Prognostic performance of 7 biomarkers compared to liver biopsy in early alcohol-related liver disease

Ditlev Nytoft Rasmussen,[1,2] Maja Thiele,[1,2,*] Stine Johansen,[1,2] Maria Kjærgaard,[1] Katrine Prier Lindvig,[1,2] Mads Israelsen,[1] Steen Antonsen,[2,3] Sönke Detlefsen,[2,4] Aleksander Krag,[1,2] , and MicrobLiver consortia

[1]Department of Gastroenterology and Hepatology, Odense University Hospital, Denmark

[2]Department of Clinical Research, Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark

[3]Department of Clinical Biochemistry, Odense University Hospital Svendborg, Denmark

[4]Department of Pathology, Odense University Hospital, Denmark

Maja Thiele: maja.thiele@rsyd.dk

*Corresponding author. Address: Center for Liver Research, Kloevervaenget 10, entrance 112, 11. floor, Odense University Hospital, Denmark. Tel.: +45 24998068. maja.thiele@rsyd.dk

Abstract

Background & Aims

Alcohol is the most common cause of liver-related mortality and morbidity. We therefore aimed to assess and compare the prognostic performance of elastography and blood-based markers to predict time to the first liver-related event, severe infection, and all-cause mortality in patients with a history of excess drinking.

Methods

We performed a prospective cohort study in patients with early, compensated alcohol-related liver disease. At baseline, we obtained a liver biopsy, transient elastography (TE), 2-dimensional shear-wave elastography (2D-SWE), enhanced liver fibrosis test (ELF), FibroTest, fibrosis-4 index (FIB-4), non-alcoholic fatty liver fibrosis score (NFS) and Forns index. We compared C-statistics and time-dependent AUC for prognostication. We used validated cut-off points to create 3 risk groups for each test: low, intermediate and high risk.

Results

We followed 462 patients for a median of 49 months (IQR 31–70). Median age was 57 years, 76% were males, 20% had advanced fibrosis. Eighty-four patients (18%) developed a liver-related event after a median of 18 months (7-34). TE had the highest prognostic accuracy, with a C-statistic of 0.876, and time-dependent AUC at 5 years of 0.889, comparable to 2D-SWE and ELF. TE, ELF and 2D-SWE outperformed FibroTest, FIB4, NFS, Forns index and biopsy-verified fibrosis stage. Compared to patients with TE <10 kPa, the hazard ratios for liver-related events for TE 10–15 kPa were 8.1 (3.2–20.4), and 27.9 (13.8–56.8) for TE >15 kPa. Periods of excessive drinking during follow-up increased the risk of progressing to liver-related events, except for patients in the low-risk groups.

Conclusion

TE, ELF and 2D-SWE are highly accurate prognostic markers in patients with alcohol-related liver disease. Easy-to-use cut-offs can distinguish between substantially different risk profiles.

Lay summary

Alcohol is the leading cause of death and illness due to liver disease. In this study, we assessed the ability of biomarkers to predict the risk of developing symptomatic liver disease in patients with early stages of alcohol-related liver disease. We found that several tests accurately predicted the risk of liver-related events such as ascites, esophageal varices and hepatic encephalopathy during an average follow-up of 4.1 years. Liver stiffness measurements by ultrasound elastography and the enhanced liver fibrosis test performed best. By using them, we were able to stratify patients into 3 groups with significantly different risks.

Graphical abstract
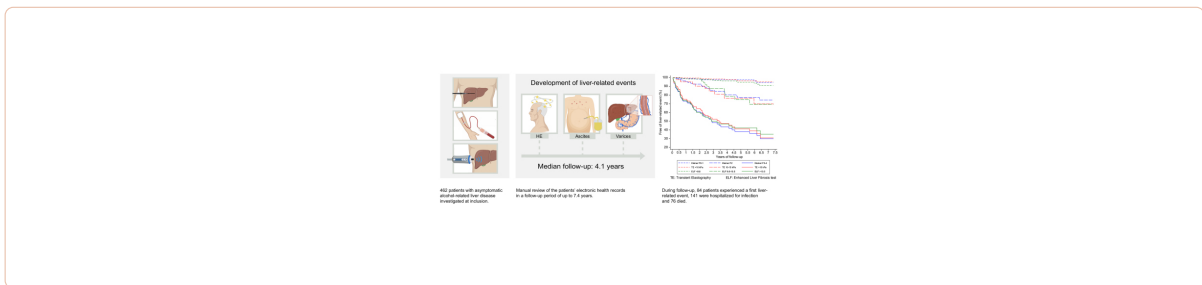
Introduction

Alcohol-related liver disease (ALD) is on the rise, with an annual growth exceeding 3% in the United States throughout the past decade.[1] More than ever, ALD remains the major cause of liver-related mortality, accounting for more than half of all deaths associated with liver disease.[1] In an effort to improve early detection of advanced fibrosis in ALD, the diagnostic utility of non-invasive tests, such as elastography and blood-based tests, has been thoroughly investigated and validated in recent years.[2,3] However, to improve risk stratification, prognostic markers are needed for patients with ALD, before development of decompensation. The presence of advanced fibrosis is an independent predictor of mortality in compensated ALD.[4] While non-invasive markers of fibrosis predict mortality and morbidity in non-alcoholic fatty liver disease, there is scarce evidence of a prognostic potential in ALD.[5] This gap in knowledge includes the most widely used non-invasive test, transient elastography (TE), particularly in comparison with competing imaging and blood-based biomarkers, patented or non-patented.

We hypothesized that non-invasive fibrosis tests may provide long-term prognostic information in patients with early, compensated ALD, and that the prognostic accuracy of non-invasive tests could potentially outperform biopsy-verified liver fibrosis stage. Our primary aim was to evaluate the prognostic ability of TE, 2-dimensional shear-wave elastography (2D-SWE), FibroTest, enhanced liver fibrosis test (ELF), fibrosis-4 index (FIB-4), non-alcoholic fatty liver fibrosis score (NFS), and Forns index, compared to biopsy-verified Kleiner fibrosis stage, to predict time to the first liver-related event in a cohort of patients with ALD without symptoms of decompensated cirrhosis. Secondary outcomes were infections requiring hospitalizations and all-cause mortality. We also assessed the ability of the tests to stratify patients into groups at low, moderate and high risk of liver-related events, according to previously reported cut-off points for each test. Finally, we evaluated the robustness of the prognostic estimates in subgroup and sensitivity analyses accounting for alcohol at inclusion and during follow-up, and other known risk factors.

Materials and methods

Study design

This is a single-center, biopsy-controlled, prospective, prognostic study. The Danish Data Protection Agency (13/8204) and the Ethics Committee for the Region of Southern Denmark (S-20120071, S-20160021) approved the study. The study is registered with Odense Patient data Explorative Network under study ID OP_40. The study adheres to the TRIPOD checklist for transparent reporting of prediction models (Table S11). All patients signed a consent form after written and oral information and the study conformed to the Declaration of Helsinki.

