Enhanced Liver Fibrosis Test Predicts Transplant-free Survival in Primary Sclerosing Cholangitis, a Multi-center Study

Elisabeth MG de Vries[1], Martti Färkkilä[2], Piotr Milkiewicz[3], Johannes R. Hov[4,5,6,7], Bertus Eksteen[8], Douglas Thorburn[9], Olivier Chazouillères[10], Albert Pares[11], Ståle Nygård[12,13,14], Odd Helge Gilja[15,16], Ewa Wunsch[17], Pietro Invernizzi[18], Marco Carbone[18], Francesca Bernuzzi[18], Kirsten Muri Boberg[4,5,7], Helge Røsjø[14,19], William Rosenberg[9], Ulrich H. Beuers[1], Cyriel Y Ponsioen[1], Tom H. Karlsen[4,5,6,7], Mette Vesterhus[4,15]

[1]Department of Gastroenterology and Hepatology, Academic Medical Center, Amsterdam, the Netherlands; [2]Helsinki University, Clinic of Gastroenterology, Helsinki University Hospital, Helsinki, Finland; [3]Liver and Internal Medicine Unit, Medical University of Warsaw, Warsaw, Poland; [4]Norwegian PSC Research Center, Division of Surgery, Inflammatory Diseases and Transplantation, Oslo University Hospital, Rikshospitalet, Oslo, Norway

[5]Institute of Clinical Medicine, University of Oslo, Oslo, Norway; [6]K.G. Jebsen Inflammation Research Centre, University of Oslo, Oslo, Norway; [7]Section of Gastroenterology, Department of Transplantation Medicine, Division of Surgery, Inflammatory Diseases and Transplantation, Oslo University Hospital, Oslo, Norway; [8]Snyder Institute of Chronic Diseases, Department of Medicine, University of Calgary, Calgary, Canada; [9]UCL Institute for Liver and Digestive Health, Division of Medicine, University College London & Royal Free London, NHS Foundation Trust, London, UK; [10]Department of Hepatology, Hôpital Saint Antoine, Paris, France; [11]Liver Unit, Hospital Clínic, IDIBAPS, CIBERehd, University of Barcelona, Barcelona, Spain; [12]Bioinformatics Core Facility, Institute for Medical Informatics, Oslo University Hospital, Oslo, Norway; [13]Institute for Experimental Medical Research, Oslo University Hospital and University of Oslo, Oslo, Norway; [14]Center for Heart Failure Research, University of Oslo, Oslo, Norway; [15]National Centre for Ultrasound in Gastroenterology, Haukeland University Hospital, Bergen, Norway; [16]Department of Clinical Medicine, University of Bergen, Bergen, Norway; [17]Department of Translational Medicine, Pomeranian Medical University, Szczecin, Poland; [18]Center for Autoimmune Liver Diseases, Humanitas Clinical and Research Center, Rozzano (MI), Italy; [19] Division of Medicine, Akershus University Hospital, Lørenskog, Norway.

Correspondence

Dr Mette Vesterhus, Department of Medicine, Haukeland University Hospital, Jonas Lies vei 49, N-5021 Bergen, Norway

Tel: +47 55976119; *Fax:* +47 55 97 29 50; *Email:* mette.vesterhus@helse-bergen.no

Electric word count manuscript: 4996

Number of tables: 3

Number of figures: 2

Authors contribution

MV designed the study, coordinated the collection of serum samples, performed interpretation of the data and drafting of the manuscript, and supervised the project. EdV collected patient data, performed the statistical analyses, interpretation of the data and prepared the first draft of the manuscript. MF, PM, BE, OC, AP, EW, PI, MC, FB, and CP identified PSC patients that were included in the study, collected clinical patient data, and contributed patient sera for ELF test. SN performed statistical analyses. JRH, OHG and KMB contributed to interpretation of the data. DT and WR contributed to ELF test analyses and interpretation of results. HR contributed to the designing, performance and interpretation of statistical analyses. THK contributed to the designing and interpretation of the study and drafting of the manuscript. All authors reviewed the manuscript for critical content, and approved the final version.

Conflicts of Interest:

William Rosenberg is among the inventors and patent holders for the ELF® Test and receives consultancy fees from Siemens.

Financial support:

MV is funded by the Norwegian PSC Research Center.

List of abbreviations:

AIH = autoimmune hepatitis

APRI = AST to platelet ratio index

AST = aspartate aminotransferase

AUC = area under the curve

CCA = cholangiocarcinoma

ELF = enhanced liver fibrosis

IBD = inflammatory bowel disease

IQR = interquartile range

OR = odds ratio

PIIINP = amino-terminal pro-peptide of type III pro-collagen

PSC = primary sclerosing cholangitis

ROC = receiver operating characteristic

TIMP-1 = tissue inhibitor of metalloproteinases-1

.

Abstract

Background and aims

Biomarkers reflecting disease activity and prognosis in primary sclerosing cholangitis (PSC) have not been firmly established. Enhanced Liver Fibrosis (ELF) test was previously reported to predict outcome in PSC. We aimed to validate the prognostic utility of ELF test in an independent, multicenter, retrospective PSC study population.

