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], Ewa Wunsch[16], Pietro Invernizzi[17], Marco Carbone[17], Francesca Bernuzzi[17], Kirsten Muri Boberg[4,5,6,7], Helge Røsjø[14,18], 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 Cancer Medicine, Surgery 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, 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, Department of Medicine, Haukeland University Hospital, Bergen, Norway

[16] Department of Clinical and Molecular Biochemistry, Pomeranian Medical University, Szczecin, Poland

[17] Center for Autoimmune Liver Diseases, Humanitas Clinical and Research Center, Rozzano (MI), Italy

[18] Division of Medicine, Akershus University Hospital, Oslo, 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: 4580 (text body + refs)

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.

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

Conflicts of Interest:

William Rosenberg and Massimo Pinzani are among the inventors and patent holders for the ELF® Test. William Rosenberg 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 (Word count: 250)

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. The aim of this study was 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, out of which 324 (61%) were male. Features of autoimmune hepatitis or concomitant inflammatory bowel disease affected 44 (8%) and 379 (71%), 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]) 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 was demonstrated to independently predict 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 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.

Key Points (4 bullet points; max 100 words)

- 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 validated as surrogate endpoints 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 end-stage liver failure.[1] In the absence of medical treatment options 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 severely hampered by the slowly progressive nature and low event rate of this disease, which results in difficulties to demonstrate treatment effects in clinical trials.[4] New developments in the design of clinical trials in PSC show a shift from measuring treatment effect by the traditionally used ‘solid endpoints’ such as death and liver transplantation, to an effort to establish and implement easier attainable ‘surrogate endpoints’ that have been proven to be predictive of clinical outcome.[4]

Liver fibrosis is a well-established predictor of 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 was shown to accurately predict significant liver fibrosis and was furthermore able to predict 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 a large single center study from Norway including two independent PSC cohorts.[17] It was demonstrated that 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

For this multicenter cohort study, PSC patients from seven centers in six European countries and Canada were included: Helsinki University Hospital, Helsinki, Finland; Medical University of Warsaw, Warsaw, Poland; University of Calgary, 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 ethic approval at the national level and all patients provided written, informed consent.

Clinical data had previously been collected in the context of the international immunochip-subphenotyping project, conducted by the International PSC Study Group *(Alberts R et al. Genotype-phenotype analysis across 130 422 genetic*

variants identifies RSPO3 as the first genome-wide significant modifier gene in primary sclerosing cholangitis. ECCO 2016). 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. We collected the following clinical data: gender, date of PSC diagnosis, PSC phenotype, concomitant inflammatory bowel disease (IBD), type of IBD and date of IBD diagnosis, autoimmune hepatitis overlap syndrome, the occurrence of hepatobiliary malignancy, liver transplantation and death or date of last follow-up. IBD diagnosis was based on findings at colonoscopy and histology. Patients with a diagnosis of small duct PSC and patients for whom successful serum analysis of ELF test was not achieved, were excluded from the study.

Deep 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 when having a skewed distribution, dichotomous variables were expressed as percentage (%) of the cohort. Since reference values of biochemical variables differed slightly in between centers according to local instrumentation and kit, all biochemical variables were expressed using a ratio of times upper or lower limit of normal. 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 the ELF test and the date of reaching the first 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 the 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 the 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.

Correlation between the ELF test and the prognostic index as calculated by the novel Amsterdam-Oxford prognostic model including PSC type, age at PSC diagnosis, albumin, platelets, aspartate aminotransferase, alkaline phosphatase and bilirubin,[19] was assessed by Spearman's rank correlation test. To assess the additive value of ELF over routine biomarkers as used in the Amsterdam-Oxford model the prediction increment of the ELF test when added as a predictor to this prognostic model for PSC was calculated by the category-free net reclassification index.

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 for adequate ELF® test assessment, and five samples were excluded following analysis 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. The final number of patients included in the statistical analysis was 534.

