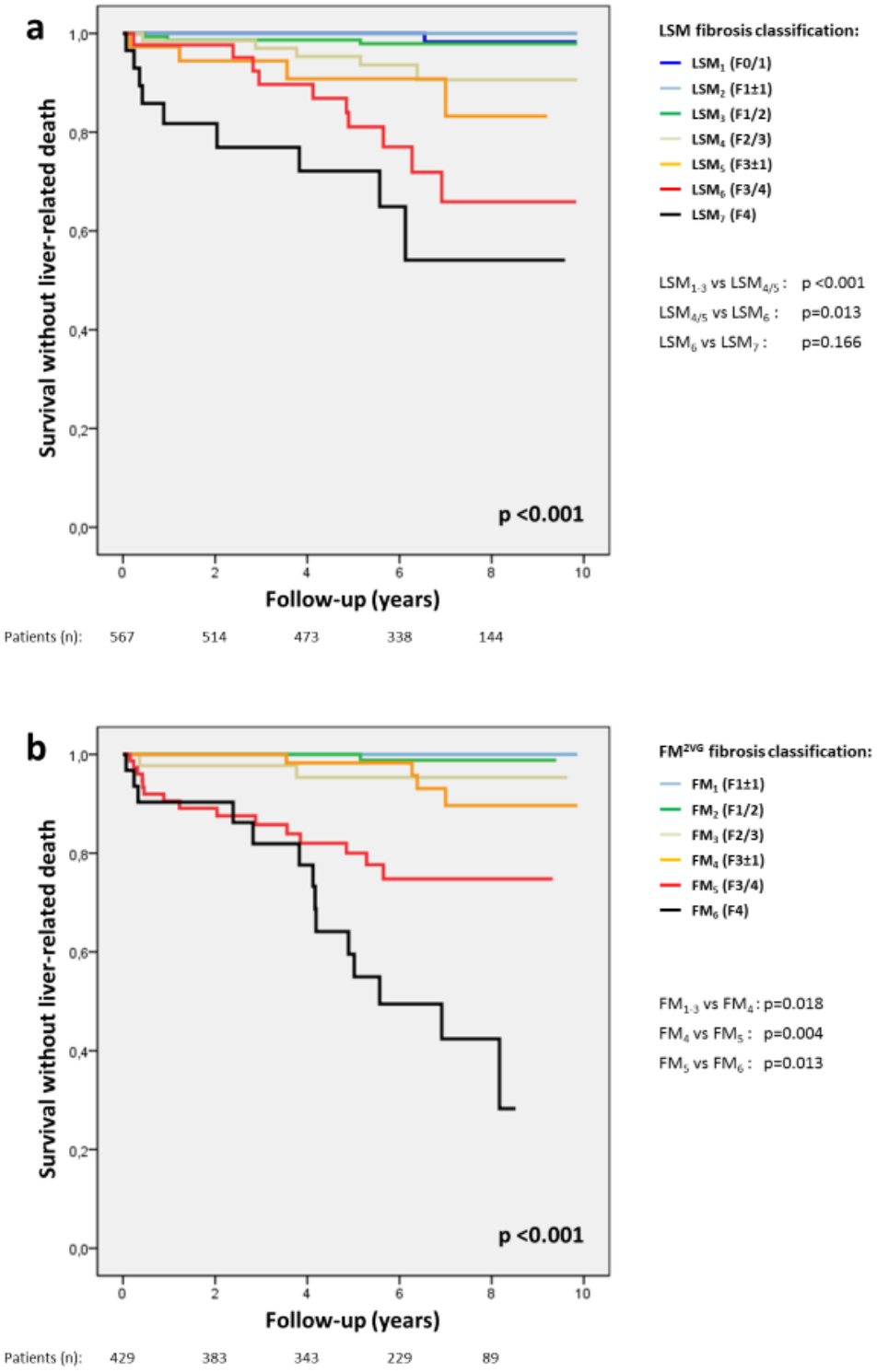

Figure s3: Survival without liver-related death as a function of the subgroups defined by the fibrosis classification developed for liver stiffness measurement by Fibroscan, (***panel s3a***) or the fibrosis classification developed for FibroMeter^{V2G} (***panel s3b***).