Methods

We collected serum samples from PSC patients from seven countries. We estimated rates of transplant-free survival by the Kaplan–Meier method, used Cox proportional hazards regression to explore the association between ELF test and clinical outcome and determined prognostic performance of ELF test by computing the area under the receiver operating characteristic (AUC-ROC) curve.

Results

The final analysis included 534 PSC patients (61% males). Features of autoimmune hepatitis or concomitant inflammatory bowel disease affected 44 (8%) and 379 (71%) patients, respectively. ELF test levels were higher in patients reaching the combined endpoint liver transplantation or death (median 10.9 [interquartile range (IQR) 9.8-12.1]; n=24 deaths, 79 liver transplantations) compared to those censored (8.8 [IQR 8.0-9.8]); $p<0.001$. ELF test expressed as mild, moderate and severe fibrosis was significantly associated with the risk of reaching the endpoint ($p<0.001$). ELF test independently predicted clinical outcome (Hazard ratio 1.31; 95% confidence interval [1.05-1.65]; $p=0.018$), and enabled good discrimination between PSC patients with and without endpoint (AUC-ROC 0.79).

Conclusion

Our retrospective data validates the predictive utility of ELF test for clinical outcomes in PSC. The clinical utility of biomarkers for fibrosis in patients with PSC should be assessed in prospective patient cohorts.

Electric word count abstract: 249

Key words: primary sclerosing cholangitis; enhanced liver fibrosis (ELF) test; risk stratification; surrogate endpoint; biomarker

Key Points

- Primary sclerosing cholangitis (PSC) is a progressive biliary disease lacking medical treatment with currently no established tools to predict prognosis in the individual patient. The lack of biomarkers for risk stratification is an important obstacle to the development of therapy.
- The Enhanced Liver Fibrosis (ELF®) test was previously reported to predict clinical outcome in two Norwegian PSC cohorts independently of clinical risk scores.
- Our data confirm, in a large, international, multicenter cohort, that ELF test predicts prognosis in PSC and may be used for risk stratification in clinical follow-up.
- Combining ELF test with clinical prognostic scores may add incremental prognostic value.

INTRODUCTION

Primary sclerosing cholangitis (PSC) is a chronic, cholestatic liver disease of unknown etiology resulting from the development of fibrotic strictures throughout the biliary tree. Eventually most patients develop fibrosis, cirrhosis and end-stage liver failure.[1] The only curative treatment modality is liver transplantation,[2] and PSC is the number one indication of liver transplantation within the spectrum of autoimmune and cholestatic liver disease.[3]

There is an unmet need for medical therapeutic options in the management of PSC patients. However, the development of new treatment strategies is hampered by the lack of prognostic markers and the overall slowly progressive nature of the disease, which results in difficulties to demonstrate treatment effects in clinical trials.[4]

Liver fibrosis is a well-established predictor of disease outcome in PSC – exemplified by the implementation of liver histology and liver elastography in several prognostic models for PSC.[5–9] Over recent years, non-invasive methods to measure liver fibrosis have gained interest, including the use of serum biomarkers. The Enhanced Liver Fibrosis (ELF) test is a promising panel, incorporating three direct serum markers of fibrosis in an algorithm: hyaluronic acid, tissue inhibitor of metalloproteinases-1 (TIMP-1), and amino-terminal pro-peptide of type III pro-collagen (PIIINP).[10,11] The ELF test accurately predicted significant liver fibrosis and furthermore predicted clinical outcome in several independent populations and in patients with various etiologies of chronic liver disease.[12–16]

Recently, the prognostic value of the ELF test in PSC was assessed in two independent Norwegian PSC cohorts.[17] The ELF test consistently predicted liver transplant-free survival in PSC patients independently of other risk factors or risk

scores.[17] In the present study, we aimed to validate the prognostic value of the ELF test in a large, multi-center PSC cohort.

PATIENTS AND METHODS

Study design, patient and tissue requirements

PSC patients from seven centers in Europe and Canada were included: Helsinki University Hospital, Finland; Medical University of Warsaw, Poland; University of Calgary, Canada; Hôpital Saint Antoine, Paris, France; Hospital Clínic, Barcelona, Spain; Humanitas Clinical and Research Center, Rozzano, Italy, and the Academic Medical Center, Amsterdam, the Netherlands. PSC diagnosis was established according to the EASL clinical practice guidelines.[18] A diagnosis of PSC with features of autoimmune hepatitis (AIH) was made in keeping with expertise of the contributing center. The individual centers received ethical approval at the national level (Suppl. Table 1). All patients provided written, informed consent.

Clinical data had previously been collected in the context of the International PSC Study Group. Where missing, additional clinical and laboratory data as well as data on liver biochemistry at time of the ELF test sample withdrawal (+/- 1 month) were retrospectively retrieved from patient files by the participating centers. IBD diagnosis was based on findings at colonoscopy and histology.

Frozen serum samples were collected from 577 PSC patients. For determination of the ELF test, serum samples were analyzed by the commercially available ELF®Test (Siemens Medical Solutions Diagnostics Inc., Tarrytown, NY, USA). The assays were performed using the Siemens ELF®Test kits and an ADVIA

Centaur XP analyzer (Siemens Medical Solutions Diagnostics Inc., Tarrytown, NY, USA).