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. 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, nor between patients with and without colorectal carcinoma (n=14), median 9.3 [IQR 8.6-11.0] and 9.1 [IQR 8.2-10.5], respectively; p=0.264)

A total of 19 (4%) patients developed hepatobiliary malignancies; 3 gallbladder carcinomas, 2 hepatocellular carcinomas and 15 cholangiocarcinomas. The ELF test was significantly higher in the 15 CCA patients compared with non CCA patients, median 10.7 [IQR 9.0-11.3] and 9.1 [IQR 8.2-10.4], respectively; p=0.024. Ten patients developed their cholangiocarcinoma after serum withdrawal for ELF test, with a median interval of 14 months [IQR 11-24]. This subgroup of patients with cholangiocarcinoma also had a significantly higher ELF test than patients without cholangiocarcinoma, median 10.7 [IQR 9.3-11.4] and 9.1 [IQR 8.2-10.4], respectively; p=0.035.

Prognostic performance of the ELF test

The manufacturer of the ELF test defines three groups of fibrosis severity based on ELF test 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 of 0.796 (95% CI 0.746-0.846) 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). Applying 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 for both, specificity 0.89 and 0.90, respectively).

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), the prognostic index (defined in Methods), and the ELF test (Table 2). Subsequent multivariable analysis demonstrated an independent prognostic value of ELF test (hazard ratio (HR) 1.31 [95% CI 1.05-1.65], p=0.018; Table 3). In addition to ELF test, total bilirubin and albumin remained independently associated with outcome in multivariable analysis.

There was a significant correlation between ELF test and the prognostic index of the Amsterdam-Oxford prognostic model [19] with a correlation coefficient of 0.537, p<0.001. The net reclassification index (NRI) of adding ELF test to this prognostic index was highly significant, when ELF test was added as a continuous variable or a categorical variable using the thresholds provided by the manufacturer (p<0.001 for both).

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, when comparing the Kaplan-Meier curves of the present study with the original results, a diminished ability to distinguish mild from moderate disease was suggested.[17] This may in part be due to the use of thresholds proposed by the manufacturer, that were not originally developed to differentiate mild, moderate and severe disease in a biliary disease 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), therefore this 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]; this yielded increased specificity at the cost of reduced sensitivity when applied to the present study population. Further studies in larger patient panels should search to define PSC-specific cut-off levels that might also robustly identify a low-risk group.

We report increased ELF test in patients diagnosed with cholangiocarcinoma (n=15) in line with previous results.[17] The ELF test was increased also in the subgroup (n=10) of patients who developed cholangiocarcinoma after serum withdrawal for ELF test analysis. However, the small numbers warrant caution. The present data do not allow conclusions regarding whether the association between increased ELF test and CCA in PSC is a result of more advanced disease in these patients or is explained by the “desmoplastic” and “scirrhous” type of CCA often found in PSC, showing an excessive fibrotic response in the surrounding tissue, which may be captured by ELF test.[20,21] Dedicated analyses seem warranted to further explore the association between ELF test and cholangiocarcinoma. Regardless, the association of ELF test with CCA does not perturb the significance of the ELF test as a risk indicator for clinical outcome. Additional analyses excluding patients with CCA showed similar results (data not shown).

Traditionally, liver biopsy has been considered the reference standard to stage the degree of liver fibrosis. The role of histology in the clinical management of PSC is not firmly established. Histologic assessment in PSC can be informative for assessment of comorbidities, and to estimate prognosis.[22–24] On the other hand, the usefulness of liver biopsy is limited by the risk of complications for the patient, sampling error due to a patchy disease distribution, and intra- and inter-observer variability.[25] Therefore, measurement of liver fibrosis through liver histology is increasingly replaced by non-invasive modalities in the most prevalent chronic liver diseases.[26] In PSC, however, there is no consensus yet concerning which of these techniques should be preferred.

The development of new non-invasive techniques to measure fibrosis, can be subdivided into two different approaches.[27] First, a ‘biological approach’, based on

serum biomarkers that represent serum levels of products arising from fibrogenesis and matrix turnover, as a derivate of degree of fibrosis.[27] Secondly, a ‘physiological approach’, assessing liver stiffness as a proxy for the degree of scarring of the liver using imaging techniques such as vibration controlled transient elastography or magnetic resonance electrography.[27] Measurement of serum biomarkers has the advantage of being easily applicable, inter-laboratory reproducible and generally available. Furthermore, serum biomarkers are not influenced by operator experience, food intake, congestion or obesity.[28–30]

