UvA-DARE (Digital Academic Repository)

Biomarkers for disease progression in primary sclerosing cholangitis

de Vries, E.M.G.

Publication date

2016

Document Version

Final published version

[Link to publication](#)

Citation for published version (APA):

de Vries, E. M. G. (2016). *Biomarkers for disease progression in primary sclerosing cholangitis*. [Thesis, fully internal, Universiteit van Amsterdam].

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

BIOMARKERS FOR DISEASE PROGRESSION IN PSC-LIESBETH DE VRIES

**Biomarkers for disease progression in
primary sclerosing cholangitis**

ISBN: 978-94-028-0380-8

Cover design: Don Soepboer

www.donsoepboer.com

Lay-out: Persoonlijk Proefschrift, persoonlijkproefschrift.nl

Printing: Ipskamp Printing

Copyright 2016, E.M.G. de Vries

All rights reserved. No parts of this publication may be reproduced, stored, or transmitted in any form or by any means, without written permission of the author or publishers of the included scientific papers.

The printing of this thesis was financially supported by: AbbVie, Academisch Medisch Centrum, Dr. Falk Pharma Benelux, Nederlandse Vereniging voor Gastroenterologie, Olympus, PSC Patients Europe, Takeda, Tramedico, US Endoscopy, Zambon.

**Biomarkers for disease progression
in primary sclerosing cholangitis**

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op vrijdag 18 november 2016, te 10.00 uur
door

Elisabeth Maria Gerarda de Vries

geboren te Leiderdorp

Promotor:

Prof. dr. U.H.W. Beuers

Universiteit van Amsterdam

Copromotor:

Dr. C.Y. Ponsioen

Universiteit van Amsterdam

Overige leden:

Prof. dr. P.L.M. Jansen

Universiteit van Amsterdam

Prof. dr. T. Roskams

Katholieke Universiteit Leuven

Prof. dr. T.M. van Gulik

Universiteit van Amsterdam

Dr. E.A.J. Rauws

Universiteit van Amsterdam

Dr. H.R. van Buuren

Erasmus Universitair Medisch Centrum

Faculteit der Geneeskunde

TABLE OF CONTENTS

Chapter 1	9
Introduction and outline	
Adapted from: Primaire scleroserende cholangitis	
<i>Critical care 2013;2:12-14</i>	
Chapter 2	19
Risk factors for primary sclerosing cholangitis	
<i>Liver International 2016;36:84-91</i>	
Chapter 3	39
Biomarkers for disease progression of primary sclerosing cholangitis	
<i>Current Opinion in Gastroenterology 2015;31:239-46</i>	
Chapter 4	57
Alkaline phosphatase at diagnosis of primary sclerosing cholangitis and one year later: an evaluation of prognostic value	
<i>Liver International 2016; epub ahead of print</i>	
Chapter 5	77
Enhanced liver fibrosis test predicts transplant-free survival in primary sclerosing cholangitis, a multi-center study	
<i>Manuscript in preparation</i>	
Chapter 6	95
Genetic variants associated with disease phenotype and course in primary sclerosing cholangitis	
<i>Submitted</i>	
Chapter 7	123
Applicability and prognostic value of histologic scoring systems in primary sclerosing cholangitis	
<i>Journal of Hepatology 2015;63:1212-9</i>	

Chapter 8	149
Validation of the prognostic value of histologic scoring systems in primary sclerosing cholangitis; an international cohort study <i>Submitted</i>	
Chapter 9	179
A novel prognostic model for transplant-free survival in primary sclerosing cholangitis <i>Manuscript in preparation</i>	
Chapter 10	203
Summary, discussion, future perspectives	
Nederlandse samenvatting	214
List of publications	219
Contributing authors	221
PhD portfolio	228
Dankwoord	230
About the editor	235

1

Introduction and outline

Adapted from:

Primaire scleroserende cholangitis

Elisabeth M.G. de Vries, Cyriel Y. Ponsioen

Critical care 2013;2:12-14

Primary sclerosing cholangitis

Primary sclerosing cholangitis (PSC) is a puzzling, chronic, cholestatic liver disease. Inflammation of the intra-and/or extrahepatic bile ducts leads to the formation of fibro-sclerotic strictures interspersed with dilatations, distributed in a beaded pattern that is typical for PSC.¹ (Figure 1) In a slowly progressive manner, destruction of the biliary tree results in cholestasis, biliary cirrhosis, and eventually liver failure.²

Two PSC subtypes can be distinguished: large duct PSC, mainly affecting the medium-/large sized intra and/or extrahepatic bile ducts, and small duct PSC which is confined to the small intrahepatic bile ducts. The typical PSC patient is a male patient (male-to female ratio 2:1) of around 40 years old.³ Approximately 70% of patients with PSC are diagnosed with concomitant inflammatory bowel disease (IBD), mainly ulcerative colitis.³ PSC can occur in association with other auto-immune diseases, such as auto-immune hepatitis (AIH), which is commonly referred to as PSC-/AIH-overlap syndrome.⁴ In addition, PSC predisposes to the development of malignancies, with an increased risk of cholangiocarcinoma, gall-bladder carcinoma, and colorectal carcinoma.³

PSC is thought to be a rare, so called 'orphan' disease, with an incidence and prevalence of 0.5, and 6 per 100.000 inhabitants, respectively, in the Netherlands.³ Incidence and prevalence are highly variable throughout the world, and seem to be increasing alongside the increased incidence and prevalence rates of IBD.⁵⁻⁷

Pathogenesis

The exact aetiopathogenesis of PSC is still largely unknown, but the process is likely to be multifactorial.⁸ There is evidence that genetic susceptibility factors are involved, with human leukocyte antigen (HLA) on chromosome 6 remaining to be the strongest link to PSC.⁹⁻¹¹ Furthermore, recent genome wide association studies have resulted in an increased insight in the genetic architecture of PSC, with the identification of 15 non-HLA risk loci at time of writing.¹²⁻¹⁴ Environmental predisposing factors are likely to play an additional role in triggering the disease. Overall, the strong association between HLA haplotypes¹⁵, and concurrent diagnosis of at least one auto-immune diseases outside the liver and the bowel in 20-25% of PSC patients¹⁶, suggest that PSC is an immune-mediated disorder, that acts by a pathways shared with other auto-immune diseases.

Elaborating on the strong association between PSC and IBD, the 'gut homing lymphocyte hypothesis' elucidates on the missing link between the liver and the gut.¹⁷ This hypothesis postulates that memory T-lymphocytes, primed in the inflamed gut, undergo enterohepatic circulation and subsequently trigger

portal inflammation via aberrantly expressed adhesion molecules in the liver and the gut. Sustained reaction may result in chronic inflammation and cholangiocyte damage, as observed in PSC.

Clinical presentation and diagnosis

PSC patients often present with complaints of fatigue, abdominal pain, itch, jaundice, or fever in case of bacterial cholangitis. However, approximately 40% of patients are asymptomatic at time of diagnosis.¹⁸ During physical examination signs of portal hypertension can be found, such as hepatomegaly or splenomegaly.¹⁹ Laboratory examination typically shows a cholestatic pattern, with an increased γ -glutamyl transferase, and increased alkaline phosphatase – the hallmark of PSC.¹⁹ In addition there can be a mild transaminitis, with aspartate aminotransferase and alanine aminotransferase elevated up to 2-3 times the upper limit of normal.¹⁹ Biochemical values can be highly fluctuating during the disease course, and an elevation in cholestatic parameters – including bilirubin which is typically normal at diagnosis – can indicate an obstruction of the bile ducts by a (dominant) stricture, cholelithiasis which is increasingly prevalent in PSC patients, or bacterial cholangitis.

To establish PSC diagnosis, additional imaging of the bile ducts is required. With the advent of high quality magnetic resonance cholangiography, yielding excellent sensitivity and specificity to diagnose PSC, this is nowadays the diagnostic imaging modality of first choice.^{19–21} However, if results of this cholangiogram are inconclusive, a diagnostic endoscopic retrograde cholangiography can be considered.^{19,21} In case of typical complaints, biochemical abnormalities, and an cholangiogram with PSC specific abnormalities, a diagnosis of large duct PSC can be confirmed. Of importance, the only way to establish a diagnosis of primary (of unknown origin) sclerosing cholangitis, is by excluding secondary causes of sclerosing cholangitis, such as choledocholithiasis, ischemic cholangitis, cholangiocarcinoma or IgG4-associated cholangitis, and other cholestatic diseases.^{19,22} In small duct PSC no abnormalities are observed on cholangiography, and this diagnosis can only be established by liver biopsy.^{19,21} Histologically, the early PSC stage is characterized by varying degrees of portal and periportal inflammation. As the disease progresses, there is loss of small and medium-sized ducts, and development of secondary changes related to chronic cholestasis, including ductular reaction and the accumulation of copper binding protein in periportal hepatocytes.²³ In a small proportion of patients periductal concentric (onion skinned) fibrosis can be found, typical for PSC diagnosis. (Figure 2)

Figure 1. PSC cholangiogram. Characteristic anomalies of the intra- and extrahepatic bile ducts.

Figure 2. PSC liver biopsy. Periductal concentric fibrosis.

Treatment

The unknown aetiopathogenesis of PSC has complicated the development of an appropriate medical therapy. Immunosuppressants, antibiotics, and therapeutic bile acids have extensively been tested, but unfortunately so far without success on event-free survival in PSC.²⁴ To date, there are no medical therapies that can alter the natural disease course. The only curative treatment option for PSC patients is liver transplantation, and PSC is one of the leading indications of liver transplantations in the Western world.²⁵ In clinical practice, patients are often treated with low-medium dose of ursodeoxycholic acid, which seems to have a beneficial effect on liver biochemistry.^{26,27}

Approximately half of the patients with PSC will develop a dominant stricture in their disease course,²¹ which is defined arbitrary as a stenosis with a lumen diameter of 1.5 mm or less of the common bile duct, and/or 1 mm or less of the right or left hepatic duct.²⁸ When a dominant stricture leads to complaints, endoscopic or percutaneous therapeutic intervention is indicated to relieve the biliary obstruction.²¹ By means of endoscopic resonance cholangiography this stricture can be reached, and treated by either balloon dilatation, or temporary stenting.

In an era of increasing attention of scientists to the pathophysiology of PSC, new clues are provided by the large scale genetics studies. Understanding the biological role that PSC risk genes have in the pathophysiology, and linking them to specific PSC disease phenotypes, guides towards the development of novel therapeutic opportunities, and may pave the way to a future personalized medicine approach.²⁴

Disease course and prognosis

The natural disease course of PSC is unpredictable, and characterized by relapses and remission which are often the result of obstructive strictures and episodes of acute cholangitis. PSC is a slowly progressive disease, leading to biliary cirrhosis in the majority of patients. Patients can suffer substantial morbidity, resulting from consequences of portal hypertension and hepatic decompensation. Cumulative risk of developing cholangiocarcinoma is reported up to 20% after 30 years, and has a high mortality.³ The time from PSC diagnosis until patients reach the combined clinical endpoint of liver transplantation and PSC related death ranges in between 13 and 21 years, as measured in referral and population based cohorts, respectively.³

Unfortunately, robust biomarkers to measure disease activity and disease complications are lacking. Many attempts have been made to predict prognosis of PSC patients, resulting in several prognostic models that have combined biomarkers for disease progression.^{18,29–36} The majority of these models were, however, based on referral center cohorts, wherefore their predictive value may be restricted to specific patient groups. At present, the most commonly used prognostic model in PSC is the Mayo risk score, composed of the variables age, bilirubin, aspartate aminotransferase, variceal bleeding and albumin.³²

Biomarkers for disease progression in PSC

In the absence of an effective medical treatment for PSC, the necessity for clinical trials that evaluate new drugs is obvious. Fortunately, an improved insight in genetic risk factors and biological pathways associated with PSC, has resulted in an increased interest of several pharmaceutical companies to develop new therapies. One of the greatest challenge in PSC research at this point is to determine which patients are suitable to be entered into clinical trials, and in addition, how to measure treatment effect in clinical trials. The slowly progressive disease course and low event rate in PSC hinders the denomination of attainable endpoints. When using the traditional ‘solid’ endpoints, such as liver transplantation or death, enormous sample sizes are required to be able to show a difference in between treatment and placebo arms.³⁷ Implementing such large studies is an impossible effort in a rare disease such as PSC, – exemplified by the largest randomized trial in PSC to date that failed to achieve its predefined sample size.³⁸ The inefficiency of solid clinical endpoints has resulted in a search for biomarkers for disease progression in PSC, that could be used as meaningful surrogate endpoints in clinical trials.³⁷ A proper surrogate endpoint should be correlated with the solid clinical endpoint,³⁹ and thereby reliably reflect the (potential) beneficial effect of the therapy under study.³⁷

Depending on the expected mechanism of action of an investigational product, investigators may want to include either early or advanced stage PSC patients in a clinical trial. A marker for disease progression could therefore – in addition to its value as a surrogate endpoint – be used to stratify patients with greatest potential benefit into clinical trials.⁴⁰ The importance of risk stratification is not restricted to clinical trials, and can be extrapolated to patient care. In a heterogeneous disease such as PSC, stratification of patient into low versus high risk of poor outcome can result in a more patient tailored counselling, and management.⁴⁰

Several categories of prognostic biomarkers can be discerned.

1. Clinical- phenotypical markers
2. Biochemical markers
3. Imaging modalities
4. Liver histology
5. Genetic markers
6. Prognostic models

With this thesis, we aimed to identify, and/or validate biomarkers with prognostic value for disease progression in PSC. Ultimately we integrated prognostic biomarkers into a novel prognostic model for patients with PSC.

REFERENCES

1. Hirschfield, G. M., Karlsen, T. H., Lindor, K. D. & Adams, D. H. Primary sclerosing cholangitis. *Lancet* 382, 1587–99 (2013).
2. Ali, A. H., Carey, E. J. & Lindor, K. D. Current research on the treatment of primary sclerosing cholangitis. *Intractable rare Dis. Res.* 4, 1–6 (2015).
3. Boonstra, K. et al. Population-based epidemiology, malignancy risk, and outcome of primary sclerosing cholangitis. *Hepatology* 58, 2045–55 (2013).
4. Trivedi, P. J. & Hirschfield, G. M. Review article: Overlap syndromes and autoimmune liver disease. *Aliment. Pharmacol. Ther.* 36, 517–533 (2012).
5. Boonstra, K., Beuers, U. & Ponsioen, C. Y. Epidemiology of primary sclerosing cholangitis and primary biliary cirrhosis: a systematic review. *J. Hepatol.* 56, 1181–8 (2012).
6. Molodecky, N. a et al. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. *Gastroenterology* 142, 46–54 (2012).
7. Molodecky, N. a et al. Incidence of primary sclerosing cholangitis: a systematic review and meta-analysis. *Hepatology* 53, 1590–9 (2011).
8. Eaton, J. E., Talwalkar, J. a, Lazaridis, K. N., Gores, G. J. & Lindor, K. D. Pathogenesis of primary sclerosing cholangitis and advances in diagnosis and management. *Gastroenterology* 145, 521–36 (2013).
9. Spurkland, a et al. HLA class II haplotypes in primary sclerosing cholangitis patients from five European populations. *Tissue Antigens* 53, 459–469 (1999).
10. Mehal, W. et al. HLA DR4 is a marker for rapid disease progression in primary sclerosing cholangitis. *Gastroenterology* 106, 106–7 (1994).
11. Boberg, K. et al. The HLA-DR3,DQ2 heterozygous genotype is associated with an accelerated progression of primary sclerosing cholangitis. *Scand J Gastroenterol* 36, 886–90. (2001).
12. Folseraas, T. et al. Extended analysis of a genome-wide association study in primary sclerosing cholangitis detects multiple novel risk loci. *J. Hepatol.* 57, 366–75 (2012).
13. Liu, J. Z. et al. Dense genotyping of immune-related disease regions identifies nine new risk loci for primary sclerosing cholangitis. *Nat. Genet.* 45, 670–5 (2013).
14. Ellinghaus, D. et al. Genome-wide association analysis in primary sclerosing cholangitis and ulcerative colitis identifies risk loci at GPR35 and TCF4. *Hepatology* 58, 1074–1083 (2013).
15. Pollheimer, M. J., Halilbasic, E., Fickert, P. & Trauner, M. Pathogenesis of primary sclerosing cholangitis. *Best Pract. Res. Clin. Gastroenterol.* 25, 727–39 (2011).
16. Saarinen, S., Olerup, O. & Broomé, U. Increased frequency of autoimmune diseases in patients with primary sclerosing cholangitis. *Am. J. Gastroenterol.* 95, 3195–3199 (2000).
17. Grant, A. J., Lalor, P. F., Salmi, M., Jalkanen, S. & Adams, D. H. Homing of mucosal lymphocytes to the liver in the pathogenesis of hepatic complications of inflammatory bowel disease. *Lancet* 359, 150–7 (2002).
18. Broomé, U. et al. Natural history and prognostic factors in 305 Swedish patients with primary sclerosing cholangitis. *Gut* 38, 610–615 (1996).
19. European Association for the Study of the Liver. EASL Clinical Practice Guidelines: management of cholestatic liver diseases. *J. Hepatol.* 51, 237–67 (2009).
20. Maneesh, D., Elmunzer, B. J., Dwame-

- na, B. & Higgins, P. Primary Sclerosing Cholangitis: Meta-Analysis of Diagnostic Performance of MR Cholangiopancreatography. *Radiology* 256, 387–396 (2010).
21. Chapman, R. et al. Diagnosis and management of primary sclerosing cholangitis. *AASLD Practice Guidelines. Hepatology* 51, 660–78 (2010).
22. Abdalian, R. & Heathcote, E. J. Sclerosing cholangitis: a focus on secondary causes. *Hepatology* 44, 1063–74 (2006).
23. Portmann, B. & Zen, Y. Inflammatory disease of the bile ducts-cholangiopathies: liver biopsy challenge and clinicopathological correlation. *Histopathology* 60, 236–48 (2012).
24. Eksteen, B. Advances and controversies in the pathogenesis and management of primary sclerosing cholangitis. *Br. Med. Bull.* 110, 89–98 (2014).
25. Adam, R. et al. Evolution of indications and results of liver transplantation in Europe. A report from the European Liver Transplant Registry (ELTR). *J. Hepatol.* 57, 675–688 (2012).
26. Wunsch, E. et al. Prospective evaluation of ursodeoxycholic acid withdrawal in patients with primary sclerosing cholangitis. *Hepatology* 60 (3), 931–40 (2014).
27. Chazouillères, O. et al. Ursodeoxycholic acid for primary sclerosing cholangitis. *J. Hepatol.* 11, 120–123 (1990).
28. Bjornsson, E., Lindqvist-Ottosson, J., Asztely, M. & Olsson, R. Dominant Strictures in Patients with Primary Sclerosing Cholangitis. *Am. J. Gastroenterol.* 99, 502–508 (2004).
29. Wiesner, R. et al. Primary sclerosing cholangitis: natural history, prognostic factors and survival analysis. *Hepatology* 10, 430–6 (1989).
30. Farrant, J. et al. Natural history and prognostic variables in primary sclerosing cholangitis. *Gastroenterology* 100, 1710–7 (1991).
31. Dickson, E. et al. Primary sclerosing cholangitis: refinement and validation of survival models. *Gastroenterology* 103, 1893–901 (1992).
32. Kim, W. et al. A revised natural history model for primary sclerosing cholangitis. *Mayo Clin Proc.* 75, 688–94 (2000).
33. Boberg, K. M. et al. Time-dependent Cox regression model is superior in prediction of prognosis in primary sclerosing cholangitis. *Hepatology* 35, 652–7 (2002).
34. Ponsioen, C. Y. et al. Natural history of primary sclerosing cholangitis and prognostic value of cholangiography in a Dutch population. *Gut* 51, 562–6 (2002).
35. Tischendorf, J. J. W., Hecker, H., Krüger, M., Manns, M. P. & Meier, P. N. Characterization, outcome, and prognosis in 273 patients with primary sclerosing cholangitis: A single center study. *Am. J. Gastroenterol.* 102, 107–14 (2007).
36. Ponsioen, C. Y., Reitsma, J. B., Boberg, K. M., Aabakken, L. & Schrupf, E. Validation of a cholangiographic prognostic model in primary sclerosing cholangitis. *Endoscopy* 742–747 (2010).
37. Ponsioen, C. Y. 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 print*
38. Olsson, R. et al. High-Dose Ursodeoxycholic Acid in Primary Sclerosing Cholangitis: A 5-Year Multicenter, Randomized, Controlled Study. *Gastroenterology* 129, 1464–1472 (2005).
39. Prentice, R. Surrogate endpoints in clinical trials: definition and operational criteria. *Stat. Med.* 8, 431–440 (1989).

40. Trivedi, P. J., Corpechot, C., Pares, A. & Hirschfield, G. M. Risk Stratification in autoimmune cholestatic liver diseases: Opportunities for clinicians and trialists. *Hepatology* 63, 644–59 (2016).

Risk factors for primary sclerosing cholangitis

Kirsten Boonstra*, Elisabeth M.G. de Vries*, Nan van Geloven, Karel J. van Erpecum, Marcel Spanier, Alexander C. Poen, Carin M. van Nieuwkerk, Ben J. Witteman, Hans A. Tuynman, Anton H. Naber, Paul J. Kingma, Ulrich H. Beuers, Cyriel Y. Ponsioen, on behalf of the Epi PSC PBC Study Group

* These authors contributed equally to this work

Liver International 2016;36:84-91

ABSTRACT

Background and Aim

Primary sclerosing cholangitis (PSC) is a progressive cholestatic liver disease of unknown cause, but strongly associated with inflammatory bowel disease (IBD). Potential risk factors triggering PSC have never been studied on a population level. The aim of this study was to evaluate smoking, appendectomy, family history and geographical distribution in a population-based cohort of PSC patients, as compared to IBD control patients, and healthy controls (HC).

Methods

For this case-control study 343 PSC patients, 370 IBD controls and 232 HC's living in a geographically defined area in the Netherlands filled-out a questionnaire concerning smoking, appendectomy and family history of IBD and auto-immune liver diseases.

Results

Smoking was associated with a lower risk of developing PSC in PSC-ulcerative colitis (UC) patients (adjusted OR 0.21; 95% CI 0.12-0.34; $p < 0.001$). Comparable results were found for PSC-Crohn's disease (CD) patients (16% former smokers) compared to CD patients (55% former smokers) (adjusted OR 0.17; 95% CI 0.08-0.39; $p < 0.001$). Frequency of appendectomy did not differ between PSC and HC, but PSC-UC patients had undergone appendectomy more often than UC patients (13% vs. 6%) (adjusted OR 2.51; 95%CI 1.04-6.07; $p = 0.041$). We found no association between family history of IBD or auto-immune liver disease and risk of PSC. Degree of urbanization was not associated with PSC incidence.

Conclusion

In this large population-based case-control study we confirm that smoking is associated with a lower risk of developing PSC, independent of its protective effect for developing UC. Appendectomy is not associated with the risk of developing PSC.