Statistical Analysis

Patient characteristics and laboratory values were expressed as median and interquartile range. Dichotomous variables were expressed as percentage (%) of the cohort. Since reference values of biochemical variables differed slightly between centers according to local instrumentation and kit, all biochemical variables were expressed as fold change the upper or lower limit of normal of each center. Biochemical values showing a skewed distribution were transformed using natural logarithmic transformation. Continuous variables were tested for normal distribution, and for comparison between groups the Student's t-test or the Mann-Whitney U test was applied, as appropriate.

Time of PSC diagnosis was defined by the first pathological cholangiogram. A composite endpoint composed by all-cause death and liver transplantation was defined.[17] Survival time was calculated as the interval between the date of serum withdrawal for ELF test and the date of reaching the composite endpoint, or, in case no endpoint was reached, date of last follow-up.

Rates of transplant-free survival were estimated for three groups of fibrosis severity: mild, moderate and severe fibrosis defined as ELF test level <7.7, ≥7.7 to <9.8, and ≥9.8, respectively, as recommended by the manufacturer; crude risk was compared using log-rank test. Due to the small number of patients with a follow-up longer than 60 months (n=37 out of 516), survival curves were truncated at 60 months.

Univariable Cox proportional hazards model was used to assess the potential association of all clinical and biochemical variables with the occurrence of the endpoint. Factors that were significantly associated (P<0.05) with outcome in the univariable analysis were entered into the multivariable model. Using stepwise forward multivariable Cox regression analysis, the independent prognostic value of ELF test was assessed. The criterion for retaining predictors was a p-value <0.05. The proportionality during follow-up for risk prediction with ELF test as a continuous variable was found acceptable for all assays and cohorts as tested by the cox.zph function in R.

The prognostic performance of ELF test was determined by computing the area under the receiver operating characteristic (ROC) curve. The optimal threshold to distinguish patients that experience an endpoint from those that do not, was calculated by Youden's index – the maximum total sensitivity and specificity.

Correlations between ELF test and other laboratory variables were assessed by Spearman's rank correlation test.

Statistical analyses were performed using SPSS version 22.0 software (SPSS, Chicago, IL); calculation of the net reclassification index and testing for the proportional hazards assumptions were performed in R (R Foundation for Statistical Computing, Vienna, Austria). p<0.05 was considered statistically significant.

RESULTS

Patient characteristics

Serum samples of 577 PSC patients were received from the participating centers. A total of 17 samples were excluded because of insufficient serum volumes and five samples were excluded due to inability to calculate the ELF test because of undetectable (<0.50 ng/mL) or high (out of range despite 1:10 dilution) PIIINP levels in repeated analyses (n=3 and 1, respectively), or (for one patient) widely discrepant results from duplicate samples (hyaluronic acid 42.02 vs 11.79, PIIINP 21.85 vs 2.94 and TIMP1 331.9 vs 58.2). In addition, 21 patients diagnosed with small duct PSC were excluded to reduce heterogeneity. The final number of patients included was 534. Out of these, 24 patients died and 79 underwent liver transplantation (Table 1).

The median age at PSC diagnosis was 34 years (IQR 25-45), and 379 (71%) patients suffered from concurrent IBD, out of which 289 (54% of the total study population) were classified as ulcerative colitis. The median disease duration at time of serum withdrawal for ELF test analysis was 57 months (IQR 28-111). An overview of baseline characteristics and laboratory values at time of ELF test sample withdrawal is provided in Table 1.

Differentiation of PSC phenotype by ELF test score

The ELF test was higher in patients reaching an endpoint than in those censored, with medians of 10.9 (IQR 9.8-12.1) and 8.8 (IQR 8.0-9.8), respectively; p<0.001. ELF was elevated in men compared to women (median 9.2 [IQR 8.3-10.7] and 8.8 [IQR 8.0-10.0], respectively; p=0.006), and associated with more advanced disease in males as illustrated by similarly elevated bilirubin levels, APRI scoreand liver

transplants in males compared to females (data not shown). The median ELF test did not differ between patients with and without inflammatory bowel disease (median 9.1 [IQR 8.2-10.5] and 9.2 [IQR 8.2-10.4], respectively; p=0.936.

A total of 19 (4%) patients developed hepatobiliary malignancies; 3 gallbladder carcinomas, 2 hepatocellular carcinomas and 15 cholangiocarcinomas (CCA). Ten patients were diagnosed with CCA after serum withdrawal for ELF test, with a median interval of 14 months [IQR 11-24]. This subgroup of patients with CCA had a significantly higher ELF test than patients without CCA, median 10.7 [IQR 9.3-11.4] and 9.1 [IQR 8.2-10.4], respectively; p=0.035. The ELF test was not significantly different between five patients who had a diagnosis of CCA at ELF test serum withdrawal and patients with no CCA (10.5 [IQR 9.2-11.8] and 9.1 [IQR 8.0-10.2], respectively, p=0.35).

Prognostic performance of the ELF test

The manufacturer of the ELF test defines three groups of fibrosis severity based on ELF scores, i.e. none to mild, moderate, and severe (ELF score <7.7, ≥7.7 to <9.8 and ≥9.8, respectively). There was a significant association between the ELF test subdivided into three groups based on these definitions (N=81 mild, 257 moderate and 178 severe fibrosis, respectively), and the risk of reaching the clinical composite endpoint all cause death and liver transplantation, p<0.001 (Figure 1). Additional Kaplan-Meier survival analysis when applying the composite endpoint PSC related death and liver transplantation showed a comparable result (Supplementary Figure 1).