In addition to the ELF test, several other established biomarkers of fibrosis have been used in other liver diseases, including the AST to platelet ratio index (APRI)[31], Fibrosis-4-score (an algorithm including age, AST, platelet count and ALT)[32], and FibroTest (an algorithm including age, Alpha-2-macroglobulin, Haptoglobin, Apolipoprotein A1, gamma-glutamyl transpeptidase, total bilirubin, and ALT).[33] The diagnostic performance of these biomarkers along with ELF test and liver histology has been assessed in a PSC patient population that were included in a randomized trial of simtuzumab.[34] They showed that ELF test could accurately diagnose advanced fibrosis and cirrhosis with a sensitivity of 97% and 79%, and a specificity of 9% and 64%, respectively whereas the main value of FibroTest (sensitivity 58% and 58; specificity 81% and 91%, respectively), APRI (sensitivity 17% and 38%; specificity 95% and 97%, respectively) and FIB-4 (sensitivity 44% and 26%; specificity 78% and 99%, respectively) was in excluding advanced fibrosis and cirrhosis.[34] 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] Our data from a large, international, multicenter cohort confirm previous findings showing surplus value of ELF test as prognosticator in PSC, when

compared with other biomarkers of fibrosis, but validation of the ELF test in a prospective cohort is an important next step before general implementation in clinical practice can be advocated.[16] Given the discrepant results regarding the ability of ELF to segregate a specific low-risk group in need of less vigilant follow-up, further investigations regarding PSC-specific optimal cut-off values is also needed to establish the optimal clinical management of this group.

In accordance with the previous study [17], we used the composite endpoint all-cause death and liver transplantation for time-to event analyses. The majority of patients died of PSC related causes, and large differences between the use of all-cause death or more specifically PSC related death in this composite endpoint were not expected. Indeed, Kaplan-Meier survival analysis assessing the association between the ELF test and the composite endpoint PSC related death and liver transplantation showed a similar result.

The most widely used prognostic model in PSC research is the Mayo Risk model. We could not compare nor assess the value of ELF test as an adjunct to the Mayo risk score because of lack of reliable data on variceal bleeding due to the retrospective nature of our study. The Mayo risk score, however, failed to predict adverse outcomes in high-dose ursodeoxycholic acid studies and has not been validated for individual prognostication. Recently, the novel Amsterdam-Oxford prognostic model for PSC was presented which was validated in an external PSC cohort, and is also applicable in early stage disease.[19] The ELF test showed a moderate correlation with the prognostic index of this model (including the variables PSC type, age at PSC diagnosis, albumin, platelets, aspartate aminotransferase, alkaline phosphatase and bilirubin), and calculation of the net reclassification index showed a significantly increased precision of predicting survival when adding ELF

test results. These findings suggest that ELF test has an independent prognostic value in addition to current routine biomarkers, and that the combination of the ELF test and clinically derived prognostic models in PSC might yield increased prognostic power, and such composite models warrant further research. Furthermore, it would be interesting to explore whether compound assessments combining ELF test with ultrasound- or MR-based liver stiffness measurements and perhaps also clinical scores, could provide incremental prognostic information.

This study was subject to some limitations. Firstly, in the absence of data on histological stage, correction for biopsy determined stage of fibrosis in the multivariable analysis was not possible. Secondly, assessment of the dynamics of the ELF test over the disease course was not feasible because of the cross-sectional design of this study. 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] The three basic components of the ELF test represent various elements of fibrogenesis and matrix turnover (PIIINP reflects both formation and degradation of pro-collagen III, TIMP1 regulates extracellular matrix degradation and HA has various roles in wound healing); thus, ELF test is not a clear-cut fibrosis formation marker. 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 PSC.

In conclusion, we confirm the prognostic value of the ELF test and its ability to stratify risk of poor outcome in PSC. The clinical utility of the ELF test should be

further investigated in prospective cohorts. Studies exploring refinements of the components of the ELF test or compound assessments combining the ELF test with clinical scores and imaging aiming at incremental prognostic information, might contribute to improved prognostication. 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, the ELF test is a promising candidate and investigations regarding this aspect 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 [Internet]. 2013 [cited 2014 May 28]; 382:1587–99.

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

4. Ponsioen CY, Chapman RW, Chazouillères O, Gideon M, Karlsen TH, Lohse AW, et al. Surrogate endpoints for clinical trials in primary sclerosing cholangitis ; review and results from an International PSC Study Group consensus process. Hepatology. Ahead of p.