Key points

- Smoking is associated with a lower risk of developing PSC, independent of its protective effect for developing UC.
- Appendectomy is not an independent risk factor for PSC.
- Family history of IBD and auto-immune liver disease are not independent risk factors for developing PSC, when compared to IBD patients.
- Residential environment does not seem to play a role in the development of PSC.

INTRODUCTION

Primary sclerosing cholangitis (PSC) is a puzzling cholestatic liver disease affecting the intra- and extrahepatic bile ducts. PSC is more common in men than in women (2:1) and can occur at any age with a peak incidence around 40 years.¹ Disease course is highly variable and there is no treatment available with proven efficacy in halting disease progression other than liver transplantation (LTx). PSC is strongly associated with inflammatory bowel diseases (IBD) and patients are at increased risk for developing colorectal and biliary malignancies.^{1–3} Although the strong association between PSC and IBD is well established, the underlying pathogenesis is not known. IBD in PSC patients represents a distinct phenotype in that pancolitis is observed in 94% of PSC-ulcerative colitis (UC) and colitis in 96% of PSC-Crohn's disease (CD) patients.⁴ PSC and IBD are considered to be complex genetic diseases, caused by a combination of genetic and environmental factors. In IBD, 163 risk loci have been identified so far and some of these risk loci have also been associated with PSC.⁵ However, the number of patients included in genetic studies in IBD is tenfold higher than in PSC.⁶ First degree relatives of PSC patients run an increased risk of PSC and UC.⁷ Besides limited knowledge of the genetic risk factors, little is known about environmental factors triggering disease. Current smoking and previous appendectomy have been shown to protect against developing UC.^{8–20} However, studies on smoking behaviour and appendectomy and the risk of developing PSC are many fold smaller, warranting larger studies.^{21–24}

The aim of the present case-control study was to evaluate smoking behaviour, history of appendectomy, family history of IBD and autoimmune liver disease, and geographical distribution in a large group of PSC patients compared to IBD disease controls, and healthy controls.

PATIENTS AND METHODS

For this case-control study, PSC patients were recruited from the Epi PSC PBC project; a large population based observational longitudinal cohort study of PSC and primary biliary cirrhosis (PBC) in the Netherlands. IBD controls and healthy controls were also included in this cohort.

Study design and ethical consideration

The protocol of the Epi PSC PBC project was approved by the Central Committee for Research Ethics in Utrecht and all 44 local ethics committees of the participating hospitals in the Netherlands (trialregister.nl, NTR2813). This Epi PSC PBC observational longitudinal cohort study was undertaken between Jan 1, 2008, and Dec 31, 2011. All PSC patients in 44 hospitals from 2000 onwards were identified in a geographically defined area of 6 adjacent provinces comprising 50% of

the Dutch population (2000: 7,342,295 – 2007: 7,758,980) (supplementary figure 1). This area consists of 169 municipalities subdivided in 5 categories by degree of urbanisation; 1 ≥ 2500 , 2; 1500-2500; 3; 1000-1500, 4; 500-1000, 5; < 500 residences per km². The 44 participating hospitals provided care to the entire population of the study area. Healthy controls (HC) and IBD patients were randomly recruited from the outpatient clinics of four participating hospitals, evenly spread throughout the study area (supplementary figure 1).

Case-finding and case-ascertainment

The case-finding and case-ascertainment methods have been described previously.¹

PSC diagnosis was based on: 1) elevated alkaline phosphatase and gamma-glutamyltransferase, not explained otherwise, 2) presence of characteristic bile duct changes with multifocal strictures and segmental dilatations on endoscopic retrograde cholangiography (ERC) or magnetic resonance cholangiography (MRC) and/or 3) liver histology, and 4) no evidence for secondary sclerosing cholangitis. When criteria 1, 3 and 4 were fulfilled in the absence of cholangiographic abnormalities on MRC or ERC, cases were diagnosed as small duct PSC.²⁵ For this case-control study, PSC patients with concomitant IBD were subdivided in PSC-UC and PSC-CD subgroups, in which PSC-IBD unspecified (IBD-U) patients were grouped under PSC-UC.

Subjects visiting an outpatient clinic without a history of liver disease or autoimmune disease were considered to be healthy controls. These volunteers were healthy individuals accompanying PSC, primary biliary cirrhosis (PBC), and IBD patients to the outpatient clinic, or visiting the outpatient clinic for regular surveillance of Barrett's esophagus or colon polyps. IBD diagnosis was based on the Lennard-Jones criteria.²⁶ Similar to the PSC patients, IBD control patients were subdivided in UC and CD patients, IBD-U patients were grouped under UC. To be able to serve as genuine control group, IBD patients could not have any diagnosis of liver disease and had to have normal routine liver tests.

Data collection

At inclusion of the Epi PSC PBC project, the following data were collected from hospital patient files of PSC patients: date and type of PSC diagnosis, zip code of residence at time of diagnosis, date of diagnosis and type of IBD, smoking behaviour, and date and type of appendectomy. Hospital files of IBD controls were reviewed in order to verify IBD diagnosis, and exclude patients with history of liver disease or abnormal routine liver tests. If available, hospital files of HC's were checked as well, if not, history of liver disease was excluded anamnestically. PSC patients, IBD patients and HC's were asked to give informed consent and to fill out

a questionnaire regarding smoking status (current smoking, defined as smoking at time of inclusion, and former smoking, defined as smoking more than one year in the past twenty years), history of appendectomy and first degree family history of IBD and auto-immune liver disease. Auto-immune liver diseases included primary sclerosing cholangitis, primary biliary cirrhosis or auto-immune hepatitis. For PSC patients, data on history of appendectomy reported in questionnaires were completed by data on history of appendectomy from hospital files. Population estimates were retrieved from the Dutch Central Office for Statistics (Centraal Bureau voor de Statistiek, Den Haag, the Netherlands; <http://statline.cbs.nl>).

Statistical analysis

Potential risk factors (sex, age, smoking behaviour, appendectomy, family history of IBD and autoimmune liver disease) for the development of PSC were analyzed by comparing the relative frequencies between PSC patients and healthy controls as well as between PSC-IBD patients and IBD controls using logistic regression models. The few cases in which response to an item from the questionnaire was missing, were excluded from the pertaining analyses. Only appendectomies performed before PSC diagnosis was established were included in the analysis. Relative frequencies were expressed as odds ratio (OR), with corresponding 95% confidence intervals (CI). To differentiate risk factors for the development of PSC from risk factors for IBD, PSC-IBD patients and IBD controls were matched for type of IBD; PSC-UC patients were compared with UC patients, PSC-CD patients were compared with CD patients. In addition, PSC patients without IBD were analysed separately compared with HC's. After assessing the association between risk factors and disease with univariable logistic regression, multivariable logistic regression modeling was used to find independent risk factors. All relevant risk factors were entered in the multivariable analysis. Differences in age at inclusion and gender between groups were adjusted for in the multivariable analysis. Given the overlap in current smoking and former smoking, only former smoking was considered for multivariable analysis.

For analysis of geographical distribution, all PSC patients from the Epi PSC PBC cohort that were diagnosed between 2000 and 2008 were included. PSC incidence was calculated for the sum of municipalities per degree of urbanization, using the mean population between 2000 and 2008. Standardized incidence ratios were calculated to describe differences between observed and expected incidence of PSC per degree of urbanization. Cochran-Armitage test for trend in proportions was used to determine if there was statistical evidence for increasing or decreasing incidence over degrees of urbanization.

Statistical analyses were performed using SPSS version 20.0 software (SPSS, Inc., Chicago, IL) and R version 2.13.1 (R Foundation for Statistical Computing, Vienna, Austria). $p < 0.05$ was considered statistically significant.

RESULTS

Study inclusion

The case-finding exercise yielded 697 PSC patients, 254 HC's and 404 IBD controls. Four-hundred and fifty PSC patients were alive at time of inclusion (between Jan 1, 2008, and Dec 31, 2011) and were asked to give informed consent and to fill out a ten-item questionnaire. In total, 343 PSC patients completed the questionnaire resulting in a response rate of 76%. A total of 232 (92%) HC's and 370 (92%) IBD controls completed the questionnaire.

At time of diagnosis 590 of the 697 identified PSC patients lived in the geographically defined area, creating the population-based Epi PSC PBC cohort as described previously.¹ The subgroup of 343 PSC patients analyzed in the present study was not different from the entire population-based cohort, and is thereby representative of the PSC population in the Netherlands. (Supplementary table 1) The subgroup was only selected on the basis of being alive at the time of inclusion in the database and receipt of the questionnaire that had been sent to them.

Analysis of geographical distribution was performed in 218 patients that were diagnosed with PSC between 2000 and 2008 while living in the geographically defined area (Supplementary figure 1).

Patient characteristics

Patient characteristics are summarized in table 1. Of the 343 PSC patients, 168 (49%) had concomitant UC, 50 (15%) CD and the remaining 125 (36%) patients did not have IBD. Sixty-five percent of PSC patients were male, the median age at time of inclusion was 48 years (IQR 38-59 years). Healthy controls had a median age at inclusion of 55 years (IQR 42-65 years) and 46% were male. Of IBD controls 40% were male, and the median age at inclusion was 43 years (IQR 31-55).

Table 1. Patient characteristics

	PSC		PSC		HC		IBD	
	PSC	PSC-no IBD	PSC-UC	PSC-CD			UC	CD
N	343	125	168	50	232		167	203
Sex (male) [n (%)]	222 (65)	76 (61)	116 (69)	30 (60)	106 (46)		73 (44)	71 (35)
Age at inclusion (years) [median (IQR)]	48 (38-59)	49 (40-60)	48 (35-57)	46 (39-59)	55 (42-65)		46 (34-58)	40 (28-52)
Small duct PSC [n (%)]	36 (11)	20 (16)	12 (7)	4(8)	NA		NA	NA
Age at PSC diagnosis (years) [median (IQR)]	39 (29-49)	42 (32-50)	37 (25-46)	36 (30-49)	NA		NA	NA
Age at IBD diagnosis (years) [median (IQR)]	31 (21-44)	NA	31 (21-44)	32 (24-45)	NA		34 (24-45)	25 (20-36)
Smoking current [n (%)]	28 (8)	14 (11)	10 (6)	4 (8)	44 (19)		32 (19)	53 (26)
Smoking former [n (%)]	79 (23)	34 (27)	37 (22)	8 (16)	84 (36)		90 (54)	112(55)
Appendectomy [n (%)]	46 (13)	15 (12)	21 (13)	10 (20)	30 (13)		10 (6)	67 (33)
First-degree relative with IBD [n (%)]	37 (11)	12 (10)	17 (10)	8 (16)	35 (15)		28 (17)	35 (17)
First-degree relative with autoimmune liver disease [n (%)]	6 (2)	3 (2)	2 (1)	1 (2)	10 (4.3)		1 (1)	0 (0)

PSC = primary sclerosing cholangitis; HC = healthy control; IBD = inflammatory bowel disease; UC = ulcerative colitis; CD = Crohn's disease; IQR = inter quartile range; NA = not applicable.

Risk factors

Smoking

As shown in table 1, there were less current smokers (8%) in the PSC group than in the HC group (19%). PSC patients were also less often former smokers (23%) compared to healthy controls (36%), with a statistically significant effect in multivariable analysis (OR 0.52 95% CI 0.35-0.75; $p=0.001$) (table 2). A similar trend was observed in the subgroup of PSC patients without IBD ($n=125$) compared to healthy controls (OR 0.65 95% CI 0.40-1.06; $p = 0.085$) (table 2).

When comparing IBD patients to PSC-IBD patients, UC patients were more often current smokers (19%), and former smokers (54%) compared to PSC-UC patients (6%, 22% respectively). Similar results were seen for CD patients compared to PSC-CD patients. When assessing the independent effect in multivariable analysis, former smoking was significantly associated with a lower risk for the development of PSC in both UC (OR 0.21; 95% CI 0.12- 0.34; $p<0.001$) and CD patients (OR 0.17; 95% CI 0.08-0.39; $p<0.001$) (table 2).

Appendectomy

An equal proportion of PSC patients and HC underwent appendectomy (13% in both groups). In one PSC patient the appendix was removed as part of an ileocecal resection. More PSC-UC patients (13%) than UC controls (6%) had undergone appendectomy (table 1). In these subgroups all appendectomies were performed for clinical suspicion of acute appendicitis. When correcting for possible interacting variables using multivariable modelling, the positive association persisted in this subgroup (OR 2.51; 95%CI 1.04-6.07; $p=0.041$) (table 2). Appendectomy as risk factor for PSC was not assessed in the subgroups PSC-CD versus CD, because appendectomy was the result of ileocecal resection in the majority of patients in these subgroups.

Table 2. Univariable and multivariable logistic regression analyses risk factors

	Univariable analysis		Multivariable analysis	
	OR (95% CI)	p-value	OR (95% CI)	p-value
PSC vs HC (n=343 vs 232)				
Sex (male)	2.18 (1.55-3.07)	<0.001	2.05 (1.45-2.90)	<0.001
Age at inclusion	1.02 (1.01-1.03)	<0.001	1.02 (1.01-1.03)	<0.001
Smoking current	0.38 (0.23-0.63)	<0.001	NA	
Smoking former	0.53 (0.37-0.76)	0.001	0.52 (0.35-0.75)	0.001
Appendectomy	1.04 (0.64-1.71)	0.87	1.17 (0.70-1.98)	0.55
PSC- no IBD vs HC (n=125 vs 232)				
Sex (male)	1.84 (1.19-2.87)	0.007	1.77 (1.13-2.78)	0.012
Age at inclusion	1.01 (1.00-1.03)	0.074	1.01 (1.00-1.03)	0.089
Smoking current	0.54 (0.28-1.03)	0.061		
Smoking former	0.065 (0.41-1.05)	0.081	0.65 (0.40-1.06)	0.085
Appendectomy	0.92 (0.47-1.78)	0.80	1.03 (0.52-2.05)	0.93
PSC-UC vs UC (n=168 vs 167)				
Sex (male)	2.87 (1.84-4.50)	<0.001	3.10 (1.90-5.04)	<0.001
Age at inclusion	1.001 (0.99-1.02)	0.90	1.00 (0.99-1.02)	0.66
Smoking current	0.27 (0.13-0.56)	0.001	NA	
Smoking former	0.25 (0.15-0.40)	<0.001	0.21 (0.12-0.34)	<0.001
Appendectomy	2.24 (1.02-4.92)	0.044	2.51 (1.04-6.07)	0.041
First-degree relative with IBD	0.56 (0.29-1.07)	0.077	0.57 (0.28-1.16)	0.12
First-degree relative with au- to-immune liver disease*	NA		NA	
PSC-CD vs CD (n=50 vs 203)				
Sex (male)	2.79 (1.48-5.26)	0.002	2.21 (1.12-4.37)	0.022
Age at inclusion	0.97 (0.95-0.99)	0.007	0.97 (0.95-1.00)	0.017
Smoking current	0.25 (0.09-0.72)	0.010	NA	
Smoking former	0.16 (0.07-0.35)	<0.001	0.17 (0.08-0.39)	<0.001
First-degree relative with IBD	0.91 (0.40-2.12)	0.83	0.94 (0.37-2.38)	0.90
First-degree relative with au- to-immune liver disease*	NA		NA	

PSC = primary sclerosing cholangitis; HC = healthy control; UC = ulcerative colitis; CD = Crohn's disease; IBD = inflammatory bowel disease; OR = odds ratio; CI = confidence interval; NA = not assessed.

*Auto-immune liver disease not assessed in univariable and multivariable analyses, numbers are too small.

Family history

Seventeen (10%) PSC-UC patients and 28 (17%) UC controls had a first-degree family member with IBD. No differences in IBD prevalence were found between families of PSC-CD patients and CD controls (16% *vs.* 17%, respectively) (table 1). Multivariable analysis showed that family history of IBD was not an independent risk factor for PSC in the PSC-UC – UC nor PSC-CD – CD subgroups (OR 0.57; 95%CI 0.28-1.16; p=0.12 and OR 0.94; 95% CI 0.37-2.38; p=0.90, respectively) (table 2). The percentages of first-degree relatives with autoimmune liver diseases were similar for PSC-UC patients (1%) and UC controls (1%). Autoimmune liver diseases were more prevalent in families of PSC-CD patients than in families of CD controls, however numbers were very small (2% *vs.* 0%) (table 1). Family history was not assessed in the subgroups PSC *vs.* HC, as the HC group was composed partly from healthy family members accompanying PSC patients to the outpatient clinic, which would create an unrepresentative high percentage of HC's with PSC family members.

Geographical distribution

When combining municipalities according to degree of urbanization, the observed number of PSC patients was very similar to the expected number of PSC patients resulting in standardized incidence ratio's close to unity in each category (table 3). There was no evidence for a trend in incidence over the degrees of urbanization (p=0.97).

Table 3. Geographical distribution

Degree of urbanization (residences per km ²)	1 ≥2500	2 1500-2500	3 1000-1500	4 500-1000	5 <500
Residences (mean)	1156301	1988164	1818281	1675589	723785
Observed PSC cases	25	61	49	43	18
Expected PSC cases	30.8	52.9	48.5	44.5	19.2
SIR (95% CI)	0.81 (0.53-1.18)	1.15 (0.88-1.47)	1.01 (0.75-1.33)	0.97 (0.70-1.29)	0.94 (0.55-1.45)

PSC = primary sclerosing cholangitis, SIR = standardized incidence ratio, CI = confidence interval

DISCUSSION

This is the largest case-control study showing that smoking behaviour is independently associated with a lower risk of developing PSC, using a population based PSC cohort. The negative association between smoking and PSC in the present study is in agreement with previous studies, showing odds ratios of having PSC for smokers versus non-smokers ranging from 0.33 to 0.37.^{21–23,27} Interestingly, a recent case-control study by Eaton *et al.* showed that smoking was negatively associated with PSC when concomitant IBD was present, but not among PSC patients without IBD.²⁸ This finding was confirmed in the present study. This may have been the result of the smaller amount of PSC patients without IBD included in this analysis, however, it could also suggest that the association between smoking and PSC is driven by a – specific – genotype and corresponding PSC-IBD phenotype. The underlying mechanisms through which smoking affects the pathogenesis of PSC or IBD are largely unknown. Tobacco contains more than 4000 chemicals, of which nicotine is studied the most.²⁹ Nicotine activates nicotinic acetylcholine receptors in the central nervous system, autonomic ganglia and neuromuscular junctions, but also on immune cells and in the mucosa and lymphoid tissue of small and large bowel.²⁹ Van der Zanden *et al.* showed that repeated exposure to nicotine or tobacco smoking elevates $\alpha 7$ nicotinic acetylcholine receptor expression on human monocytes which may lead to sensitization of the ‘cholinergic anti-inflammatory pathway’.³⁰ More recently, nicotine has been shown to elevate mRNA expression of the pro-fibrogenic factors collagen 1- $\alpha 2$ and TGF- β (transforming growth factor beta) and induces human hepatic stellate cell proliferation.³¹ The effect of smoking on the progression of PSC has never been studied, yet in patients with primary biliary cirrhosis smoking has been associated with advanced histological disease at presentation.³² These findings may suggest that smoking is associated with a lower risk of developing PSC, but once liver fibrosis is present smoking might accelerate disease progression.

Although several studies have shown a protective effect of appendectomy on developing UC, a meta-analysis combining 342 PSC patients from four studies showed no association with PSC.^{21,23,24} This was in accordance with the present study, where we observed more appendectomies in PSC-UC patients and HC than in UC controls. These findings match the previously described protective effect of appendectomy in UC. However, appendectomy does not seem to result in a lower risk of PSC, since there is no difference in frequency of appendectomy between the PSC-UC subgroup and HC. The higher frequency of appendectomy in the PSC-UC group compared to the UC group indicates that appendectomy may be a risk factor for PSC in UC patients. This observation could be explained by a distinct genotype and IBD phenotype of PSC-UC patients.⁴ Possible immune-modulating effects of the appendix have been the focus of research for many years. A

large study from Sweden performed between 1964 and 1993 including 212,963 patients showed that patients who underwent an appendectomy for an appendicitis or mesenteric lymphadenitis but not for non-specific abdominal pain before the age of 20 had a low risk of UC.³³ These findings suggest that the inflammation rather than the absence of the appendix has a protective effect. Although the human biological pathways involved in the protective effect of appendicitis or complete removal of the appendix are just starting to be unravelled, an appendicitis mouse model showed down-regulation of 14 genes including IBD-associated genes after appendectomy.³⁴

Bergquist *et al.* showed in another large Swedish study that first-degree family members of PSC patients have a 3.8-fold increased risk of PSC and a 3.3-fold increased risk of developing UC.⁷ A drawback of the present study design is that relatives and friends accompanying PSC and IBD patients to the outpatient clinic were included as HC, and therefore we were unable to study the prevalence of PSC and IBD in families of healthy volunteers. In addition, this composition of the HC group could have introduced selection bias regarding the other environmental factors that were studied. On the other hand, using household members to some extent controls for unknown environmental factors, such as diet.

The balanced geographical distribution of PSC patients at time of diagnosis over cities and villages suggests that residential environment does not play a role in the etiopathogenesis of PSC. In agreement with our findings, Ala *et al.* found no association between standardized prevalence ratios of PSC patients listed for liver transplantation in New York City and proximity to super fund toxic waste sites.³⁵

This study has several limitations. First of all, cases and controls could unfortunately not be matched evenly for number of patients, age and gender. PSC patients were derived from 44 hospitals in the Netherlands, while control patients were derived from only 4 of these hospitals, explaining the relatively lower numbers. To rule out differences among subjects based on site of recruitment, a sensitivity analysis was performed including only cases and controls from the same site of recruitment. Observations remained consistent. (supplementary table 2) Due to the unbalanced gender and age distribution of PSC patients at diagnosis, one-on-one sex and age matching was not possible.¹ The differences in age and gender were accounted for in the multivariable regression analysis. Secondly, on average, three percent of all IBD patients have concomitant PSC, 40% of these will develop their PSC after the diagnosis of IBD.⁴ Therefore 1.2% of IBD controls may develop PSC in the future, which would disqualify them as proper controls. Lastly, several confounding factors may play a role when studying smoking behaviour. Suffering from a life threatening chronic illness such as PSC

may influence tobacco use. Notably, controls were recruited from outpatient clinics, which could have influenced the number of smokers. However, 27% of the general population (> 14 years of age) in the Netherlands was smoking in 2010, implying that the percentage of smokers at time of inclusion in the patient and control groups did not exceed the national average.³⁶ When interviewing patients about their past smoking behaviour, there is the risk of recall bias. By keeping questions on smoking behaviour simple, we attempted to minimize this risk. The downside is that our questionnaire did not capture intermittent smokers or smokers who quit more than 20 years ago. In the general Dutch population, - studied since 1989 - the percentage smokers is highest in the age group 25-44 years for males, and in the age group 16-24 years for females.³⁷ Thereafter, with each age increment the percentage of smokers decreases.³⁷ Therefore, most PSC patients will have commenced smoking before the age of 44 (male) or 24 (female). Since the median age at diagnosis of the PSC patients was 39 years (IQR 29-49) the vast majority of PSC patients will have started smoking before their PSC diagnosis.