Patients

The patients were included as part of a previously published study of diagnostic markers in ALD.[2,6] We included patients from 3 sources: secondary care patients who were referred to 3 outpatient liver clinics, primary care patients recruited from 2 municipal alcohol rehabilitation centers, and patients who responded to our online advertisement. The inclusion criteria were age 18–75 years, prior or current chronic alcohol overuse defined as >24 g/d for women and >36 g/d for men for >1 year, and informed consent to a research liver biopsy. Exclusion criteria were: evidence of decompensated cirrhosis (clinically obvious ascites, overt hepatic encephalopathy, large esophageal varices with or without variceal bleeding); concurrent liver disease other than ALD (chronic viral hepatitis, autoimmune liver and bile duct diseases, or heritable diseases related to deposition of iron, copper or α-1-antitrypsin); cancer or other debilitating diseases with an expected survival <12 months; severe alcoholic hepatitis scored by Glasgow Alcoholic Hepatitis Score; hepatic congestion or bile duct dilation evidenced by ultrasound; and contraindication to percutaneous liver biopsy.

Non-invasive tests and liver biopsy

We performed all investigations and sampled blood for biobanking at inclusion within 1 week. Investigations included 7 non-invasive tests: TE, 2D-SWE, ELF, FibroTest, FIB-4, NFS, and Forns index. We thawed stored, frozen serum once for ELF and FibroTest measurements. Experienced nurse operators performed TE with the FibroScan 502 Touch (Echosens, France) while experienced ultrasound operators performed 2D-SWE with Aixplorer (Supersonic Imagine, France) using standard operating procedures and quality criteria.[7] We measured the ELF test on an Advia Centaur XP (Siemens Healthcare, Erlangen, Germany), and FibroTest (BioPredictive, Paris, France) according to the manufacturer's instructions. We calculated FIB-4, NFS and Forns index from routine blood tests, in accordance with published equations (Table S10).[8], [9], [10] Finally, we performed a percutaneous suction needle liver biopsy in the same intercostal space as the elastography (17 G Menghini needle; Hepafix, Braun, Germany). An experienced liver pathologist assigned Kleiner fibrosis stage while blinded to all other study data (previously reported, Kappa 0.71 for intra-observer variance).[11] We considered a biopsy to be of adequate quality if it had a length of at least 10 mm and at least 6 portal tracts, or presence of regeneration nodules. In 2016 we changed the study protocol, since no patient with TE <6.0 kPa had advanced fibrosis. We therefore refrained from a liver biopsy in 96 patients with TE below 6.0 kPa. These patients were not used in analyses of the prognostic ability of fibrosis stage but contributed to the analyses of non-invasive markers.

Risk groups

For each of the 7 non-invasive tests of fibrosis, we used validated cut-offs to stratify patients into 3 risk groups for comparisons between no or minimal fibrosis (F0-1), moderate fibrosis (F2), and severe fibrosis or cirrhosis (F3-4). For TE, we labelled <10.0 kPa as low risk, 10.0–15.0 kPa as intermediate risk, and >15.0 kPa as high risk, in accordance with the Baveno VI guideline.[12,13] For 2D-SWE, we used 10.0 kPa and 16.4 kPa, similar to the optimal cut-offs for diagnosing significant fibrosis (≥F2) and advanced fibrosis (≥F3).[2,6] For ELF, we used 9.8 and 10.5 as cut-offs, according to published literature.[2,14,15] For FibroTest, we used manufacturer-recommended thresholds of 0.31 and 0.58.[16,17] For FIB-4 and NFS we used different cut-offs for people below or above the age of 65, as previously suggested.[18] For FIB-4, the low cut-off points were 1.3 (age ≤65 years) and 2.0 (age >65 years), and the high cut-off point was 2.67 (for all). For NFS, the cut-off points were -1.455 (age ≤65 years) or 0.12 (age >65 years), and 0.76 (for all). For Forns index we used 4.2 and 6.9 as suggested in a recent population-based study.[19]

Follow-up and outcome assessment

We manually reviewed the participants' electronic health records for every encounter with the health care system in the Region of Southern Denmark. Follow-up started on inclusion and ended if the patients died, moved out of the region, or by October 2020.

We defined liver-related events following guideline definitions as occurrence of any of the following: Alcoholic hepatitis, varices needing treatment, variceal bleeding, ascites, spontaneous bacterial peritonitis, hepatic encephalopathy (HE), hepatocellular carcinoma (HCC), hepatorenal syndrome, upper gastrointestinal bleeding, or jaundice due to liver failure (see supplementary information for detailed definitions). We chose these liver-related events because of their negative consequences for the patient in terms of need for medical care/interventions, reduced quality of life, and/or hospitalization. Secondary outcomes were infection requiring hospitalization, and all-cause mortality.

Subgroup and sensitivity analyses

We investigated whether the prognostic estimates changed in subgroup analyses of patients stratified according to episodes of excessive drinking during follow-up, type 2 diabetes, obesity defined as a BMI $\geq$30 kg/m^2, age $\geq$65 years, and smoking. We also tested whether the non-invasive tests independently predicted liver-related events in multivariable Cox regression analyses adjusting for age, sex, BMI, drinking at inclusion and during follow-up, fibrosis stage, and presence of steatohepatitis (ballooning, lobular inflammation and steatosis). We conducted sensitivity analyses for prognostic accuracies using a more restrictive endpoint of decompensation, defined as overt HE, variceal bleeding, ascites requiring paracentesis, or jaundice. Finally, we analyzed the robustness of our findings for the primary endpoint in analyses using multiple imputation of missing values and competing risk Cox regression with all-cause mortality as a competing event.

Statistics

We report quantitative data as medians (IQR), and categorical data as counts and percentages, with Wilcoxon rank sum or Fischer's exact test for difference. We used Cox regression with Harrell's C to assess overall prognostic ability.[20] Harrell's C of 1.0 denotes perfect ability to predict the order of occurrence of events, whereas 0.5 is a random prediction. We also assessed the ability of each of the tests to predict outcomes at 5 years of follow-up using time-dependent AUC statistics. We used the log rank test to test for significant differences between the low-, intermediate- and high-risk groups, and we illustrated the time-to-event in these risk groups with Kaplan-Meier plots. If people died or moved out of the country, we treated it as random censoring, and in case of missing values, we made complete case analyses. We used STATA 16 (Statacorp TX, US) for statistical analyses and considered p <0.05 as statistically significant.

Results

Patients

We included 462 consecutive patients between April 2013 and September 2018, 222 from secondary care, 143 from alcohol rehabilitation centers, and 97 from online advertisement. Median age at inclusion was 57 years, 351 patients (76%) were men, and 94 (20%) had advanced fibrosis ($\geq$F3) at inclusion (Table 1).