5. Wiesner R, Grambsch P, Dickson E, Ludwig J, MacCarty R, Hunter 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, Karani J, Portmann B, Westaby D, et al. Natural history and prognostic variables in primary sclerosing cholangitis. Gastroenterology. 1991; 100:1710–7.

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

8. Broomé U, Olsson R, Lööf L, Bodemar G, Hultcrantz R, Danielsson A, 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, Kemgang A, Wendum D, Poupon R, 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 [Internet]. 2014 [cited 2014 Oct 22]; 146:970-9–6.

10. Rosenberg WMC, Voelker M, Thiel R, Becka M, Burt A, Schuppan D, et al. Serum markers detect the presence of liver fibrosis: A cohort study. Gastroenterology [Internet]. 2004 [cited 2014 Oct 30]; 127:1704–1713.

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 [Internet]. 2014 [cited 2014 Oct 28]; 9:e92772.

12. Parkes J, Roderick P, Harris S, Day C, Mutimer D, Collier J, et al. Enhanced liver fibrosis test can predict clinical outcomes in patients with chronic liver disease. Gut [Internet]. 2010 [cited 2014 Oct 30]; 59:1245–51.

13. Mayo MJ, Parkes J, Adams-Huet B, Combes B, Mills a S, Markin RS, et al. Prediction of clinical outcomes in primary biliary cirrhosis by serum enhanced liver fibrosis assay. Hepatology [Internet]. 2008 [cited 2014 Oct 31]; 48:1549–57.

14. Guha IN, Parkes J, Roderick P, Chattopadhyay D, Cross R, Harris S, et al. Noninvasive markers of fibrosis in nonalcoholic fatty liver disease: Validating the European Liver Fibrosis Panel and exploring simple markers. Hepatology [Internet].

2008 [cited 2014 Oct 10]; 47:455–60.

15. Parkes J, Guha IN, Roderick P, Harris S, Cross R, Manos MM, et al. Enhanced Liver Fibrosis (ELF) test accurately identifies liver fibrosis in patients with chronic hepatitis C. J. Viral Hepat. [Internet]. 2011 [cited 2014 Oct 29]; 18:23–31.

16. Irvine KM, Wockner LF, Shanker M, Fagan KJ, Horsfall LU, Fletcher LM, 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, Schrumpf E, Nygård S, Godang K, 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. [Internet]. 2009 [cited 2014 Aug 11]; 51:237–67.

19. de Vries E, Wang J, Williamson KD, Leeflang MM, Boonstra K, Beuers U, et al. A novel prognostic model for primary sclerosing cholangitis. Hepatol. :244A–245A.

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

21. Kajiyama, Kiyoshi M.D.; Maeda, Takashi M.D., Ph.D.; Takenaka, Kenji M.D., Ph.D.; Sugimachi, Keizo M.D., Ph.D.; Tsuneyoshi, Masazumi M.D. PD. 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. Sebastiani G, Alberti A. Non invasive fibrosis biomarkers reduce but not substitute the need for liver biopsy. World J. Gastroenterol. 2006; 12:3682–3694.

23. De Vries EMG, Verheij J, Hubscher SG, Leeflang MMG, Boonstra K, Beuers U, et al. Applicability and prognostic value of histologic scoring systems in primary sclerosing cholangitis. J. Hepatol. 2015; 63:1212–1219.

24. de Vries E, Verheij J, Farkkila M, Arola J, Goeppert B, Schirmacher P, et al. Validation of the prognostic value of histologic scoring systems in primary sclerosing cholangitis; a multicenter cohort study. J. Hepatol. 2016; EASL Barce.

25. Portmann B, Zen Y. Inflammatory disease of the bile ducts-cholangiopathies: liver biopsy challenge and clinicopathological correlation. Histopathology [Internet]. 2012 [cited 2014 Aug 6]; 60:236–48.

26. Corpechot C. Utility of Noninvasive Markers of Fibrosis in Cholestatic Liver Diseases. Clin. Liver Dis. 2016; 20:143–158.

27. Castera L. Noninvasive Assessment of Liver Fibrosis. Dig. Dis. 2015; :325–335.doi:10.1002/hep.24013

28. Ehlken H, Lohse AW, Schramm C. Transient elastography in primary sclerosing cholangitis-the value as a prognostic factor and limitations. Gastroenterology [Internet]. 2014 [cited 2014 Nov 4]; 147:542–3.