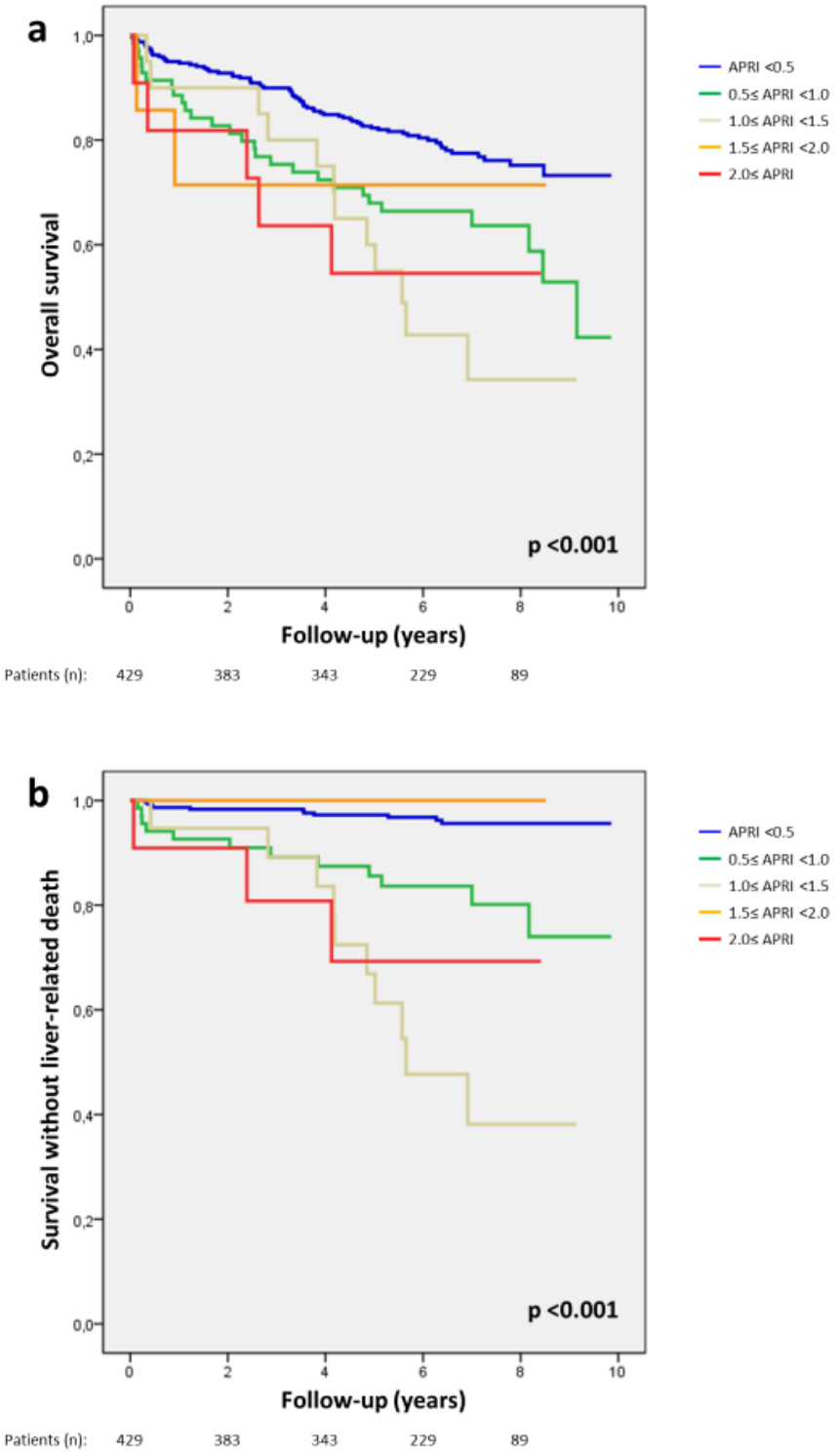

Figure s4: Overall survival (***panel s4a***) and survival without liver-related death (***panel s4b***) as a function of patient subgroups defined by the previously published diagnostic cut-offs for APRI (Wai, Hepatology 2003;38:518-26).