Follow-up

During a median follow-up of 49 months (IQR 31–70) and 1,878 person years, 84 (18%) patients experienced a first liver-related event, 141 (31%) were hospitalized for infection and 76 (16%) died. Ten patients were lost to follow-up. The first liver-related event occurred after a median of 18 months (range 1–77, IQR 7–34). Patients who experienced a liver-related event were on average more likely to have type 2 diabetes and less likely to be obese (Table 1). All the investigated non-invasive tests, liver fibrosis stage and presence of steatohepatitis showed higher results in patients with liver-related events (Table 1). Ascites, covert HE and varices needing treatment were the most frequent liver-related events with 31, 15

and 15 cases, respectively (Table S1).

Prognostic markers of liver-related events

TE had the highest C-statistic for prediction of liver-related events with a value of 0.876 (0.833–0.918), comparable to 2D-SWE and ELF with values at 0.868 (0.832–0.903) and 0.859 (0.822–0.897) (Table 2, Fig. 1). TE, 2D-SWE and ELF significantly outperformed all other non-invasive tests (Table S2).

Similarly, TE, ELF and 2D-SWE had the highest AUCs for prediction of liver-related events over time (Fig. 2; integrated AUCs in Tables S3-S4 and Fig. S1). FibroTest, Forns index, NFS and FIB-4 had good prognostic accuracies, with C-statistics around 0.80 and no statistically significant differences.

Three risk groups for prediction of liver-related events

For 4 of the non-invasive tests (TE, 2D-SWE, ELF and FibroTest), the pre-selected cut-off points divided patients into 3 significantly different risk groups with low, intermediate and high hazard ratios for liver-related events (Table 3, Fig. 3), while hazard ratios for Forns, NFS and FIB-4 did not statistically differ between low- and intermediate-risk groups (Fig. S2). Patients with TE between 10 and 15 kPa in the intermediate-risk group had an 8-fold higher likelihood of liver-related events than the low-risk patients with TE below 10 kPa (hazard ratio 8.07; 95% CI 3.2–20.4), while those with TE above 15 kPa had a 28-fold higher likelihood of liver-related events than low-risk patients (hazard ratio 28.0; 95% CI 13.8–56.8). Of the 443 patients with a reliable TE measurement, 303 (68%) were in the low-risk group with TE <10 kPa, and 9 of those (3%) developed a liver-related event. Forty-two patients with ALD (9%) were in the intermediate group, of whom 9 (21%) experienced an event, and 98 (22%) were in the high-risk group, of whom 53 (54%) experienced an event.

The prognostic performance of fibrosis stage compared to non-invasive tests

Fibrosis stage had C-statistics of 0.819 for prediction of liver-related events, and a time-dependent 5-year AUC of 0.786 (Table 2, Fig. 2). Consequently, fibrosis stage had a lower prognostic accuracy than TE, 2D-SWE and ELF, albeit only significantly so for TE (Tables S2-4). Fibrosis stage is however an ordinal outcome, and therefore at a disadvantage compared to the continuous non-invasive markers in C-statistic and ROC curve analyses. When dividing into 3 risk groups, fibrosis stage performed similarly to TE, 2D-SWE, and ELF (Table 3, Fig. 3).

Subgroup and sensitivity analyses

Excessive drinking was the most prominent predictor of liver-related events in subgroup analyses, independent of the non-invasive tests and of fibrosis stage. We were able to obtain data on alcohol drinking pattern during follow-up in 405 patients. Sixty-seven (80%) of those who developed liver-related events drank excessively on at least 1 occasion during follow-up, compared to 198 (62%) of those who did not develop events, resulting in a hazard ratio of 2.20 (1.29–3.74, $p = 0.004$) in univariable Cox regression (Fig. 4).

Controlling for excessive drinking did not change the C-statistics of any of the tests significantly. However, when using TE, ELF, and fibrosis stage to divide patients into 3 risk groups, those in the intermediate- and high-risk groups with excessive drinking during follow-up had significantly higher hazard ratios for liver-related events compared to patients without drinking episodes (Fig. 4). For the other subgroup analyses, we observed only minor changes to Harrell's C when stratifying patients according to obesity, smoking, age ≥65 years, and type 2 diabetes. Similarly, all of the prognostic tests independently predicted liver-related events in a multivariable Cox regression analysis adjusted for age, sex, BMI, drinking at inclusion, drinking during follow-up, fibrosis stage, and presence of steatohepatitis. Steatohepatitis predicted liver-related events in univariable analysis, but the correlation disappeared when controlling for fibrosis stage. Adjusting hazard ratios for the 3 risk groups also led to only slight changes (Tables S6-7). Finally, our results only changed minimally when we used a more restrictive endpoint of decompensation (Tables S8-9), and when using multiple imputation of missing values and competing risk Cox regression with all-cause mortality as a competing event (not reported).

All-cause mortality and infections requiring hospitalization

Six of the 7 non-invasive tests predicted all-cause mortality with moderate accuracy and C-statistics at 0.7 or above. The exception was NFS, with a C-statistic of 0.66 (Table 2). For infections requiring hospitalization, we found poor C-statistics in the range of 0.594 to 0.677, with TE, 2D-SWE and ELF significantly outperforming Forns index, FIB-4 and NFS (Table 2, Table S5).

Discussion

In this prospective, single-etiology cohort of 462 patients with biopsy-proven ALD and up to 7 years of follow-up, we found that elastography and the ELF test predict liver-related events with excellent prognostic accuracy, superior to fibrosis stage. Other, cheaper, blood-based tests are valid alternatives with good prognostic accuracies. Our findings suggest that widely known cut-offs for TE can be used to separate patients into 3 groups of distinctly different risks profiles: Compared to patients with a liver stiffness below 10 kPa, patients with liver stiffness between 10 and 15 kPa had an 8-fold higher hazard for liver-related events, and those with liver stiffness >15 kPa had a 28-fold higher hazard. Our results furthermore highlight the importance of excessive drinking, which significantly worsened prognosis in patients at high and moderate risk of progressing to liver-related events. Finally, our data indicate that liver biopsy can be substituted with non-invasive testing for prognostication purposes.

This is the first study to comprehensively report on the prognostic ability of several non-invasive tests to predict liver-related events, all-cause mortality, and infections requiring hospitalizations in a population of patients with biopsy-verified ALD at mostly early stages of fibrosis. By ascertainment of events from manual review of the electronic health records of every study participant, we were able to obtain a detailed follow-up with a negligible number of losses to follow-up and extensive subgroup analyses.