When re-classifying PSC patients in low-risk and high-risk groups based on the cut-off of ≥9.8 for severe fibrosis, there were 178 (34%) high risk and 338 (66%)

low risk patients and PSC patients. There were significantly more endpoints in the high compared to the low risk group (67 [37.6%] vs 23 [6.8%]; odds ratio (OR) 6.72 [95%CI 4.14-10.90]), and this difference persisted if patients with hepatobiliary malignancy were excluded (n=58 vs 21 endpoints, OR 8.13 [4.71-14.03]). The risk of liver transplantation alone was also higher in the high risk compared to low risk group (n=54 vs 19, OR 5.85 (95%CI 3.47-9.86). The high risk group had longer median PSC duration at ELF test withdrawal compared to the low risk group, i.e. 76 [interquartile range, (IQR) 30-121] and 51 [IQR 28-103] months, respectively; p=0.039).

The ELF test had a good discriminative ability to distinguish patients that reach an endpoint from those that do not, with an area under the curve (AUC) of 0.80 (95% CI 0.75-0.85) p<0.001 (Figure 2). The optimal threshold of the ELF test to discriminate between patients that do, and do not reach an endpoint was 9.85 (sensitivity 0.74 [0.64, 0.83], specificity 0.75 [0.71, 0.79], Youden's index: 0.50). Application of the previously identified cut-off levels for ELF test in PSC of 11.1 yielded increased specificity at the cost of reduced sensitivity (sensitivity 0.43, specificity 0.90, respectively). The discriminatory ability of the ELF test was not significantly different from that of bilirubin (AUC 0.83) or APRI score (AUC 0.80) but significantly better than Fib4 (AUC 0.73, p=0.02) and albumin (AUC 0.67, p=0.005) (Suppl. Fig. 2).

Clinical and biochemical prognostic indicators of transplant-free survival

Univariable Cox regression analysis showed a significant association between transplant-free survival and the following variables: sex, aspartate aminotransferase, alanine aminotransferase, alkaline phosphatase, total bilirubin, albumin, international

normalized ratio, platelet count, AST to platelet ratio index (APRI), and the ELF test (Table 2). Subsequent multivariable analysis including all of the variables showing associations in the univariable analyses, demonstrated an independent prognostic value of the ELF test (hazard ratio (HR) 1.31 [95% CI 1.05-1.63], p=0.016; Table 3). In addition to the ELF test, total bilirubin and albumin remained independently associated with outcome in multivariable analysis (Table 3).

DISCUSSION

This study confirms the prognostic value of ELF test in the prediction of clinical outcome in PSC, in a large, well characterized, multicenter PSC cohort. We found that the ELF test was a strong predictor of clinical outcome as defined by liver transplantation or death independent of other clinical and laboratory variables associated with outcome. One unit increase in the ELF test was associated with a 1.31-fold increased risk of death or liver transplantation.

By subdividing ELF test results into three groups of fibrosis severity based on cut-off levels provided by the manufacturer, we showed that patients with PSC can be stratified into low, intermediate and high risk groups for the composite endpoint of death or liver transplantation. Although the difference between these three groups was statistically significant, a comparison of the Kaplan-Meier curves of the present study with the original results, suggested a suboptimal ability to distinguish mild from moderate disease.[17] This may in part be explained by the use of thresholds not originally developed to differentiate disease stages in a biliary disorder with a porto-portal fibrosis pattern like PSC. However, the manufacturer's optimal cut-off to discriminate between patients with and without severe fibrosis (9.8) was similar to the optimal cut-off value to discriminate between patients that do and do not reach an

endpoint as estimated by the Youden's index in our study population (9.85) and seems to be a robust cut-off level to identify high-risk patients. Previously, higher optimal cut-off values for ELF of 11.1 and 11.2 were identified in two PSC populations;[17] application of any of these cut-off levels in the present study population yielded increased specificity at the cost of reduced sensitivity compared to a cut-off of 9.8. Further studies should aim to define clinically meaningful PSC-specific cut-off levels that might also robustly identify a low-risk group.

We report increased ELF test in patients later diagnosed with CCA (n=10) in line with previous results.[17] The ELF test was not significantly increased in five patients who had a diagnosis of CCA at serum withdrawal for ELF test analysis. Excluding patients with CCA from the analyses did not alter the association of ELF test with clinical outcome (data not shown). The present data cannot resolve the question of whether the association between ELF test and CCA in PSC reflects more advanced disease in these patients or results from the excessive fibrotic response in the surrounding tissue of the "scirrhous" type of CCA often found in PSC, potentially an early risk sign for CCA.[20,21] Dedicated analyses seem warranted to further explore the association between ELF test and CCA.

