A Model to Identify Heavy Drinkers at High Risk for Liver Disease Progression

Claire Delacôte, Pierre Bauvin, Alexandre Louvet, Flavien Dautrecque, Line Carolle Ntandja Wandji, Guillaume Lassailly, Cosmin Voican, Gabriel Perlemuter, Sylvie Naveau, Philippe Mathurin, et al.

TITLE: A modelling approach to predict liver disease progression in heavy drinkers

SHORT TITLE: A model to predict ALD progression

AUTHORS: Claire Delacôte[1] (master degree), Pierre Bauvin[1]* (master degree), Alexandre Louvet[1,2]* (MD-PhD), Flavien Dautrecque[2] (MD), Line Carolle Ntandja Wandji[2] (MD), Guillaume Lassailly[1,2] (MD), Cosmin Voican[3,4,5] (MD), Gabriel Perlemuter[3,4,5] (MD-PhD), Sylvie Naveau[3,4,5] (MD-PhD), Philippe Mathurin[1,2]† (MD-PhD), Sylvie Deuffic-Burban[1,6]† (PhD). *,† equal contribution

AFFILIATION:

[1]Univ. Lille, Inserm, CHU Lille, U995 - LIRIC – Lille Inflammation Research International Center, F-59000 Lille, France.

[2]Service des maladies de l'appareil digestif et de la nutrition, CHU Lille, Lille, France.

[3]Service d'hépato-gastro-entérologie et nutrition, Hôpital Antoine Béclère, Hôpitaux universitaires Paris-Sud, Assistance Publique-Hôpitaux de Paris, Clamart, France.

[4]Inserm U996, DHU Hepatinov, Labex LERMIT, Clamart, France.

[5]Faculté de médecine Paris-Sud, Univ. Paris-Sud, Université Paris-Saclay, Le Kremlin-Bicêtre, France.

[6]Université de Paris, IAME, INSERM, F-75018 Paris, France.

FUNDING: none

CORRESPONDING AUTHOR:

Sylvie Deuffic-Burban

IAME INSERM U1137, Université de Paris, Faculté de Médecine Site Xavier Bichat, 16 rue Henri Huchard, 75018 Paris, France

sylvie.burban@inserm.fr

Tel: +33 1 57 27 75 29

Fax: +33 1 57 27 75 21

CONFLICT OF INTEREST

None: CD, PB, AL, FD, LCNW

GL has received lecture fees from Gilead and Novartis.

PM was paid speaking at symposiums for Abbvie, Gilead Sciences, Ipsen, Eisai, MSD, Bayer Healthcare pharmaceutical companies. He is member of the French boards of experts in Hepatology for Abbvie, MSD, Gilead Sciences, Ipsen, Eisal and Bayer Healthcare pharmaceutical companies. He is consulting for the MSD, Ipsen, Eisal, Abbvie, Sanofi, Gilead Sciences, Servier, Bayer Healthcare.

SDB has received research grants and consultancy honoraria from MSD, and lecture fees from Abbvie and Gilead.

The others authors did not answer about their potential conflicts of interest at the time of submission.

AUTHOR CONTRIBUTIONS

PM and SDB had the idea of the study. CD, PB, AL, PM and SDB contributed to the conception and design of the analysis. CV, GP and SN gave access to the raw data. CD performed the analysis. All authors contributed to the interpretation. CD, PB, AL, PM and

SDB drafted the work. All authors revised it critically for important intellectual content and approved the final version to be published. All authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

ABSTRACT

Background and aims: Alcohol-related liver disease (ALD) is a major cause of chronic liver disease worldwide. The development of a predictive tool could improve understanding of the natural history of the disease, target the therapeutic strategy and support medical decision-making.

Design: A Markov model, simulating steatosis, fibrosis, alcoholic steatohepatitis (ASH) and liver complications, was fitted to data of 1,599 heavy drinkers (≥50 g/d) with liver biopsy. It integrates the main risk factors and the duration of exposure to heavy drinking. Once validated, it simulated 5-year disease progression.

Results: Independent cofactors of progression were the age at onset of heavy drinking, sex, BMI and amount of alcohol intake. The 5-year weighted risk of complications are illustrated in 40-year old patients who started to drink at the age of 25, and drank 150 g/d, with a BMI of 22 kg/m². When baseline assessment is F0-F2, the probabilities of having a normal liver, steatosis and ASH are 31.8%, 61.5% and 6.7% for men. Accordingly, the 5-year weighted risk of liver complications was 1.9%, varying from 0.2% in a normal liver to 10.3% for ASH. This risk increases in women (3.2%, from 0.5% to 14.7%). When baseline assessment is F3-F4, the weighted risk of complications was 24.5% (from 20.2% to 34.5%) for men and 30.1% (from 24.7% to 41.0%) for women.

Conclusion: This model integrating duration of alcohol exposure may assist medical decision-making by targeting patients with a high risk of progression. This approach provides new insight in the understanding of disease progression in ALD patients can play a role in adapting therapeutic strategies.

KEY WORDS : alcohol-related liver disease; Markov modelling; prediction; liver complications

INTRODUCTION

Alcohol use is a leading cause of disease and mortality worldwide[1]. The daily amount of alcohol, duration of exposure to excessive drinking, environmental and individual factors are key drivers of the consequences of alcohol-induced liver injury, including fatty liver, fibrosis and alcoholic steatohepatitis (ASH). ASH, steatosis and the severity of fibrosis influences the progression of fibrosis with the highest risk of disease progression in heavy drinkers with ASH [2-4]. Determining the respective contribution of these factors to disease progression is crucial to predict individual risk based on patient characteristics. This approach can help identify patients with a higher risk of progression to cirrhosis and this subgroup should be the target population for therapeutic intervention aimed at reducing disease progression.

Liver biopsy is the gold standard to assess disease stage in ALD [5]. Its limitations, including invasiveness, poor acceptability and cost, explain the difficulty of getting large cohorts with liver samples. The Liver Unit of the Antoine-Béclère Hospital, Clamart, France, has collected a unique cohort of around 1,600 heavy drinkers with histological assessments for each [6]. This cohort makes it possible to develop a predictive model by providing relevant information on cofactors of disease progression. Markov modelling has often been used in these settings to model the natural history of chronic diseases and estimate state-specific progression rates [7-9]. It allows including cofactors that affect the state-specific transition rates [10].Once this model becomes available it can be used to simulate patient outcome at an individual level. Since a liver biopsy cannot be offered to all patients who are heavy drinkers, an accurate baseline evaluation using non-invasive methods such as blood tests or transient elastography [5, 11, 12] can be used as entry data for the Markov model. This approach provides new perspectives for the prediction of patient outcome.

The aims of this study using the Antoine-Béclère cohort were to a/ describe steatosis, fibrosis, ASH progression and liver complications in heavy drinkers, b/ identify independent disease progression cofactors, c/ predict expected disease progression, and d/ propose a medical decision-making tool to predict individual risk based on a Markov model.

MATERIALS AND METHODS

The Markov model

The progression of ALD to liver complications was modelled according to the severity of fibrosis according to METAVIR score and the development of ASH (Figure 1 and Supplementary Methods). The 7 relevant defined stages of ALD were: stage 1, normal liver (no steatosis and no fibrosis); stage 2, steatosis and F0 to F2; stage 3, ASH and F0-F2; stage 4, steatosis and F3-F4; stage 5, ASH and F3-F4; stage 6, liver complications without ASH; and stage 7, liver complications and ASH. Liver complications were defined as the presence of hepatocellular carcinoma (HCC) and/or liver decompensation (defined as bilirubin ≥50 µmol/L and/or gastrointestinal haemorrhage and/or ascites).

We hypothesized that all patients had a normal liver when alcohol abuse began and we assessed the progression of ALD from one biopsy from each patient in relation to the known duration of exposure to excessive alcohol consumption. Recovery from lesions or the disappearance of ASH were considered to be impossible. The effects of potential covariates on the progression of ALD were evaluated (more details below).