We did not register the indication of appendectomy, and were not able to assess if inflammation rather than the absence of the appendix could be a risk factor for PSC, as was suggested for UC.³³ Given the high frequency of ileocecal resection in CD patients, appendectomies were not compared between the PSC-CD and CD subgroups. Only appendectomies performed before PSC diagnosis were included in the risk factor analysis, so IBD controls and HC's had more time at risk for undergoing appendectomy. This could have resulted in a somewhat higher number of appendectomies in the HC and IBD control groups compared with the PSC group, making our observation of relatively high appendectomy rate in the PSC groups even more notable.

In conclusion, we confirm in a large population-based cohort that smoking is associated with a lower risk of developing PSC, independent of the protective effect in UC. Appendectomy is not an independent risk factor for PSC.

REFERENCES

1. Boonstra, K. et al. Population-based epidemiology, malignancy risk, and outcome of primary sclerosing cholangitis. *Hepatology* 58, 2045–55 (2013).
2. Burak, K. et al. Incidence and risk factors for cholangiocarcinoma in primary sclerosing cholangitis. *Am J Gastroenterol* 99, 523–6 (2004).
3. Claessen, M. M. H., Vleggaar, F. P., Tytgat, K. M. a J., Siersema, P. D. & van Buuren, H. R. High lifetime risk of cancer in primary sclerosing cholangitis. *J Hepatol* 50, 158–64 (2009).
4. Boonstra, K. et al. Primary sclerosing cholangitis is associated with a distinct phenotype of inflammatory bowel disease. *Inflamm Bowel Dis* 18, 2270–6 (2012).
5. Anderson, C. et al. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat Genet* 43, 246–52 (2011).
6. Karlsen, T. H. A lecture on the genetics of primary sclerosing cholangitis. *Dig Dis* 30 Suppl 1, 32–8 (2012).
7. Bergquist, A. et al. Increased risk of primary sclerosing cholangitis and ulcerative colitis in first-degree relatives of patients with primary sclerosing cholangitis. *Clin. Gastroenterol. Hepatol.* 6, 939–943 (2008).
8. Harries, A., Jones, L., Heatley, R. & Rhodes, J. Non-smoking: a feature of ulcerative colitis. *Br Med J (Clin Res Ed)* 284, 706 (1982).
9. Lindberg, E., Tysk, C., Andersson, K. & Järnerot, G. Smoking and inflammatory bowel disease. A case control study. *Gut* 29, 352–7 (1988).
10. Russel, M. G. et al. Appendectomy and the risk of developing ulcerative colitis or Crohn's disease: results of a large case-control study. *South Limburg Inflammatory Bowel Disease Study Group. Gastroenterology* 113, 377–382 (1997).
11. Koutroubakis, I. E. & Vlachonikolis, I. G. Appendectomy and the development of ulcerative colitis: results of a metaanalysis of published case-control studies. *Am J Gastroenterol* 95, 171–6 (2000).
12. Reif, S. et al. Lack of Association Between Smoking and Crohn's Disease But the Usual Association With Ulcerative Colitis in Jewish Patients in Israel : A Multicenter Study. *Am J Gastroenterol* 95, 474–8 (2000).
13. Naganuma, M. et al. Appendectomy Protects Against the Development of Ulcerative Colitis and Reduces Its Recurrence : Results of a Multicenter Case-Controlled Study in Japan. *Am J Gastroenterol* 96, 1123–6 (2001).
14. Abraham, N., Selby, W., Lazarus, R. & Solomon, M. Is smoking an indirect risk factor for the development of ulcerative colitis ? An age- and sex-matched case – control study. *J Gastroenterol Hepatol* 18, 139–46 (2003).
15. Regueiro, M., Kip, K. E., Hegazi, R. A. & Plevy, S. Cigarette Smoking and Age at Diagnosis of Inflammatory Bowel Disease. *Inflamm Bowel Dis* 11, 42–47 (2005).
16. Jiang, L. et al. Risk Factors for Ulcerative Colitis in a Chinese Population: an age-matched and sex-matched case-control study. *J Clin Gastroenterol* 41, 280–284 (2007).
17. Gearry, R. B., Richardson, A. K., Frampton, C. M., Dodgshun, A. J. & Barclay, M. L. Population-based cases control study of inflammatory bowel disease risk factors. *J Gastroenterol Hepatol* 25, 325–33 (2010).
18. Hansen, T. S. et al. Environmental factors in inflammatory bowel disease: a case-control study based on a Danish inception cohort. *J Crohns Colitis* 5, 577–84 (2011).
19. Castiglione, F. et al. Risk factors for inflammatory bowel diseases according

- to the 'hygiene hypothesis': a case-control, multi-centre, prospective study in Southern Italy. *J Crohns Colitis* 6, 324–9 (2012).
20. Lakatos, P. L. et al. Is current smoking still an important environmental factor in inflammatory bowel diseases? Results from a population-based incident cohort. *Inflamm Bowel Dis* 19, 1010–7 (2013).
 21. van Erpecum, K. J. et al. Risk of primary sclerosing cholangitis is associated with nonsmoking behavior. *Gastroenterology* 110, 1503–6 (1996).
 22. Loftus, E. V et al. Primary sclerosing cholangitis is associated with nonsmoking; a case-control study. *Gastroenterology* 110, 1496–1502 (1996).
 23. Mitchell, S. et al. Cigarette smoking, appendectomy, and tonsillectomy as risk factors for the development of primary sclerosing cholangitis: a case control study. *Gut* 51, 567–73 (2002).
 24. Florin, T., Pandeya, N. & Radford-Smith, G. Epidemiology of appendectomy in primary sclerosing cholangitis and ulcerative colitis: its influence on the clinical behaviour of these diseases. *Gut* 53, 973–979 (2004).
 25. European Association for the Study of the Liver. EASL Clinical Practice Guidelines: management of cholestatic liver diseases. *J. Hepatol.* 51, 237–67 (2009).
 26. Lennard-Jones, J. Classification of inflammatory bowel disease. *Scand J Gastroenterol* 170, 16–9 (1989).
 27. Andersen, I. M. et al. Effects of coffee consumption, smoking, and hormones on risk for primary sclerosing cholangitis. *Clin. Gastroenterol. Hepatol.* 12, 1019–28 (2014).
 28. Eaton, J. E. et al. A comprehensive assessment of environmental exposures among 1000 North American patients with primary sclerosing cholangitis, with and without inflammatory bowel disease. *Aliment. Pharmacol. Ther.* 41, 980–990 (2015).
 29. Thomas, G., Rhodes, J. & Ingram, J. Mechanisms of disease: nicotine—a review of its actions in the context of gastrointestinal disease. *Nat Clin Pr. Gastroenterol Hepatol* 2, 536–44 (2005).
 30. van der Zanden, E. et al. Nicotinic acetylcholine receptor expression and susceptibility to cholinergic immunomodulation in human monocytes of smoking individuals. *Neuroimmunomodulation* 19, 255–65 (2012).
 31. Soeda, J. et al. Nicotine induces fibrogenic changes in human liver via nicotinic acetylcholine receptors expressed on hepatic stellate cells. *Biochem Biophys Res Commun* 417, 17–22 (2012).
 32. Zein, C. O. et al. Smoking and increased severity of hepatic fibrosis in primary biliary cirrhosis: A cross validated retrospective assessment. *Hepatology* 44, 1564–71 (2006).
 33. Andersson, T., Olaison, G., Tysk, C. & Ekbo, A. Appendectomy and Protection against Ulcerative Colitis. *N. Engl. J. Med.* 344, 808–814 (2001).
 34. Cheluvappa, R., Luo, a S., Palmer, C. & Grimm, M. C. Protective pathways against colitis mediated by appendicitis and appendectomy. *Clin. Exp. Immunol.* 165, 393–400 (2011).
 35. Ala, A. et al. Increased prevalence of primary biliary cirrhosis near Superfund toxic waste sites. *Hepatology* 43, 525–31 (2006).
 36. Zeegers, T. & Zantinge, E. Hoeveel mensen roken? [Internet]. In: *Volksgezondheid Toekomst Verkenning, Nationaal Kompas Volksgezondheid*. Available from: <http://www.nationaalkompas.nl>.
 37. Bruggink, J.-W. Trends in the share of smokers in the Netherlands since 1989. *Tijdschr. voor gezondheidswetenschappen* 91, 234–240 (2013).

Supplementary table 1. Cases included in the Epi PSC PBC cohort compared to cases included in the risk factor cohort

	Epi PSC PBC cohort (n=590)	Risk factor cohort (n=343)	p-value
*Sex (male) [n (%)]	375 (64)	222 (65)	0.77
*Small duct [n (%)]	58 (10)	36 (11)	0.74
*Concomitant IBD	401 (68)	218 (64)	0.17
*Ulcerative Colitis	324 (81)	168 (77)	0.35
*Crohn's disease	78 (19)	50 (23)	0.35
**Age at PSC diagnosis (years) [mean, SD]	39 (15)	39 (14)	0.90
**Age at IBD diagnosis (years) [mean, SD]	32 (15)	32 (15)	0.49

IBD = inflammatory bowel disease, SD = standard deviation

* Difference between groups assessed by Chi square test

** Difference between groups assessed by Unpaired t-test

Supplementary table 2. Sensitivity analysis including only cases and controls from the same site of recruitment

	Univariable analysis		Multivariable analysis	
	OR (95% CI)	p-value	OR (95% CI)	p-value
PSC vs HC (n=163 versus 232)				
Sex (male)	2.93 (1.92-4.49)	<0.001	2.89 (1.86-4.51)	<0.001
Age at inclusion	1.03 (1.01-1.04)	<0.001	1.03 (1.01-1.04)	<0.001
Smoking current	0.25 (0.12-0.53)	<0.001	NA	
Smoking former	0.43 (0.27-0.70)	0.001	0.40 (0.25-0.66)	<0.001
Appendectomy	1.16 (0.65-2.07)	0.61	1.31 (0.69-2.47)	0.41
PSC-no IBD vs HC (n= 59 versus 232)				
Sex (male)	2.71 (1.47-2.71)	0.001	2.66 (1.42-4.97)	0.002
Age at inclusion	1.02 (1.00-1.04)	0.033	1.02 (1.00-1.04)	0.040
Smoking current	0.40 (0.15-1.05)	0.062		
Smoking former	0.32 (0.15-0.67)	0.003	0.30 (0.14-0.65)	0.002
Appendectomy	0.76 (0.30-1.93)	0.57	0.79 (0.30-2.12)	0.65
PSC-UC vs UC (n=81 versus 167)				
Sex (male)	3.68 (2.05-6.59)	<0.001	4.02 (2.15-7.53)	<0.001
Age at inclusion	1.01 (0.99-1.03)	0.27	1.01 (0.99-1.03)	0.23
Smoking current	0.22 (0.08-0.64)	0.006	NA	
Smoking former	0.29 (0.16-0.52)	<0.001	0.23 (0.12-0.43)	<0.001
Appendectomy	2.73 (1.13-6.62)	0.026	3.17 (1.16-8.70)	0.025
First-degree relative with IBD	0.62 (0.28-1.39)	0.244	0.71 (0.30-1.71)	0.45
First-degree relative with auto-immune liver disease*	NA		NA	
PSC-CD vs CD (n=23 versus 203)				
Sex (male)	3.49 (1.41-8.62)	0.007	2.62 (1.01-6.80)	0.047
Age at inclusion	0.97 (0.94-1.00)	0.02	0.97 (0.94-1.00)	0.043
Smoking current	NA		NA	
Smoking former	0.12 (0.04-0.42)	0.001	0.14 (0.04-0.49)	0.002
First-degree relative with IBD	0.72 (0.20-2.56)	0.61	0.73 (0.18-2.89)	0.73
First-degree relative with auto-immune liver disease*	NA		NA	

PSC = primary sclerosing cholangitis; HC = healthy control; IBD = inflammatory bowel disease; UC = ulcerative colitis; CD = Crohn's disease; OR = odds ratio; CI = confidence interval; NA = not assessed.

*Auto-immune liver disease not assessed in univariable and multivariable analyses, numbers are too small.

Supplementary figure 1. Geographically defined study area.
Black dots: 44 hospitals from which PSC patients were derived
White dots: 4 hospitals from which IBD controls and healthy controls were derived

Biomarkers for disease progression of primary sclerosing cholangitis

Elisabeth M.G. de Vries, Ulrich H. Beuers, Cyriel Y. Ponsioen

Current Opinion in Gastroenterology 2015;31:239-46

ABSTRACT

Purpose of review

Primary sclerosing cholangitis (PSC) is a chronic cholestatic liver disease of unknown origin. There is no medical treatment of proven benefit on survival ; once patients have progressed to end-stage liver disease the only treatment option is liver transplantation.

Recent findings

Over the last years, some progress has been made in identifying biomarkers of PSC disease progression. Categories that can be distinguished include clinical and biochemical biomarkers, histology, imaging, prognostic modelling and genetics. With this review we summarize biomarkers for progression of PSC from these six categories, that have been studied to date.

Summary

Biomarkers for the progression of PSC disease course can be used for several purposes. First of all, they can be implemented as surrogate endpoints for clinical trials. Secondly, biomarkers of disease progression form the basis of prognostic modelling, which is needed for proper patient counselling and management. Lastly, these biomarkers may yield a better understanding of PSC pathogenesis.

Key Points

- Biomarkers for disease progression are important to enable a better understanding of pathogenesis, to improve patient management and counselling, and to serve as surrogate endpoint in the design of clinical trials.
- There are currently no validated biomarkers for disease progression of both early and late stage PSC.
- Serum biomarker alkaline phosphatase and histology may serve as biomarkers during early and late disease stage.
- Exploratory biomarkers such as vibration-controlled transient elastography, magnetic resonance cholangiopancreatography , magnetic resonance elastography, and ELF® are promising and should be evaluated in multi-center cohort studies as well as incorporated as exploratory variables in clinical trials.

INTRODUCTION

Primary sclerosing cholangitis (PSC) is a slowly progressive cholestatic liver disease affecting the intra- and extrahepatic bile ducts. Chronic inflammation induces destruction of hepatic bile ducts, leading to biliary fibrosis, which will ultimately result in cirrhosis of the liver and liver failure.[1] Incidence and prevalence rates of PSC vary widely throughout the literature, and have been reported from 0-1.3 per 100.000 inhabitants per year and 0-16.2 per 100.000 inhabitants, respectively. [2] Survival time differs in population based versus transplant center cohorts, with a median of 21.3 years versus 13.2 years, respectively.[3]

The etiology and pathogenesis of PSC is still largely unknown and an effective, causal drug therapy is lacking.[1,4–6] The only therapeutic option for patients that have progressed to end stage liver disease, is liver transplantation (LTx). [1,4,5] The need for developing new therapeutic options is thereby obvious. An important challenge in designing drug intervention studies in PSC is constituted by the delineation of clinical endpoints. The chronic disease course hampers the use of solid endpoints such as LTx or PSC related death, as they simply take too long to occur. The only two clinical trials that attempted to use clinical solid endpoints as primary outcome measure, employing a 5 years follow-up (FU), demonstrated an incidence of clinical endpoints of 3-4% in the placebo arms. [7,8] Given this small percentage even with as large as a 50% reduction in hazard rate as a result of a new drug, more than 2000 patient-years are necessary to attain a sufficient power. This is simply not feasible with an orphan disease such as PSC.

Surrogate endpoints such as alkaline phosphatase (ALP) and liver histology have been frequently used in trials, but were never properly validated. [7,9–15] To identify biomarkers with an accurate predictive value for the progression of PSC is important for multiple reasons, including enabling a better understanding of the pathogenesis, to improve patient management and counselling, and to serve as surrogate endpoint in the design of clinical trials.[15] With this review we aim to provide a comprehensive overview of biomarkers for progression of PSC that have been investigated to date.

An overview of the most relevant studies that looked into the association between biomarkers and outcome is provided in table 1. The six different categories of biomarkers that will be addressed are: clinical symptoms, biochemical biomarkers, liver histology, imaging, prognostic models and genetics serving as biomarker for progression of PSC.

Table 1. Overview of biomarkers for disease progression of primary sclerosing cholangitis and association with outcome

Biomarker	Authors	Year	N (PSC)
Clinical biomarkers			
Fatigue	Ter Borg <i>et al</i> [16]	2005	27
	Björnsson <i>et al</i> [17]	2004	93
Health-related quality of life	Benito de Valle <i>et al</i> [18]*	2012	160
	Gotthardt <i>et al</i> [19]	2014	113
Biochemical biomarkers			
Alkaline phosphatase			
Normalization	Stanich <i>et al</i> [20]*	2011	87
Reduction <1.5xULN	Al Mamari <i>et al</i> [21]*	2013	139
Normalization/reduction ≥40%	Lindström <i>et al</i> [22]*	2013	198
1. Normalization, reduction 50%, reduction <1.5xULN in 6mnd 2. Reduction <1.5xULN in 12 months	Rupp <i>et al</i> [23]*	2014	215
Normalization in 12 months after diagnosis	De Vries <i>et al</i> [24]	2014	298
ELF	Vesterhus <i>et al</i> [25]	2014	305
AST/ALT ratio	Nyblom <i>et al</i> [26]*	2007	154
Bilirubin	Kim <i>et al</i> [27]*	2000	405
	Broomé <i>et al</i> [28]	1996	305
Auto-antibodies (ACA)	Angelo <i>et al</i> [29]	2000	73
IgG4	Benito de Valle <i>et al</i> [30]	2014	345
Histology			
	de Vries <i>et al</i> [31]	2014	64
Imaging			
ERCP	Craig <i>et al</i> [32]	1991	129
MRC	Ruiz <i>et al</i> [33]*	2014	64
VTCE	Corpechot <i>et al</i> [34]*	2014	73 ^a /168 ^b
Prognostic models			
Mayo risk score	Kim <i>et al</i> [27]*	2000	405
Amsterdam cholangiographic model	Ponsioen <i>et al</i> [35]*	2002	174
Validation Amsterdam model	Ponsioen <i>et al</i> [36]	2010	111

ACA, anticardiolipin antibody; AST/ALT ratio, aspartate to Alanine aminotransferase ratio; CCA, cholangiocarcinoma; nance cholangiopancreatography; MRS, Mayo risk score; LTx, liver transplantation; PSC, primary sclerosing cholangitis; analysis. ^b Survival analysis.

Outcome	
Association with disease severity Endpoints: biochemical tests, MRS, liver histology	Association with clinical outcome Endpoints: (liver related) death, LTx, CCA, liver decompensation
No association (biochemical tests)	-
No association (biochemical tests, histologic stage)	-
No association (MRS, cirrhosis)	-
No association (MRS)	-
-	Association with endpoint-free survival (CCA, LTx, death)
-	Association with endpoint-free survival (liver related death, LTx, liver decompensation). Reduced risk CCA
-	Association with endpoint-free survival (death, LTx, CCA)
-	Association with transplant-free survival
-	Association with endpoint-free survival (liver related death, LTx, CCA)
-	Associated with transplant-free survival
Association with cirrhosis	Associated with liver related death and LTx
-	Association with survival (all cause death)
-	Association with transplant-free survival
Correlation MRS, histologic stage	-
-	No association with endpoint free survival (liver related death, LTx or CCA)
No association (biochemical tests)	Association of staging systems with transplant-free survival and time to LTx
-	<i>Association with transplant-free survival</i>
-	-
Associated with histological stage	Association with endpoint-free survival (death, LTx, hepatic complication)
-	Association with survival (all cause death)
-	Association with transplant-free survival
-	Association with transplant-free survival

ERCP, endoscopic retrograde cholangiopancreatography; IBD, inflammatory bowel disease; MRC, magnetic resonance VCTE, vibration-controlled transient elastography; xULN, times upper limit of normal; -, not studied. *Diagnostic

1. Clinical symptoms

At time of diagnosis only 56% of PSC patients are symptomatic.[28] Common clinical symptoms include fatigue and pruritus, which could be of influence on health related quality of life.

1.1 Fatigue

Fatigue can be a serious symptom, which is poorly understood. A role for abnormalities in amino acid metabolism in the pathophysiology of fatigue has been proposed, but could not be confirmed for PSC patients. [16,17] Fatigue often coincides with depression and the latter was shown to be an independent predictor of fatigue score in PSC. [17] However, attempts to treat fatigue with antidepressants were of no beneficial effect for PSC patients.[37] A significant correlation between fatigue and disease severity measured by liver tests and liver histology has never been demonstrated.[16–18]* In fact, Björnsson *et al.* showed that PSC patients had significantly lower fatigue scores than sex and age matched controls from the general population. [17]

1.2 Pruritus

Pruritus is a common and debilitating symptom, experienced by 70-80% of PSC patients at any time during their disease course.[38] Recently, a role for the enzyme forming lysophosphatic acid, autotaxin, was proposed, as serum autotaxin activity was found to correlate with itch intensity and response to antipruritic treatment in patients with cholestatic pruritus.[38] Several studies have shown that pruritus is related to a reduced HRQL, and that the intensity of pruritus is associated with a significantly lower HRQL.[18,19,39]* However a correlation between degree of pruritus and disease severity has never been shown.

1.3 Health related quality of life

Health related quality of life (HRQL) was shown to be impaired in multiple studies focusing on chronic liver diseases.[40–42] In a study investigating the HRQL of an unselected group PSC patients from a combined Swedish and English cohort, an impaired HRQL was demonstrated compared with the general population.[18]* Older age, large-duct disease, and systemic symptoms seemed to reduce HRQL. [18]* This was in contrast with a publication of Gotthardt *et al.* who showed that HRQL of PSC patients was comparable with the range recorded for healthy population.[19] HRQL is related to severity of disease in the sense that cirrhotic patients were found to have a significantly lower HRQL than non-cirrhotic patients. [18]* However, liver disease severity measured by Mayo risk score (MRS) and the presence of liver cirrhosis was not found to be a predictor of HRQL, indicating that

liver disease severity is not a major determinant if HRQL.[18]* Gotthardt *et al.* measured disease severity according to the presence of dominant strictures, MRS and coexisting IBD, and could also not demonstrate an association with HRQL.[19]

2. Biochemical markers

Several biochemical markers such as alkaline phosphatase, bilirubin, as well as composite markers have been studied.

2.1 Alkaline phosphatase

An elevated alkaline phosphatase (ALP) is the most common biochemical abnormality in PSC, which often prompts clinicians to establish a diagnosis.[43] The effect of reduction in alkaline phosphatase level on clinical outcome in terms of PSC related death, LTx and development of cholangiocarcinoma has been investigated in several cohorts using different cut-off levels to define reduction, including normalization, reduction below 1.5 times upper limit of normal and normalization or reduction of 40%.[20–24]* All these studies have confirmed that a reduction of ALP over time is associated with a favorable prognosis, independent of the use of UDCA, the presence of dominant strictures or endoscopic interventions.[20–24]* The chosen cut-off values remain arbitrary, and the most specific and sensitive cut-off value is yet to be determined.