One prior study reported the ability of non-invasive liver tests to predict mortality in 218 patients with ALD.[21] However, this study only reported on serum markers, assessed death as the only outcome, and included almost one-third with cirrhosis at baseline, not clearly excluding those with evidence of decompensated disease. We did however confirm their reported cut-off points for FibroTest of 0.31 and 0.58, since the same cut-offs in our cohort reliably stratified patients at low, moderate and high risk of liver-related events.[21]

Our study validates the use of 7 non-invasive fibrosis markers for prognostication in ALD. We thereby expand the findings of other studies, which have investigated the same markers for predicting clinical outcomes in patients with liver fibrosis of other etiologies.[22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32] For example, we find a similar prognostic performance of ELF and FibroTest, as others have reported in patients with chronic hepatitis C.[25,29] And the 9.8 and 10.5 ELF cut-off points to assign patients into 3 risk groups with distinct hazard ratios are comparable to cut-off points reported in studies of chronic hepatitis B and mixed-etiology patients.[30,31]

One study found that TE below 20 kPa excluded decompensation, HCC or liver-related death in a population of mixed etiology, including 65% with viral hepatitis.[33] This was not the case in our study, where 18 patients with TE below 15 kPa developed decompensation. In general, our study demonstrated a surprisingly high event rate for patients with moderate fibrosis (F2) at baseline and/or test results in the intermediate risk groups. This probably reflects the well-known excess morbidity and faster progression rate in ALD compared to other etiologies like non-alcoholic fatty liver disease.[34] In line with this, episodes of excessive alcohol use during follow-up worsened the hazard ratios for liver-related events in intermediate- and high-risk groups, but not in the low-risk groups.

In a large Swedish registry study in patients from the general population, FIB-4, Forns index and NFS accurately stratified the population into 3 groups with different risk of liver-related outcomes.[19] Our data confirms this finding in a clinical cohort. However, while FIB-4, Forns index and NFS may provide a robust estimate of the risk for liver-related events, their prognostic abilities are lower than ELF and elastography, making them less suited for individual risk stratification.

Our study's unique strength is the head-to-head comparison of both elastography, patented and non-

patented serum markers, and the traditional gold standard of liver biopsy. We evaluated well-established cut-off points to make the prognostic predictions more manageable for the clinician, and thereby easier to incorporate into everyday patient care. By validating previously reported cut-off points, we also avoided cut-off points that were fitted to our data, which would lead to over-optimistic prognostic accuracies. Another strength is the number of patients and the length of follow-up, which enabled us to study the prognosis of liver disease in a population of patients where decompensating events take longer to develop and are relatively infrequent.

In comparison to population-based studies, however, our study cohort is small and the follow-up period short. This particularly affects estimates for mortality. Yet, 16% of patients died during a median follow-up of 4.1 years, underlining the severity of ALD, even for patients from outpatient clinics and primary care who were free of decompensated cirrhosis and had low levels of inflammatory activity. Another limitation is the lack of repeated testing to evaluate the role of changes from baseline of the non-invasive markers. This is in contrast to most real-world settings, where patients are repeatedly monitored. Our data indicates that especially those in the intermediate-risk groups could benefit from a monitoring program.

The non-invasive markers of fibrosis predicted all-cause mortality and severe infections with poorer prognostic accuracy than liver-related events and decompensation. This is probably because patients have high competing risks of alcohol-associated extrahepatic disease. For example, harmful use of alcohol alone increases the risk of severe infections.

In conclusion, our study endorses the use of elastography and the ELF test as highly accurate prognostic markers in patients with early stages of alcohol-related liver fibrosis or compensated cirrhosis, with easy-to-use cut-offs to distinguish between patients at substantially different risks of liver-related events.

Abbreviations

2D-SWE, 2-dimensional shear-wave elastography; ALD, alcohol-related liver disease; ELF, enhanced liver fibrosis test; FIB-4, fibrosis-4 index; HCC, hepatocellular carcinoma; HE, hepatic encephalopathy; HR, hazard ratio; NFS, non-alcoholic fatty liver fibrosis score; TE, transient elastography.

Financial support

Authors' contributions

AK and MT designed the study; DNR, MT, MK, KPL, SA, SD, SJ, MI collected the data; DNR, SJ and MT performed the data analyses; DNR and MT drafted the manuscript; AK and MT secured funding for the study; all authors critically revised the manuscript and approved of the final version.

Data availability statement

The full dataset is available by contact to open@rsyd.dk, after approval from the Danish Data Protection Agency.

GALAXY consortium

Ema Anastasiadou, Manimozhiyan Arumugam, Peer Bork, Torben Hansen, Christina Hartoft, Hans Israelsen, Morten Karsdal, Cristina Legido-Quigley, Hans Olav Melberg, Maja Thiele, Jonel Trebicka,

Aleksander Krag (coordinator).

MicrobLiver consortium

Peer Bork, Mathias Mann, Jelle Matthijnssens, Aleksander Krag, Torben Hansen (coordinator).

Conflicts of interest

Acknowledgements

Footnotes

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2021.05.037.

Contributor Information

, on behalf of the GALAXY:

Ema Anastasiadou, Manimozhian Arumugam, Peer Bork, Torben Hansen, Christina Hartoft, Hans Israelsen, Morten Karsdal, Cristina Legido-Quigley, Hans Olav Melberg, Maja Thiele, Jonel Trebicka, and Aleksander Krag

MicrobLiver consortia:

Peer Bork, Mathias Mann, Jelle Matthijnssens, Aleksander Krag, and Torben Hansen

Supplementary data

The following are the supplementary data to this article:

> **Multimedia component 1:**

> **Multimedia component 2:**

> **Multimedia component 3:**

References

1. Kim D., Li A.A., Gadiparthi C., Khan M.A., Cholankeril G., Glenn J.S. Changing trends in etiology-based annual mortality from chronic liver disease, from 2007 through 2016. *Gastroenterology.* 2018;**155**:1154–1163. [PMCID: PMC6467699] [PubMed: 30009816]

2. Thiele M., Madsen B.S., Hansen J.F., Detlefsen S., Antonsen S., Krag A. Accuracy of the enhanced liver fibrosis test vs fibrotest, elastography and indirect markers in detection of advanced fibrosis in patients with alcoholic liver disease. *Gastroenterology.* 2018;**154**:1369–1379. [PubMed: 29317276]

3. Nguyen-Khac E., Thiele M., Voican C., Nahon P., Moreno C., Boursier J. Non-invasive diagnosis of liver fibrosis in patients with alcohol-related liver disease by transient elastography: an individual patient data meta-analysis. *Lancet Gastroenterol Hepatol.* 2018;**3**:614–625. [PubMed: 29983372]

4. Lackner C., Spindelboeck W., Haybaeck J., Douschan P., Rainer F., Terracciano L. Histological parameters and alcohol abstinence determine long-term prognosis in patients with alcoholic liver disease. *J Hepatol.* 2017;**66**:610–618. [PubMed: 27894795]

5. Rhodes F.A., Trembling P., Panovska-Griffiths J., Tanwar S., Westbrook R.H., Rodger A. Systematic review: investigating the prognostic performance of four non-invasive tests in alcohol-related liver disease. *J Gastroenterol Hepatol.* 2020 [PubMed: 33171534]

6. Thiele M., Detlefsen S., Møller L., Madsen B.S., Hansen J.F., Fialla A.D. Transient and 2-dimensional shear-wave elastography provide comparable assessment of alcoholic liver fibrosis and cirrhosis. *Gastroenterology.* 2016;**150**:123–133. [PubMed: 26435270]