Data

The initial cohort included all patients admitted to the Hepatogastroenterology Unit of the Antoine Béclère Hospital between January 1982 and December 1997 in Clamart, France [6]. For the aim of our study, we identified 2,334 heavy drinkers (≥ 50 g/d) having persistently abnormal liver test results and a recorded duration of alcohol abuse during their first admission. None of them had known chronic viral hepatitis. All data were collected during hospitalization using a standardized questionnaire. The population was divided into two groups according to the availability of the histological assessment which was systematically

offered to each patient and performed during routine practice and not for the purpose of a research study (Supplementary Figure 1).

A first group included 1,599 patients who underwent a liver biopsy during their first admission. These patients were used to estimate the model parameters and formed the Development cohort. The METAVIR score was not available in 795 of these patients but 542 of them had an available disease stage (normal liver, pure steatosis, compensated cirrhosis without ASH, compensated cirrhosis with ASH, complicated liver disease with ASH, complicated liver disease without ASH) that could be translated into a METAVIR score. For the remaining 253, the METAVIR score was imputed by predictive mean matching by sex, age, BMI, alcohol consumption and the absence/presence of ASH (Supplementary Table 1).

The second group included 735 patients who did not undergo liver biopsy, but were classified at hospitalization according to the presence of liver complications. This cohort was used as a Validation cohort to determine the adequacy of the model for the prediction of liver complications.

Covariates

Covariates were incorporated into the model through the proportional hazards assumption. We tested four covariates likely to be associated with the progression of the disease based on the literature[3, 6, 13-19] and available in our database: sex; age when alcohol abuse began in years; body mass index (BMI in kg/m²); and daily alcohol consumption (g/d) over the 5 years before hospitalization. Missing BMI and daily alcohol intake values were imputed using predictive mean matching by sex and age. Variations in alcohol intake and BMI over time

were not considered. Univariate and multivariate analyses were conducted on the basis of likelihood ratio statistics (details in Supplementary Methods).

Procedure

There were two steps to this study. In the first step, the model was developed and fitted to Development cohort data. The resulting model produced parameter estimates (transition rates and covariate effects) from maximum likelihood methods. Transition probabilities could be directly computed from these parameter estimates (details in Supplementary Methods). The reproducibility of the model was then evaluated in the Validation cohort.

In the second step, the model was used to predict future disease progression for specific patient profiles based on the available information on current disease stage (details in Supplementary Methods). In the results section, examples of the prediction are provided for patients with an evaluation of fibrosis by non-invasive methods. The model is also a helpful tool for the identification of populations at high risk of disease progression and decision-making for the patient's care pathway. This was illustrated by arbitrarily defining a high risk population as those with a 5-year weighted risk of liver complications > 5%.

RESULTS

Patients characteristics

The patient characteristics of the two cohorts are summarized in Table 1. There were no differences in gender or the number of smokers between the two groups. Patients in the Validation cohort were older with a higher BMI while they reported a lower daily alcohol intake than those in the Development cohort. Patients in the Validation cohort had a significantly higher risk of liver complications at admission (29% vs 22%, $p<0.001$).

Development of the model

Adequacy with data. The model satisfactorily predicted the observed stages of the patients in the Development cohort (Supplementary Table 2). For example, it predicted that 200 patients would move from normal liver to steatosis-F3-F4, based on the declared duration of heavy drinking by each individual, and close to 209 transitions were observed.

Parameter estimates. The transition rates between disease stages differ according to individual characteristics in relation to the estimated effect of the covariates incorporated in the model. Table 2 reports the estimated baseline transition rates λ_{ij} (and their 95% confidence intervals (CI)) corresponding to patients with mean individual characteristics: men with a BMI of 22 kg/m², who began abusing alcohol at the age of 25 and drank 150 g/d. They can be expressed as the number of annual transitions from stage *i* to stage *j*. For example, the baseline λ_{24}=3% means that 3 out of 100 patients with steatosis-F0-F2 will progress to steatosis-F3-F4 in 1 year. As expected, the fibrosis progression rates and the rates of the occurrence of liver complications are increased in the presence of ASH (14% vs. 3%) and (8.4% vs. 4.3%), respectively.

Concerning the independent effect of each incorporated covariate (Table 2), women were found to have a 24.8% greater risk of progression than men (HR=1.248) and a 1-year increment of age at the beginning of heavy drinking increased the risk of progression by 3.8% independently of disease stage disease. Compared to a person with a BMI of 22 kg/m², the risk of disease progression in an obese person (BMI=30 kg/m²) is increased by 11.8%. The impact of an additional standard drink per day (10 g in France) were found to have less influence. Thus transition rates in women with a more unfavourable profile (older when heavy drinking began -30 years-, higher BMI -35 kg/m²-, higher alcohol intake – 180 g/d) are increased by 83% compared to the baseline transitions provided in Table 2.

The average time spent in each stage can be obtained from the transition rates. This provides a better understanding of the impact of covariates on disease progression as illustrated by 4 arbitrary profiles (Table 3).

Men who began drinking at 25, drinking 150 g/d with a BMI of 22 kg/m² have an estimated mean 10.8 years from the onset of alcohol abuse to the development of alcohol-induced steatosis. In women with all other characteristics being equal steatosis will develop in 8.7 years.

In individuals who remain in the steatosis stage the estimated mean time from F0-F2 to the occurrence of liver complications is 35.8 years in men and 28.7 years in women. In patients with ASH, this mean time is reduced to 19.0 years in men and 15.2 years in women.

All of the above mean times were reduced by 1.45 when heavy drinking began 10 years later, all other things being equal.

Reproducibility of the model. The accuracy of the model was assessed in the Validation cohort for the number of liver complications and found to be good (Supplementary Table 3).

The model predicted 34 liver complications less than the observed number (182 instead of 216).

Prediction of disease progression in patients with a baseline assessment using non-invasive methods

As explained in the Methods section, non-invasive methods were considered to have been used to define baseline fibrosis and the model predicts the weighted probability of developing complications based on this baseline assessment of fibrosis. The usefulness of the model was confirmed by providing examples of disease progression predictions in the next 5 years.

In patients with baseline F0-F2 who are 40-year-old men, and have been drinking 150 g/d for 15 years, with BMI 22 kg/m², the current probability of having a normal liver, steatosis and ASH are 31.8%, 61.5% and 6.7%, respectively (Supplementary Figure 2A). Based on this distribution, the 5-year weighted risk of liver complications is 1.9%, minimal (0.2%) for a normal liver and maximal (10.3%) in the presence of ASH (Figure 2A).

The current probabilities were found to vary according to individual characteristics. The prevalence of ASH is higher in older men (9.6%, Supplementary Figure 2B) and in women (Supplementary Figure 2C and 2D). According to these estimations, the weighted risk of complications at 5 years increases in older men (Figure 2B) and is even more pronounced in women, reaching 7.3% for 50-year-old women, varying from 1.3% to 25.4% (Figures 2C and 2D). Based on these predictions, 50-year old women are a high-risk subgroup of disease progression and should receive close follow-up.

As progression to severe fibrosis is an endpoint in clinical studies, Supplementary Figure 3 shows the risk of progression to F3-F4 (with or without liver complications) for the above described profiles.

The risk of complications in patients with baseline F3-F4 are shown in Figure 3. In 40-year-old patients who have abused alcohol for 15 years, considering the estimated prevalence of ASH (30.0% for men and 33.3% for women; Supplementary Figure 2A-2C), the 5-year weighted risk of complications is 24.5% for men and 30.1% for women (ranging from 20.2% to 34.5% and 24.7% to 41.0%, respectively in absence and presence of ASH) (Figures 3A and 3C). In the subpopulation of 50 year olds, where the prevalence of ASH is higher (Supplementary Figures 2B and 2D), these risks are increased by 1.3-1.4-fold (Figures 3B and 3D). Obviously, all patients with stage F3-F4 fibrosis, are at high-risk of disease progression at 5 years.