2.2 Enhanced liver fibrosis test

The enhanced liver fibrosis test (ELF®) test is an algorithm incorporating three serum markers of liver fibrosis; hyaluronic acid (HA), amino-terminal pro-peptide of type III pro-collagen (PIIINP), and tissue inhibitor of metalloproteinase 1 (TIMP-1), in order to detect hepatic fibrosis.[44,45] ELF® test was shown to be accurate and reproducible in predicting significant liver fibrosis in independent populations of patients with chronic hepatitis C, nonalcoholic fatty liver disease (NAFLD) and PBC.[45–47] Furthermore, ELF® test was shown to be a predictor of clinical outcome in patients with chronic liver disease of different etiologies.[48] The usefulness of ELF® in PSC has been assessed only recently, with promising first result, showing an independent prognostic value of ELF® in predicting LTx and all cause death in a tertiary referral cohort.[25] However, the applicability of ELF® in PSC requires validation in a large population-based PSC cohort.

2.3 Aspartate to Alanine aminotransferase ratio

Aspartate aminotransferase (AST) and alanine aminotransferase (ALT) are markers of hepatocellular injury or necrosis, and elevation of AST or ALT is present in case of parenchymal liver damage.[49] In PSC, transaminases can be normal, but typically an elevation of AST and ALT up to 2 or 3 times upper limit of normal is ob-

served. [43] A high AST/ALT ratio was shown to be an indicator of liver cirrhosis in alcoholic liver disease, chronic hepatitis C and PBC.[50–53] On this basis Nyblom *et al.* investigated the use of AST/ALT ratio as an indicator for cirrhosis measured by histological staging according the Ludwig system, and clinical outcome defined as liver related death and LTx, in PSC patients.[26]* It was shown that a high AST/ALT ratio is indeed an indicator of cirrhosis and poor clinical outcome in patients with PSC.[26]* The prognostic value of AST/ALT ratio has not been confirmed in other cohorts.

2.4 Other biochemical markers

Bilirubin: Bilirubin is a routinely used in clinical practice as a measure of extrahepatic obstruction or advanced disease. Bilirubin is a proven prognostic marker in late PSC, and has been used in several studies on prognostic modelling in PSC. [27,28,54–57]* However, since an increased bilirubin is only appreciated once there is late disease stage, a dominant stricture, or suppurative cholangitis, its usefulness as biomarker for progression measured earlier in the disease course of PSC is limited.

Auto-antibodies: In the majority of PSC at least one non-organ specific serum autoantibody can be detected.[29] However there is a large variation of frequency between different studies, and the specificity of autoantibodies is often low. [29,58] Auto-antibodies most commonly identified in PSC patients include anti-nuclei antibody (ANA), anti-neutrophil cytoplasm antibody (ANCA) anticardiolipin antibody (ACA), antibody against biliary epithelial cells (Anti-BEC), and smooth muscle antibody (SMA).[58] In a study of Angulo *et al.* a positive correlation between serum levels of ACA and disease severity measured by MRS and histologic stage of disease was shown.[29] This finding has so far not been reproduced.

IgG4: Recently, the emerging entity of immunoglobulin G4-associated cholangitis (IAC) has been described, a cholangiopathy characterized by a high serum IgG4 level and responsiveness to corticosteroids, by which it can be distinguished from PSC.[59] An elevation of serum level IgG4 is not specific to IAC, as it is also measured 9-15% of PSC patients.[30,60–62] PSC patients with elevated IgG4 were shown to have a high prevalence of cirrhosis, and a shorter time to LTx, suggesting a more severe liver disease course.[60,61] However in a study of Benito de Valle *et al.*, an elevated IgG4 level was shown not to be associated with an increased risk of LTx, liver related death or cholangiocarcinoma.[30]

3. Histology

At present, liver biopsy in PSC is only indicated in case of suspicion of small duct PSC or auto-immune sclerosing cholangitis hepatitis overlap (PSC auto-immune hepatitis overlap syndrome).[15] Liver biopsy has been used frequently as surrogate endpoint in clinical trials, it is, however, subject to several limitations including risk for the patient, a reported sampling variability of at least 20% and subjectivity of interpretation.[15,63] Nevertheless, Angulo *et al.* has described progression of liver biopsies over 2 year time in 53% of PSC-patients with an initial Ludwig stage of I-III.[64]* Conversely, 14% of Ludwig stage III-IV had stage I-II on repeat liver biopsy after 2 years, which may reflect the sampling variability, assuming that spontaneous regression of stage does not occur.[64]* These findings indicate that the evaluation of drug effect on disease progression in clinical trials for PSC may well be assessed using histology.[64]* Furthermore, the prognostic value of liver biopsy in PSC has recently been shown by a strong association between disease stage, examined by the Ishak, Nakanuma and Ludwig histologic scoring systems and transplant-free survival and time to LTx.[31] These three histologic systems were not designed primarily to assess disease grade and-or stage in PSC, but for chronic hepatitis (Ludwig) and PBC (Ishak and Nakanuma).[65–67] Since the Nakanuma grading and staging system incorporates histological scoring components that may also be considered PSC specific, this scoring system merits further evaluation in PSC.[67]

Despite its drawbacks liver histology appears to be a useful biomarker for progression of PSC and is indeed still the gold standard for (lack of) disease progression for regulatory authorities.

4. Imaging

The advent of high quality non-invasive imaging techniques such as magnetic resonance cholangiopancreatography and vibration-controlled transient elastography, constitute promising non-invasive biomarkers for disease progression.

4.1 Endoscopic retrograde cholangiopancreatography

Cholangiographic features characteristic for PSC include diffuse multifocal strictures of the intra- and extrahepatic bile ducts, alternating with normal or minimally dilated segments, creating the so called “beaded” appearance.[68] The value of endoscopic retrograde cholangiopancreatography (ERCP) in determining prognosis was studied by Craig *et al.* who found high grade, and diffuse strictures of the intrahepatic bile ducts to be indicators of poor prognosis.[32] In 1991 Majoie *et al.* designed a modified classification system of abnormalities of the intra and extrahepatic bile ducts.[69] Scoring by this classification system was shown to increase over time, suggesting correlation with disease severity. [70]

4.2 Magnetic resonance cholangiopancreatography (MRC)

Magnetic resonance imaging (MRI) with magnetic resonance cholangiopancreatography (MRC) is nowadays the standard of reference in establishing PSC diagnosis, largely replacing the more invasive ERCP. [71,72] MRI features of PSC have been mainly described using two dimensional (2D) MRC, and have shown no correlation with disease severity measured by MRS. [73,74] Using the newer three-dimensional (3D) MRC, visibility of the biliary tree has improved. [75] Systematic analysis of bile ducts and liver parenchyma over time visualized by 3D MRC has recently been used to discriminate two radiologic PSC disease patterns; stability and worsening.[33]* Using three radiologic features, two radiologic scores were created which can be used to obtain information on risk of radiologic progression.[33]* Whether radiologic progression is associated with the occurrence of clinical outcome has not yet been determined.

4.3 Vibration-controlled transient elastography (FibroScan®)

Vibration-controlled transient elastography (VCTE) using FibroScan® is a noninvasive method to assess the degree of liver fibrosis, using liver stiffness measurement (LSM).[76] Corpechot *et al.* analyzed the usefulness and prognostic value of VCTE in PSC patients, and found an independent association between LSM and degree of fibrosis, histologically staged by a Metavir derived scoring system (F0-4).[34]* For those patients with an advanced fibrosis stage (F3) at baseline, LSM was predictive of adverse clinical outcomes.[34]* Furthermore, the progression rate of LSM was found to be independently associated with clinical outcome. [34]* Measurements of LSM can be influenced by several confounding factors such as obesity, inflammation, cholestasis, congestion, and food intake, which need to be taken into account when interpreting the results.[76,77] Albeit promising, the applicability and prognostic value of VCTE awaits validation in a large international cohort study (ongoing FICUS study).

5. Prognostic modelling

Combining multiple variables of disease progression in mathematical models may allow for an improved estimation of survival. Several prognostic models have been created for PSC, here we describe the most commonly used Mayo risk score and the Amsterdam cholangiographic scoring system.

5.1 MAYO risk score

The prognostic model most commonly used for purpose of patient counselling as well as surrogate endpoint in clinical trials, is the Mayo risk score.[27]*MRS was based on a cohort of 405 PSC patients and validated in a cohort of 124 PSC

patients. [27]* Outcome was death; for patients who underwent LTx an assumption of survival time had they not undergone LTx was made. [27]* MRS was finally composed of variables age, bilirubin, AST, variceal bleeding and albumin. [27]* Despite that MRS was validated and constructed from a large cohort, the model has several shortcomings. Twenty-five percent of included patients were derived from an ursodeoxycholic acid study which was justified by the negative result of this study. [27]* However, ursodeoxycholic acid does have effects on liver biochemistry. [78] Furthermore, the model has a horizon of only 4 years, so performs best in end-stage patients, and only estimates death and not time to LTx. [27]*

5.2 Amsterdam cholangiographic scoring system

The only validated prognostic model that considers the clinical endpoint transplant free-survival, was published by Ponsioen *et al.* [35]* They demonstrated an inverse correlation of the Amsterdam cholangiographic classification system with survival, and together with age at ERCP a model was created, strongly predictive of survival. [35,69]* In 2010, the performance of this prognostic model was externally validated in a Norwegian cohort, showing adequate performance. [36] Using data from both cohorts the model was updated. [36]

6. Genetic biomarkers

Recent genome-wide association studies and candidate gene studies have identified 16 non-HLA genetic susceptibility loci in PSC. [79–82] It is still unknown whether individual differences in PSC sub-phenotypes are driven by genetic background. The PSC Immunochip sub-phenotyping project, a study initiated by the International PSC study group (IPSCSG) is currently investigating, these possible associations between PSC risk loci and PSC sub-phenotype.

CONCLUSION

There is a big unmet need for accurate, preferably noninvasive biomarkers for assessing progression of PSC disease course. Clinical symptoms are important when treatment efficacy aimed at relieving complaints is tested in a clinical trial. However, since clinical symptoms are poorly associated with progression of PSC they appear not useful as biomarkers of disease progression. [16–19]* Biochemical biomarkers such as ALP and ELF® were found to be of strong prognostic value at least in stratifying patients in favorable and unfavorable disease course, and may be useful biomarkers of progression in PSC. [20–25]* Furthermore, despite its invasive nature and risk of sampling variability, there is still a place for histology as biomarker for disease progression in PSC. [31] However, exploratory biomarkers such as MRC, VCTE, MR-elastography, and ELF® are currently being

validated in larger cohorts, and these biomarkers may replace histology in the future.[25,33,34]*Combining biomarkers of progression in a prognostic model may improve estimation of survival, yielding even better tools for patient counseling and management, but also better stratifiers and surrogate endpoints for clinical trials.[83]

REFERENCES

1. Hirschfield GM, Karlsen TH, Lindor KD, et al. Primary sclerosing cholangitis. *Lancet* 2013;382:1587–99.
2. Boonstra K, Beuers U, Ponsioen CY. Epidemiology of primary sclerosing cholangitis and primary biliary cirrhosis: a systematic review. *J Hepatol* 2012;56:1181–8.
3. Boonstra K, Weersma RK, van Erpecum KJ, et al. Population-based epidemiology, malignancy risk, and outcome of primary sclerosing cholangitis. *Hepatology* 2013;58:2045–55.
4. Karlsen TH, Schrumpf E, Boberg KM. Update on primary sclerosing cholangitis. *Dig Liver Dis* 2010;42:390–400.
5. Eaton JE, Talwalkar J a, Lazaridis KN, et al. Pathogenesis of primary sclerosing cholangitis and advances in diagnosis and management. *Gastroenterology* 2013;145:521–36.
6. Trauner M, Halilbasic E, Kazemi-Shirazi L, et al. Therapeutic role of bile acids and nuclear receptor agonists in fibrosing cholangiopathies. *Dig Dis* 2014;32:631–6.
7. Lindor KD, Kowdley K V, Luketic V a C, et al. High-dose ursodeoxycholic acid for the treatment of primary sclerosing cholangitis. *Hepatology* 2009;50:808–14.
8. Olsson R, Boberg KM, de Muckadell OS, et al. High-dose ursodeoxycholic acid in primary sclerosing cholangitis: a 5-year multicenter, randomized, controlled study. *Gastroenterology* 2005;129:1464–72.
9. Lindor KD. Ursodiol for primary sclerosing cholangitis. *NEJM* 1997;336:691–5.
10. Mitchell S a., Bansil DS, Hunt N, et al. A preliminary trial of high-dose ursodeoxycholic acid in primary sclerosing cholangitis. *Gastroenterology* 2001;121:900–7.
11. Färkkilä M, Karvonen A-L, Nurmi H, et al. Metronidazole and ursodeoxycholic acid for primary sclerosing cholangitis: a randomized placebo-controlled trial. *Hepatology* 2004;40:1379–86.
12. Hommes DW, Erkelens W, Ponsioen C, et al. A double-blind, placebo-controlled, randomized study of infliximab in primary sclerosing cholangitis. *J Clin Gastroenterol* 2008;42:522–6.
13. Thiel DH Van, Carroll P, Abu-elmagd K, et al. Tacrolimus (FK), a Treatment for Primary Sclerosing Cholangitis; Results of an Open-Label Preliminary Trial. *Am J Gastroenterol* 2010;90:455–9.
14. Angulo P, Batts KP, Jorgensen R a, et al. Oral budesonide in the treatment of primary sclerosing cholangitis. *Am J Gastroenterol* 2000;95:2333–7.
15. Karlsen TH, Vesterhus M, Boberg KM. Review article: controversies in the management of primary biliary cirrhosis and primary sclerosing cholangitis. *Aliment Pharmacol Ther* 2014;39:282–301.
16. Ter Borg PCJ, Fekkes D, Vrolijk JM, et al. The relation between plasma tyrosine concentration and fatigue in primary biliary cirrhosis and primary sclerosing cholangitis. *BMC Gastroenterol* 2005;5:11.
17. Björnsson E, Simren M, Olsson R, et al. Fatigue in patients with primary sclerosing cholangitis. *Scand J Gastroenterol* 2004;39:961–8.
- *18. Benito de Valle M, Rahman M, Lindkvist B, et al. Factors that reduce health-related quality of life in patients with primary sclerosing cholangitis. *Clin Gastroenterol Hepatol* 2012;10:769–75.e2. This is the first publication that assessed health related quality of life in PSC patients and investigated the association with disease severity.

- 19 Gotthardt DN, Rupp C, Bruhin M, et al. Pruritus is associated with severely impaired quality of life in patients with primary sclerosing cholangitis. *Eur J Gastroenterol Hepatol* 2014;:1–6.
- *20 Stanich PP, Björnsson E, Gossard A a, et al. Alkaline phosphatase normalization is associated with better prognosis in primary sclerosing cholangitis. *Dig Liver Dis* 2011;43:309–13.
- *21 Al Mamari S, Djordjevic J, Halliday JS, et al. Improvement of serum alkaline phosphatase to <1.5 upper limit of normal predicts better outcome and reduced risk of cholangiocarcinoma in primary sclerosing cholangitis. *J Hepatol* 2013;58:329–34.
- *22 Lindström L, Hultcrantz R, Boberg KM, et al. Association between reduced levels of alkaline phosphatase and survival times of patients with primary sclerosing cholangitis. *Clin Gastroenterol Hepatol* 2013;11:841–6.
- *23 Rupp C, Rössler a, Halibasic E, et al. Reduction in alkaline phosphatase is associated with longer survival in primary sclerosing cholangitis, independent of dominant stenosis. *Aliment Pharmacol Ther* 2014;:1–10.
Of the 4 publications covering the prognostic value of reduction of alkaline phosphatase over a certain time frame, the publications of Lindström et al. and Rupp et al. may be most illustrative, as they truncated time to normalization at 1 year, thereby minimizing the chance of looking at effects of self fulfilling prophecy rather than the true prognostic value of alkaline phosphatase.
- 24 De Vries E, Boonstra K, Leeftang M, et al. Normalization of alkaline phosphatase within the first year after diagnosis is of prognostic value in primary sclerosing cholangitis. *Gastroenterology; Conference Digestive Disease Week 2014, DDW 2014 Chicago, United States* 2014;:907.
- 25 Vesterhus M, Hov J, Holm A, et al. Enhanced Liver Fibrosis score predicts transplant-free survival in PSC independently of the Mayo risk score. *J Hepatol; Conference 49th Annual Meet European Association Study Liver, International Liver Congress London United Kingdom* 2014;:188.
- *26 Nyblom H, Nordlinder H, Olsson R. High aspartate to alanine aminotransferase ratio is an indicator of cirrhosis and poor outcome in patients with primary sclerosing cholangitis. *Liver Int* 2007;27:694–9.
This is the only study that investigates the prognostic value of AST/ALT ratio on disease severity and clinical outcome in PSC patients.
- *27 Kim W, Therneau T, Wiesner R, et al. A revised natural history model for primary sclerosing cholangitis. *Mayo Clin Proc* 2000;75:688–94.
The only validated prognostic model including clinical and biochemical variables of PSC disease progression.
- 28 Broome U, Olsson R, Loof L, et al. Natural history and prognostic factors in 305 Swedish patients with primary sclerosing cholangitis. 1996;:610–5.
- 29 Angulo P, Peter JB, Gershwin ME, et al. Serum autoantibodies in patients with primary sclerosing cholangitis. 2000;32:182–7.
- 30 Benito de Valle M, Müller T, Björnsson E, et al. The impact of elevated serum IgG4 levels in patients with primary sclerosing cholangitis. *Dig Liver Dis* 2014;46:903–8.
- 31 De Vries E., Verheij J, Hubscher SG, et al. Applicability and Prognostic Value of Histologic Scoring Systems in Primary Sclerosing Cholangitis. *Hepatology; Conference 65th Annual Meeting American Association Study Liver Disease Liver Meeting Boston, MA United States* 2014;:349A.

- 32 Craig DA, Larusso F. Primary Sclerosing Value of Cholangiography in Determining the Prognosis. *AJR* 1991;157:959–64.
- *33 Ruiz A, Lemoine S, Carrat F, et al. Radiologic course of primary sclerosing cholangitis: assessment by three-dimensional magnetic resonance cholangiography and predictive features of progression. *Hepatology* 2014;59:242–50.
The first study that assessed the disease course of PSC by means of changes of findings on 3D MRC over time.
- *34 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;146:970–9; quiz e15–6.
This is the first study to investigate the usefulness of VCTE in PSC, showing a association with disease severity as well as association with clinical outcome.
- *35 Ponsioen CY, Vrouenraets SME, Prawirodirdjo W, et al. Natural history of primary sclerosing cholangitis and prognostic value of cholangiography in a Dutch population. *Gut* 2002;51:562–6.
This is the only validated prognostic model, incorporating transplant-free survival as clinical outcome measure.
- 36 Ponsioen CY, Reitsma JB, Boberg KM, et al. Validation of a cholangiographic prognostic model in primary sclerosing cholangitis. 2010;:742–7.
- 37 Ter Borg PCJ, van Os E, van den Broek WW, et al. Fluvoxamine for fatigue in primary biliary cirrhosis and primary sclerosing cholangitis: a randomised controlled trial [ISRCTN88246634]. *BMC Gastroenterol* 2004;4:13.
- 38 Kremer AE, Bolier R, van Dijk R, et al. Advances in pathogenesis and management of pruritus in cholestasis. *Dig Dis* 2014;32:637–45.
- 39 Younossi Z, Kiwi M, Boparai N, et al. Cholestatic Liver Diseases and Health-Related Quality of Life. *Am J Gastroenterol* 2000;95:0–5.
- 40 Younossi Z, Boparai N, Price LL, et al. Health-Related Quality of Life in Chronic Liver Disease : The Impact of Type and Severity of Disease. *Am J Gastroenterol* 2001;96:2199–205.
- 41 Hauser W, Holtmann G, Grandt D. Determinants of Health-Related Quality of Life in Patients With Chronic Liver Diseases. *Clin Gastroenterol Hepatol* 2004;2:157–63.
- 42 Afendy a, Kallman JB, Stepanova M, et al. Predictors of health-related quality of life in patients with chronic liver disease. *Aliment Pharmacol Ther* 2009;30:469–76.
- 43 Association E. EASL Clinical Practice Guidelines: management of cholestatic liver diseases. *J Hepatol* 2009;51:237–67.
- 44 Rosenberg WMC, Voelker M, Thiel R, et al. Serum markers detect the presence of liver fibrosis: A cohort study. *Gastroenterology* 2004;127:1704–13.
- 45 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;18:23–31.
- 46 Guha IN, Parkes J, Roderick P, et al. Noninvasive markers of fibrosis in non-alcoholic fatty liver disease: Validating the European Liver Fibrosis Panel and exploring simple markers. *Hepatology* 2008;47:455–60.
- 47 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;48:1549–57.

- 48 Parkes J, Roderick P, Harris S, et al. Enhanced liver fibrosis test can predict clinical outcomes in patients with chronic liver disease. *Gut* 2010;59:1245–51.
 - 49 Giboney PT. Mildly Elevated Liver Transaminase Levels in the Asymptomatic Patient. *Am Fam Physician* 2005;71:1105–10.
 - 50 Nyblom H, Berggren U, Balldin J, et al. High AST/ALT ratio may indicate advanced alcoholic liver disease rather than heavy drinking. *Alcohol Alcohol* 2004;39:336–9.
 - 51 Park GJ, Lin BPC, Ngu MC, et al. Aspartate aminotransferase : alanine aminotransferase ratio in chronic hepatitis C infection : Is it a useful predictor of cirrhosis ? 2000;;386–90.
 - 52 Sheth SG, Flamm SL, Gordon FD, et al. AST / ALT Ratio Predicts Cirrhosis in Patients With Chronic Hepatitis C Virus Infection. 1998;93:44–8.
 - 53 Nyblom H, Björnsson E, Simrén M, et al. The AST/ALT ratio as an indicator of cirrhosis in patients with PBC. *Liver Int* 2006;26:840–5.
 - 54 Wiesner R, Grambsch P, Dickson E, et al. Primary sclerosing cholangitis: natural history, prognostic factors and survival analysis. *Hepatology* 1989;10:430–6.
 - 55 Dickson E, Murtaugh P, Wiesner R, et al. Primary sclerosing cholangitis: refinement and validation of survival models. *Gastroenterology* 1992;103:1893–901.
 - 56 Boberg KM, Rocca G, Egeland T, et al. Time-dependent Cox regression model is superior in prediction of prognosis in primary sclerosing cholangitis. *Hepatology* 2002;35:652–7.
 - 57 Tischendorf JJW, Hecker H, Krüger M, et al. Characterization, outcome, and prognosis in 273 patients with primary sclerosing cholangitis: A single center study. *Am J Gastroenterol* 2007;102:107–14.
 - 58 Hov J-R, Boberg KM, Kalsen TH. Autoantibodies in primary sclerosing cholangitis. *World J Gastroenterol* 2008;14:3781.
 - 59 Björnsson E, Chari ST, Smyrk TC, et al. Immunoglobulin G4 associated cholangitis: description of an emerging clinical entity based on review of the literature. *Hepatology* 2007;45:1547–54.
 - 60 Björnsson E, Chari S, Silveira M, et al. Primary Sclerosing Cholangitis Associated with Elevated ImmunoglobulinG4 : Clinical Characteristics and Response to Therapy. *Am J Ther* 2011;205:198–205.
 - 61 Mendes FD, Jorgensen R, Keach J, et al. Elevated serum IgG4 concentration in patients with primary sclerosing cholangitis. *Am J Gastroenterol* 2006;101:2070–5.
 - 62 Boonstra K, Culver EL, de Buy Wenniger LM, et al. Serum immunoglobulin G4 and immunoglobulin G1 for distinguishing immunoglobulin G4-associated cholangitis from primary sclerosing cholangitis. *Hepatology* 2014;59:1954–63.
 - 63 Olsson R, Hägerstrand I, Broomé U, et al. Sampling variability of percutaneous liver biopsy in primary sclerosing cholangitis. *J Clin Pathol* 1995;48:933–5.
 - *64 Angulo P, Larson DR, Thorneau TM, et al. Time course of histological progression in primary sclerosing cholangitis. *Am J Gastroenterol* 1999;94:3310–3.
- With this study the change in histologic stage over time, independent of effects of sampling variability as demonstrated, implicating histology can be used to assess PSC disease progression over time.
- 65 Ludwig J, Dickson E, McDonald G. Staging of chronic nonsuppurative destructive cholangitis (syndrome of primary biliary cirrhosis). *Virchows Arch A Pathol Anat Histol* 1978;22:103–12.