In addition to the ELF test, several other established biomarkers of fibrosis have been used in other liver diseases, including the APRI score,[22] Fibrosis-4-score,[23] and FibroTest.[24] The diagnostic performance of these biomarkers along with ELF test and liver histology was assessed in a PSC patient population that was included in a randomized trial of simtuzumab.[25] The ELF test accurately diagnosed advanced fibrosis and cirrhosis (sensitivity 97% and 79%, specificity 9% and 64%, respectively) whereas FibroTest, APRI and FIB-4 scores all had lower sensitivities (ranging 17-58%) and their main value was in excluding advanced fibrosis and

cirrhosis.[25] These results corroborate previous findings showing that baseline APRI and FIB-4 did not identify patients with higher risk of developing liver related events while ELF test did.[16]

The most widely used prognostic model in PSC research is the Mayo Risk model; however, this model notably failed to predict adverse outcomes in high-dose ursodeoxycholic acid studies. We could not compare ELF test to the Mayo risk score because of lack of reliable data on variceal bleeding. However, our data show that the ELF test predicted clinical outcome independently of all individual biochemistries identified as relevant through univariable analyses. These findings suggest that ELF test has an independent prognostic value, and that the combination of the ELF test and clinically derived prognostic models in PSC might increase prognostic power. Such composite models warrant further research. As PSC in its early stages primarily is an inflammatory disease of the biliary epithelium, it would be interesting to assess whether addition of an inflammatory marker would improve prognostication in the low-risk groups defined by ELF. Furthermore, it would be interesting to explore whether compound assessments combining ELF test with ultrasound- or MR-based liver stiffness measurements, could provide incremental prognostic information.

Whether the ELF test reflects merely fibrosis stage or also disease intensity has not been firmly established. The original paper on the development of the ELF test describes excellent correlation between ELF test and degree of fibrosis, but only moderate correlation with histological grade, suggesting that it is mostly a stage marker.[10] Exploring the dynamics of ELF test results over time, as well as its ability to measure treatment effect in terms of fibrosis regression is warranted to establish the ELF test's applicability in clinical practice and its usefulness to function as a potential surrogate endpoint in clinical trials in PSC.

Proving the clinical value of a new test, and deciding when and how to implement a new test in clinical practice are important challenges. To be clinically useful, a biomarker should be measurable by available, reliable analytical methods, add new information compared to existing markers, and guide patient management.[26] The pivotal criterion is the consistency and strength of the association between the biomarker and the outcome, and the extent to which the new marker improves prognostication by addition to or replacing established tools. External validation in at least two adequately-sized prospective studies is advised for prognostic markers in cardiac disease.[26] The ELF test is commercially available and well validated for other liver diseases.[14–16] Furthermore, ELF test has shown consistent, strong association with clinical outcome independent of clinical risk models in two independent monocenter PSC panels, and now in a large, multicenter PSC patient panel.[17] The ELF test has shown incremental value when added to the clinically based Mayo risk score. However, prospective validation is lacking and further comparisons and combinations with other biomarkers and risk scores merit investigation before implementation of ELF in clinical practice.

The retrospective nature and the lack of radiological or histological staging represent limitations to the present study. The choice of all-cause death and liver transplantation as combined end-point may also introduce elements of uncertainty based on variable indications for liver transplantation. However, in lack of gold standards, clinical outcome is a valid variable against which ELF can be benchmarked. Assessment of the dynamics of the ELF test over the disease course was not feasible because of the cross-sectional design of this study.

In conclusion, our data from a large, international, multicenter cohort confirm the prognostic value of the ELF test and its ability to stratify risk of poor outcome in

PSC. Further investigations of the clinical utility of the ELF test in prospective cohorts, is an important next step before general implementation in clinical practice can be advocated. Clinically meaningful PSC-specific cut-off values for risk stratification should be established to facilitate clinical use. Further refinements of the components of the ELF test, or compound assessments combining the ELF test with clinical scores and imaging, are avenues that should be explored in order to optimize our ability to capture risk. In an era with a considerable clinical need to identify surrogate markers of liver fibrosis and prognosis to measure treatment effect in clinical trials for PSC, investigations aiming at exploring the potential utility of the ELF test in this regard should be integrated in clinical trials.

ACKNOWLEDGEMENTS

The authors thank Liv Wenche Thorbjørnsen for assistance in the collection of serum samples.

REFERENCES

1. Williamson K, Chapman C. Primary sclerosing cholangitis: a clinical update. Br. Med. Bull. 2015; 114:56–64.

2. Hirschfield GM, Karlsen TH, Lindor KD, Adams DH. Primary sclerosing cholangitis. Lancet. 2013; 382:1587–99.

3. Adam R, Karam V, Delvart V, et al. Evolution of indications and results of liver transplantation in Europe. A report from the European Liver Transplant Registry (ELTR). J. Hepatol. 2012; 57:675–688.

4. Ponsioen CY, Chapman RW, Chazouillères O, et al. Surrogate endpoints for clinical trials in primary sclerosing cholangitis ; review and results from an International PSC Study Group consensus process. Hepatology. 2016; 63:1357–67.

5. Wiesner R, Grambsch P, Dickson E, et al. Primary sclerosing cholangitis: natural history, prognostic factors and survival analysis. Hepatology. 1989; 10:430–6.

6. Farrant J, Hayllar K, Wilkinson M, et al. Natural history and prognostic variables in primary sclerosing cholangitis. Gastroenterology. 1991; 100:1710–7.

7. Dickson E, Murtaugh P, Wiesner R, et al. Primary sclerosing cholangitis: refinement and validation of survival models. Gastroenterology. 1992; 103:1893–901.

8. Broomé U, Olsson R, Lööf L, et al. Natural history and prognostic factors in 305 Swedish patients with primary sclerosing cholangitis. Gut. 1996; 38:610–615.