DISCUSSION

By exploring the complex relationship between the duration of exposure to alcohol, age at the onset of heavy drinking, amount of alcohol intake, sex and obesity and the progression of ALD, our model provides a robust tool to predict disease progression according to patient characteristics. This model integrates baseline fibrosis to predict disease progression and identifies different courses among patients with the same severity of baseline fibrosis. This new insight in the understanding of disease progression in ALD patients could be used to adapt therapeutic strategies.

The Antoine-Béclère cohort is unique since a liver biopsy was systematically offered to all heavy drinkers at their first admission. Despite the large size of this cohort, newer machine learning algorithms could not be applied, as they often require tens of thousands of observations to obtain pertinent results. With the structure of our database and with the restrictions of machine learning, the Markov model was the method of choice.

For the first time our model quantifies the annual risk of the progression of fibrosis and the development of ASH in a cohort of heavy drinkers in relation to the impact of independent covariables such as female gender, the presence of ASH, BMI, age and alcohol intake. As previously suggested[17, 20, 21], our model also shows that women have a higher risk of disease progression than men with the same declared level of alcohol. Although we cannot exclude that women underestimated their alcohol consumption more than men which could contribute to this sex difference, published studies suggest that women are more susceptible to the hepatotoxic effects of alcohol[5]. In addition, it was not surprising that the presence of necro-inflammatory activity was associated with more rapid disease progression. The extent of the effect in our study is similar to that observed in NAFLD disease[22] or in HCV-patients[23]. As age at the start of alcohol abuse is a cofactor of disease progression, the risk of progression

increased when alcohol abuse started later (all other cofactors being equal). Nevertheless, this risk increased according to duration of exposure. For men with a BMI of 22 kg/m² who drink 150 g/d, at the same age (45), the risk of liver complications is 10.6% after 20 years (starting at 25) vs. 4.7% after 10 years (starting at 35). The impact of ethnicity could not be assessed because French law prohibits recording this variable. Data on diabetes mellitus was not available in the cohort. However, we analysed the effect of high glycaemia (>7.8 mmol/L) on disease progression and this did not change the adequacy of the model (not shown). Although, most of the independent covariates of the model have been previously described [3, 6, 13-19], none of those studies developed a model because large cohorts of biopsy-proven ALD were not available. Thus, our model, which is based on non-invasive methods of diagnosis, represents significant progress for clinicians.

One interesting result of our study was an estimation of the mean time until the development of the first alcohol-induced liver injury (i.e. steatosis). For patients who begin abusing alcohol at the age of 25 and drink 150g/d with a BMI of 22 kg/m², the interval free of liver disease is around 11 years. For the same profile severe fibrosis (F3-F4) will develop in 20 years. In patients with same profile who have developed ASH advanced fibrosis will develop in around 7 years. In comparison, in chronic hepatitis C (CHC), a previous model estimated that in patients with heavy alcohol consumption and CHC since the age of 30, the mean time to cirrhosis was around 22-23 years for both sexes while in case of CHC alone this was around 40 years for men and 58 years for women[24]. Our model confirms that ALD patients have a higher risk of disease progression than patients with other causes of liver chronic diseases.

This study has certain limitations. First, we predict disease progression in patients assuming that they will continue to drink the same amount of alcohol. We consider that there will be no

disease progression after alcohol is stopped in patients who become abstinent. Alcohol consumption may have varied over time and we cannot make predictions based on fluctuations in alcohol intake over time. Although this creates a potential bias, our only option was to assume that the declared consumption was representative of the patient's past history. Although the duration and amount of alcohol consumption was based on the patients' self-reported history which can be unreliable, the standardized questionnaire to obtain this information was constructed to limit this bias. Indeed, patient's relative were interviewed in case of doubt and alcohol consumption was measured not only for the days preceding the hospitalization but over a 5-year period to reduce potential understatement and memory bias. Second, our model applies to patients with biological abnormalities because liver biopsies were only performed in patients with at least one abnormal liver test. Imputations of data for the Metavir score, BMI and alcohol intake might have introduced some bias in certain cases. However, the accuracy of the model was not affected by a sensitivity analysis limited to patients with all available data (results not shown). Third, determination of the progression of fibrosis was based on F0-F2 to F3-F4. Thus the model cannot predict disease progression from one stage of fibrosis to the next. Nevertheless, this progression is not pertinent for the evaluation of fibrosis by non-invasive methods, and liver biopsy is no longer used due to its invasiveness. Finally, although it would have been interesting to validate the model in different cohort of patients in different countries, these data were not available at the time of this study.

This model helps discriminate at-risk subgroups of heavy drinkers with at least one abnormal liver test and determines the consequences for medical management for general practitioners or hepatologists. When patients identified as high-risk have had an assessment of fibrosis by a hepatologist, follow-up should be reinforced (with an annual assessment of fibrosis) and

consultation with an addiction specialist should be encouraged. If patients are identified as being heavy drinkers by the general practitioner with no evaluation of fibrosis, these patients should be referred to a hepatologist. Nevertheless, we think that the threshold defining the high risk population which has been arbitrarily fixed at 5% should be discussed by experts because it affects the patient's care pathway. An online application is being developed to help clinicians and general practitioners in their daily practice.

Like in other fields of medical research, it is urgent to create optimal scientific conditions for the development of new drugs and strategies in patients with compensated ALD. Adequate sample sizes must be calculated based on clear assessment of disease stage and a rational evaluation of disease progression is needed to reach these goals.
Our model is an interesting approach to predict the risk of disease progression in untreated patients and help calculate the expected number of events.

In conclusion, the present model could play a role in adapting ALD patient management and assist for medical decision-making.

REFERENCES

1. Rehm J, Samokhvalov AV, Shield KD. Global burden of alcoholic liver diseases. J Hepatol. 2013;59(1):160-8.
2. Galambos JT. Natural history of alcoholic hepatitis. 3. Histological changes. Gastroenterology. 1972;63(6):1026-35.
3. Marbet UA, Bianchi L, Meury U, et al. Long-term histological evaluation of the natural history and prognostic factors of alcoholic liver disease. J Hepatol. 1987;4(3):364-72.
4. Mathurin P, Beuzin F, Louvet A, et al. Fibrosis progression occurs in a subgroup of heavy drinkers with typical histological features. Aliment Pharmacol Ther. 2007;25(9):1047-54.
5. European Association for the Study of the Liver. Electronic address eee, European Association for the Study of the L. EASL Clinical Practice Guidelines: Management of alcohol-related liver disease. J Hepatol. 2018;69(1):154-81.
6. Naveau S, Giraud V, Borotto E, et al. Excess weight risk factor for alcoholic liver disease. Hepatology. 1997;25(1):108-11.
7. Longini IM, Jr., Clark WS, Gardner LI, et al. The dynamics of CD4+ T-lymphocyte decline in HIV-infected individuals: a Markov modeling approach. J Acquir Immune Defic Syndr. 1991;4(11):1141-7.
8. Marshall G, Jones RH. Multi-state models and diabetic retinopathy. Stat Med. 1995;14(18):1975-83.
9. Deuffic-Burban S, Poynard T, Valleron AJ. Quantification of fibrosis progression in patients with chronic hepatitis C using a Markov model. J Viral Hepat. 2002;9(2):114-22.
10. Marshall G, Guo W, Jones RH. MARKOV: a computer program for multi-state Markov models with covariables. Comput Methods Programs Biomed. 1995;47(2):147-56.
11. Nguyen-Khac E, Thiele M, Voican C, et al. 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(9):614-25.
12. Naveau S, Gaude G, Asnacios A, et al. Diagnostic and prognostic values of noninvasive biomarkers of fibrosis in patients with alcoholic liver disease. Hepatology. 2009;49(1):97-105.
13. Teli MR, Day CP, Burt AD, et al. Determinants of progression to cirrhosis or fibrosis in pure alcoholic fatty liver. Lancet. 1995;346(8981):987-90.
14. Pares A, Caballeria J, Bruguera M, et al. Histological course of alcoholic hepatitis. Influence of abstinence, sex and extent of hepatic damage. J Hepatol. 1986;2(1):33-42.
15. Krasner N, Davis M, Portmann B, et al. Changing pattern of alcoholic liver disease in Great Britain: relation to sex and signs of autoimmunity. Br Med J. 1977;1(6075):1497-500.
16. Klatsky AL, Morton C, Udaltsova N, et al. Coffee, cirrhosis, and transaminase enzymes. Arch Intern Med. 2006;166(11):1190-5.
17. Becker U, Deis A, Sorensen TI, et al. Prediction of risk of liver disease by alcohol intake, sex, and age: a prospective population study. Hepatology. 1996;23(5):1025-9.
18. Becker U, Gronbaek M, Johansen D, et al. Lower risk for alcohol-induced cirrhosis in wine drinkers. Hepatology. 2002;35(4):868-75.
19. Hart CL, Morrison DS, Batty GD, et al. Effect of body mass index and alcohol consumption on liver disease: analysis of data from two prospective cohort studies. BMJ. 2010;340:c1240.
20. Dam-Larsen S, Becker U, Franzmann MB, et al. Final results of a long-term, clinical follow-up in fatty liver patients. Scand J Gastroenterol. 2009;44(10):1236-43.
21. Klatsky AL, Armstrong MA. Alcohol, smoking, coffee, and cirrhosis. Am J Epidemiol. 1992;136(10):1248-57.