7. Dietrich C.F., Bamber J., Berzigotti A., Bota S., Cantisani V., Castera L. EFSUMB guidelines and recommendations on the clinical use of liver ultrasound elastography, update 2017 (long version) *Ultraschall in der Medizin.* 2017;**38**:e16–e47. [PubMed: 28407655]

8. Sterling R.K., Lissen E., Clumeck N., Sola R., Correa M.C., Montaner J. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology.* 2006;**43**:1317–1325. [PubMed: 16729309]

9. Angulo P., Hui J.M., Marchesini G., Bugianesi E., George J., Farrell G.C. The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. *Hepatology.* 2007;**45**:846–854. [PubMed: 17393509]

10. Forns X., Ampurdanes S., Llovet J.M., Aponte J., Quinto L., Martinez-Bauer E. Identification of chronic hepatitis C patients without hepatic fibrosis by a simple predictive model. *Hepatology.* 2002;**36**:986–992. [PubMed: 12297848]

11. Brunt E.M., Kleiner D.E., Wilson L.A., Belt P., Neuschwander-Tetri B.A. Nonalcoholic fatty liver disease (NAFLD) activity score and the histopathologic diagnosis in NAFLD: distinct clinicopathologic meanings. *Hepatology.* 2011;**53**:810–820. [PMCID: PMC3079483] [PubMed: 21319198]

12. de Franchis R. Expanding consensus in portal hypertension: report of the Baveno VI Consensus Workshop: stratifying risk and individualizing care for portal hypertension. *J Hepatol.* 2015;**63**:743–752. [PubMed: 26047908]

13. Papatheodoridi M., Hiriart J.B., Lupsor-Platon M., Bronte F., Boursier J., Elshaarawy O. Refining the Baveno VI elastography criteria for the definition of compensated advanced chronic liver disease. *J Hepatol.* 2021;**74**(5):1109–1116. doi: 10.1016/j.jhep.2020.11.050. [PubMed: 33307138] [CrossRef: 10.1016/j.jhep.2020.11.050]

14. Glen J., Floros L., Day C., Pryke R., Guideline Development G. Non-alcoholic fatty liver disease (NAFLD): summary of NICE guidance. *BMJ.* 2016;**354**:i4428. [PubMed: 27605111]

15. Fagan K.J., Pretorius C.J., Horsfall L.U., Irvine K.M., Wilgen U., Choi K. ELF score >/=9.8 indicates advanced hepatic fibrosis and is influenced by age, steatosis and histological activity. *Liver Int.* 2015;**35**:1673–1681. [PubMed: 25495373]

16. Lemoine M., Assoumou L., De Wit S., Girard P.M., Valantin M.A., Katlama C. Diagnostic accuracy of noninvasive markers of steatosis, NASH, and liver fibrosis in HIV-monoinfected individuals at risk of nonalcoholic fatty liver disease (NAFLD): results from the ECHAM study. *J Acquir Immune Defic Syndr.* 2019;**80**:e86–e94. [PubMed: 30570529]

17. Ngo Y., Munteanu M., Messous D., Charlotte F., Imbert-Bismut F., Thabut D. A prospective analysis of the prognostic value of biomarkers (FibroTest) in patients with chronic hepatitis C. *Clin Chem.* 2006;**52**:1887–1896. [PubMed: 16931569]

18. McPherson S., Hardy T., Dufour J.F., Petta S., Romero-Gomez M., Allison M. Age as a confounding factor for the accurate non-invasive diagnosis of advanced NAFLD fibrosis. *Am J Gastroenterol.* 2017;**112**:740–751. [PMCID: PMC5418560] [PubMed: 27725647]

19. Hagstrom H., Talback M., Andreasson A., Walldius G., Hammar N. Ability of noninvasive scoring systems to identify individuals in the population at risk for severe liver disease. *Gastroenterology.* 2020;**158**:200–214. [PubMed: 31563624]

20. Harrell F.E., Jr., Califf R.M., Pryor D.B., Lee K.L., Rosati R.A. Evaluating the yield of medical tests. *Jama.* 1982;**247**:2543–2546. [PubMed: 7069920]

21. Naveau S., Gaudé G., Asnacios A., Agostini H., Abella A., Barri-Ova N. Diagnostic and prognostic values of noninvasive biomarkers of fibrosis in patients with alcoholic liver disease. *Hepatology.* 2009;**49**:97–105. [PubMed: 19053048]

22. Mayo M.J., Parkes J., Adams-Huet B., Combes B., Mills A.S., Markin R.S. Prediction of clinical outcomes in primary biliary cirrhosis by serum enhanced liver fibrosis assay. *Hepatology.* 2008;**48**:1549–1557. [PMCID: PMC2597274] [PubMed: 18846542]

23. Berenguer J., Zamora F.X., Aldamiz-Echevarria T., Von Wichmann M.A., Crespo M., Lopez-Aldeguer J. Comparison of the prognostic value of liver biopsy and FIB-4 index in patients coinfected with HIV and hepatitis C virus. *Clin Infect Dis.* 2015;**60**:950–958. [PubMed: 25422386]

24. Sebastiani G., Alshaalan R., Wong P., Rubino M., Salman A., Metrakos P. Prognostic value of non-invasive fibrosis and steatosis tools, hepatic venous pressure gradient (HVPG) and histology in nonalcoholic steatohepatitis. *PloS One.* 2015;**10** [PMCID: PMC4471234] [PubMed: 26083565]

25. Boursier J., Brochard C., Bertrais S., Michalak S., Gallois Y., Fouchard-Hubert I. 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. [PubMed: 24889599]

26. Grgurevic I., Bokun T., Mustapic S., Trkulja V., Heinzl R., Banic M. Real-time two-dimensional shear wave ultrasound elastography of the liver is a reliable predictor of clinical outcomes and the presence of esophageal varices in patients with compensated liver cirrhosis. *Croat Med J.* 2015;**56**:470–481. [PMCID: PMC4655932] [PubMed: 26526884]

27. Kim S.U., Lee J.H., Kim D.Y., Ahn S.H., Jung K.S., Choi E.H. Prediction of liver-related events using fibroscan in chronic hepatitis B patients showing advanced liver fibrosis. *PloS One.* 2012;**7** [PMCID: PMC3344942] [PubMed: 22574212]

28. Onnerhag K., Hartman H., Nilsson P.M., Lindgren S. Non-invasive fibrosis scoring systems can predict future metabolic complications and overall mortality in non-alcoholic fatty liver disease (NAFLD) *Scand J Gastroenterol.* 2019;**54**:328–334. [PubMed: 30907181]

29. Parkes J., Roderick P., Harris S., Day C., Mutimer D., Collier J. Enhanced liver fibrosis test can predict clinical outcomes in patients with chronic liver disease. *Gut.* 2010;**59**:1245–1251. [PubMed: 20675693]