9. Corpechot C, Gaouar F, El Naggar A, et al. Baseline values and changes in liver stiffness measured by transient elastography are associated with severity of fibrosis and outcomes of patients with primary sclerosing cholangitis. Gastroenterology. 2014 Apr; 146(4):970-9–6.

10. Rosenberg WMC, Voelker M, Thiel R, et al. Serum markers detect the presence

of liver fibrosis: A cohort study. Gastroenterology. 2004 Dec; 127(6):1704–1713.

11. Xie Q, Zhou X, Huang P, Wei J, Wang W, Zheng S. The performance of enhanced liver fibrosis (ELF) test for the staging of liver fibrosis: a meta-analysis. PLoS One. 2014 Apr; 9(4):e92772.

12. Parkes J, Roderick P, Harris S, et al. Enhanced liver fibrosis test can predict clinical outcomes in patients with chronic liver disease. Gut. 2010 Sep; 59(9):1245–51.

13. Mayo MJ, Parkes J, Adams-Huet B, et al. Prediction of clinical outcomes in primary biliary cirrhosis by serum enhanced liver fibrosis assay. Hepatology. 2008 Nov; 48(5):1549–57.

14. Guha IN, Parkes J, Roderick P, et al. Noninvasive markers of fibrosis in nonalcoholic fatty liver disease: Validating the European Liver Fibrosis Panel and exploring simple markers. Hepatology. 2008 Feb; 47(2):455–60.

15. Parkes J, Guha IN, Roderick P, et al. Enhanced Liver Fibrosis (ELF) test accurately identifies liver fibrosis in patients with chronic hepatitis C. J. Viral Hepat. 2011 Jan; 18(1):23–31.

16. Irvine KM, Wockner LF, Shanker M, et al. The Enhanced liver fibrosis score is associated with clinical outcomes and disease progression in patients with chronic liver disease. Liver Int. 2016; 36:370–7.

17. Vesterhus M, Hov JR, Holm A, et al. Enhanced Liver Fibrosis Score Predicts Transplant-Free Survival in Primary Sclerosing Cholangitis. Hepatology. 2015; 62:188–197.

18. European Association for the Study of the Liver. EASL Clinical Practice Guidelines: management of cholestatic liver diseases. J. Hepatol.2009 Aug;

51(2):237–67.

19. de Vries E, Wang J, Williamson KD, et al. A novel prognostic model for primary sclerosing cholangitis. Hepatology. 2015; 62:244A–245A.

20. Lazaridis KN, Gores GJ. Primary Sclerosing Cholangitis and Cholangiocarcinoma. Semin. Liver Dis. 2006; 26:42–51.

21. Kajiyama K, Maeda T, Takenaka K, Sugimachi K, Tsuneyoshi M. The Significance of Stromal Desmoplasia in Intrahepatic Cholangiocarcinoma: A Special Reference of “Scirrhous-type” and “Nonscirrhous-type” Growth. Am. J. Surg. Pathol. 1999; 23:892–902.

22. Lin ZH, Xin YN, Dong QJ, et al. Performance of the aspartate aminotransferase-to-platelet ratio index for the staging of hepatitis C-related fibrosis: An updated meta-analysis. Hepatology. 2011; 53:726–736.

23. Sterling RK, Lissen E, Clumeck N, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology. 2006; 43:1317–1325.

24. 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; 357:1069–1075.

25. Bowlus CL, Muir P, Guha IN, et al. Validation of serum fibrosis marker panels in patients with primary sclerosing cholangitis (PSC) in a randomized trial of simtuzumab. Hepatology. 2015; 62:519A.

26. Morrow DA, De Lemos JA. Benchmarks for the assessment of novel cardiovascular biomarkers. Circulation. 2007; 115:949–952.

Table 1. Patient characteristics and clinical outcomes

	Total panel		Risk groups defined by ELF test using manufacturer's cutoff levels					
			<7.7		≥7.7 - <9.8		≥9.8	
Patient characteristics		*N*		*N*		*N*	*Median(IQR) or N (%)*	*N*
N	534		85		266		183	
Male [n (%)]	324 (61)		42 (49)		156 (59)		126 (69)	
Age at ELF withdrawal (years) [median (IQR)]	40 (30-52)	516	37 (29-48)	81	42 (31-53)	257	41 (30-53)	178
Age at diagnosis PSC (years) [median (IQR)]	34 (25-45)	534	31 (23-42)	85	35 (26-45)	266	34 (24-44)	183
AIH overlap [n (%)]	44 (8)	534	6 (7)	85	24 (9)	266	14 (8)	183
Inflammatory bowel disease [n (%)]	379 (71)	534	59 (69)	85	185 (70)	266	135 (74)	183
Ulcerative colitis [n (%)]	289 (54)	534	42 (49)	85	139 (52)	266	108 (59)	183
Crohn's disease [n (%)]	63 (12)	534	15 (18)	85	31 (12)	266	17 (9)	183
Unspecified [n (%)]	27 (5)	534	2 (2)	85	15 (6)	266	10 (6)	183
Disease duration at ELF withdrawal (months) [median (IQR)]	57 (28-111)	516	45 (29-86)	81	53 (28-111)	257	76 (29.8-121.25)	178
Follow up time from ELF withdrawal (months) [median (IQR)]	23 (5-39)	516	26 (0-36)	81	24 (5-41)	257	17 (5-38)	178
Death [n (%)]	24 (5)	534	0 (0)	85	7 (3)	266	17 (9)	183
PSC related death [n (%)]	15 (3)	534	0 (0)	85	4 (2)	266	11 (6)	183
Liver transplantation [n (%)]	79 (15)	534	2 (2)	85	22 (8)	266	55 (30)	183
Liver transplantation for end-stage liver disease	35 (7)	534	1 (1)	85	12 (5)	266	22 (12)	183
Liver transplantation for CCA or high-grade dysplasia	4(1)	534	0 (0)	85	2 (1)	266	2 (1)	183
Liver transplantation for intractable symptoms	12 (2)	534	0 (0)	85	4 (2)	266	8 (4)	183