22. Singh S, Allen AM, Wang Z, et al. Fibrosis progression in nonalcoholic fatty liver vs nonalcoholic steatohepatitis: a systematic review and meta-analysis of paired-biopsy studies. Clin Gastroenterol Hepatol. 2015;13(4):643-54 e1-9; quiz e39-40.
23. Deuffic-Burban S, Babany G, Lonjon-Domanec I, et al. Impact of pegylated interferon and ribavirin on morbidity and mortality in patients with chronic hepatitis C and normal aminotransferases in France. Hepatology. 2009;50(5):1351-9.
24. Deuffic-Burban S, Deltenre P, Buti M, et al. Predicted effects of treatment for HCV infection vary among European countries. Gastroenterology. 2012;143(4):974-85 e14.

LEGENDS OF TABLES

Table 1: Patient characteristics

Table 2: Parameter estimates: baseline transition rates (λ_{ij}) and regression coefficient (β_z)

Table 3: Mean number of years in each stage, and their 95% confidence interval in each non-absorbing state of the disease for four profiles of patients

LEGENDS OF FIGURES

Figure 1: Markov model of alcohol related-liver disease

Figure 2: Prediction of risk of liver complications over 5 years in four profiles of individuals classified with either a normal liver, steatosis-F0-F2 or ASH-F0-F2 by non-invasive tests: A) Men seen at the age of 40, B) Men seen at the age of 50, C) Women seen at the age of 40, D) Women seen at the age of 50. All patients have been exposed to alcohol abuse for 15 years, have a BMI of 22 kg/m² and drink 150 g/d.

Figure 3: Prediction of risk of liver complications over 5 years in four profiles of individuals classified with either steatosis-F3-F4 or ASH-F3-F4 by non-invasive tests: A) Men seen at the age of 40, B) Men seen at the age of 50, C) Women seen at the age of 40, D) Women seen at the age of 50. All patients have been exposed to alcohol abuse for 15 years, have a BMI of 22 kg/m² and drink 150 g/d.

Table 1. Patient characteristics

	Development cohort	Validation cohort	P-value
Number of patients	1599	735	
Male (%)	1204 (75)	548 (75)	0.70
Age at admission, mean (SD), years	50.2 (11.9)	53.1 (13.8)	<0.001
Smoker (%)	1135 (71)	535 (73)	0.37
BMI[a], mean (SD), kg/m²	22 (4)	22 (4)	0.010
Duration of alcohol abuse, mean (SD), years	22 (13)	25 (15)	<0.001
Alcohol intake over the last 5 years, mean (SD), g/d	153 (94)	144 (96)	0.03
Compensated liver disease (%)	1250 (78)	519 (71)	<0.001
Normal liver	224	NA	
Steatosis-F0-F2	607	NA	
ASH-F0-F2	101	NA	
Steatosis-F3-F4	209	NA	
ASH-F3-F4	109	NA	
Liver complications[b] (%)	349 (22)	216 (29)	<0.001
Without ASH	185	NA	
With ASH	164	NA	

Abbreviations: BMI, body mass index; ASH, alcoholic steatohepatitis; SD, standard deviation; NA, not available; F0, no fibrosis; F1, portal fibrosis without septa; F2, few septa; F3, numerous septa without cirrhosis; and F4, cirrhosis.

[a]BMI was calculated from the minimum weight in the last 10 years.

[b]Hepatocellular and/or clinical liver decompensation.

Table 2. Parameter estimates: baseline transition rates ($\boldsymbol{\lambda_{ij}}$) and regression coefficient ($\boldsymbol{\beta_z}$)

	Considering transition	Estimate (95% CI)
Baseline $\boldsymbol{\lambda_{12}}$	from normal liver to steatosis-F0-F2	9.2% (9.2%-9.3%)
Baseline $\boldsymbol{\lambda_{24}}$	from steatosis F0-F2 to steatosis F3-F4	3.0% (2.7%-3.3%)
Baseline $\boldsymbol{\lambda_{46}}$	from steatosis-F3-F4 to liver complications	4.3% (3.9%-4.8%)
Baseline $\boldsymbol{\lambda_{35}}$	from ASH-F0-F2 to ASH-F3-F4	14.0% (13.9%-14.1%)
Baseline $\boldsymbol{\lambda_{57}}$	from ASH-F3-F4 to liver complications	8.4% (8.4%-8.5%)
Baseline $\boldsymbol{\lambda_{23} = \lambda_{45}}$	from steatosis to ASH	2.0% (1.8%-2.2%)
β_{Women}		0.222 (0.221-0.222)
$\beta_{\text{Age when drinking began}}$		0.037 (0.037-0.037)
β_{BMI}		0.014 (0.014-0.014)
$\beta_{\text{Alcohol intake}}$		0.003 (0.003-0.003)

λ_{ij} are the transition rates between stage *i* and j. Baseline λ values correspond to men drinking since the age of 25, 150 g/d with a BMI of 22 kg/m². Hazard ratio (HR), expressing the impact of the covariate z on the transition rate, can be directly obtained from the regression coefficient β_z: HR= $e^{\beta z}$.

Table 3. Mean number of years in each stage, and their 95% confidence interval in each non-absorbing state of the disease for four profiles of patients

	Mean number of years in				
	Normal liver	**Steatosis-F0-F2**	**Steatosis-F3-F4**	**ASH-F0-F2**	**ASH-F3-F4**
Men, drinking since the age of 25, BMI 22, 150 g/d	10.8 (10.8-10.8)	20.0[a] (19.0-21.1)	15.8 (14.7-17.1)	7.2 (7.1-7.2)	11.8 (11.7-11.9)
Men, drinking since the age of 35, BMI 22, 150 g/d	7.5 (7.5-7.5)	13.8 (13.1-14.6)	10.9 (10.1-11.8)	4.9 (4.9-5.0)	8.2 (8.1-8.2)
Women, drinking since the age of 25, BMI 22, 150 g/d	8.7 (8.6-8.7)	16.0 (15.2-16.9)	12.7 (11.7-13.7)	5.7 (5.7-5.8)	9.5 (9.4-9.5)
Women, drinking since the age of 35, BMI 22, 150 g/d	6.0 (6.0-6.0)	11.1 (10.5-11.7)	8.8 (8.1-9.4)	4.0 (3.9-4.0)	6.5 (6.5-6.6)

[a]For this profile, transition rates to leave the stage "steatosis-F0-F2" are respectively 3.0% (to steatosis-F3-F4) and 2.0% (to ASH-F0-F2) therefore transition rate to stay in this stage is –(3%+2%) = - 5%. By definition, mean amount of time in steatosis-F0-F2 is -1/ (-5%) = 20.0 years. The other mean amounts of time were calculated in the same way.