- 66 Ishak K, Baptista a, Bianchi L, et al. Histological grading and staging of chronic hepatitis. *J Hepatol* 1995;22:696–9.
- 67 Nakanuma Y, Zen Y, Harada K, et al. Application of a new histological staging and grading system for primary biliary cirrhosis to liver biopsy specimens: Interobserver agreement. *Pathol Int* 2010;60:167–74.
- 68 Maccarty R, Larusso N, Wiesner R, et al. Primary Sclerosing Cholangitis: Findings on Cholangiography and Pancreatography. *Radiology* 1983;149:39–44.
- 69 Majoie CBL., Reeders WAJ, Huibregtse K, et al. Primary Sclerosing Cholangitis : Cholangiographic Findings A Modified Classification. *AJR* 1991;157:495–7.
- 70 Rajaram R, Ponsioen CY, Majoie CBM, et al. Evaluation of a modified cholangiographic classification system for primary sclerosing cholangitis. *Abdom Imaging* 2001;26:43–7.
- 71 Maneesh D, Elmunzer BJ, Dwamena B, et al. Primary Sclerosing Cholangitis: Meta-Analysis of Diagnostic Performance of MR Cholangiopancreatography. *Radiology* 2010;256:387–96.
- 72 Weber C, Kuhlencordt R, Grotelueschen R, et al. Magnetic resonance cholangiopancreatography in the diagnosis of primary sclerosing cholangitis. *Endoscopy* 2008;40:739–45.
- 73 Petrovic BD, Nikolaidis P, Hammond N a, et al. Correlation between findings on MRCP and gadolinium-enhanced MR of the liver and a survival model for primary sclerosing cholangitis. *Dig Dis Sci* 2007;52:3499–506.
- 74 Fulcher AS, Turner MA, Franklin KJ, et al. Gastrointestinal Imaging Primary Sclerosing Cholangitis: Evaluation with MR Cholangiography — A Case-Control Study. *Radiology* 2000;:71–80.
- 75 Sodickson A, Koenraad MJ, Barish MA, et al. Three-dimensional Fast- Recovery Fast Spin-Echo MRCP: Comparison with Two-dimensional Single-Shot Fast Spin-Echo Techniques. *Radiology* 2006;238:549–59.
- 76 Tapper EB, Castera L, Afdhal NH. Fibro-Scan (Vibration-Controlled Transient Elastography): Where Does It Stand in the United States Practice. *Clin Gastroenterol Hepatol* Published Online First: 5 June 2014.
- 77 Ehlken H, Lohse AW, Schramm C. Transient elastography in primary sclerosing cholangitis-the value as a prognostic factor and limitations. *Gastroenterology* 2014;147:542–3.
- 78 Wunsch E, Trottier J, Milkiewicz M, et al. Prospective evaluation of ursodeoxycholic acid withdrawal. *Hepatology* 2014;:1–33.
- 79 Liu JZ, Hov JR, Folseraas T, et al. Dense genotyping of immune-related disease regions identifies nine new risk loci for primary sclerosing cholangitis. *Nat Genet* 2013;45:670–5.
- 80 Karlsen TH, Franke A, Melum E, et al. Genome-wide association analysis in primary sclerosing cholangitis. *Gastroenterology* 2010;138:1102–11.
- 81 Janse M, Lamberts LE, Franke L, et al. Three ulcerative colitis susceptibility loci are associated with primary sclerosing cholangitis and indicate a role for IL2, REL, and CARD9. *Hepatology* 2011;53:1977–85.
- 82 Melum E, Franke A, Schramm C, et al. Genome-wide association analysis in primary sclerosing cholangitis identifies two non-HLA susceptibility loci. *Nat Genet* 2011;43:17–9.
- 83 Talwalkar J a, Lindor KD. Natural history and prognostic models in primary sclerosing cholangitis. *Best Pract Res Clin Gastroenterol* 2001;15:563–75.

Alkaline phosphatase at diagnosis of primary sclerosing cholangitis and one year later: evaluation of prognostic value

Elisabeth M.G. de Vries, Junfeng Wang, Mariska M.G. Leeflang, Kirsten Boonstra, Rinse K. Weersma, Ulrich H. Beuers, Ronald B. Geskus, Cyriel Y. Ponsioen

Liver International 2016; epub ahead of print

ABSTRACT

Background

Primary sclerosing cholangitis (PSC) is a slowly progressive liver disease. Reliable biomarkers to predict outcome are urgently needed to serve as surrogate end-points and/or stratifiers in clinical trials. Reduction of serum alkaline phosphatase (ALP) has been proposed as prognostic surrogate marker in PSC. The aim of this study was to assess if ALP at diagnosis (T0), 1 year later (T1), and percentage change between both time points holds prognostic value, and to determine the optimal threshold.

Methods

We retrospectively collected ALP levels at T0 and T1 for patients included in a large PSC cohort. The association of ALP at T0, T1, and percentage change with the combined endpoint (PSC-related death, liver transplantation) was analysed. Predictive value was determined using C-statistics.

Results

366 patients were included, of whom 66 (18%) reached an endpoint: 26 (7%) PSC-related death, 40 (11%) liver transplantation. At T0 and T1, 84% used ursodeoxycholic acid. A positive association was observed between level of ALP at T0 and T1 and the hazard of reaching an endpoint, up to values around 2.5 times upper limit of normal (xULN). A larger decrease in ALP between T0 and T1 decreased the event rate. A range of thresholds (0.5-3xULN) with about similar C-statistics was found. In this cohort, the optimal threshold was 1.3xULN at T1.

Conclusion

ALP can be used to discriminate between PSC patients with a good and a poor prognosis. These findings indicate that ALP can serve as stratifier, and potentially as surrogate endpoint for clinical trials in PSC.

Key points

- Levels of alkaline phosphatase at diagnosis (T0) and after one year (T1)
- are associated with clinical outcome.
- A range of ALP thresholds from 0.5 to 3 xULN have close to equal C-statistic, the optimal threshold in this mostly population based cohort is 1.3xULN at T1.
- The biomarker ALP may be used as a disease outcome stratifier.
- ALP could be implemented as a biomarker to be incorporated in prognostic modelling in PSC

INTRODUCTION

Primary sclerosing cholangitis (PSC) is a rare cholestatic liver disease, causing chronic inflammation and destruction that can affect the entire biliary tree.¹ This results in characteristic multifocal bile duct strictures interchanged with dilations, biliary fibrosis, and ultimately end stage liver failure.¹ PSC is a pre-malignant condition, predisposing to the development of colorectal carcinoma (CRC), cholangiocarcinoma (CCA) and gallbladder carcinoma.^{2,3} Currently, the median transplant-free survival as assessed on a population-based level is 21 years.³

Over the last decades, several drug therapies such as immunomodulatory and bile acid agents have been tested for PSC. Unfortunately, none of these therapies have proved successful in halting disease progression.⁴ Once patients have progressed to end stage liver disease, the only treatment option is liver transplantation (LTx), but optimal timing for listing of LTx is often difficult to determine.⁵ It is obvious that there is a big unmet need for an effective medical therapy for PSC patients. However, investigating the effectiveness of potential new treatments in clinical trials is complicated by the slow progression, and low event-rate of the disease, and the resulting difficulty to pick up differences between placebo and treatment arms by the traditionally used ‘solid endpoints’, such as PSC related death and LTx. If a biomarker with the ability to accurately assess disease activity and progression in PSC within a reasonable time frame can be identified, this may serve as a stratifier, and surrogate endpoint to measure treatment effect in clinical trials.

In primary biliary cholangitis (PBC), a cholestatic liver disease showing a range of histologic changes similar to PSC, biochemical response after 1 year to treatment with ursodeoxycholic acid (UDCA) was repeatedly shown to be associated with a favorable effect on survival, posing the idea that these biochemical parameters could serve as biomarkers of long-term prognosis.^{6–8} A recently published large meta-analysis confirmed that ALP and bilirubin can predict outcome in PBC, and may therefore be used as surrogate endpoints in clinical trials.⁹ Similar to PBC, an increased serum ALP is the most common liver test abnormality in PSC.^{10,11}

Several recent publications have reported that a reduction of ALP in a certain time frame is of prognostic value for PSC disease course (Table 1).^{12–15} The first two of these studies by Stanich et al. and Al Mamari et al., were limited by the absence of a limit for time to ALP reduction, thereby affecting the validity of the ALP measurement to predict time to event, – patients that approach an endpoint will likely have an increased ALP – instead of analyzing the true prognostic value of ALP.^{12,13} In addition, all studies were undertaken in relatively small cohorts, and time until reduction, as well as thresholds for reduction of ALP differed between studies (Table 1).^{12–15}

The aim of the present study was to assess the prognostic value of ALP at diagnosis, at one-year follow-up, and percentage change between both time points in a large, PSC cohort. Furthermore we aimed to determine the optimal threshold value to distinguish favorable from unfavorable disease course in our cohort.

Table 1. Publications on the prognostic value of alkaline phosphatase in PSC

First Author, year	Year of publication	Cohort	Patient cohort n; FU time, years, median	ALP threshold: Definition	Composite endpoint
Stanich et al.	2011	Mayo clinic Royester, USA	n= 87; FU: 7 (IQR 5-9)	Normalization (1xULN): Normalization: one/ more ALP values of 1xULN or less	- CCA - LTx - Death (all cause)
Al Mamari et al.	2013	Oxford, UK	n=139; FU: 10 (IQR 6-15)	1.5xULN: At least 3 ALP values <1.5 ULN, at least 6 months apart, continued to be <1.5xULN for the study follow-up period.	- Liver decompensation - LTx - Death (liver related, incl death from CCA)
Lindström et al.	2012	Scandinavian UDCA trial	n=195; FU: 11 (range 0-13)	Normalization (1xULN) or reduced ALP by >40%: after 1 year in the trial	- CCA - LTx - Death (all cause)
Rupp et al.	2014	Heidelberg, Germany	n=215; FU: 10	ALP normalization, 50% reduction compared with baseline values, or reduction < 1.5 xULN: within 6 months. ALP reduction < 1.5 xULN: within 12 months	- LTx - Death (all cause)

ALP = alkaline phosphatase; FU = follow up; xULN = times upper limit of normal; CCA = cholangiocarcinoma; LTx = liver transplantation

MATERIALS AND METHODS

Study design

This retrospective cohort study is part of the Epi PSC PBC project, a large observational longitudinal cohort study of PSC and PBC in the Netherlands. All PSC patients that were treated from 2000 onwards, were identified in 44 hospitals situated in a geographically defined area of 6 adjacent provinces, comprising 50% of the Dutch population (2007: 7,758,980 inhabitants). The protocol was approved by the central Committee for Research Ethics in Utrecht and all 44 local ethics committees of the participating hospitals in the Netherlands (trialregister.nl number, NTR2813). In addition, patients were included from one academic medical centre, that served as referral centre for LTx. Anonymous processing of these patient data precluded the need for patient informed consent.

Case-finding and case-ascertainment methods have been described previously.³

Data collection

Clinical data were collected from medical records, and included: gender, date of PSC diagnosis, PSC type, concomitant IBD, date of IBD diagnosis, IBD type, auto-immune hepatitis overlap (AIH-overlap), and the use of UDCA. Biochemical values at time of PSC diagnosis (T0) and after 1 year (T1) were collected for aspartate aminotransferase (AST), alanine aminotransferase (ALT), ALP and total bilirubin. Gamma-glutamyl transpeptidase (γ GT) was only collected at T0. Biochemical values within a range of 3 months around T0 and T1 were accepted as entries. If a patient had more than one biochemical value within the range of 3 months, the value taken at the date closest to T0 or T1 was used.

Endpoint

To analyze the association of ALP with outcome, a composite endpoint was defined, composed of LTx and PSC related death (death from end-stage liver failure, liver surgery, CCA). Since the risk of developing CCA in PSC is 400-fold increased when compared with the general population, CCA related mortality was treated as PSC-related death.³ Individuals that received LTx before 2000 were included, but those that died before 2000 were not.

Survival time was calculated as the interval between date of PSC diagnosis and date of first occurring clinical endpoint, or, in case of no endpoint reached, date of last follow-up was used as censoring date. Patients with non-PSC related death were censored at time of death.

Statistical analysis

Patient characteristics were expressed as mean $\pm$ standard deviation (SD), and median and interquartile range (IQR) when having asymmetrical and skewed distribution, respectively. Dichotomous variables were expressed as percentage (%) of the cohort. As assays can vary between hospitals and over time, biochemistry values were expressed as ratio of the upper limit of normal (xULN). Patients that reached a clinical endpoint within one year were excluded from the analysis at T1. The effect of ALP level at T0, T1, and percentage change $(\text{ALP(T1)} - \text{ALP(T0)}) / \text{ALP(T0)}$ on the hazard of reaching the combined clinical endpoint was assessed using Cox models with restricted cubic spline functions, which models smoothly varying non-linear trends in a flexible way. The knots were selected based on Harrell's recommended percentiles at 0.05, 0.35, 0.65 and 0.95.¹⁶ Significance of the effect was assessed using likelihood ratio test.

To assess both the short-term and long-term predictive value of ALP, the time-dependent C-statistic as proposed by Uno et al. was calculated for each month in the follow-up period.¹⁷

Subsequently, the overall predictive value at T0 and T1 was calculated for each threshold between 0.5xULN to 3.0xULN with a step of 0.1, using Harrell's overall C-statistic, from which the optimal threshold could be determined.¹⁸ Similar analysis was performed for percentage change, using thresholds ranging from -100% to 100%, with a step of 5%.¹⁸

Statistical analyses were performed using R version 3.0.1, with packages rms, Hmisc and timeROC. A p-value < 0.05 was considered statistically significant.

RESULTS

Study inclusion

A total of 692 PSC patients have been included in the Epi PSC PBC cohort. Serum ALP values at T0 and T1 were available and collected for 366 and 345 patients, respectively. Of the 366 patients, 53 came from the transplant center. Seven patients reached an endpoint within the first twelve months of follow-up. Thus, 366 patients could be included in the analysis at T0, 338 patients at T1, and 338 patients to analyze percentage change.

Patient characteristics

Patient characteristics are summarized in Table 2. There were no large differences in patient characteristics between patients with available data at T0, T1 and patients with missing values at T1.

When comparing biochemistry values at T0 with T1; AST, ALT ALP and total bilirubin all showed a large decrease the first year after diagnosis. During follow-up (median 100 months (IQR 67-150)) a total of 66 patients (18%) reached an endpoint; 26 (7%) PSC related death and 40 (11%) liver transplantations; 12 (18%) of these endpoints occurred within the subset of patients retrieved from the referral center. A total of 8 (2%) patients died of causes unrelated to PSC.

Table 2. Patient characteristics

	T0	T1	Missing T1
N	366	338	21
Male [n (%)]	241 (66)	220 (65)	15 (71)
Age at diagnosis PSC (years) [mean (SD)]	40 (14)	40 (14)	36 (16)
Large duct PSC [n (%)]	328 (89)	301 (89)	19 (91)
AIH overlap	22 (6)	22 (7)	0 (0)
Inflammatory bowel disease [n (%)]	252 (69)	237 (70)	10 (48)
Ulcerative colitis [n (%)]	194 (53)	182 (54)	8 (38)
Crohn's disease [n (%)]	50 (14)	48 (14)	1 (5)
Unspecified [n (%)]	8 (2)	7 (2)	1 (5)
IHD score at diagnosis [mean]*	1.46		
IHD score 0 [n (%)]	2 (3)		
IHD score 1 [n (%)]	44 (56)		
IHD score 2 [n (%)]	26 (33)		
IHD score 3 [n (%)]	6 (8)		

Table 2. Continued Patient characteristics

	T0	T1	Missing T1
EHD score at diagnosis [mean]]*	1.71		
EHD score 0 [n (%)]	11 (44)		
EHD score 1 [n (%)]	18 (23)		
EHD score 2 [n (%)]	36 (46)		
EHD score 3 [n (%)]	9 (12)		
EHD score 4 [n (%)]	4 (5)		
Ursodeoxycholic acid use [n (%)]*	249 (84)	232 (84)	15 (71)
Follow up time (months) [median (IQR)]	100 (67-150)	102 (68-153)	101 (72-163)
AST xULN [median (IQR)]*	1.45 (0.95-2.68)	0.88 (0.67-1.47)	
ALT xULN [median (IQR)]*	2.11 (1.22-3.97)	1.00 (0.62-1.88)	
ALP xULN [median (IQR)]*	1.99 (1.38-3.37)	1.22 (0.80-1.89)	
γGT xULN [median (IQR)]*	5.84 (3.42-10.72)	-	
Total bilirubin xULN [median (IQR)]*	0.81 (0.53-1.59)	0.65 (0.47-1.00)	

PSC = primary sclerosing cholangitis; SD = standard deviation; IQR = inter quartile range; AIH = auto immune hepatitis; IHD = intrahepatic disease; EHD = extrahepatic disease AST = aspartate aminotransferase; ALT = alanine aminotransferase; ALP = alkaline phosphatase; γGT = gamma-glutamyl transferase; xULN = times upper limit of normal;

* Data available for: T0: IHD score n =78 EHD score n =78 Ursodeoxycholic acid use n= 295 AST n=359, ALT n= 357, ALP n=366, γGT n=321 total bilirubin n= 335. T1: Ursodeoxycholic acid use n= 275 AST n=319, ALT n= 328, ALP n=338, total bilirubin n= 279

Predictive value of alkaline phosphatase at T0, T1 and percentage change

The association of ALP on the combined clinical endpoint was significant at both T0 ($p=0.002$) and T1 ($p<0.001$) (Fig. 1A,B). Both at T0 and T1, an increased ALP level was associated with an increased hazard of reaching a clinical endpoint in the range between 0.5-2.5xULN. At levels >2.5 xULN the hazard of reaching an endpoint seemed to level off.

Percentage change in ALP was not significantly associated with the clinical endpoint ($p=0.078$). However, a reduction in ALP level from T0 to T1 (a negative percentage change in ALP) lowered the hazard of reaching a clinical endpoint. Patients with a greater reduction in ALP level had a lower hazard (Fig. 1C).

The time specific C-statistics of ALP measurements at T1 were generally higher than those of the measurements at T0, as well as those of the percentage change over time (Figure 2 A-C). Therefore, ALP at T1 was found to hold the best predictive value. When taking into account the wide 95% confidence intervals (CI), the differences between T0, T1 and percentage change were relatively small. Notwithstanding, both percentage change and ALP at T1 had C-statistics significantly larger than 0.5 for most of the time range.

Fig. 1. Association between alkaline phosphatase level and hazard of reaching a clinical endpoint, depicted with 95% CI. (A) T0; (B): T1; (C) percentage change. Reference values were set at 1.0 xULN for T0 and T1, and 0% (no change) for percentage change.

Fig. 2. Course of Uno's C-statistic over follow-up for ALP, depicted with 95% CI. (A) T0; (B): T1; (C) percentage change.

Predictive value of alkaline phosphatase: the optimal threshold at T1

In Fig 3. the predictive value of all possible thresholds, ranging from 0.5xULN-3.0xULN is shown. See also Supplementary Table 1,2.

Of the 176 patients with ALP of <1.3xULN at T1, 161 (91%) did not reach an endpoint during follow-up, compared to 119 (73%) of the 162 patients with an ALP >1.3xULN. This difference in endpoint-free survival is illustrated in Fig. 4. Patients with ALP >1.3xULN at T1 had a 2.04 times higher hazard of reaching an endpoint compared with patients with ALP <1.3xULN at T1. The predictive power of 1.3xULN at T1 remained fairly constant over the course of follow-up (Fig. 5). The sensitivity and specificity of ALP at 1.3xULN at T1 are depicted in supplementary Figure 1. This showed a sensitivity of ALP that varied between 100% for a prediction early in the disease course (0-10 months) which decreased along the course of the FU to a sensitivity of approximately 65% at 200-250 months. The

specificity was relatively low over the whole range, indicating many false positive results. Both sensitivity and specificity decreased over time, indicating that the performance of ALP to predict prognosis is better for events closer to the time of diagnosis.

Fig. 3. Harrell's C- statistic per threshold, depicted with 95% CI. (A) T0; (B) T1; (C) percentage change. (B): The asterisk at 1.3xULN represents the optimal threshold of ALP at T1, holding highest predictive power for reaching an endpoint in our dataset. (Harrell's C-statistic 0.65 (95%CI 0.55-0.68)).

Fig 4. Kaplan Meier survival curve optimal threshold: 1.3xULN at T1.

Fig 5. Uno's C-statistic over follow-up for ALP with a threshold of 1.3xULN at T1, depicted with 95% CI.

DISCUSSION

In this large PSC cohort, we demonstrate that ALP levels at T0, T1 and percentage change are of prognostic value to predict endpoint-free survival in PSC, with an optimal predictive value at a threshold of 1.3xULN at T1.

This study was performed in a mostly population based cohort, composed by patients from the previously described population-based Epi PSC PBC cohort ³, and in addition patients that were derived from one LTx referral center. The cohort was not truly population based, however, additional association analyses- performed in the subset of patients coming from the true population based cohort- showed similar results (data not shown). We chose to include all patients in order to increase the number of events, and thereby the power of the study. In addition, using a blended population based- as well as referral cohort may give a better reflection of all degrees of disease severity, and results may therefore be more comparable to other PSC cohorts worldwide.