30. Irvine K.M., Wockner L.F., Shanker M., Fagan K.J., Horsfall L.U., Fletcher L.M. The Enhanced liver fibrosis score is associated with clinical outcomes and disease progression in patients with chronic liver disease. *Liver Int.* 2016;**36**:370–377. [PubMed: 26104018]

31. Kim B.K., Kim H.-S., Yoo E.J., Oh E.J., Park J.Y., Kim D.Y. Risk assessment of clinical outcomes in Asian patients with chronic hepatitis B using enhanced liver fibrosis test. *Hepatology.* 2014;**60**:1911–1919. [PubMed: 25142433]

32. Angulo P., Bugianesi E., Bjornsson E.S., Charatcharoenwitthaya P., Mills P.R., Barrera F. Simple noninvasive systems predict long-term outcomes of patients with nonalcoholic fatty liver disease. *Gastroenterology.* 2013;**145** 782-789.e784. [PMCID: PMC3931256] [PubMed: 23860502]

33. Pang J.X., Zimmer S., Niu S., Crotty P., Tracey J., Pradhan F. Liver stiffness by transient elastography predicts liver-related complications and mortality in patients with chronic liver disease. *PloS One.* 2014;**9** [PMCID: PMC3995722] [PubMed: 24755824]

34. Ratib S., Fleming K.M., Crooks C.J., Aithal G.P., West J. 1 and 5 year survival estimates for people with cirrhosis of the liver in England, 1998–2009: a large population study. *J Hepatol.* 2014;**60**:282–289.

[PubMed: 24128415]

Figures and Tables

Table 1

Baseline patient characteristics.

Medians (IQR) or counts (proportions). *p* value for difference between groups by Wilcoxon rank sum test or Fischer's exact test. Liver-related event is a composite endpoint of alcoholic hepatitis, varices needing treatment, variceal bleeding, any ascites, spontaneous bacterial peritonitis, covert or overt hepatic encephalopathy, hepatocellular carcinoma, hepatorenal syndrome, upper gastrointestinal bleeding, and jaundice due to liver failure.

2D-SWE, 2-dimensional shear-wave elastography; ALT, alanine aminotransferase; AST, aspartate aminotransferase; ELF, enhanced liver fibrosis test; FIB-4, fibrosis-4 index; GGT, gamma glutamyltransferase; NFS non-alcoholic fatty liver disease fibrosis score; TE, transient elastography.

*Two missing alcohol status at baseline, 31 missing alcohol history.

**Fibrosis stage missing in 98 patients: we refrained from a biopsy from 2016 and onwards if TE <6 kPa (n = 96), 1 patient had an inconclusive biopsy, and the liver biopsy procedure was not technically possible in 1 patient due to emphysema. Steatosis, ballooning and lobular inflammation subscores missing in 106 (97 missing, 9 inconclusive).

***TE missing in 19 (5 unreliable measurements, 6 attempts failed, 8 because of service to the equipment).

Table 2

Harrell's C prognostic performance of 7 non-invasive tests and biopsy-verified fibrosis stage in 462 patients with alcohol-related liver disease.

	Liver-related event	All-cause mortality	Infection requiring hospitalization
TE	0.876 (0.833–0.918)	0.757 (0.702–0.812)	0.677 (0.629–0.724)
ELF	0.859 (0.822–0.897)	0.758 (0.703–0.813)	0.672 (0.624–0.720)
2D-SWE	0.868 (0.832–0.903)	0.714 (0.647–0.780)	0.660 (0.609–0.710)
FT	0.808 (0.764–0.853)	0.721 (0.660–0.781)	0.645 (0.592–0.699)
Forns	0.783 (0.729–0.838)	0.701 (0.636–0.766)	0.621 (0.572–0.670)
NFS	0.794 (0.742–0.847)	0.657 (0.579–0.735)	0.594 (0.543–0.645)
FIB-4	0.821 (0.772–0.870)	0.705 (0.636–0.774)	0.608 (0.557–0.660)
Fibrosis stage	0.819 (0.779–0.859)	0.699 (0.641–0.757)	0.638 (0.588–0.688)

Prognostic accuracy with 95% CIs for 7 non-invasive tests and histological fibrosis stage. Survival analyses with C-statistics reporting Harrell's C from univariable Cox regression for liver-related events, all-cause mortality and infections requiring hospitalization.

2D-SWE, 2-dimensional shear-wave elastography; ELF, enhanced liver fibrosis test; FIB-4, fibrosis-4 index; NFS non-alcoholic fatty liver disease fibrosis score; TE, transient elastography.

Fig. 1
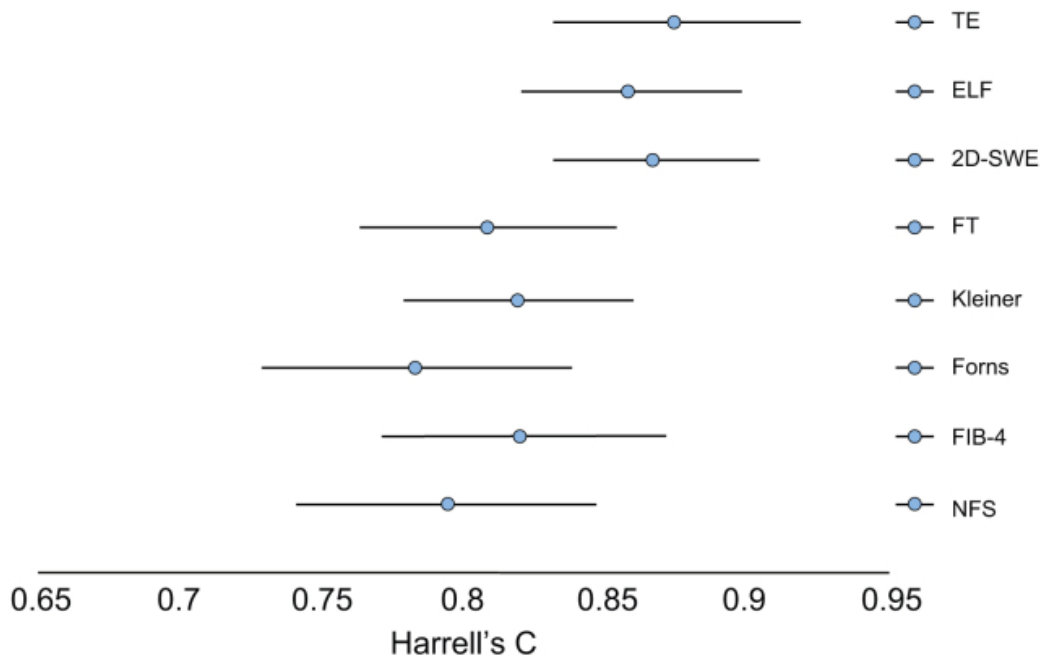
Plot of the prognostic accuracy of 7 non-invasive diagnostic tests and histological fibrosis stage.