Figure 1: Markov model of alcohol related-liver disease

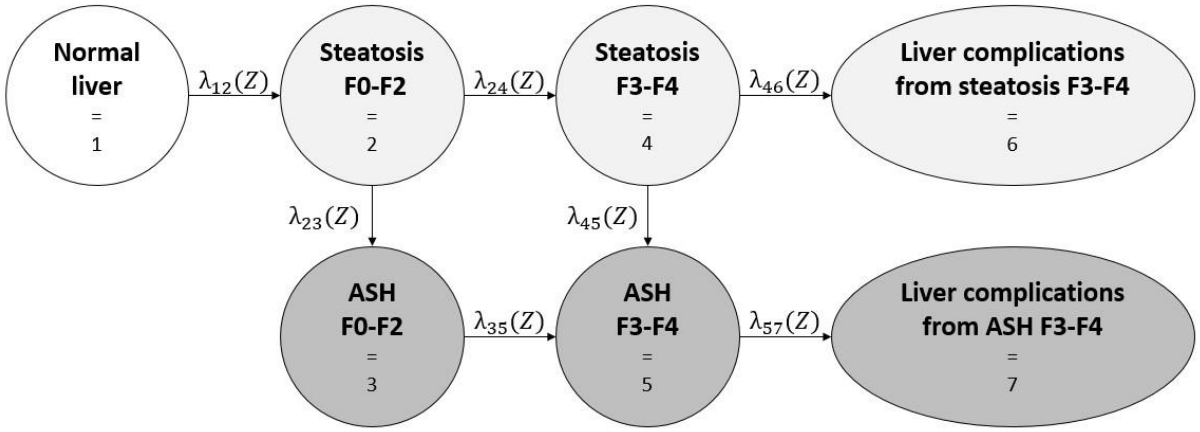


Figure 2: Prediction of risk of liver complications over 5 years in four profiles of individuals classified with either a normal liver, steatosis-F0-F2 or ASH-F0-F2 by non-invasive tests: A) Men seen at the age of 40, B) Men seen at the age of 50, C) Women seen at the age of 40, D) Women seen at the age of 50. All patients have been exposed to alcohol abuse for 15 years, have a BMI of 22 kg/m² and drink 150 g/d.

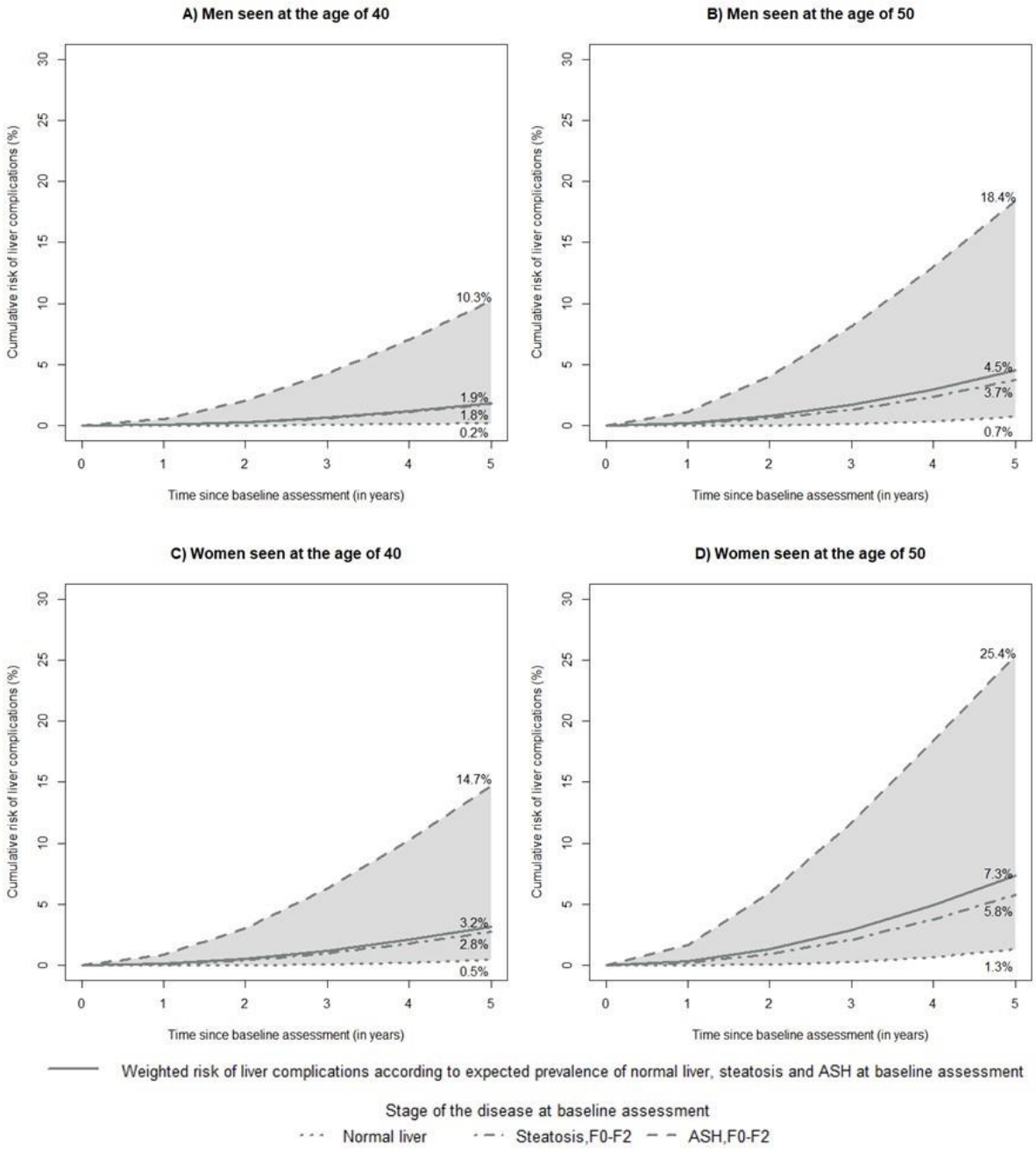
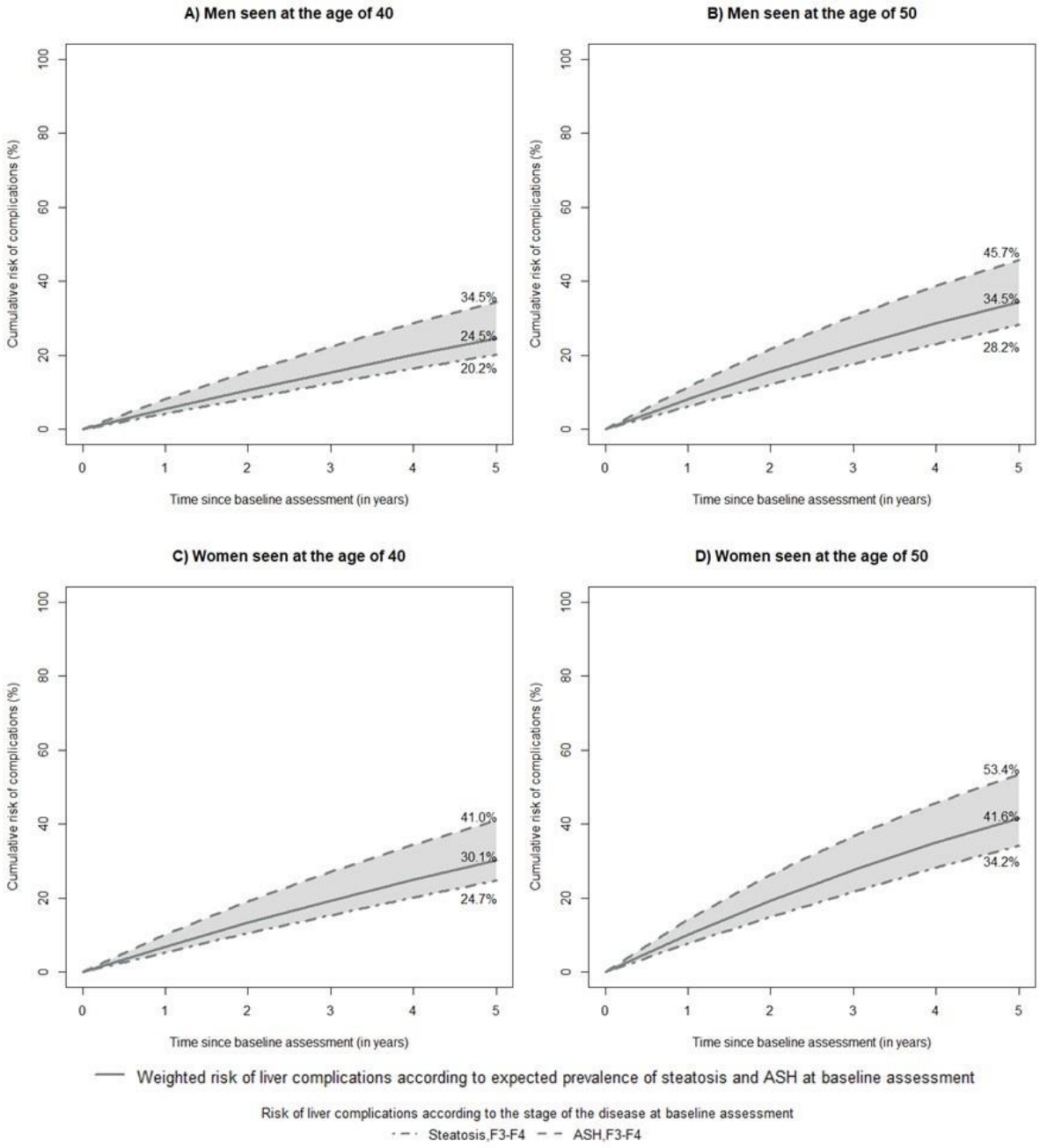