29. Tapper EB, Castera L, Afdhal NH. FibroScan (Vibration-Controlled Transient Elastography): Where Does It Stand in the United States Practice. Clin. Gastroenterol. Hepatol. [Internet]. 2014 [cited 2014 Nov 4]; doi:10.1016/j.cgh.2014.04.039

30. Casťera L, Foucher J, Bernard PH, Carvalho F, Allaix D, Merrouche W, et al. Pitfalls of liver stiffness measurement: A 5-year prospective study of 13,369 examinations. Hepatology. 2010; 51:828–835.

31. Lin ZH, Xin YN, Dong QJ, Wang Q, Jiang XJ, Zhan SH, 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.

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

33. 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.

34. Bowlus CL, Muir P, Guha IN, Chapman RW, Chazouillères O, Chalasani NP, et al. Validation of serum fibrosis marker panels in patients with primary sclerosing cholangitis (PSC) in a randomized trial of simtuzumab. 2015; 1.

FIGURE LEGENDS

Figure 1. Prediction of transplant-free survival by the ELF test, patients divided into groups of mild, moderate and severe fibrosis based on ELF levels using Siemens' cutoffs.

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 0.796 (95% CI 0.746-0.846), 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.

Table 1. Patient characteristics

N	534
Male [n (%)]	324 (61)
Age at diagnosis PSC (years) [median (IQR)]	34 (25-45)
AIH overlap [n (%)]	44 (8)
Inflammatory bowel disease [n (%)]	379 (71)
Ulcerative colitis [n (%)]	289 (54)
Crohn's disease [n (%)]	63 (12)
Unspecified [n (%)]	27 (5)
Disease duration at ELF withdrawal (months) [median (IQR)]	57 (28-111)
Follow up time from ELF withdrawal (months) [median (IQR)]	23 (5-39)
Death [n (%)]	24 (5)
PSC related death [n (%)]	15 (3)
Liver transplantation [n (%)]	79 (15)
Laboratory values at time of ELF withdrawal	
AST xULN, [median (IQR)]	1.04 (0.69-2.05)
ALT xULN, [median (IQR)]	1.15 (0.66-2.26)
ALP xULN, [median (IQR)]	1.35 (0.81-2.52)
Total bilirubin xULN [median (IQR)]	1.23 (0.59-2.73)
Albumin xLLN [median (IQR)]	1.14 (1.04-1.24)
INR [median (IQR)]	1.00 (1.00-1.10)
Platelet count xLLN [median (IQR)]	1.57 (1.21-2.01)
Creatinine xULN [median (IQR)]	0.65 (0.57-0.76)
APRI [median (IQR)]	0.44 (0.28-1.12)
PI Amsterdam=Oxford model PSC [median (IQR)]	1.68 (1.34-2.28)
ELF test [median (IQR)]	9.11 (8.19-10.48)

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

* Data available for: AST n= 338, ALT n= 351, ALP n=362, total bilirubin n=339, albumin n=302 INR n = 257, platelet count n=331, creatinine n = 299, APRI n =304, PI = 270; Follow up time from ELF withdrawal = 516

Table 2. Univariable Cox regression analysis, assessing predictors of transplant-free survival

	Data Available **N**	**Univariable analysis** **HR (95% CI)**	***p*-value**
Sex	534	0.53 (0.33, 0.84)	*0.007*
Age at PSC diagnosis	534	0.99 (0.98, 1.01)	0.301
Co-existing IBD Y_N	534	1.34 (0.83, 2.17)	0.226
Co-existing IBD type	534	0.99 (0.77, 1.29)	0.956
Disease duration at ELF withdrawal (months)	516	1.00 (1.00, 1.00)	0.896
Auto-immune hepatitis overlap	534	0.79 (0.34, 1.80)	0.570
Center of inclusion	534	1.12 (0.99, 1.26)	0.066
AST	338	3.19 (2.24, 4.54)	*<0.005*
ALT	351	2.18 (1.61, 2.96)	*<0.005*
ALP	362	2.92 (2.03, 4.18)	*<0.005*
Total bilirubin	339	4.28 (2.81, 6.53)	*<0.005*
Albumin	302	0.11 (0.04, 0.28)	*<0.005*
International normalized ratio	257	4.08 (1.99, 8.38)	*<0.005*
Platelet count	331	0.40 (0.25, 0.63)	*<0.005*
Creatinine	299	0.37 (0.05, 2.51)	0.306
APRI	304	1.45 (1.27, 1.65)	*<0.005*
PI Amsterdam-Oxford model	270	2.53 (1.83, 3.50)	*<0.005*
HA	534	1.001 (1.00, 1.00)	*<0.005*
PIIINP	534	1.04 (1.03, 1.05)	*<0.005*
TIMP1	534	1.002 (1.00, 1.00)	*<0.005*
ELF test	534	1.77 (1.58-1.99)	*<0.005*