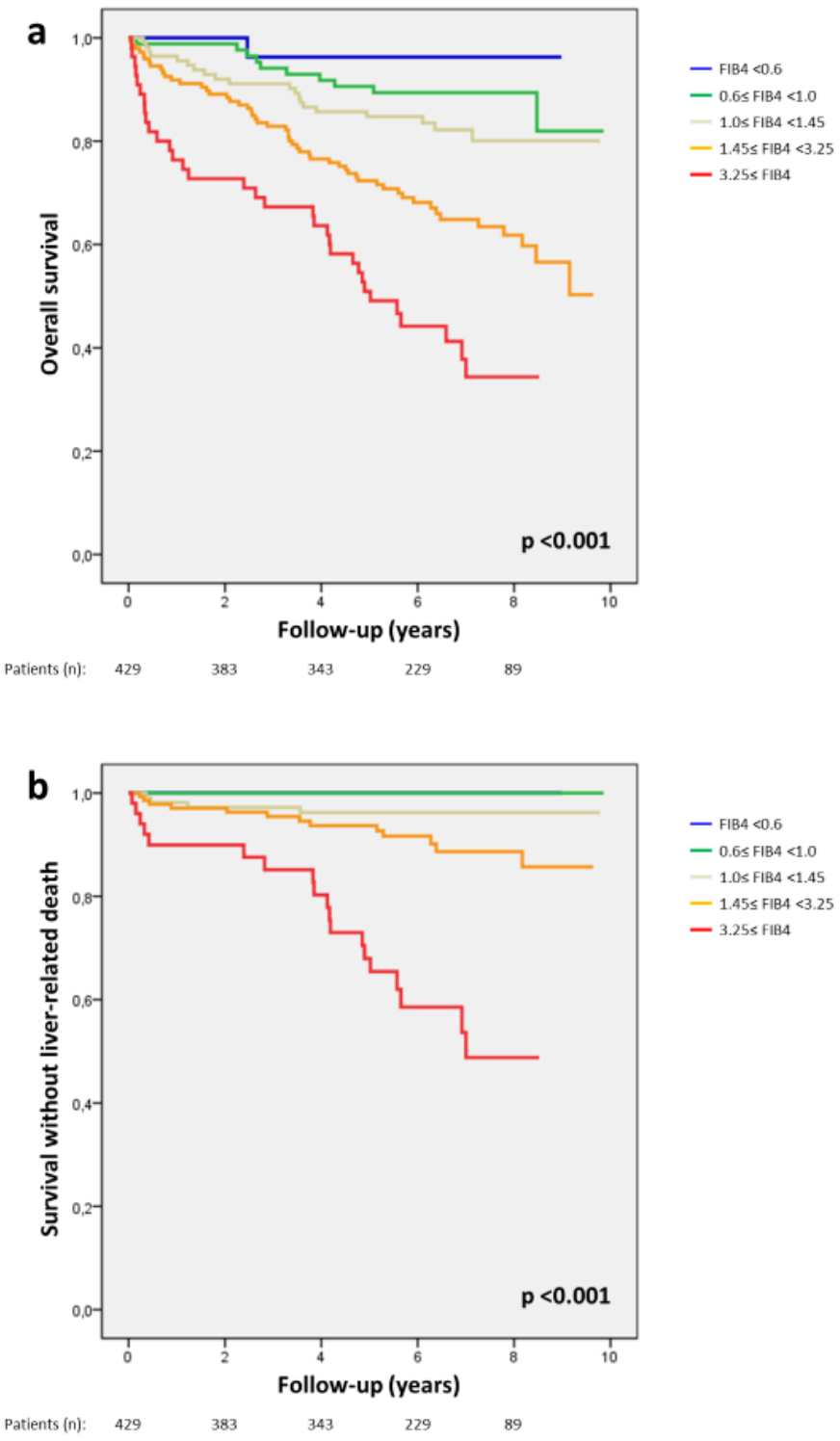

Figure s5: Overall survival (***panel s5a***) and survival without liver-related death (***panel s5b***) as a function of patient subgroups defined by the previously published diagnostic cut-offs for FIB4 (Sterling, Hepatology 2006;43:1317-25).