The plot lists Harrell's C with 95% CIs, for the ability of each test to predict a liver-related event in 462 patients with alcohol-related liver disease during a median follow-up of 49 months (IQR 31–70) and 1,878 person years. Between-test comparisons of Harrell's C showed no statistical significance between 2D-SWE, ELF, and TE (p >0.33). TE had a significantly higher Harrell's C than FT, Forns, NFS, FIB-4, and Kleiner fibrosis stage (p <0.05). The same was true for 2D-SWE, except for fibrosis stage (p = 0.14). ELF had a significantly higher Harrell's C than FT, Forns, and NFS, but not than fibrosis stage (p = 0.14) and FIB-4 (p = 0.12). 2D-SWE, 2-dimensional shear-wave elastography; ELF, enhanced liver fibrosis test; FIB-4, fibrosis-4 index; FT, FibroTest; NFS non-alcoholic fatty liver disease fibrosis score; TE, transient elastography.

Fig. 2
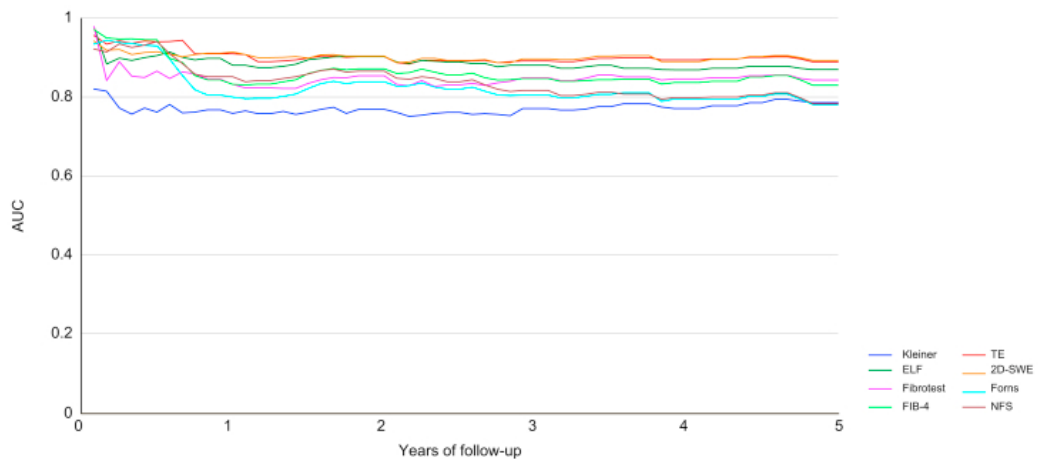
Model performance measures	Kleiner	ELF	TE	2D-SWE	FibroTest	FIB-4	Forns	NFS
Time-dependent AUC	0.786	0.870	0.889	0.893	0.843	0.830	0.781	0.783
Integrated AUC	0.868	0.890	0.890	0.909	0.859	0.855	0.810	0.812
Harrell's C	0.819	0.859	0.876	0.868	0.808	0.821	0.783	0.794

Time-dependent AUC for prediction of liver-related events during 5 years of follow-up for 7 non-invasive tests and liver fibrosis stage.

2D-SWE, 2-dimensional shear-wave elastography; AUC, area under the receiver operating characteristics curve; ELF, enhanced liver fibrosis test; FIB-4, fibrosis-4 index; FT, FibroTest; NFS non-alcoholic fatty liver disease fibrosis score; TE, transient elastography. (This figure appears in color on the web.)

Table 3

Risk of liver-related events for 3 risk groups defined by test-specific cut-offs.

| Risk groups	Events/patients in group (%)	Hazard ratio	*p* value
TE (kPa)			
<10	9/303 (3%)	1	—
10-15	9/42 (21%)	8.07 (3.20–20.35)	<0.001
>15	53/98 (54%)	27.94 (13.75–56.78)	0.001
ELF			
<9.8	15/300 (5%)	1	—
9.8-10.5	11/49 (22%)	4.53 (2.08– 9.87)	<0.001
>10.5	57/108 (53%)	16.94 (9.56–30.04)	<0.001
2D-SWE (kPa)			—
<10	12/222 (5%)	1	
10-16.4	9/60 (15%)	3.4 (1.43– 8.09)	0.006
>16.4	53/83 (64%)	21.6 (11.47–40.67)	<0.001
FibroTest			
<0.31	12/157 (8%)	1	—
0.31-0.58	19/59 (32%)	5.33 (2.58–10.98)	<0.001
>0.58	37/67 (55%)	10.93 (5.68–21.02)	0.011
Forns index			
<4.2	11/135 (8%)	1	—
4.2-6.9	22/214 (10%)	1.36 (0.66– 2.81)	0.403
>6.9	51/106 (48%)	9.19 (4.77–17.71)	<0.001
NFS			
Low[*]	16/191 (8%)	1	—
Intermediate[*]	22/175 (13%)	1.6 (0.84– 3.05)	0.151
>0.676	42/66 (64%)	13.44 (7.53–24.00)	<0.001
FIB-4			
Low[*]	11/170 (6%)	1	—
Intermediate[*]	15/156 (10%)	1.62 (0.74– 3.53)	0.224
>2.67	56/103 (54%)	13.39 (7.00–25.64)	<0.001

[Open in a separate window](#)

Hazard ratios from univariable Cox regression for prediction of liver-related events according to 3 groups of low, intermediate and high risk, with *p* values for between-group difference in hazard ratios. We used previously published cut points to define the risk groups.

2D-SWE, 2-dimensional shear-wave elastography; ELF, enhanced liver fibrosis test; FIB-4, fibrosis-4 index; NFS non-alcoholic fatty liver disease fibrosis score; TE, transient elastography.

[*]Age-dependent cut-off points for NFS and FIB-4. For NFS, the cut point between low and intermediate was -1.455 (for age ≤65 years) and 0.12 (for age >65 years). For FIB-4, the cut point between low and intermediate was 1.3 (for age ≤65 years) and 2.0 (for age >65 years).