Liver transplantation, indication not available	28 (5)	534	1 (1)	85	4 (2)	266	23 (13)	183
Laboratory values at time of ELF withdrawal								
AST xULN, [median (IQR)]	1.04 (0.69-2.05)	338	0.71 (0.57-0.98)	54	0.84 (0.67-1.25)	172	2.11 (1.48-3.09)	112
ALT xULN, [median (IQR)]	1.15 (0.66-2.26)	351	0.66 (0.42-1.00)	59	0.94 (0.60-1.74)	178	2.09 (1.23-3.39)	114
ALP xULN, [median (IQR)]	1.35 (0.81-2.52)	362	0.86 (0.64-1.28)	59	1.03 (0.73-1.81)	183	2.55 (1.62-3.81)	120
Total bilirubin xULN [median (IQR)]	1.23 (0.59-2.73)	339	0.60 (0.40-1.10)	51	1.00 (0.55-2.15)	175	2.93 (1.44-2.12)	113
Albumin xLLN [median (IQR)]	1.14 (1.04-1.24)	302	1.14 (1.05-1.22)	49	1.17 (1.08-1.26)	151	1.07 (0.97-1.20)	102
INR [median (IQR)]	1.00 (1.00-1.10)	257	1.00 (1.00-1.10)	36	1.00 (0.96-1.10)	128	1.10 (1.00-1.20)	93
Platelet count xLLN [median (IQR)]	1.57 (1.21-2.01)	331	1.77 (1.37-2.02)	54	1.64 (1.39-2.05)	169	1.28 (0.69-1.74)	108
Creatinine xULN [median (IQR)]	0.65 (0.57-0.76)	299	0.69 (0.61-0.81)	45	0.67 (0.59-0.78)	152	0.60 (0.50-0.70)	102
APRI [median (IQR)]	0.44 (0.28-1.12)	304	0.28 (0.22-0.39)	49	0.35 (0.26-0.61)	155	1.42 (0.79-2.42)	100
ELF test [median (IQR)]	9.11 (8.19-10.48)	534	7.10 (0.62-0.74)	85	8.68 (8.29 - 9.23)	266	11.01 (10.45-11.90)	183

AIH, autoimmune hepatitis; ALT, alanine aminotransferase; ALP, alkaline phosphatase; ALT, alanine aminotransferase; APRI, AST to platelet ratio index; AST, aspartate aminotransferase; CI, confidence interval; ELF, enhanced liver fibrosis score; HR, Hazard ratio; INR, International normalized ratio; IQR, inter quartile range; PSC, primary sclerosing cholangitis; xLLN = times lower limit; xULN = times upper limit of normal.

Table 2. Predictors of transplant-free survival in PSC assessed by univariable Cox regression analysis

	Univariable analysis		Data Available
	HR (95% CI)	*p*-value	N
Sex	0.53 (0.33, 0.84)	*0.007*	534
Age at ELF withdrawal	0.99 (0.98, 1.01)	0.27	516
Age at PSC diagnosis	0.99 (0.98, 1.01)	0.30	534
Co-existing IBD	1.34 (0.83, 2.17)	0.23	534
Co-existing IBD phenotype	0.99 (0.77, 1.29)	0.96	534
PSC duration at ELF withdrawal (months)	1.00 (1.00, 1.00)	0.90	516
Auto-immune hepatitis overlap	0.79 (0.34, 1.80)	0.57	534
Center of inclusion	1.12 (0.99, 1.26)	0.07	534
AST	3.19 (2.24, 4.54)	*<0.005*	338
ALT	2.18 (1.61, 2.96)	*<0.005*	351
ALP	2.92 (2.03, 4.18)	*<0.005*	362
Total bilirubin	4.28 (2.81, 6.53)	*<0.005*	339
Albumin	0.11 (0.04, 0.28)	*<0.005*	302
International normalized ratio	4.08 (1.99, 8.38)	*<0.005*	257
Platelet count	0.40 (0.25, 0.63)	*<0.005*	331
Creatinine	0.37 (0.05, 2.51)	0.31	299
APRI	1.45 (1.27, 1.65)	*<0.005*	304
ELF test	1.77 (1.58, 1.99)	*<0.005*	534

AST, ALT ALP and total bilirubin were transformed by the natural logarithm prior to regression analyses due to a right-skewed distribution. AIH, autoimmune hepatitis; ALT, alanine aminotransferase; ALP, alkaline phosphatase; APRI, AST to platelet ratio index; AST, aspartate aminotransferase; CI, confidence interval; ELF, enhanced liver fibrosis score; HR, Hazard ratio; INR, International normalized ratio; IQR, inter quartile range; PSC, primary sclerosing cholangitis.