Figure 3: Prediction of risk of liver complications over 5 years in four profiles of individuals classified with either steatosis-F3-F4 or ASH-F3-F4 by non-invasive tests: A) Men seen at the age of 40, B) Men seen at the age of 50, C) Women seen at the age of 40, D) Women seen at the age of 50. All patients have been exposed to alcohol abuse for 15 years, have a BMI of 22 kg/m² and drink 150 g/d.



Supplementary Methods.

Generality on Markov models

Time homogeneous Markov models were used to analyze the data. This method is used to describe the process by which an individual moves through a series of states in continuous time. It considers the data representing observations of the process at arbitrary times. Interval censoring is considered in the calculations. At time *t* the individual is in state *S(t)*. Transitions allowed between states are defined. The state to which the individual moves and the time of change are governed by a set of transition rates $\lambda_{ij}(t,z(t))$ for each pair of states *i* and *j*. Transition rates between two stages of the disease do not depend on the time or the duration of time spent in a given stage. However, the rates may depend on a set of individual-specific explanatory variables *z(t)*. Covariates are incorporated into the model through the proportional hazards assumption.

The rate represents the instantaneous risk of moving from state *i* to state *j* : $\lambda_{ij}\big(t, z(t)\big) = \lim_{\delta t \to 0} \frac{P(S(t+\delta t)=j | S(t)=i)}{\delta t}$. The rates form a matrix Λ whose rows sum to zero. Under the Markov assumption $\lambda_{ij}(t,z(t))$ is independent of the observation history of the process up to the time preceding *t.* A form of proportional hazards model is used to study covariates in which the transition rate matrix elements λ_{ij} can be replaced by $\lambda_{ij}\big(z(t)\big) = \lambda_{ij}{}^{(0)} \exp\left(\beta_{ij}^{T} \mathrm{z(t)}\right)$.

This matrix of transition rates, Λ, yields a transition probability matrix, *P(t)*, through the Kolmogorov relationship $P(t) = Exp(t\Lambda)$. A time-homogeneous process is considered therefore the *(i,j)* entry of *P(t)*, $p_{ij}(t)$, is the probability of being in state *j* at a time $t+u$, given the state at time *u* is *i.* It does not say anything about the time of transition from *i* to *j*, indeed the process may have entered other states between time *u* and $t+u$.

A likelihood function is calculated from the transition probability matrix *P(t)*, therefore, from the transition rate matrix Λ. To fit a multistate model to observed data , Λ is estimated using likelihood maximization.

Once the model is fitted, it is also possible to predict the theoretical course of any individual using *P(t).*

Estimated mean times spent in each transient state *i* are calculated as $-1/\lambda_{ii}$ where λ_{ii} is the *i*th diagonal entry of the estimated transition rates matrix (*Λ).*

A detailed user manual covering the above methodology, and enriching it for some aspects, and providing an example on the multistate model is available online[1].

Natural history of ALD

Certain stages of the METAVIR[2, 3] score were gathered for the purpose of our model because the number of observations in certain stages of this score was insufficient to obtain reliable estimates of transition rates to these stages: stage 1, normal liver (no steatosis and no fibrosis); stage 2, steatosis and F0-F2; stage 3, ASH and F0-F2; stage 4, steatosis and F3-F4; stage 5, ASH F3-F4; stage 6, liver complications without ASH; and stage 7, liver complications and ASH. In this structure direct transitions can occur from one stage to the same stage with ASH or to the next stage indicated by an arrow (Figure 1).

Covariates

The individual effect of each covariate (sex, age when alcohol abuse began, BMI, daily alcohol consumption) on the transition rates was assessed on univariate analysis. All covariates except daily alcohol intake had a p-value < 0.20 for likelihood ratio statistics. The four covariates were retained for multivariate analysis, considering that a threshold effect could be reached in this cohort with very high alcohol intake.

Preliminary model investigations showed that it could be assumed that each variable had the same effect on all transitions.

Calculation of the transitions

For the previously describe model and for the baseline profile (men, drinking since the age of 25, BMI 22 kg/m², 150 g/d), $\Lambda =$

$$\begin{pmatrix} -0.092 & 0.092 & 0 & 0 & 0 & 0 & 0 \\ 0 & -0.050 & 0.020 & 0.030 & 0 & 0 & 0 \\ 0 & 0 & -0.140 & 0 & 0{,}140 & 0 & 0 \\ 0 & 0 & 0 & -0.063 & 0.020 & 0.043 & 0 \\ 0 & 0 & 0 & 0 & -0.084 & 0 & 0.084 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Analytical expression of the probability for non-diagonal entries are not simple for this type of structure. Considering the simple case, having a normal liver the annual probability of remaining with a normal liver is $e^{-0.092} = 0.912$ and the 5-year probability is $e^{-0.092*5} = 0.631$. On the other hand, the annual probability of leaving this stage is (1-0.912) =0.088 and the 5-year probability is (1-0.631) =0.369. Finally, the matrix of 1-year and 5-year probabilities of this profile are respectively:

$$\begin{pmatrix} 0.912 & 0.086 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.951 & 0.018 & 0.028 & 0 & 0 & 0 \\ 0 & 0 & 0.869 & 0 & 0.125 & 0 & 0 \\ 0 & 0 & 0 & 0{,}939 & 0.019 & 0.042 & 0 \\ 0 & 0 & 0 & 0 & 0.919 & 0.081 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \text{ and}$$

$$\begin{pmatrix} 0.630 & 0.324 & 0.025 & 0.002 & 0.015 & 0.004 & 0 \\ 0 & 0.779 & 0.063 & 0.113 & 0.028 & 0.013 & 0.004 \\ 0 & 0 & 0.497 & 0 & 0.400 & 0 & 0.103 \\ 0 & 0 & 0 & 0.729 & 0.069 & 0.185 & 0.017 \\ 0 & 0 & 0 & 0 & 0.655 & 0 & 0.345 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Some rows are not equal to 1 because of rounded number.

Statistical analysis

Patient characteristics in each group were compared using parametric tests (χ^2 for qualitative variables and Student *t* tests for quantitative variables).

All analyses were performed with the statistical software R using the msm analysis routine for the Markov model and the package MICE for imputations.