The hazard on reaching an endpoint approached a plateau at an ALP level beyond 2.0xULN.(Figure 1) There were only a few patients with an ALP >2.0 xULN in the cohort, so this plateau may be explained by an insufficient sample size. The wide CI's, however, implicate that an increase in prognostic value is possible beyond an ALP value of 2.0 xULN.

ALP level at T1 performed better in predicting more favorable outcome compared with ALP at T0, or percentage change between T0 and T1. Interestingly, when comparing ALP levels at T0 and T1, median ALP at T1 was significantly lower ($p<0.001$). This overall improvement in ALP during the first year of follow-up may have been the result of starting treatment (e.g. UDCA treatment or therapeutic endoscopic retrograde cholangiography (ERC)). It could be speculated that those patients who respond with a sizeable decrease in their ALP level to the introduction of UDCA within the first year after diagnosis might have a more favorable prognosis than those who have no or only a minor decrease, because this may reflect responsive versus refractory disease. This might explain the superior performance of T1-levels as compared with T0-levels, because the vast majority of patients were put on UDCA after diagnosis. This potential effect modifier property of UDCA was assessed with an additional sensitivity analysis of the prognostic value of percentage change in UDCA users only. Results were comparable with the prognostic value of the cohort as a whole. (Supplementary Figure 2). These findings are in concordance with those of the post-hoc analysis of the 5-year Scandinavian UDCA trial, where survival was clearly associated with >40% reduction or normalization of ALP level at one year irrespective of the treatment allocation. ¹⁴

In clinical practice thresholds are often used as guidance for decision making. In general, the stratification of patients by risk on outcome, using a threshold holds several drawbacks. First of all, the categorization of a continuous variable will result in the loss of information, increasing the risk of failing to detect a real association.^{19,20} Secondly, different studies may find different thresholds due to random variability in different data sets. Previous studies reporting on the association between a reduced ALP level and prognosis, have used various thresholds to define reduction in ALP (table1). As opposed to these previous studies, the present study used a systematic approach to determine the ‘optimal’ threshold by using the C-statistic as measure of predictive discrimination. The optimal threshold of 1.3 xULN at T1 comes close to the thresholds previously reported, which all are within the 95% CI associated with 1.3xULN in our dataset. The disadvantage of using a particular threshold is emphasised by the observation that the dichotomized optimal ALP threshold of 1.3xULN yielded a lower predictive value compared with ALP as a continuous value (see Fig. 5 as compared with Fig. 2B).

Analyses of the association of ALP with outcome were performed using the composite endpoint LTx and PSC related death (death from: end-stage liver failure, liver surgery, CCA). The risk of developing CCA in PSC is approximately 400 times increased when compared with the general population.³ Chronic inflammation may predispose to neoplasia as in other chronic inflammatory disorders. Therefore we felt that death from CCA is part of the disease behaviour and choose not to censor patients who died as a result of CCA at time of death. To assess the influence of death from CCA on this endpoint, a sensitivity analysis was performed in which the effect of ALP at T0 on the hazard of reaching the endpoint was re-calculated, whilst censoring the 19 patients that had developed CCA in this cohort. The splines in this sensitivity analysis did not differ from the splines in Fig 1A (Supplementary figure 3A). Similar sensitivity analysis was performed for the patients that developed CCA only, whilst censoring all other patients. In this subgroup of PSC-CCA patients, the level of ALP at diagnosis was not predictive of time to develop CCA. (Supplementary figure 3B) This result should, however, be interpreted with caution since the number of PSC-CCA patients was small. Overall the additional sensitivity analyses suggest that occurrence of CCA did not have a major influence on the prognostic value of ALP in our dataset. In accordance with the composite endpoint used in previous studies, which included CCA diagnosis instead of death from CCA in the composite endpoint^{12,14,21}, all analyses were repeated while using the composite endpoint: diagnosis of CCA, LTx, and PSC related death (death from end-stage liver failure, liver surgery, CRC). No difference in results were found (data not shown).

A limitation of this study is that due to its retrospective nature, biochemical values were unfortunately not available for all patients that have been included in the Epi PSCPBC cohort. However, we think that missingness was completely at random, because the majority of missing values concerned patients that were diagnosed in the late 80's- early 90's of whom patient records were not always retrievable. Since treatment with regards to the use of UDCA or indications for LTx in the Netherlands have not changed over the last decades, this is not expected to induce bias. Second, the level of ALP may be influenced by several factors, such as an episode of acute cholangitis, or intervention by therapeutic ERC. Unfortunately, data on these items, including symptoms were incomplete, which precluded assessment of this issue. However, Rupp et al. reported that the association between reduction in ALP and longer survival was independent of dominant strictures.¹⁵ Lastly, PSC diagnosis is made in different histological disease stages. As liver biopsy is no longer routinely performed for PSC diagnosis, liver histology was available in only a subgroup of patients. Of those patients, data on actual histologic stage were not reported in all pathology reports. Therefore it was not possible to stratify patients into early and advanced disease stage.

Individuals that were diagnosed and died before 2000 were excluded, but we included LTx events before 2000 in order to increase power. We checked whether this would induce left truncation bias by comparing the overall Kaplan-Meier time-to-event curve with the one that only includes follow-up from 2000 onwards. Results were practically the same.

ALP has been widely used as surrogate endpoint, in over 40% of clinical trials in PSC, but was never validated.²² This is the first study in which the performance of ALP to predict prognosis in PSC is evaluated by means of a robust measure of prognostic value, the c-statistic. The c-statistic between 0.5 and 0.7 and the wide corresponding confidence intervals indicated that the predictive accuracy of ALP at T0 and T1 is relatively low, but statistically significant. Therefore, ALP is an important candidate surrogate endpoint/stratifier in PSC. Future randomized controlled clinical trials should demonstrate that substantial effects on ALP reliably reflect clinically meaningful effects (e.g. PSC related death and LTx).²³

In conclusion, we confirm in this large PSC cohort that ALP is of prognostic value for disease course. Our data indicate that values at 1 year after diagnosis bear somewhat more predictive power than at diagnosis, warranting further studies to explore the use of ALP as surrogate endpoint for clinical trials.

REFERENCES

- Williamson, K. D. & Chapman, R. W. Primary sclerosing cholangitis. *Dig. Dis.* 32, 438–45 (2014).
- Hirschfield, G. M., Karlsen, T. H., Lindor, K. D. & Adams, D. H. Primary sclerosing cholangitis. *Lancet* 382, 1587–99 (2013).
- Boonstra, K. et al. Population-based epidemiology, malignancy risk, and outcome of primary sclerosing cholangitis. *Hepatology* 58, 2045–55 (2013).
- Cullen, S. & Chapman, R. The medical management of primary sclerosing cholangitis. *Semin. Liver Dis.* 26, 52–61 (2008).
- Karlsen, T. H., Vesterhus, M. & Boberg, K. M. Review article: controversies in the management of primary biliary cirrhosis and primary sclerosing cholangitis. *Aliment. Pharmacol. Ther.* 39, 282–301 (2014).
- Parés, A., Caballería, L. & Rodés, J. Excellent long-term survival in patients with primary biliary cirrhosis and biochemical response to ursodeoxycholic acid. *Gastroenterology* 130, 715–720 (2006).
- Corpechot, C. et al. Biochemical response to ursodeoxycholic acid and long-term prognosis in primary biliary cirrhosis. *Hepatology* 48, 871–877 (2008).
- Kuiper, E. M. M. et al. Improved Prognosis of Patients With Primary Biliary Cirrhosis That Have a Biochemical Response to Ursodeoxycholic Acid. *Gastroenterology* 136, 1281–1287 (2009).
- Lammers, W. J. et al. Levels of alkaline phosphatase and bilirubin are surrogate end points of outcomes of patients with primary biliary cirrhosis: an international follow-up study. *Gastroenterology* 147, 1338–49.e5; quiz e15 (2014).
- Chapman, R. et al. Diagnosis and management of primary sclerosing cholangitis. *AASLD Practice Guidelines. Hepatology* 51, 660–78 (2010).
- Poupon, R. Liver alkaline phosphatase: A missing link between cholestasis and biliary inflammation. *Hepatology* (2015).
- Stanich, P. P. et al. Alkaline phosphatase normalization is associated with better prognosis in primary sclerosing cholangitis. *Dig. Liver Dis.* 43, 309–13 (2011).
- Al Mamari, S., Djordjevic, J., Halliday, J. S. & Chapman, R. W. Improvement of serum alkaline phosphatase to <1.5 upper limit of normal predicts better outcome and reduced risk of cholangiocarcinoma in primary sclerosing cholangitis. *J. Hepatol.* 58, 329–34 (2013).
- Lindström, L., Hultcrantz, R., Boberg, K. M., Friis-Liby, I. & Bergquist, A. Association between reduced levels of alkaline phosphatase and survival times of patients with primary sclerosing cholangitis. *Clin. Gastroenterol. Hepatol.* 11, 841–6 (2013).
- Rupp, C. et al. Reduction in alkaline phosphatase is associated with longer survival in primary sclerosing cholangitis, independent of dominant stenosis. *Aliment. Pharmacol. Ther.* 1–10 (2014).
- Harrell, F. E., Jr. *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis.* (2001).
- Uno, H., Cai, T., Tian, L. & Wei, L. J. Evaluating Prediction Rules for t-Year Survivors With Censored Regression Models. *J. Am. Stat. Assoc.* 102, 527–537 (2007).
- Harrell, F. J., Lee, K. & Mark, D. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat. Med.* 15, 361–87 (1996).

19. Altman, D. G., Lausen, B., Sauerbrei, W. & Schumacher, M. Dangers of Using 'Optimal' Cutpoints in the Evaluation of Prognostic Factors. *J. Natl. Cancer Inst.* 86, 829–835 (1994).
20. Buettner, P., Garbe, C. & Guggenmoos-Hofzmann, I. Problems in defining cutoff points of continuous prognostic factors: Example of tumor thickness in primary cutaneous melanoma. *J. Clin. Epidemiol.* 50, 1201–1210 (1997).
21. De Vries, E. M. G. et al. Applicability and prognostic value of histologic scoring systems in primary sclerosing cholangitis. *J. Hepatol.* 63, 1212–1219 (2015).
22. Ponsioen, C. Y. 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*,
23. Fleming, T. R. & Powers, J. H. Biomarkers and surrogate endpoints in clinical trials. *Stat. Med.* 31, 2973–2984 (2012).

Supplementary Table 1. Performance of ALP thresholds for predicting endpoint-free survival at T0 and T1.

Threshold ALP	C-statistics T 0	95% CI	C-statistics T 1	95% CI
0.5	0.506	0.49-0.51	0.521	0.49-0.54
0.6	0.514	0.49-0.53	0.532	0.47-0.59
0.7	0.520	0.49-0.54	0.560	0.48-0.63
0.8	0.529	0.50-0.55	0.578	0.48-0.66
0.9	0.532	0.47-0.58	0.602	0.51-0.69
1	0.533	0.46-0.60	0.608	0.50-0.71
1.1	0.550	0.47-0.62	0.623	0.50-0.73
1.2	0.560	0.48-0.63	0.644	0.52-0.76
1.3	0.574	0.49-0.65	0.649	0.52-0.77
1.4	0.558	0.45-0.65	0.640	0.51-0.77
1.5	0.574	0.47-0.67	0.627	0.48-0.76
1.6	0.562	0.44-0.67	0.637	0.49-0.77
1.7	0.570	0.45-0.68	0.633	0.49-0.77
1.8	0.574	0.45-0.69	0.629	0.49-0.76
1.9	0.570	0.44-0.69	0.612	0.47-0.75
2	0.540	0.40-0.67	0.620	0.48-0.75
2.1	0.542	0.40-0.67	0.622	0.48-0.76
2.2	0.539	0.40-0.67	0.616	0.47-0.75
2.3	0.557	0.42-0.68	0.595	0.46-0.72
2.4	0.540	0.40-0.67	0.597	0.46-0.73
2.5	0.541	0.40-0.67	0.588	0.45-0.72
2.6	0.518	0.38-0.64	0.584	0.45-0.71
2.7	0.518	0.39-0.64	0.574	0.44-0.70
2.8	0.490	0.36-0.61	0.575	0.44-0.70
2.9	0.494	0.37-0.61	0.578	0.44-0.70
3	0.495	0.37-0.61	0.582	0.45-0.71

Supplementary Table 2. Performance of ALP thresholds for predicting endpoint-free survival at percentage change.

Threshold	C-statistics	95% CI	Threshold	C-statistics	95% CI
-0.95	0.508	0.49-0.52	0.05	0.532	0.41-0.65
-0.9	0.519	0.49-0.54	0.1	0.536	0.41-0.65
-0.85	0.533	0.50-0.56	0.15	0.546	0.42-0.66
-0.8	0.546	0.50-0.58	0.2	0.557	0.43-0.67
-0.75	0.554	0.50-0.60	0.25	0.550	0.43-0.66
-0.7	0.565	0.49-0.63	0.3	0.548	0.44-0.65
-0.65	0.563	0.47-0.65	0.35	0.549	0.44-0.65
-0.6	0.557	0.45-0.66	0.4	0.550	0.44-0.65
-0.55	0.559	0.43-0.67	0.45	0.557	0.45-0.66
-0.5	0.560	0.43-0.68	0.5	0.559	0.45-0.66
-0.45	0.557	0.42-0.69	0.55	0.552	0.45-0.65
-0.4	0.568	0.43-0.70	0.6	0.553	0.45-0.65
-0.35	0.585	0.44-0.72	0.65	0.555	0.45-0.65
-0.3	0.608	0.46-0.74	0.7	0.535	0.44-0.62
-0.25	0.607	0.46-0.74	0.75	0.539	0.45-0.62
-0.2	0.589	0.44-0.73	0.8	0.539	0.45-0.62
-0.15	0.591	0.45-0.73	0.85	0.540	0.45-0.62
-0.1	0.584	0.44-0.72	0.9	0.540	0.45-0.62
-0.05	0.573	0.43-0.71	0.95	0.509	0.45-0.56
0	0.558	0.42-0.69	1	0.509	0.45-0.56

Prognostic value of alkaline phosphatase in PSC

Supplementary Figure 1. Sensitivity and specificity of ALP at 1.3xULN at T1.

Supplementary Figure 2. Course of Uno's C-statistic over follow-up for ALP at percentage change. Measured in UDCA users only: sensitivity analysis.

Supplementary Figure 3A. Association between alkaline phosphatase level and the hazard of reaching the clinical endpoint with censoring CCA patients, depicted with 95% CI.

Supplementary Figure 3B. Association between alkaline phosphatase level and the hazard of having CCA with censoring patients meeting other endpoints, depicted with 95% CI.

Enhanced liver fibrosis test predicts transplant - free survival in primary sclerosing cholangitis, a multi-center study

Elisabeth M.G. de Vries, Martti Färkkilä, Piotr Milkiewicz, Johannes R. Hov, Bertus Eksteen, Douglas Thorburn, Olivier Chazouillères, Albert Pares, Odd Helge Gilja, Ewa Wunsch, Pietro Invernizzi, Marco Carbone, Francesca Bernuzzi, Kirsten Muri Boberg, Helge Røsjø, William Rosenberg, Ulrich H. Beuers, Cyriel Y. Ponsioen, Tom H. Karlsen, Mette Vesterhus

Manuscript in preparation

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. The aim of this study was to validate the prognostic utility of ELF test in an independent, multicenter, retrospective PSC study population.

Methods

Serum samples of PSC patients were collected. The composite endpoint death or liver transplantation was defined, and transplant-free survival was estimated by the Kaplan–Meier method. Cox proportional hazards regression was used to explore the association between ELF test and clinical outcome. The prognostic performance of ELF test was determined by computing the area under the receiver operating characteristic (AUC-ROC) curve.

Results

The final analysis included 534 PSC patients, of which 324 (61%) were male. 44 (8%) showed features of autoimmune hepatitis and 379 (71%) concomitant inflammatory bowel disease. ELF test levels were higher in patients reaching an endpoint (median 10.9 [interquartile range (IQR) 9.8–12.1]) than in those censored (8.8 [IQR 8.0–9.8]) ; $p < 0.001$. There was a significant association between ELF test expressed as mild, moderate and severe fibrosis, and 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 a 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.

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 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 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 ELF test sample withdrawal (+/- 1 month) were retrospectively collected from patient files by the participating centers. Clinical data were collected of the variables: 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 of 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 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 or liver biopsy. 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 the first endpoint reached, 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 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 ELF test and the prognostic index as calculated by a previously presented algorithm including PSC type, age at PSC diagnosis, albumin, platelets, aspartate aminotransferase, alkaline phosphatase and bilirubin¹⁹ was assessed by Spearman's rank correlation test. The prediction increment of ELF test when adding this as a predictor to this novel prognostic model for PSC¹⁹ 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 was 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 ELF test analysis because of widely out-of-range results. 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 ELF test withdrawal was 57 months (IQR 28-111). An overview of patient characteristics and laboratory values at time of ELF test sample withdrawal is provided in Table 1.

Differentiation of PSC phenotype by ELF test score

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. 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 for ELF test was withdrawn, 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 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 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)

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)
<i>Laboratory values at time of ELF withdrawal</i>	
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 novel prognostic 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; ELF = enhanced liver fibrosis score

* 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

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 338 (66%) low risk patients and 178 (34%) high risk 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]). The median PSC duration at ELF test withdrawal was higher in the high risk compared to the low risk group, i.e. 76 [interquartile range, (IQR) 30-121] and 51 [IQR 28-103] months, respectively; $p=0.039$).

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 ELF test to discriminate between patients that do, and do not reach an endpoint was 9.85 (sensitivity 0.74, specificity 0.75, Youden’s index: 0.50).

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.

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 PI of a novel prognostics model¹⁹, and 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 a novel prognostic PSC model¹⁹ 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 ($p<0.001$ and $p=2e-0.4$, respectively).

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

	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 with- drawal (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)	0.000
ALT	351	2.18 (1.61-2.96)	0.000
ALP	362	2.92 (2.03-4.18)	0.000
Total bilirubin	339	4.28 (2.81-6.53)	0.000
Albumin	302	0.11 (0.04-0.28)	0.000
International normalized ratio	257	4.08 (1.99-8.38)	0.000
Platelet count	331	0.40 (0.25-0.63)	0.000
Creatinine	299	0.37 (0.05-2.51)	0.306
APRI	304	1.45 (1.27-1.65)	0.000
PI novel prognostic model PSC19	270	2.53 (1.83-3.50)	0.000
ELF test	534	1.77 (1.58-1.99)	0.000

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 score; 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 score; HR = Hazard ratio; CI = confidence Interval.

Multivariable Cox analysis; n=516 patients with follow-up.

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.

An increased ELF test in patients diagnosed with cholangiocarcinoma – although this analysis included only 15 patients – was demonstrated, 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. This may be explained by the “desmoplastic” and “scirrhous,” type of CCA in PSC.^{20,21} The neoplastic cells of these carcinoma show an excessive fibrotic response in the surrounding tissue, which may be captured by ELF test. However, the PSC patients that developed cholangiocarcinoma included in this study did not seem to effect the main observation – additional analyses whilst excluding these patients showed similar results (data not shown). Dedicated analyses seem warranted to further explore the association between ELF test and cholangiocarcinoma.

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 possibly to estimate prognosis.^{22,23} 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.²⁴ Therefore, measurement of liver fibrosis through liver histology is increasingly replaced by non-invasive modalities in the most prevalent chronic liver diseases.²⁵ 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.²⁶ First, a 'biological approach', based on serum biomarkers that represent serum levels of products arising from fibrogenesis and matrix turnover, as a derivative of degree of fibrosis.²⁶ 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 elastography.²⁶ 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.^{27–29}

In addition to ELF test, several other established biomarkers of fibrosis have been used in other liver diseases, including the AST to platelet ratio index (APRI)³⁰, Fibrosis-4-score (an algorithm including age, AST, platelet count and ALT)³¹, and FibroTest (an algorithm including age, Alpha-2-macroglobulin, Haptoglobin, Apolipoprotein A1, gamma-glutamyl transpeptidase, total bilirubin, and ALT)³². 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.³³ 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.³³ 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 Given the apparent surplus value of ELF test as prognosticator in PSC, when compared with other biomarkers of fibrosis, the current validation has potential implications for clinical practice.¹⁶

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 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, has a hori-

zon of only 4 years and performs best in advanced disease patients. Recently, a novel 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 the ELF test results. These findings suggest that ELF test could be a valuable adjunct to clinically derived prognostic models in PSC.

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 ELF test over the disease course was not feasible because of the cross-sectional design of this study. 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 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 ELF test and its ability to stratify risk of poor outcome in PSC. 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, ELF test is a promising candidate.