Table 3. Multivariable Cox regression analysis, assessing independent predictors of transplant-free survival in PSC patients. The analysis included all of the variables showing significant associations with transplant-free survival in the univariable analyses (sex, AST, ALT, ALP, albumin, bilirubin, INR, thrombocytes and ELF test) for n=219 patients with complete data available. Omitting INR from the analysis in order to increase the number of patients with available data, yielded the same final model with similar HRs in n=256 patients (data not shown). Laboratory values were entered using value times the upper or lower limit of normal as appropriate. AST, ALT ALP and total bilirubin were transformed by the natural logarithm prior to regression analyses due to a right-skewed distribution. ALT, alanine aminotransferase; ALP, alkaline phosphatase; AST, aspartate aminotransferase; CI, confidence interval; ELF, enhanced liver fibrosis score; HR, Hazard ratio; INR, International normalized ratio; PSC, primary sclerosing cholangitis.

	Multivariable analysis	
	HR (95% CI)	***p*-value**
Total bilirubin	2.91 (1.50-5.64)	*0.002*
Albumin	0.12 (0.03-0.50)	*0.004*
ELF test	1.31 (1.05-1.63)	*0.016*

FIGURES

Figure 1. Prediction of transplant-free survival by the ELF test.

The figure shows Kaplan-Meier curves of time to transplantation or death for PSC patients (n=516) stratified into groups of mild, moderate and severe fibrosis defined as ELF <7.7, ≥7.7 to <9.8, and ≥9.8, respectively, as recommended by the manufacturer; illustrating shorter survival in patients in the group with severe fibrosis as defined by the ELF test compared to patients with intermediate and low ELF levels. ELF test, enhanced liver fibrosis test; PSC, primary sclerosing cholangitis.

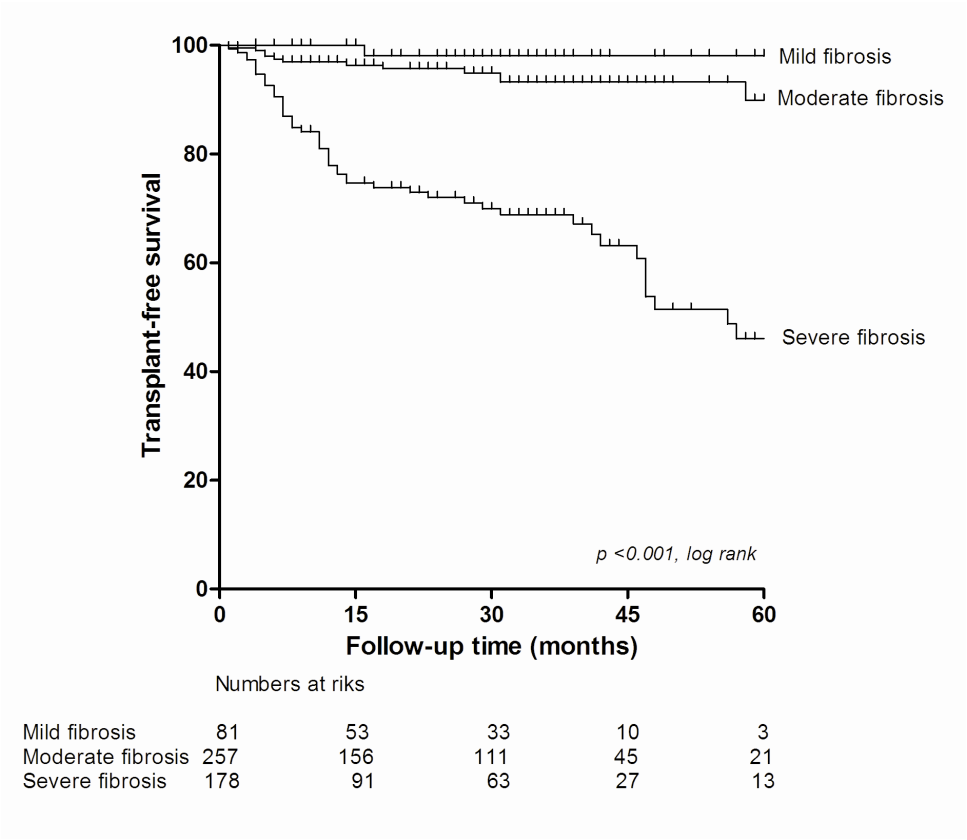

Figure 2. Prognostic performance of the ELF test.

The prognostic performance of the ELF test was assessed by analysis of the area under the curve of the receiver operator characteristics curve (AUC-ROC). The ELF test distinguished patients that reached liver transplantation or death from those that did not with an area under the curve of 0.80 (95% CI [0.75, 0.85]), p<0.001, demonstrating a good discriminatory ability. The optimal threshold of the ELF test to discriminate between patients that did, and did not reach an endpoint was 9.85 (sensitivity 0.74, specificity 0.75). AUC-ROC, area under the curve of the receiver operator characteristics curve; ELF test, enhanced liver fibrosis test; PSC, primary sclerosing cholangitis.