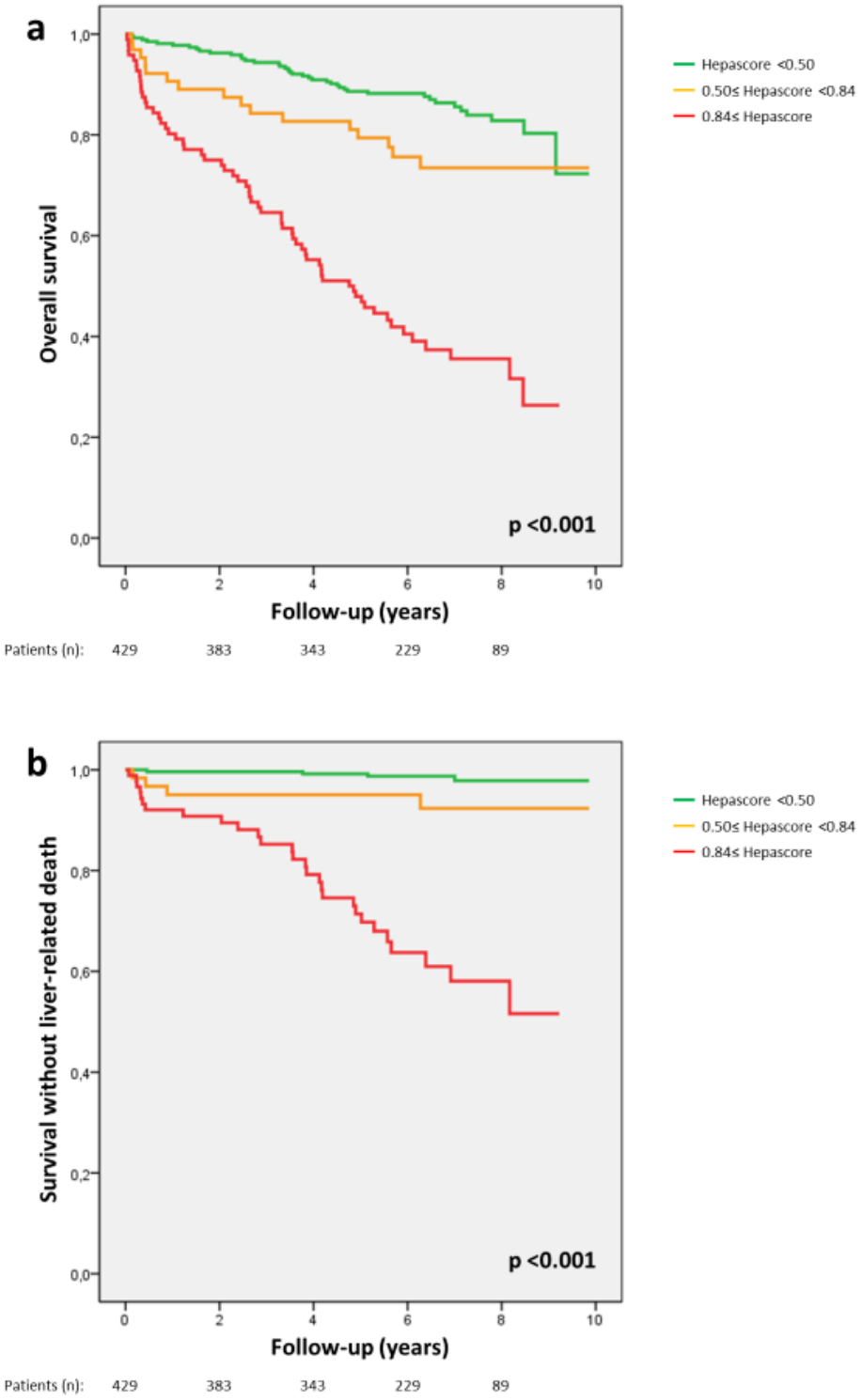

Figure s6: Overall survival (***panel s6a***) and survival without liver-related death (***panel s6b***) as a function of patient subgroups defined by the previously published diagnostic cut-offs for Hepascore (Adams, Clin Chem 2005;51:1867-73).

LANNES Adrien

PERFORMANCE DIAGNOSTIQUE ET PRONOSTIQUE DES TESTS SANGUINS DE FIBROSE ET DE LA MESURE DE L'ELASTICITE HEPATIQUE PAR FIBROSCAN DANS LA MALADIE HEPATIQUE STEATOSIQUE NON ALCOOLIQUE

RESUME

Objectif: La prévalence de la NAFLD est élevée, cependant peu de patients développent une fibrose hépatique sévère avec altération du pronostic hépatique. Notre but était de comparer la performance de la mesure de l'élasticité hépatique par Fibroscan et par des tests sanguins de fibrose, pour le diagnostic de fibrose hépatique et le pronostic dans la NAFLD.

Matériel et méthodes : La performance diagnostique a été évaluée dans une étude transversale incluant 588 patients NAFLD avec biopsies hépatique (stade de fibrose NASH-CRN), Fibroscan et 8 tests sanguins de fibrose (BARD, NAFLD Fibrosis Score, FibroMeter^{S}, APRI, FIB4, Fibrotest, Hepascore, FibroMeter^{V2G}). La performance pronostique a été évaluée dans une étude longitudinale incluant 626 patients NAFLD.