REFERENCES

1. Williamson, K. & Chapman, C. Primary sclerosing cholangitis: a clinical update. *Br. Med. Bull.* 114, 56–64 (2015).
2. Hirschfield, G. M., Karlsen, T. H., Lindor, K. D. & Adams, D. H. Primary sclerosing cholangitis. *Lancet* 382, 1587–99 (2013).
3. Adam, R. et al. Evolution of indications and results of liver transplantation in Europe. A report from the European Liver Transplant Registry (ELTR). *J. Hepatol.* 57, 675–688 (2012).
4. Ponsioen, C. Y. et al. Surrogate end-points 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. et al. Primary sclerosing cholangitis: natural history, prognostic factors and survival analysis. *Hepatology* 10, 430–6 (1989).
6. Farrant, J. et al. Natural history and prognostic variables in primary sclerosing cholangitis. *Gastroenterology* 100, 1710–7 (1991).
7. Dickson, E. et al. Primary sclerosing cholangitis: refinement and validation of survival models. *Gastroenterology* 103, 1893–901 (1992).
8. Broomé, U. et al. Natural history and prognostic factors in 305 Swedish patients with primary sclerosing cholangitis. *Gut* 38, 610–615 (1996).
9. Corpechot, C. 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* 146, 970–9; quiz e15–6 (2014).
10. Rosenberg, W. M. C. et al. Serum markers detect the presence of liver fibrosis: A cohort study. *Gastroenterology* 127, 1704–1713 (2004).
11. Xie, Q. et al. The performance of enhanced liver fibrosis (ELF) test for the staging of liver fibrosis: a meta-analysis. *PLoS One* 9, e92772 (2014).
12. Parkes, J. et al. Enhanced liver fibrosis test can predict clinical outcomes in patients with chronic liver disease. *Gut* 59, 1245–51 (2010).
13. Mayo, M. J. et al. Prediction of clinical outcomes in primary biliary cirrhosis by serum enhanced liver fibrosis assay. *Hepatology* 48, 1549–57 (2008).
14. Guha, I. N. et al. Noninvasive markers of fibrosis in nonalcoholic fatty liver disease: Validating the European Liver Fibrosis Panel and exploring simple markers. *Hepatology* 47, 455–60 (2008).
15. Parkes, J. et al. Enhanced Liver Fibrosis (ELF) test accurately identifies liver fibrosis in patients with chronic hepatitis C. *J. Viral Hepat.* 18, 23–31 (2011).
16. Irvine, K. M. et al. The Enhanced liver fibrosis score is associated with clinical outcomes and disease progression in patients with chronic liver disease. *Liver Int.* 36, 370–7 (2016).
17. Vesterhus, M. et al. Enhanced Liver Fibrosis Score Predicts Transplant-Free Survival in Primary Sclerosing Cholangitis. *Hepatology* 62, 188–197 (2015).
18. European Association for the Study of the Liver. EASL Clinical Practice Guidelines: management of cholestatic liver diseases. *J. Hepatol.* 51, 237–67 (2009).
19. de Vries, E. et al. A novel prognostic model for primary sclerosing cholangitis. *Hepatol.* 244A–245A
20. Lazaridis, K. N. & Gores, G. J. Primary Sclerosing Cholangitis and Cholangiocarcinoma. *Semin. Liver Dis.* 26, 42–51 (2006).
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., P. D. The Significance of Stromal Desmoplasia in Intrahepatic Cholangiocarcinoma: A Special Reference of 'Scirrhus-type' and 'Nonscirrhous-type' Growth. *Am. J. Surg. Pathol.* 23, 892–902 (1999).
22. Sebastiani, G. & Alberti, A. Non invasive fibrosis biomarkers reduce but not substitute the need for liver biopsy. *World J. Gastroenterol.* 12, 3682–3694 (2006).
 23. De Vries, E. M. G. et al. Applicability and prognostic value of histologic scoring systems in primary sclerosing cholangitis. *J. Hepatol.* 63, 1212–1219 (2015).
 24. Portmann, B. & Zen, Y. Inflammatory disease of the bile ducts-cholangiopathies: liver biopsy challenge and clinico-pathological correlation. *Histopathology* 60, 236–48 (2012).
 25. Corpechot, C. Utility of Noninvasive Markers of Fibrosis in Cholestatic Liver Diseases. *Clin. Liver Dis.* 20, 143–158 (2016).
 26. Castera, L. Noninvasive Assessment of Liver Fibrosis. *Dig. Dis.* 325–335 (2015).
 27. Ehlken, H., Lohse, A. W. & Schramm, C. Transient elastography in primary sclerosing cholangitis-the value as a prognostic factor and limitations. *Gastroenterology* 147, 542–3 (2014).
 28. Tapper, E. B., Castera, L. & Afdhal, N. H. FibroScan (Vibration-Controlled Transient Elastography): Where Does It Stand in the United States Practice. *Clin. Gastroenterol. Hepatol.* (2014).
 29. Castera, L. et al. Pitfalls of liver stiffness measurement: A 5-year prospective study of 13,369 examinations. *Hepatology* 51, 828–835 (2010).
 30. Lin, Z. H. et al. Performance of the aspartate aminotransferase-to-platelet ratio index for the staging of hepatitis C-related fibrosis: An updated meta-analysis. *Hepatology* 53, 726–736 (2011).
 31. Sterling, R. K. et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology* 43, 1317–1325 (2006).
 32. Imbert-Bismut, F. et al. Biochemical markers of liver fibrosis in patients with hepatitis C virus infection: a prospective study. *Lancet* 357, 1069–1075 (2001).
 33. Bowlus, C. L. et al. Validation of serum fibrosis marker panels in patients with primary sclerosing cholangitis (PSC) in a randomized trial of simtuzumab. 1, (2015).

ELF test predicts transplant-free survival in PSC

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.

Genetic variants associated with disease phenotype and course in primary sclerosing cholangitis

Rudi Alberts*, Elisabeth M.G. de Vries*, Elizabeth C. Goode, Xiaojun Jiang, Fotis Sampaziotis, Krista Rombouts, Katrin Schölzel, Trine Folseraas, Tobias J. Weismüller, Andrew L. Mason, Weiwei Wang, Graeme Alexander, Domenico Alvaro, Annika Bergquist, Niklas K. Björkström, Francesca Bernuzzi, Ulrich Beuers, Einar Björnsson, Kirsten Muri Boberg, Christopher L. Bowlus, Maria C. Bragazzi, Olivier Chazouillères, Angela Cheung, Georgios Dalekos, John Eaton, Bertus Eksteen, David Ellinghaus, Martti Färkkilä, Eleonora A.M. Festen, Annarosa Floreani, Irene Franceschet, Daniel Nils Gotthardt, Gideon M. Hirschfield, Bart van Hoek, Kristian Holm, Simon Hohenester, Johannes Roksund Hov, Floris Imhann, Pietro Invernizzi, Brian D. Juran, Henrike Lenzen, Wolfgang Lieb, Jimmy Z. Liu, Hanns-Ulrich Marschall, Marco Marzioni, Espen Melum, Piotr Milkiewicz, Tobias Müller, Albert Pares, Christian Rupp, Christian Rust, Richard N. Sandford, Christoph Schramm, Stefan Schreiber, Erik Schruppf, Mark Silverberg, Brijesh Srivastava, Martina Sterneck, Andreas Teufel, Ludovic Vallier, Joanne Verheij, Arnau Vich Vila, Boudewijn de Vries, Kalliopi Zachou, The International PSC Study Group, The UK PSC Consortium, Roger W. Chapman, Michael P. Manns, Massimo Pinzani, Simon M. Rushbrook, Konstantinos N. Lazaridis, Andre Franke, Carl A. Anderson, Tom H. Karlsen, Cyriel Y. Ponsioen#, Rinse K. Weersma1#

* These authors contributed equally to this work

Shared senior authors

Submitted

Primary sclerosing cholangitis (PSC) is a genetically complex, inflammatory bile duct disease of largely unknown etiology leading to liver transplantation or death of often young patients. Using a cohort of 3,402 PSC patients we identified rs853974 nearby *RSPO3* to be associated with liver transplant-free survival ($P = 6.07 \times 10^{-9}$). We demonstrate *RSPO3* expression in key liver-resident effector cells, such as human and murine cholangiocytes and human hepatic stellate cells.

Primary sclerosing cholangitis (PSC) is a genetically complex, cholestatic liver disease, in which chronic biliary inflammation and bile duct destruction lead to biliary fibrosis and liver cirrhosis.¹ Peak incidence is approximately 40 years² and estimated median survival until death or liver transplantation ranges broadly between 9.3 and 20.6 years.² There is currently no effective medical therapy, and PSC is one of the leading indications for liver transplantation in many Western countries.³ Up to 80% of PSC patients have concomitant inflammatory bowel disease (IBD), and PSC predisposes to the development of cholangiocarcinoma (CCA) and colorectal carcinoma (CRC).¹ The major histocompatibility complex (MHC) on chromosome 6p21 has been identified as the predominant genetic risk factor^{4,5} with additional genetic susceptibility imposed by a number of non-MHC genes.⁵

In Mendelian traits like cystic fibrosis and hemochromatosis, consortia efforts have led to the identification of robust and important modifier genes.^{6,7} However, little is known about the role of genetics on disease behavior or disease progression in complex diseases in general, and PSC in particular.⁸ The large variability of disease severity in PSC led us to hypothesize the existence of modifier effects in PSC. To identify genetic variants that would provide insight into biological mechanisms underlying variability of PSC disease course and progression, we performed association analyses in the largest PSC cohort available. This cohort consists of 3,789 patients recruited from 13 countries and was previously used to identify PSC risk loci.⁹ We successfully collected standardized clinical phenotypes for 3,402 patients within this cohort and combined these with their Illumina Immunochip genotype data. This customized genotyping array interrogates 195,806 single nucleotide polymorphisms (SNPs).

Clinical characteristics of the PSC cohort are described in Supplementary Tables 1 and 2. In short, 2,185 (64.7%) of patients were male and the median age at diagnosis was 38.6 years. Concomitant IBD was diagnosed in 2,390 (74.5%) patients. In total 874 patients underwent liver transplantation and 181 patients died because of PSC related causes (Supplementary Table 1). Over 11% of patients developed a malignancy, most often CCA (5.6%) or CRC (4.3%).

We developed a framework to perform Immunochip-wide Cox proportional hazards analyses to identify genetic variants associated with liver trans-

plant-free survival. We defined liver *transplant-free survival* as time from PSC diagnosis until liver transplantation, or death due to the following PSC related causes: CCA, cholangiosepsis, gallbladder carcinoma and liver failure. Univariate Cox regression for clinical parameters showed significant associations with transplant-free survival for sex, country, time point of PSC diagnosis, established IBD diagnosis and smoking status. Next, we tested 130,422 SNPs that remained after Quality Control (Online Methods) for association with liver transplant-free survival using multivariate Cox proportional hazards regression, including the genotype effect alongside the significant clinical co-variables (Online Methods). We found SNP rs853974 to be associated with liver transplant-free survival of PSC patients at genome wide significance ($P = 6.07 \times 10^{-9}$). Kaplan-Meier survival analysis (Fig. 1a) showed a 50.9% (95% CI 41.5-59.5%) transplant-free survival for homozygous AA allele carriers of rs853974 compared to 72.8% (95% CI 69.6-75.7%) for GG carriers at ten years after PSC diagnosis. Homozygotes AA have a 2.14 increased hazard compared to GG (95% CI = 1.66-2.76). Figure 1b shows a regional plot of the observed association. Associations to other time-to-event phenotypes (Supplementary Table 3) and other phenotypes are given in Supplementary Tables 4-6 (Online Methods).

SNP rs853974 is located on chromosome 6, next to *R-Spondin 3 (RSPO3)*. We did not identify a direct functional effect of this SNP (Supplementary Note). However, we hypothesized that *RSPO3* is the most likely positional candidate gene in the locus. R-spondin 3 is a member of the R-spondin protein family (R-spondin 1-4).^{10,11} These proteins are secreted agonists of the canonical Wnt/ β -catenin signaling pathway.¹² They activate the pathway leading to induced transcription of Wnt target genes. Wnt/ β -catenin signaling plays a central role in embryogenesis, organogenesis and adult homeostasis, and is a critical regulator of stem cell maintenance.^{11,13} Of note, Wnt signaling is involved in both progression and regression of liver fibrosis, either by inhibiting or promoting activation and survival of hepatic stellate cells (HSCs).¹⁴⁻¹⁷

To assess whether *RSPO3* is expressed in disease relevant cells (cholangiocytes and HSCs), we performed RNA-sequencing on healthy and cholestatic cholangiocytes and multiple organs derived from normal C57BL/6 mice. Cholestatic cholangiocytes were derived from 0.1%3,5-diethoxycarbonyl-1,4-dihydrocollidine (DDC)-fed mice, a mouse model for sclerosing cholangitis.¹⁸ *RSPO3* expression was 7 to 20 folds higher in cholangiocytes as compared to any of the other cell-types or organs, and expression was higher in healthy cholangiocytes than in cholestatic cholangiocytes (Fig. 2a). Next, we assessed expression of *RSPO3* using microarrays comparing human induced pluripotent stem cells (hiPSCs), hiPSC-derived cholangiocyte-like cells (CLCs) and human primary bile duct (PBD) samples. *RSPO3* was significantly higher expressed in CLCs and PBD com-

pared to hiPSCs (Supplementary Fig. 1). This was confirmed by qPCR (Fig. 2b). Since activated HSCs are the main cells involved in liver fibrosis¹⁹, we investigated expression of *RSPO3* in human culture-activated HSCs. We isolated, cultured and activated human HSCs of three patients without PSC.²⁰ qPCR showed expression of *RSPO3* in HSCs in all three subjects (Supplementary Fig. 2). *RSPO3* expression was downregulated upon activation of the HSCs (Fig. 2c). Also, *RSPO3* expression in human biliary epithelium was lower in PSC and primary biliary cholangitis compared to non-cholangitis control samples (Fig. 2d). We did not observe *RSPO3* expression in human CD4 and CD8 T lymphocytes (data not shown). These results show that *RSPO3* is expressed and regulated in the key effector cells involved in the pathogenesis of PSC and suggest that *RSPO3* plays a role in fibrogenesis of the liver.

In conclusion, we here report rs853974 to be genome-wide significantly associated with liver transplant-free survival in PSC. Findings of *RSPO3* being expressed and regulated in both mouse and human cholangiocytes and human activated HSCs warrant further assessments of the role of this potential key PSC modifier gene.

Figure 1

Association of genetic variants on chromosome 6 with transplant-free survival of PSC patients. **(A)** Kaplan-Meier curves of transplant-free survival. Patients are stratified according to their genotype for SNP rs853974. The *P*-value for genotype effect in the Cox proportional hazards model is $P = 6.07 \times 10^{-9}$. **(B)** Regional association plot for transplant-free survival. The Y-axis shows the $-\log_{10}(P\text{-value})$ for genotype effect in the Cox proportional hazards model.

Figure 2

RSPO3 expression in mouse cholangiocytes and in human cholangiocyte-like-cells, hepatic stellate cells, primary bile duct and biliary epithelium. **(A)** RNAseq analysis of *RSPO3* expression in DDC induced cholestatic cholangiocytes, healthy cholangiocytes and multiple organs of normal C57BL/6 mice. **(B)** Quantitative real time PCR (qPCR) analysis demonstrating the expression of *RSPO3* in human induced pluripotent stem cells (hiPSCs) derived cholangiocyte-like cells (CLCs) and primary bile duct (PBD) samples compared to expression in hiPSCs. **(C)** *RSPO3* mRNA expression normalized to *GAPDH* in one sample in human primary activated hepatic stellate cells after different passages. **(D)** RNA-seq analysis of *RSPO3* expression in biliary epithelial cells derived from liver transplant recipients with PSC, PBC (primary biliary cholangitis) and non-biliary end stage liver disease.

REFERENCES

- Hirschfield, G. M., Karlsen, T. H., Lindor, K. D. & Adams, D. H. Primary sclerosing cholangitis. *Lancet* 382, 1587–99 (2013).
- Boonstra, K., Beuers, U. & Ponsioen, C. Y. Epidemiology of primary sclerosing cholangitis and primary biliary cirrhosis: a systematic review. *J. Hepatol.* 56, 1181–8 (2012).
- Eksteen, B. Advances and controversies in the pathogenesis and management of primary sclerosing cholangitis. *Br. Med. Bull.* 110, 89–98 (2014).
- Mehal, W. et al. HLA DR4 is a marker for rapid disease progression in primary sclerosing cholangitis. *Gastroenterology* 106, 106–7 (1994).
- Liu, J. Z. et al. Dense genotyping of immune-related disease regions identifies nine new risk loci for primary sclerosing cholangitis. *Nat. Genet.* 45, 670–5 (2013).
- Bartlett, J. R. et al. Genetic modifiers of liver disease in cystic fibrosis. *JAMA* 302, 1076–83 (2009).
- Stickel, F. et al. Evaluation of genome-wide loci of iron metabolism in hereditary hemochromatosis identifies PCSK7 as a host risk factor of liver cirrhosis. *Hum. Mol. Genet.* 23, 3883–3890 (2014).
- Boberg, K. et al. The HLA-DR3,DQ2 heterozygous genotype is associated with an accelerated progression of primary sclerosing cholangitis. *Scand J Gastroenterol* 36, 886–90. (2001).
- Liu, J. Z. et al. Dense fine-mapping study identifies new susceptibility loci for primary biliary cirrhosis. *Nat. Genet.* 44, 1137–41 (2012).
- Chen, J. et al. Cloning and identification of a cDNA that encodes a novel human protein with thrombospondin type I repeat domain, hPWTSR. *Mol Biol Rep* 29, 287–92 (2002).
- de Lau, W. B., Snel, B. & Clevers, H. C. The R-spondin protein family. *Genome Biol.* 13, 242 (2012).
- Nam, J. S., Turcotte, T. J., Smith, P. F., Choi, S. & Jeong, K. Y. Mouse cristin/R-spondin family proteins are novel ligands for the frizzled 8 and LRP6 receptors and activate β -catenin-dependent gene expression. *J. Biol. Chem.* 281, 13247–13257 (2006).
- van Amerongen, R. & Nusse, R. Towards an integrated view of Wnt signaling in development. *Development* 136, 3205–3214 (2009).
- Jiang, F., Parsons, C. J. & Stefanovic, B. Gene expression profile of quiescent and activated rat hepatic stellate cells implicates Wnt signaling pathway in activation. *J. Hepatol.* 45, 401–409 (2006).
- Myung, S. J. et al. Wnt signaling enhances the activation and survival of human hepatic stellate cells. *FEBS Lett.* 581, 2954–2958 (2007).
- Cheng, J. H. et al. Wnt antagonism inhibits hepatic stellate cell activation and liver fibrosis. *Am. J. Physiol. Gastrointest. Liver Physiol.* 294, G39–G49 (2008).
- Kordes, C., Sawitzka, I. & Häussinger, D. Canonical Wnt signaling maintains the quiescent stage of hepatic stellate cells. *Biochem. Biophys. Res. Commun.* 367, 116–123 (2008).
- Fickert, P. et al. A new xenobiotic-induced mouse model of sclerosing cholangitis and biliary fibrosis. *Am. J. Pathol.* 171, 525–36 (2007).
- Mederacke, I. et al. Fate tracing reveals hepatic stellate cells as dominant contributors to liver fibrosis independent of its aetiology. *Nat. Commun.* 4, 2823 (2013).
- Jalan, R. et al. Ammonia produces pathological changes in human hepatic

- stellate cells and is a target of therapy of portal hypertension. *J. Hepatol.* 64, 823–833 (2016).
21. Lennard-Jones, J. Classification of inflammatory bowel disease. *Scand J Gastroenterol* 170, 16–9 (1989).
 22. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575 (2007).
 23. Bellenguez, C. et al. A robust clustering algorithm for identifying problematic samples in genome-wide association studies. *Bioinformatics* 28, 134–5 (2012).
 24. Jia, X. et al. Imputing Amino Acid Polymorphisms in Human Leukocyte Antigens. *PLoS One* 8, (2013).
 25. Han, B. et al. Fine Mapping Seronegative and Seropositive Rheumatoid Arthritis to Shared and Distinct HLA Alleles by Adjusting for the Effects of Heterogeneity. *Am. J. Hum. Genet.* 94, 522–532 (2014).
 26. Patsopoulos, N. A. et al. Fine-Mapping the Genetic Association of the Major Histocompatibility Complex in Multiple Sclerosis: HLA and Non-HLA Effects. *PLoS Genet.* 9, e1003926 (2013).
 27. Material, S. O., Web, S., Press, H., York, N. & Nw, A. The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science* 306, 636–640 (2004).
 28. Rosenbloom, K. R. et al. ENCODE whole-genome data in the UCSC Genome Browser. *Nucleic Acids Res.* 38, D620–5 (2010).
 29. Westra, H.-J. et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat. Genet.* 45, 1238–43 (2013).
 30. The GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* 45, 580–5 (2013).
 31. Marzioni, M. et al. PDX-1/Hes-1 interactions determine cholangiocyte proliferative response to injury in rodents: Possible implications for sclerosing cholangitis. *J. Hepatol.* 58, 750–756 (2013).
 32. Sampaziotis, F. et al. Cholangiocytes derived from human induced pluripotent stem cells for disease modeling and drug validation. *Nat. Biotechnol.* 33, 845–852 (2015).
 33. Mederacke, I., Dapito, D. H., Affò, S., Uchinami, H. & Schwabe, R. F. High-yield and high-purity isolation of hepatic stellate cells from normal and fibrotic mouse livers. *Nat. Protoc.* 10, 305–15 (2015).
 34. Rombouts, K. & Carloni, V. Lipid Signaling Protocols. *Methods Mol. Biol.* 1376, 203–212 (2016).
 35. Schmittgen, T. D. & Livak, K. J. Analyzing real-time PCR data by the comparative CT method. *Nat. Protoc.* 3, 1101–1108 (2008).
 36. Joplin, R. Isolation and culture of biliary epithelial cells. *Gut* 35, 875–878 (1994).
 37. Xu, L. et al. Does a betaretrovirus infection trigger primary biliary cirrhosis? *Proc. Natl. Acad. Sci. U. S. A.* 100, 8454–8459 (2003).
 38. Andersson, R. et al. An atlas of active enhancers across human cell types and tissues. *Nature* 507, 455–61 (2014).

ONLINE METHODS

PSC patient cohort and genotype data

Cohort description and data collection.

PSC patients previously recruited throughout Europe, North America and Canada and genotyped using the Immunochip by Liu et al.²¹ were included (**Supplementary Tables 1 and 2**). Subject recruitment was approved by the ethics committees or institutional review boards of all participating centers. Written informed consent was obtained from all participants. Patients of whom PSC diagnosis was revised since they were genotyped were excluded.

The phenotypic data that were collected for patient and disease characteristics included: sex, date of birth, PSC type, date of PSC diagnosis, intra and/or extra hepatic disease, dominant strictures, concomitant IBD and type of IBD, date of IBD diagnosis, and smoking status. Furthermore, follow up data were collected for: date and cause of death, date and indication of liver transplantation (LTx), the occurrence and date of diagnosis of hepatocellular carcinoma (HCC), cholangiocarcinoma (CCA), gallbladder carcinoma, colorectal carcinoma (CRC) and the occurrence and date of a colectomy.

PSC diagnosis was based on clinical, biochemical, cholangiographic and histologic criteria, as formulated by the EASL guidelines.²² IBD diagnosis was scored based on accepted endoscopic, radiologic and histologic criteria.²²¹

Immunochip genotyping and quality control.

Here, we used genotype data of 3,789 PSC patients described by Liu et al.²¹ DNA was extracted from whole blood, transformed lymphocytes or liver tissue. Illumina Immunochip was used to genotype the DNA samples. As quality control, SNPs with call rate below 80% were removed. Next, using PLINK²²², per sample genotype call rate and heterozygosity rate were determined. Outlying samples were identified using Aberrant.²²³ Also, for each pair of individuals with estimated identity by descent ≥ 0.9 , the sample with the lower call rate was removed. Related individuals with identity by descent > 0.1875 and < 0.9 were kept. Samples of non-European ancestry were identified using PCA analysis.²⁶ Using population samples from the 1000 Genomes Project²⁷ principal components were calculated and then projected into cases and controls.^{21,28,29} 3,789 PSC cases and 25,079 controls remained after quality control. As quality control SNPs were excluded that (i) had a MAF below 0.1%, (ii) had a Hardy-Weinberg equilibrium $P < 1 \times 10^{-5}$, (iii) had a call rate below 98% or (iv) failed the non-random differential missing data rate test of PLINK v1.07 between cases and controls at a threshold of $P < 1 \times 10^{-5}$. After cluster plot inspection, a total of 130,422 SNPs remained for analysis. A combined dataset of genotype data and phenotype data was composed for 3,402 PSC cases.

Statistical and computational analyses

Imputation of the HLA region.

Based on the ImmunoChip genotypes, we imputed classical alleles for HLA-A, HLA-B, HLA-C, HLA-DRB1, HLA-DQA1, HLA-DQB1, HLA-DPA1 and HLA-DPB1 and their corresponding amino acid polymorphisms, using a reference panel of 5,225 individuals of European descent, collected by the Type 1 Diabetes Genetics Consortium and using SNP2HLA with default settings.³⁰ This reference panel showed high imputation quality for the HLA region.^{30–32}

Statistical analyses.