PSC = primary sclerosing cholangitis; IQR = inter quartile range; AIH = autoimmune hepatitis; mm = millimeter; AST = aspartate aminotransferase; ALT = alanine aminotransferase; ALP = alkaline phosphatase; INR =International normalized ratio; APRI = AST to platelet ratio index; PI = prognostic index; ELF = enhanced liver fibrosis; HR = Hazard ratio; CI = confidence interval.
AST, ALT ALP and total bilirubin were transformed by the natural logarithm prior to regression analyses due to a right-skewed distribution.
Univariable Cox analyses; n=516 patients with follow-up.

Table 3. Multivariable Cox regression analysis, assessing independent predictors of transplant-free survival in PSC patients

	Multivariable analysis **HR (95% CI)**	***p*-value**
Total bilirubin	3.44 (1.79-6.63)	*0.000*
Albumin	0.12 (0.03-0.50)	*0.003*
ELF test	1.31 (1.05-1.65)	*0.018*

ELF = enhanced liver fibrosis ; HR = Hazard ratio; CI = confidence Interval.
Multivariable Cox analysis; n=516 patients with follow-up.

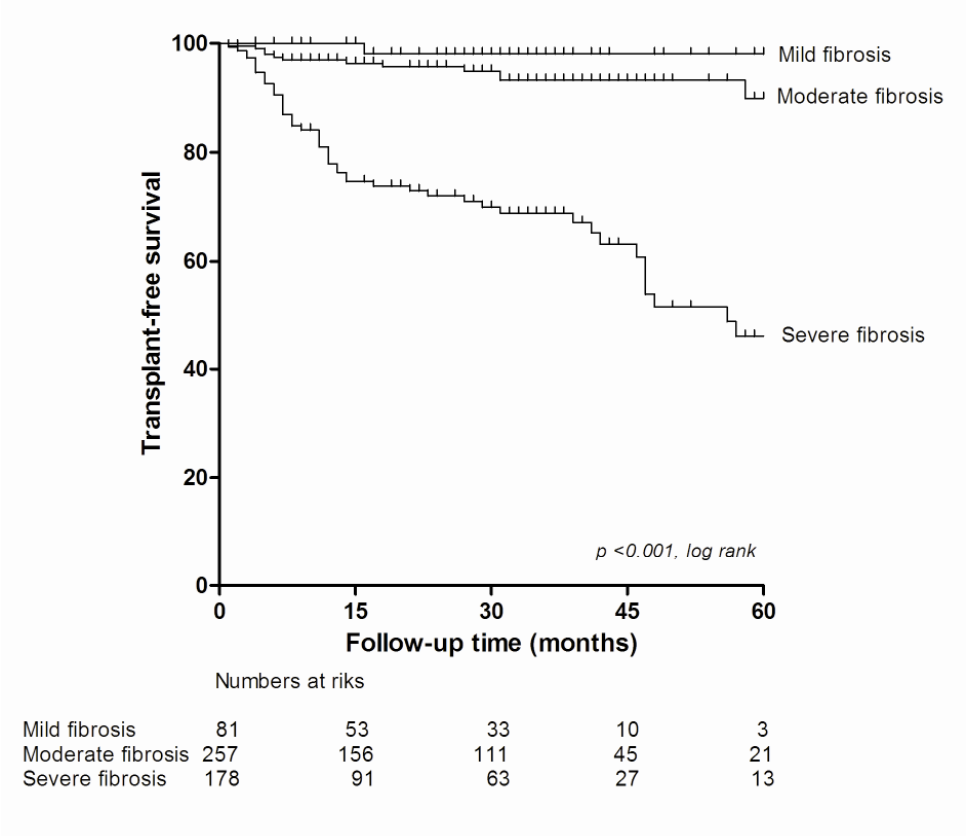


Figure 1. Prediction of transplant-free survival by the ELF test, patients divided into groups of mild, moderate and severe fibrosis based on ELF levels using Siemens' cutoffs.
Kaplan Meier Curves includes n=516 patients with follow-up.