Résultats : Le Fibroscan et le FibroMeter^{V2G} étaient les 2 tests les plus performants dans l'étude transversale. L'AUROC pour la fibrose avancée F3/4 était, respectivement : 0.831±0.019 et 0.817±0.020 ($p \leq 0.041$ vs autres tests), la proportion de patients avec une valeur prédictive positive et négative ≥ 90% pour F3/F4 : 56.4% et 46.7% ($p < 0.001$ vs autres tests), et l'index d'Obuchowski : 0.834±0.014 et 0.798±0.016 ($p \leq 0.036$ vs autres tests). Deux classifications de fibrose ont été développées pour estimer précisément les stades histologiques de fibrose à partir des résultats du Fibroscan ou du FibroMeter^{V2G} sans avoir recours à une biopsie hépatique (performance diagnostique, respectivement : 80.8% vs 77.4%, p=0.190). L'étude longitudinale a montré que les classifications de fibrose du Fibroscan ou du FibroMeter^{V2G}, permettaient de catégoriser les patients NAFLD en plusieurs sous-groupes ayant des pronostics différents ($p < 0.001$).

Conclusion: Parmi les 9 tests de fibrose évalués, le Fibroscan et le FibroMeter^{V2G} sont les plus performants pour le diagnostic non invasif de la fibrose hépatique dans la NAFLD. La classification de fibrose du Fibroscan et du FibroMeter^{V2G} aide le praticien en pratique clinique à estimer le stade de fibrose et le pronostic du patient.

MOTS-CLES

NAFLD, fibrose, tests sanguins, élastométrie hépatique, pronostic, diagnostic

FORMAT

Article[1] **:** à soumettre

Suivi par : Docteur BOURSIER

[1] statut au moment de la soutenance