Details on the second step of the procedure

First, the model estimated the probability of currently being in each of the 7 previously defined stages of the disease. There were two situations. In case 1, the patient did not have any evaluation of the severity of the fibrosis by non-invasive methods, and the estimated probabilities at time *t* of being in one of the different stage were directly used as input to predict the 5-year weighted risk of liver complications. In case 2, the patient had an evaluation of the severity of fibrosis. If the patient is classified with non-severe fibrosis (F0-F2), the probabilities of being in the advanced stages of the disease were set equal to 0 (stage 4-7 in Figure 1). If the patient is classified with severe fibrosis (F3-F4), the probabilities of being in non-advanced stages were set to 0 (stage 1-3 in Figure 1). The set of non-null probabilities were then renormalized and were used as input data for the prediction 5-year weighted risk of liver complications. The model also predicts the extreme values (slowest and fastest) of this 5-year risk based on the patient's current stage.

Supplementary Table 1. Histological stage for the biopsied group before and after missing partial information imputation

	Before imputation	After imputation
Normal liver (%)	224 (14)	224 (14)
Pure steatosis – F0 (%)	347 (22)	347 (22)
Fibrosis <F4 without ASH (%)	335 (21)	335 (21)
F1 (%)	76 (23)	171 (51)
F2 (%)	47 (14)	89 (27)
F3 (%)	41 (12)	75 (22)
Missing METAVIR score	171 (51)	-
Fibrosis <F4 with ASH (%)	123 (8)	123 (8)
F0 (%)	13 (11)	43 (35)
F1 (%)	8 (7)	27 (22)
F2 (%)	12 (10)	31 (25)
F3 (%)	8 (7)	22 (18)
Missing METAVIR score	82 (67)	-
Compensated cirrhosis – F4 without ASH (%)	134 (8)	134 (8)
Compensated cirrhosis – F4 with ASH (%)	87 (5)	87 (5)
Complications of cirrhosis without ASH (%)	185 (12)	185 (12)
Complications of cirrhosis with ASH (%)	164 (10)	164 (10)

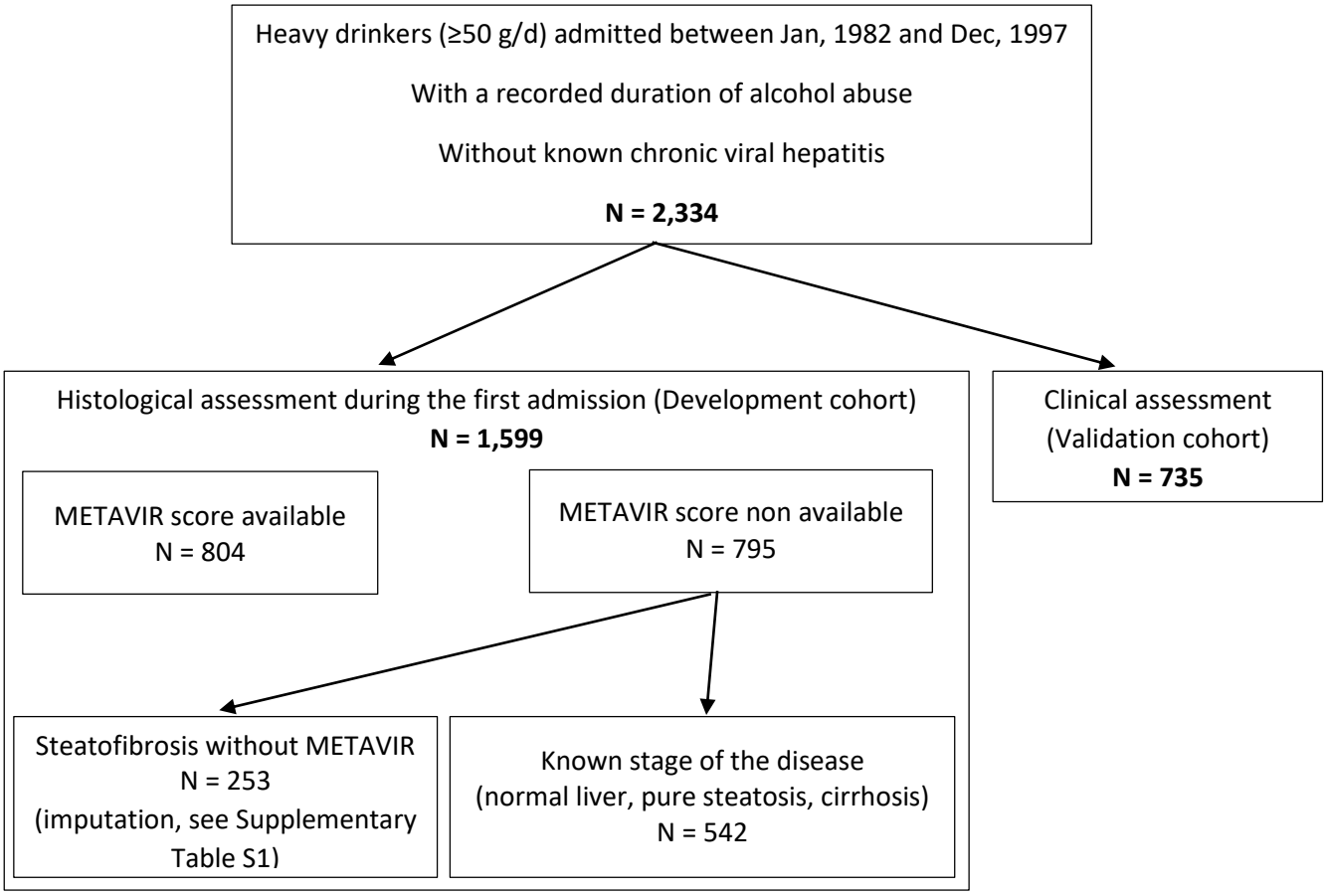
Supplementary Table 2. Adequacy between observed and predicted stage in the Development cohort

	Observed (n=1599)	**Predicted (n=1599)**
Normal liver	224	296
Steatosis-F0-F2	607	587
ASH-F0-F2	101	76
Steatosis-F3-F4	209	200
ASH-F3-F4	109	120
Liver complications from steatosis-F3-F4	185	157
Liver complications from ASH-F3-F4	164	163

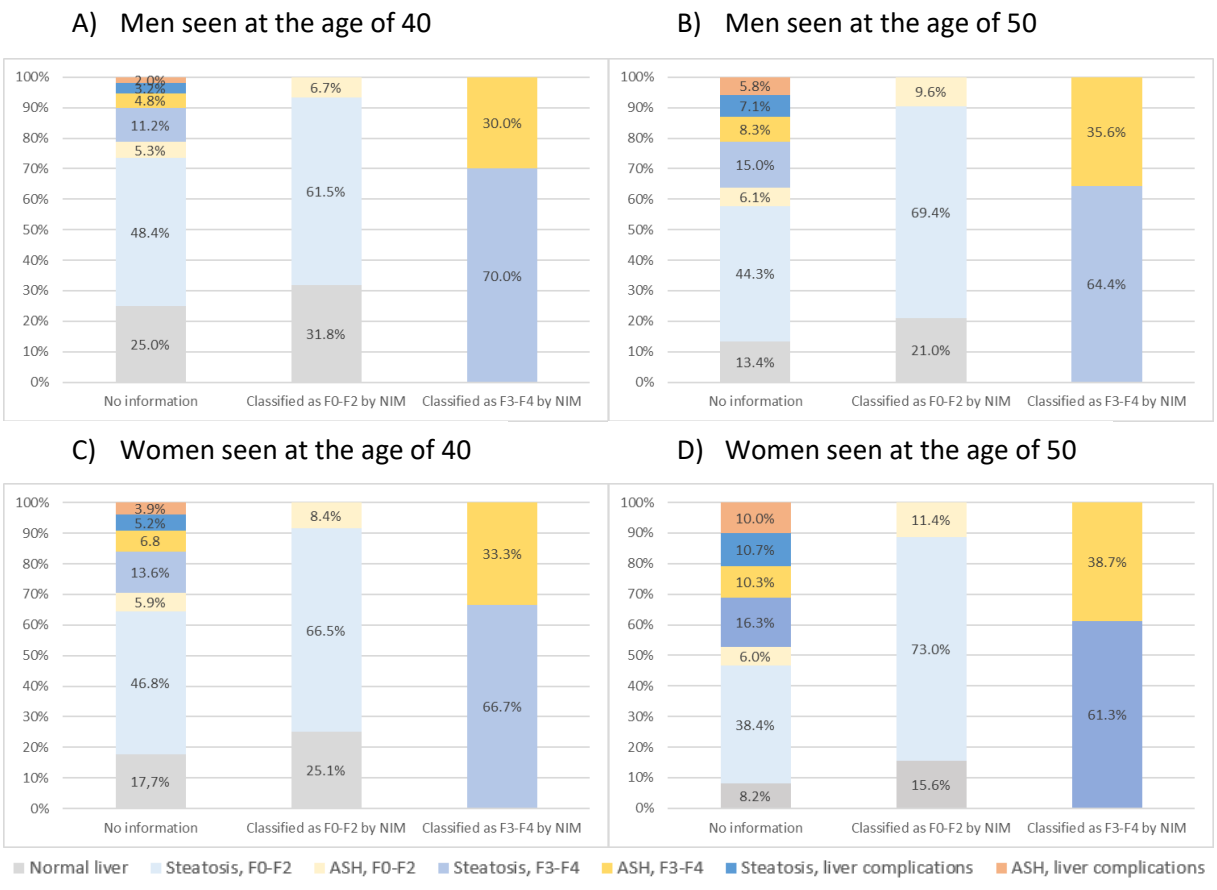
Supplementary Table 3. Adequacy between observed and predicted stage in the Validation cohort