Fig. 3
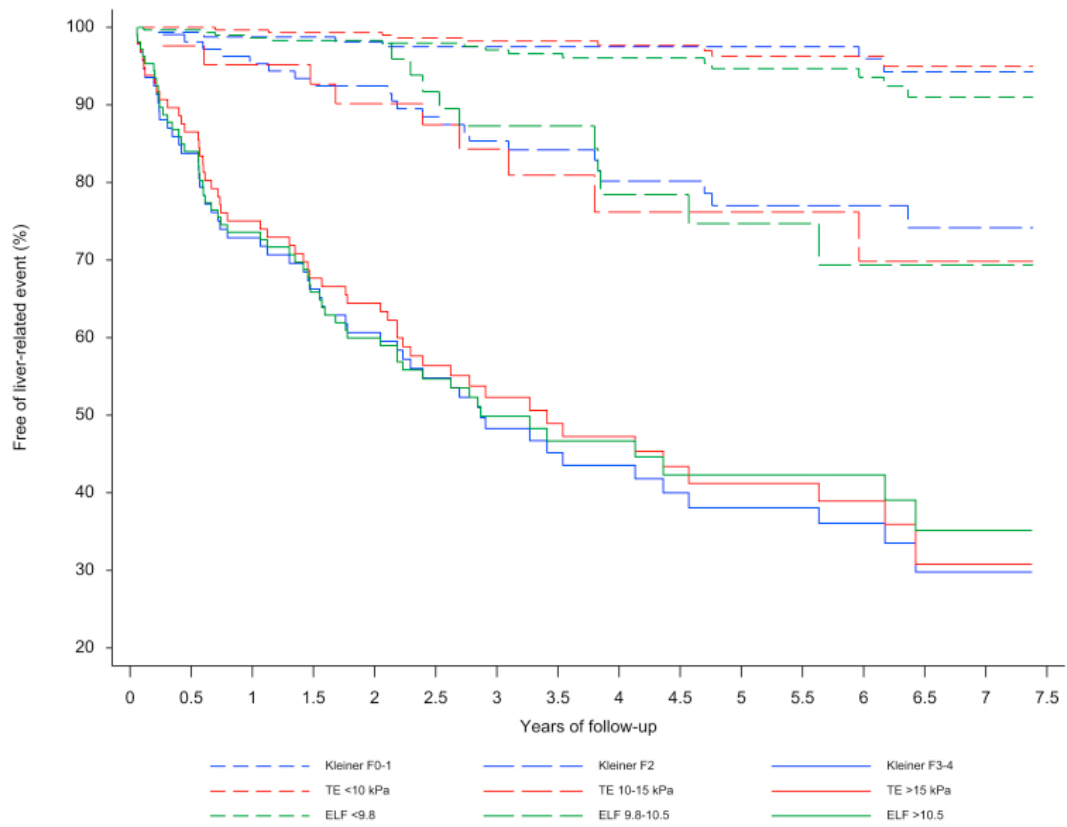
Kaplan-Meier survival curves from univariable Cox regressions for liver-related events.

Patients were stratified into 3 groups of high (solid lines), intermediate (long dash) and low (short dash) risk. All 3 tests stratified groups into significantly different hazard ratios (all *p* <0.001; Table 3). ELF, enhanced liver fibrosis test; TE, transient elastography. (This figure appears in color on the web.)

Fig. 4
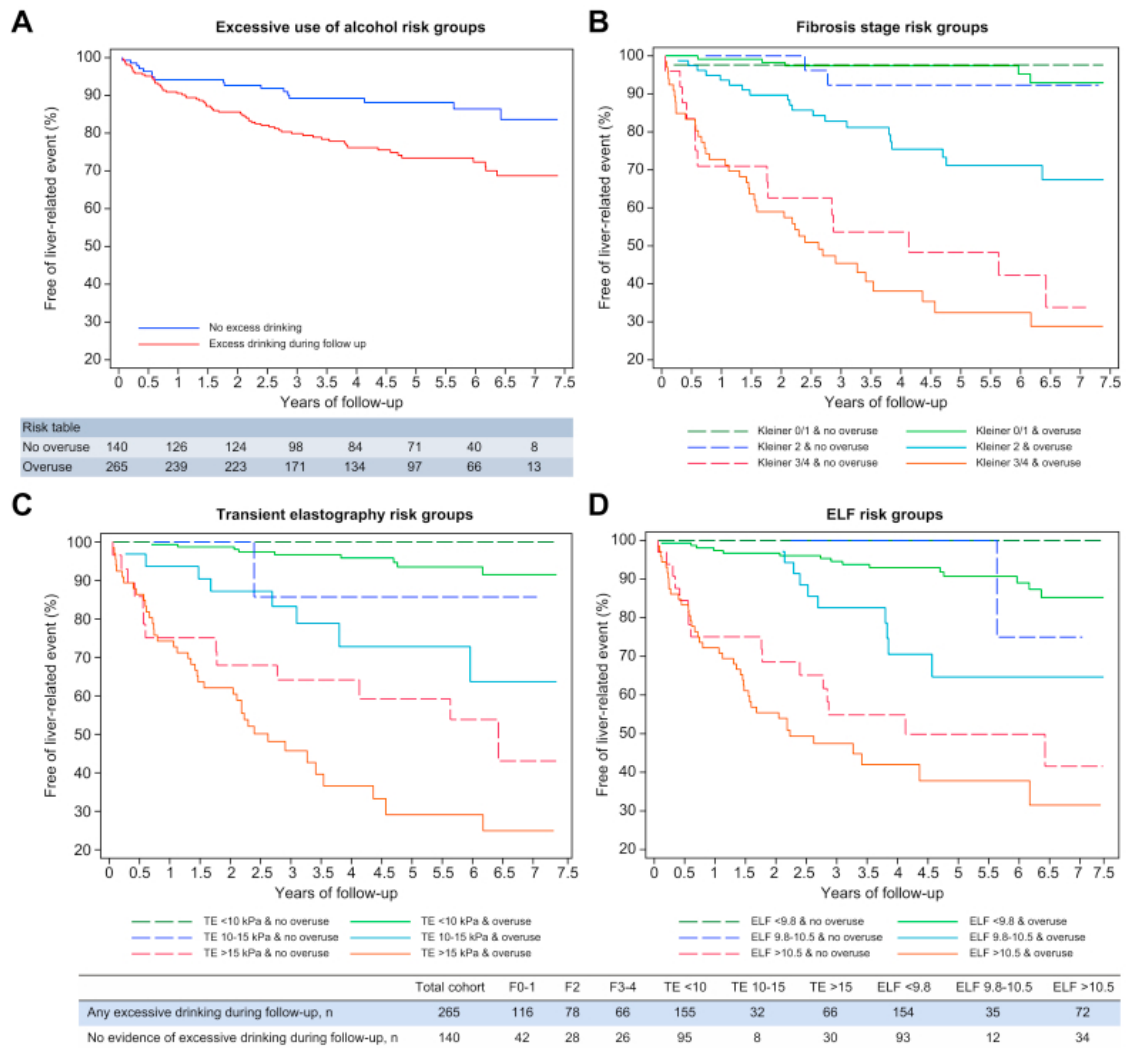
Kaplan-Meier survival curves from univariable Cox regression for liver-related events, stratified according to reports of excessive drinking episodes or not during follow-up in 405 patients.

(A) Overall cohort (hazard ratio = 2.20; 95% CI 1.29–3.74; p = 0.004). (B) According to fibrosis stage. No significant difference between F0-1 and F2 & no overuse, while significantly different hazard ratios for F2 & overuse, F3-4 & no overuse, and F3-4 & overuse. This pattern was repeated for TE (C) and ELF test (D). ELF, enhanced liver fibrosis; TE, transient elastography. (This figure appears in color on the web.)