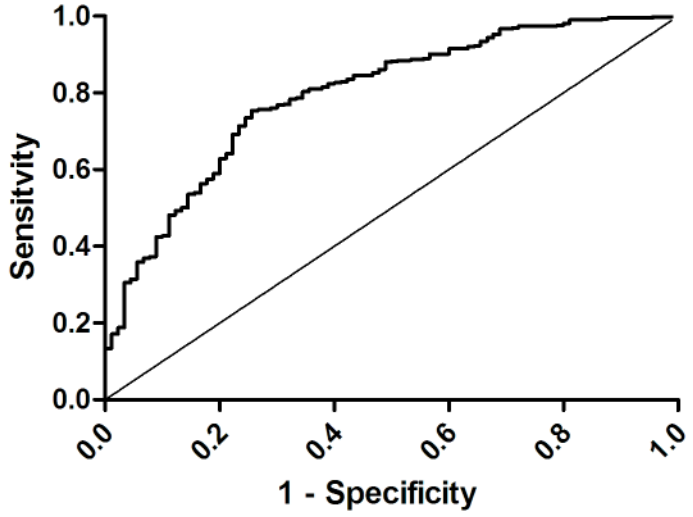
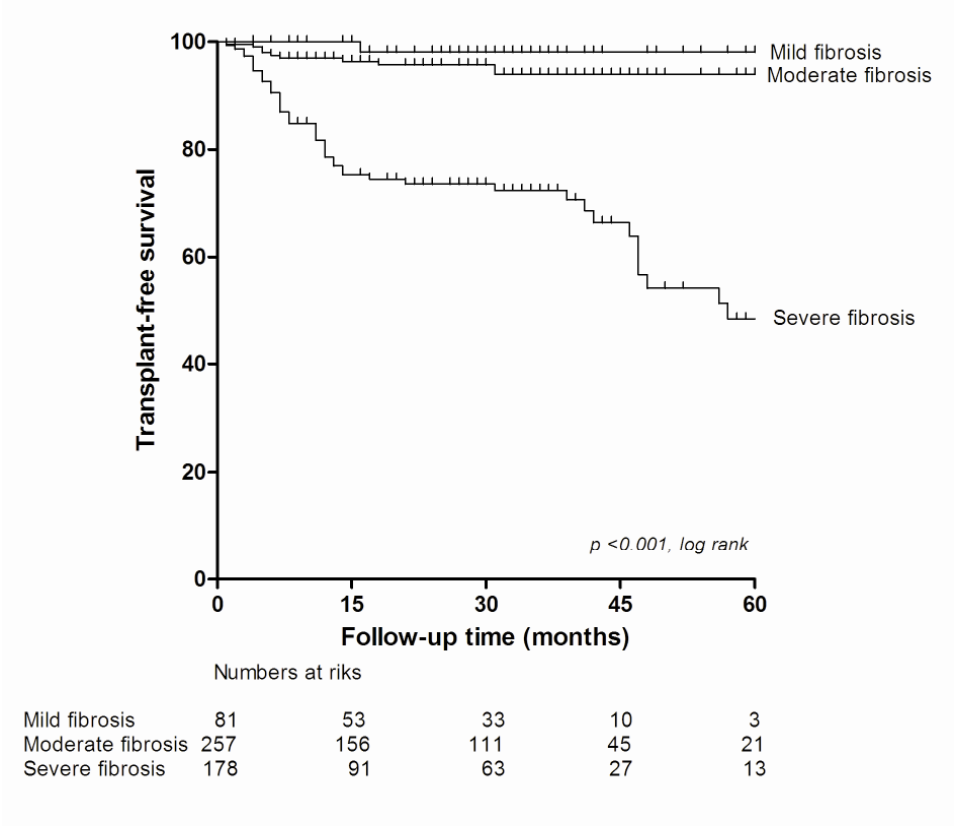


Figure 2 Receiver operator curve, ELF test.



Supplementary Figure 1. Prediction of transplant-free survival by the ELF test, patients divided into groups of mild, moderate and severe fibrosis based on ELF levels using Siemens' cutoffs.
Endpoint PSC related death and liver transplantation;
Kaplan Meier Curves includes n=516 patients with follow-up.