	Observed (n=735)	**Predicted (n=735)**
Compensated liver disease	519	553
Liver complications	216	182

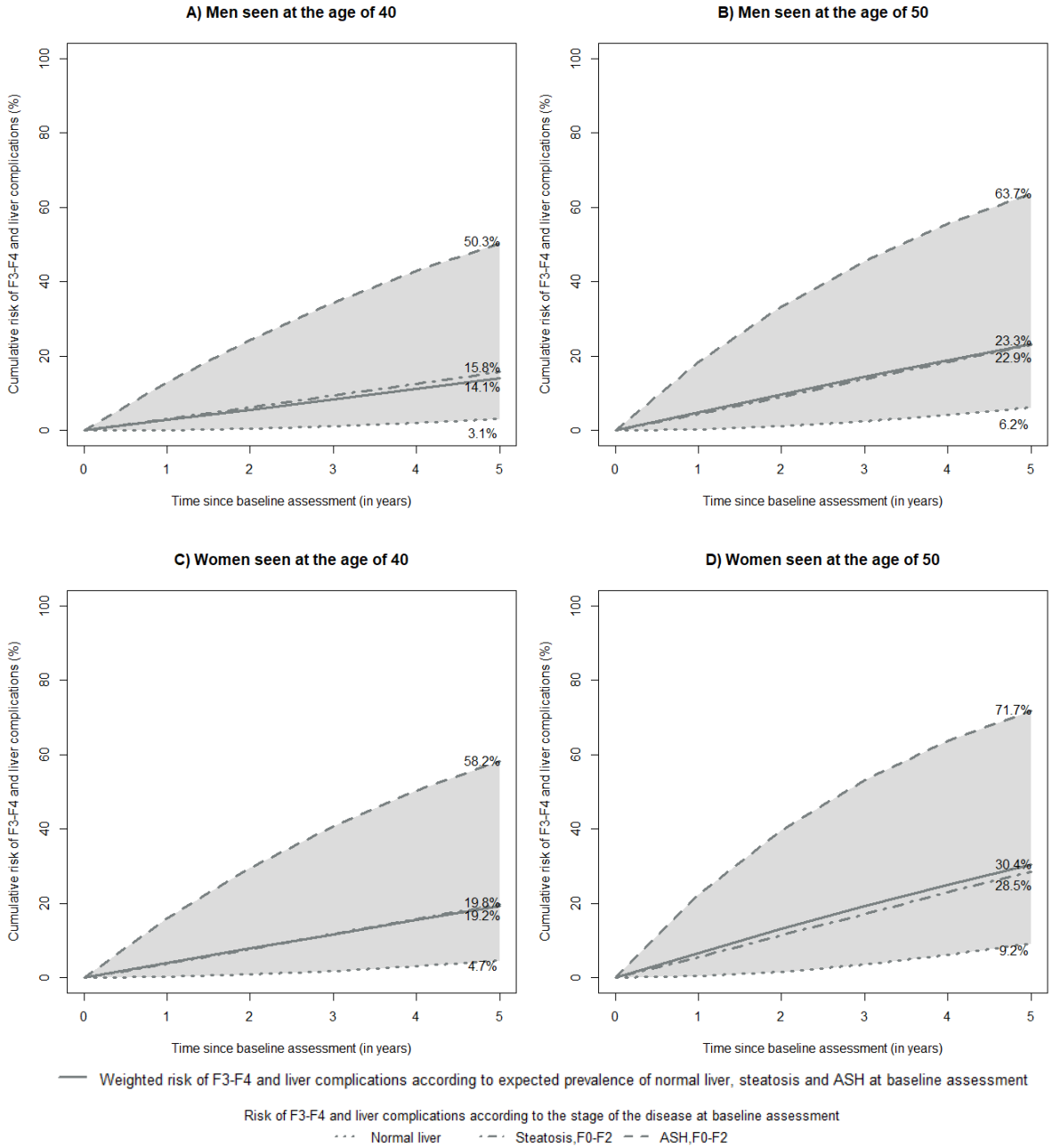
Supplementary Figure 1. Flow chart concerning the distribution of the 2,334 included patients



Supplementary Figure 2. Estimated disease progression according to three different levels of information of baseline stage: without information, classified as F0-F2 by non-invasive methods (NIM) and classified as F3-F4 by NIM. Four profiles of individuals are illustrated: A) Men seen at the age of 40, B) Men seen at the age of 50, C) Women seen at the age of 40, D) Women seen at the age of 50. All of them have a BMI of 22 kg/m² and have been drinking 150 g/d for 15 years. The distribution of the barplot "no information" is directly predicted by the model. The procedure to obtain prevalence according to the result of the non-invasive methods (NIM) is detailed for profile A as follows. When the NIM classifies the patient as non-severe fibrosis and considering this result to be reliable, the patient currently has either a normal liver, steatosis-F0-F2 or ASH-F0-F2. Considering the model's prediction, that is, 25.0% to have a normal liver, 48.4% to have steatosis-F0-F2 and 5.3% to have ASH-F0-F2, based on the rule of three, the recalculated probability of having a normal liver is 31.8% (25.0/(25.0+48.4+5.3)), steatosis-F0-F2 is 61.5% (48.4/(25.0+48.4+5.3)) and ASH-F0-F2 is 6.7% (5.3/(25.0+48.4+5.3)). In the same way, when the NIM classifies patients with severe fibrosis, the recalculated probabilities of having steatosis-F3-F4 and ASH-F3-F4 are 70.0% and 30.0%, respectively.



Supplementary Figure 3. Prediction of risk of F3-F4 (with or without ASH) and liver complications over 5 years for four profiles of individuals classified either with a normal liver, steatosis-F0-F2 or ASH-F0-F2 by non-invasive tests: A) Men seen at the age of 40, B) Men seen at the age of 50, C) Women seen at the age of 40, D) Women seen at the age of 50. All of them have a BMI of 22 kg/m² and have been drinking 150 g/d for 15 years.



Additional References

1. Available from: https://cran.r-project.org/web/packages/msm/vignettes/msm-manual.pdf.
2. Intraobserver and interobserver variations in liver biopsy interpretation in patients with chronic hepatitis C. The French METAVIR Cooperative Study Group. Hepatology. 1994;20(1 Pt 1):15-20.
3. Bedossa P, Poynard T. An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group. Hepatology. 1996;24(2):289-93.