We defined liver transplant-free survival as combined endpoint of LTx or PSC related death (death from liver failure, cholangiosepsis, cholangiocarcinoma or gall-bladder carcinoma). To determine whether clinical parameters have an effect on liver transplant-free survival, univariate Cox proportional hazards models were fitted including clinical parameters as covariates. The covariates sex, country, date of PSC diagnosis, IBD type and smoking status were significantly associated with liver transplant-free survival. Next, to determine associations between genotype and liver transplant-free survival, for all SNPs on the ImmunoChip after QC, we fitted a Cox regression model including the significant clinical variables as well as genotype. Associations with a P -value $< 5 \times 10^{-8}$ were considered significant, whereas associations with a P -values between 1×10^{-5} and 5×10^{-8} were considered suggestive. Proportional hazard assumptions were judged using Therneau and Grambsch tests. After stratifying the IBD type and categorizing time of IBD diagnosis into 5 groups the proportional hazard assumptions were met. To visualize the effect of genotype on time-to-event phenotypes Kaplan-Meier survival estimates were plotted. Cox regression analyses were also performed for time to death, time to LTx, time to CCA, CRC or HCC, time to gall bladder carcinoma and time to colectomy. All time intervals started at the time of PSC diagnosis.

Association tests were performed in PLINK version 1.07²⁴. Survival analyses were done in R software version 3.1.3 (<http://www.r-project.org>).

In silico functional analyses.

The Encyclopedia of DNA Elements (ENCODE)³³ was searched using the UCSC Genome Browser.³⁴ Specifically, overlaps of SNPs with the following regulatory features were searched: DNaseI hypersensitivity sites, transcription factor binding sites, histone modification and DNA-polymerase sites. We tested whether associated variants showed an effect on gene expression levels of genes (eQTLs). For this we used an expression quantitative trait loci (eQTL) study of non-transformed peripheral blood in 5,311 individuals (<http://www.genenetwork.nl/bloodeqtl->

browser/)³⁵ and we queried Single Tissue and Multi-tissue eQTL on the GTEx Portal (<http://gtexportal.org>), an online resource and associated tissue bank for the scientific community for studying the relationship between genetic variation and gene expression in human tissues.³⁶

RSPO3 expression in murine cholangiocytes and normal tissue

Mice.

C57BL/6 mice were purchased from Charles River (Milan, Italy). The mice were housed in a Minimal Disease Unit at the animal facility at the Università Politecnica delle Marche, Ancona, Italy. All animal experiments were performed in compliance with local institution guidelines (Italian Ministry of Health).

Tissues and cholangiocytes preparation.

Normal C57BL/6 males were sacrificed at the age of 8-10 weeks. Brain, kidney, lung, pancreas, salivary-gland-parotid, skin, skeletal, spleen, stomach, vagina, colon, duodenum, esophagus and ileum were harvested and washed by cold phosphate buffered saline (PBS). Cholangiocytes were isolated both from normal mice (n=3) and from mice (n=3) fed 0.1%3,5-diethoxycarbonyl-1,4-dihydrocollidine (DDC) for 4 weeks, as a model of sclerosing cholangitis.³⁷ Purification of cholangiocytes from mice was performed by immune-bead purification, as previously reported.³⁸

RNA sequencing.

Total RNA was extracted by TRIzol Reagent (Life Technologies Corporation, Woburn, MA). Libraries were prepared using TruSeq RNA (Illumina) reagents, and sequenced on 4 lanes of an Illumina HiSeq 2000 machine, using SBS v3 reagents (Illumina) to generate 50 bp single end reads. The reads then were aligned to the mouse genome reference sequence and annotations from Illumina built for UCSC mm10 (version of 2014_05_23) with the Tophat software (version 2.1.0) using default parameters without gene/transcript discovery. Gene expression levels were quantified and normalized using cuffquant and cuffnorm with default parameters from the software cufflinks (version 2.2.1). Fragments Per Kilobase of exon per Million mapped reads (FPKM) values were used as gene expression levels.

RSPO3 expression in human primary biliary tissue and cholangiocyte-like-cells

Human primary biliary tissue and cholangiocyte-like-cells culture.

Primary biliary tissue (bile duct, PBD) was obtained from a liver and pancreas organ donor after obtaining informed consent from the donor's family (REC refer-

ence number: 09/H0306/73). A section of the bile duct was excised and homogenized using a tissue homogenizer. RNA was extracted using a kit (Sigma-Aldrich), according to the manufacturer's instructions. Cholangiocyte-like-cells (CLCs) were generated from human induced Pluripotent Stem Cells (hPSCs) and cultured as previously described.³⁹

Real-time PCR (qPCR).

500ng of cellular RNA was reverse transcribed using Superscript II Reverse Transcriptase (Invitrogen). qPCR reaction mixtures were prepared using the SensiMix™ SYBR® & Fluorescein Kit (Bioline), according to the manufacturer's instructions. The cDNA was denatured at 94°C for 5 minutes, followed by 40 cycles at 94°C for 30 seconds, 60°C for 30 seconds, and 72°C for 30 seconds, and a final extension step at 72°C for 10 minutes. A Stratagene Mx3005P was used for all qPCR reactions.

Microarray data.

Previously published raw and processed microarray data characterizing the transcriptomic profile of PBD and CLCs were used³⁹. This data is available on ArrayExpress (Accession number: E-MTAB-2965).

In vitro studies in human HSC

Isolation, culture and experimental treatment of human hepatic stellate cells.

Primary human hepatic stellate cells (HSC) were isolated from wedge sections of liver tissue, obtained from patients undergoing surgery at the Royal Free Hospital after giving informed consent (EC01.14-RF). Cells were isolated according to Mederacke et al.⁴⁰, with modifications for human liver^{41,42}. Briefly, 10 g of total human liver tissue was digested with 0.01% Collagenase, 0.05% Pronase and 0.001% DNase I without performing perfusion. The homogenate was filtered through a 100µm cell strainer (BD Falcon) and the flow-through was centrifuged at 50xg for 2 minutes at 4°C. After washing the supernatant, gradient centrifugation was performed at 1400xg for 17 minutes at 4°C using an 11.5% Optiprep gradient (Sigma). Finally, the interface was collected and washed. Purity of hHSC was established by detection of CD140b (PDGFRbeta), CD29 (Integrin beta 1) and Cytoglobin B (CYGB).

The obtained HSCs were cultured in IMDM supplemented with 20% foetal bovine serum (FBS), Glutamine, nonessential amino acids 1X, 1.0 mM sodium pyruvate, 1X antibiotic-antimycotic (all Life Technologies), referred to as complete HSC medium hereinafter. Experiments described in this study were performed on HSC of at least three independent cell preparations, between passage 3 and 8.

Real-time qPCR.

Total cellular RNA was extracted and then cleaned up by using QIAzol Lysis Reagent and RNeasy Kit (Qiagen, CA, USA), respectively, according to the manufacturer's protocol. After quantification using NanoDrop1000 System (Thermo Scientific, USA), RNA was retro-transcribed into cDNA using the QuantiTect Reverse Transcription Kit (Qiagen, CA, USA) and 6.7 µg of the cDNA sample was used to set up real-time quantitative PCRs reactions using TaqMan gene expression assays RSPO3_TaqMan® Gene Expression Assays_Hs01072567_m1, GAPDH_TaqMan® Gene Expression Assays_Hs02758991_g1 (Life technologies, CA, USA) and 7500 Fast Real-Time PCR System following the manufacturer's protocol. Each sample was tested in duplicate. Target genes were normalized using GAPDH as endogenous control and their relative expression was carried out with the $2^{-\Delta Ct}$ method (where Ct represents the threshold cycle).⁴³ The amplification efficiency of target and reference genes was approximately the same (slope < 0.1).

RSPO3 expression in human biliary epithelial cells

Human biliary epithelial cell (BEC) extraction.

Liver samples were obtained from liver transplant recipients with end stage liver disease secondary to PSC (n=9), PBC (n=10) and non-biliary disease (n=15; from 4 patients with non-alcoholic fatty liver disease, 7 with alcoholic cirrhosis and one each with cryptogenic cirrhosis, autoimmune hepatitis, glycogen storage disease and congenital absence of portal vein). Biliary epithelial cell extraction was performed within 24 hours of collection using previously established protocols employing dentistry gradient centrifugation and immuno-selection with anti-HEA125, as previously described.^{44,45}

RNA sequencing.

Total RNA was extracted by TRIzol (Life Technologies, Carlsbad, CA) and libraries were prepared using the TruSeq RNA V2 kit (Illumina, San Diego, CA). Paired end 125bp sequencing was performed using two lanes from a HiSeq 2500 on the high output mode to generate ~ 4GB data per library. Illumina adaptors and low quality bases (quality cutoff Q15) were removed from raw reads. High quality reads were aligned against human genome (GRCh38) using TopHat2; data were normalized and differential expression analyses were carried out among PBC, PSC and control groups with Cufflinks package. Fragments per Kilobase of exon per Million mapped reads (FPKM) values were used for gene expression levels.

Supplementary Figure 1.

RSPO3 expression in human induced pluripotent stem cells (hiPSCs), cholangiocyte-like-cells (CLCs) and primary bile duct (PBD). *RSPO3* expression is significantly increased in CLCs and in PBD compared to hiPSCs. $n=3$; error bars, standard deviation. Asterisks represent statistical significance (****adjusted $P<0.0001$, ***adjusted $P<0.001$, Benjamini and Hochberg corrected P -values).

Supplementary Figure 2.

mRNA expression of RSPO3 and Cytochrome B in three patients without PSC measured by qPCR. Cytochrome B mRNA expression was evaluated as specific HSC marker.

Supplementary Figure 3.

ImmunoChip signal intensities for the reported SNPs. The clusterplots show signal intensities for PSC patients, homozygous AA or BB (red or blue) and heterozygous AB (green).

rs17102823

rs17649817

rs17839542

rs261537

rs2675647

rs34985176

rs3769839

rs5030977

rs7356671

rs743998

rs74602571

rs75948976

rs7731017

rs853974

rs989494

Supplementary Table 1.

Clinical characteristics of the PSC cohort consisting of 3,402 patients. Age at diagnosis is given as median years (interquartile range). All other quantitative data are expressed as absolute counts and percentages excluding missing data.

Variable	Number (%)
Age at PSC diagnosis	38.6y (28.0y- 50.1y)
Sex	
male	2185 (64.7%)
female	1193 (35.3%)
missing	24 (0.7%)
Main diagnosis	
PSC	3159 (94.6%)
small duct PSC	75 (2.2%)
PSC with AIH overlap	107 (3.2%)
missing	61 (1.8%)
Liver transplantation	
Yes	874 (26.3%)
No	2444 (73.7%)
missing	84 (2.5%)
Colectomy	
Yes	419 (12.6%)
No	2897 (87.4%)
missing	86 (2.5%)
IBD	
No IBD	816 (25.5%)
Ulcerative colitis	1940 (60.5%)
Crohn's disease	357 (11.1%)
IBD-U	93 (2.9%)
missing	196 (5.8%)
Cancer	
Cholangiocarcinoma	
Yes	188 (5.6%)
No	3147 (94.4%)
missing	67 (2.0%)

Supplementary Table 1. Continued

Variable	Number (%)
Colorectal carcinoma	
Yes	127 (4.3%)
No	2822 (95.7%)
missing	453 (13.3%)
Gall bladder carcinoma	
Yes	30 (1.0%)
No	2977 (99.0%)
missing	395 (11.6%)
Hepatocellular carcinoma	
Yes	22 (0.7%)
No	2984 (99.3%)
missing	396 (11.6%)
Smoking status	
Smoker	140 (6.0%)
Ex-Smoker	529 (22.7%)
Non-smoker	1657 (71.2%)
missing	1076 (31.6%)
Death	
Non PSC related	47 (1.5%)
Liver failure	66 (2.1%)
Cholangiosepsis	18 (0.6%)
Gallbladder carcinoma	12 (0.4%)
Cholangiocarcinoma	85 (2.6%)
Hepatocellular carcinoma	6 (0.2%)
Colorectal carcinoma in case of coexisting	3 (0.1%)
IBD	
Alive	2977 (92.6%)
missing	188 (5.5%)

Supplementary Table 2.**Amounts of PSC patients included in this study.** Number of patients per cohort is indicated.

Country	City	No. of patients
Canada	Calgary	46
Canada	Toronto	112
Finland	Helsinki	290
France	Paris	34
Germany	Heidelberg	176
Germany	Mainz, Hannover and Berlin	316
Germany	Munich	66
Greece	Larissa	10
Italy	Milan	51
Norway	Oslo	397
Poland	Warsaw	42
Spain	Barcelona	18
Sweden	Gothenburg	67
Sweden	Stockholm	153
Netherlands	Amsterdam	134
Netherlands	Groningen	53
Netherlands	Leiden	41
United Kingdom	Norwich	1033
United States	Rochester, Minnesota	331
United States	Sacramento	32
total		3402

Supplementary Table 3.

Time-to-event phenotypes of PSC patients. For all patients, the date of PSC diagnosis was collected. Dates of liver transplantations, death, colectomy and diagnosis of several cancer types were recorded. All time intervals start at the time of PSC diagnosis and are given as median years (interquartile range).

Phenotype	No. of patients	Years (IQR)
Time to death	358	10.4 (4.3- 17.3)
Time to first liver transplantation	816	7.0 (2.9- 12.0)
Time to colectomy	405	2.0 (-4.0- 7.7)
Time to colorectal carcinoma	117	6.1 (1.2- 10.6)
Time to gallbladder carcinoma	27	10.0 (2.9- 16.9)
Time to cholangiocarcinoma	183	5.0 (0.4- 12.6)
Time to hepatocellular carcinoma	20	15.1 (7.8- 23.7)

Supplementary Note

In silico analyses for SNP rs853974

We assessed whether rs853974 or SNPs in high LD ($r^2 > 0.8$) have known functional consequences or regulatory features. We found no eQTL effects using an expression quantitative trait loci (eQTL) study of non-transformed peripheral blood in 5,311 individuals²⁹, and single tissue and multi-tissue eQTLs in the GTEx Portal³⁰ (Online Methods). Next, querying the ENCODE Encyclopedia of DNA Elements and FANTOM5 enhancers³⁸ did not reveal an overlap with any of the regulatory features (Online Methods). Of note, in the Immunochip analyses of the international IBD genetics consortium including over 75,000 individuals, an intronic SNP rs9491697 in *RSPO3* (which is not in LD with rs853974, $r^2 = 0.014$) was identified to be associated with Crohn's disease ($P = 3.79 \times 10^{-10}$, OR = 1.077) but not with ulcerative colitis.

Clusterplots

Clusterplots of Immunochip signal intensities for the reported SNPs are given in **Supplementary Figure 3**.

Applicability and prognostic value of histologic scoring systems in primary sclerosing cholangitis

Elisabeth M.G. de Vries, Joanne Verheij, Stefan G. Hubscher, Mariska M.G. Leeflang, Kirsten Boonstra, Ulrich H. Beuers, Cyriel Y. Ponsioen

Journal of Hepatology 2015;63:1212-9

ABSTRACT

Background

Primary sclerosing cholangitis (PSC) is a chronic cholestatic liver disease. At present there is no appropriate histologic scoring system available for PSC, evaluating both degree of necroinflammatory activity (grade) and fibrosis (stage). Aim of this study was to assess if three scoring systems, commonly used in different liver diseases could be applied for grading and/or staging of PSC.

Methods

Sixty-four PSC patients from a Dutch cohort, who underwent diagnostic liver biopsy, were included. Staging was scored using Ishak, Nakanuma, and Ludwig systems. Grading was scored using Ishak and Nakanuma systems. Three measures of outcome were defined; transplant-free survival, time to liver transplantation (LTx) and occurrence of cirrhosis related symptoms (CRS). Association of grade and stage with outcome was estimated using Kaplan Meier-log-rank test, and Cox regression analysis. Correlation with biochemistry was assessed by Spearman's rank test.

Results

There were strong associations between disease stage measured by Ishak, Nakanuma, and Ludwig staging systems with both outcome measure transplant-free survival (Hazard ratio (HR) 2.56; 95%CI 1.11-5.89, HR 6.53; 95%CI 2.01-21.22, HR 1.94; 95%CI 1.00-3.79, respectively), and time to LTx (HR 4.18; 95%CI 1.51-11.56, HR 7.05; 95%CI 1.77-28.11, HR 3.13; 95%CI 1.42-6.87, respectively). Ishak and Nakanuma grading systems were not associated with CRS. Weak correlations between histopathology and liver biochemistry were shown.

Conclusion

Applying the Nakanuma, Ishak, and Ludwig histopathological staging systems is feasible and clinically relevant given their association with transplant-free survival and time to LTx. This suggests these staging systems may be likely candidates for surrogate endpoints and stratification purposes in clinical trials in PSC.

INTRODUCTION

Primary sclerosing cholangitis (PSC) is a chronic cholestatic liver disease, characterized by progressive bile duct scarring and destruction, leading to biliary fibrosis and eventually progression to end-stage liver cirrhosis.¹ PSC diagnosis is established by means of cholangiography, performed by magnetic resonance cholangiopancreatography (MRCP) which is implemented as “golden standard”, and has replaced the more invasive endoscopic retrograde cholangiopancreatography (ERCP).^{2,3} In case of large duct PSC, cholangiogram typically shows biliary strictures interchanged with dilatations creating the “beaded” appearance, and routine liver biopsy is not necessary to confirm diagnosis.⁴ However, more subtle caliber changes can easily be missed on MRCP, and conversely, due to the moderate resolution of MRCP, false-positive findings may occur. In case of doubt, suspicion of small duct PSC, or auto-immune hepatitis (AIH) overlap syndrome, liver biopsy is indicated and essential to confirm diagnosis.^{5,6} Furthermore liver biopsies have been used for the evaluation of treatment efficacy in therapeutic trials.^{7–13}

Histologic changes seen in PSC include the characteristic periductal concentric fibrosis leading to bile duct obliteration, infiltration of inflammatory cells in portal tracts, loss of bile ducts, bile ductular reaction and focal accumulation of copper binding protein (CBP).^{1,14,15} Using liver histology, disease severity and progression can be assessed in terms of grade and stage.¹⁶ Grade is usually used to describe the degree of necroinflammatory activity as measurement of the severity of the underlying disease process, while stage generally reflects the degree of fibrosis and cirrhosis as measurement of disease progression.¹⁶

If a PSC appropriate histologic scoring system would have prognostic significance in terms of predicting the occurrence of solid clinical endpoints, liver histology may be an important candidate for the evaluation of treatment efficacy in therapeutic trials. Evaluation of treatment efficacy by solid clinical endpoints such as death or liver transplantation (LTx), is hindered by the chronic disease course and the low prevalence of PSC.¹⁷ Therefore, surrogate endpoints, including clinical biomarkers, biochemical biomarkers, Mayo risk score, and liver histology have often been used, but never been validated.^{7–13} Currently, the lack of properly validated surrogate endpoints for clinical trials is one of the major challenges in PSC research. An important asset of liver histology is its face-validity, meaning that liver biopsy directly measures the degree of disease severity in the affected organ. However, it is currently unknown if face-validity is maintained in PSC livers where the patchy distribution may give rise to confounding sampling variability.¹⁸

At present there is no specific PSC histologic scoring system with clinical significance, to evaluate both disease grade and stage. Commonly, the Ludwig and Ishak systems have been used to grade and stage histologic disease severity

in PSC.^{19,20} A drawback of the Ludwig staging system is that it was designed primarily to assess disease progression of primary biliary cirrhosis (PBC).²⁰ Furthermore, the Ludwig system does not separately score disease grade, and instead incorporates features such as portal and periportal inflammation, which are probably better regarded as manifestation of disease grade rather than stage.

Recently, Nakanuma et al, have proposed a new grading and staging system for primary biliary cirrhosis (PBC), which takes into account particular features that are shared with PSC, such as the presence of copper binding protein and loss of bile ducts.²¹

The aim of this study was to determine the prognostic value of three different scoring systems, designed primarily to assess disease grade and/or stage in chronic hepatitis (Ishak *et al.*) or PBC (Ludwig *et al.*, Nakanuma *et al.*), for grading and/or staging of PSC.^{19–22}

PATIENTS AND METHODS

Study design

This cohort study is part of the ‘Epi PSC PBC project’, a large population-based cohort study of PSC and PBC in the Netherlands. All PSC patients alive on January 1st 2000 and living in a geographically defined area of 6 adjacent provinces comprising 50% of the Dutch population were included in this study between January 1st 2008 and December 31st 2011. The case-finding and case-ascertainment methods have been described previously.¹⁷ The protocol was approved by the central Committee for Research Ethics in Utrecht and all 44 local ethics committees of the participating hospitals in the Netherlands (trialregister.nl number, NTR2813). PSC diagnosis was based on: 1) elevated alkaline phosphatase and gamma-glutamyl transferase, not explained otherwise, 2) presence of characteristic bile duct changes with multifocal strictures and segmental dilatations on ERC or magnetic resonance cholangiography (MRC) and/or 3) liver histology and 4) no evidence for secondary sclerosing cholangitis. When criteria 1, 3, and 4 were fulfilled in the absence of cholangiographic abnormalities on MRC or ERC, cases were diagnosed as small duct PSC.⁶ Autoimmune hepatitis (AIH) overlap syndrome (PSC- AIH) is ill defined. A diagnosis of PSC-AIH was made in patients with a characteristic cholangiogram who, in addition, met the simplified AIH criteria.²³

PSC patients from the Epi PSC PBC cohort, who underwent diagnostic liver biopsy or liver biopsy to assess disease severity at time of diagnosis between 1978 and 2011, were included. Patients with PSC-AIH overlap syndrome were excluded. Biochemical values of aspartate aminotransferase (AST), alanine aminotransferase (ALT), alkaline phosphatase (ALP), gamma-glutamyl transferase (γGT) and total bilirubin at time of liver biopsy (range two months before until

one month after liver biopsy) were retrieved from hospital databases and Mayo risk score was calculated. Clinical data reflecting liver cirrhosis related symptoms (CRS) at follow up were collected from patient files. CRS included gastro-esophageal varices and variceal bleeding, both assessed by gastrointestinal endoscopy, ascites, and splenomegaly assessed clinically and/or by imaging, and hepatic encephalopathy.

Tissue preparation and histologic evaluation

Original liver specimens, fixed in formalin and embedded in paraffin, as well as original liver stained sections were collected from the pathology department diagnostic archives. From each paraffin block thin sections were cut for haematoxylin and eosin (HE), connective tissue (Sirius red) and orcein stainings. Orcein staining was used to assess degree of copper binding protein (CBP) deposition in hepatocytes. If well preserved and available, original stained sections were used for histologic evaluation of biopsies. Otherwise, new stains were carried out. Grade and stage of biopsy specimens were evaluated using the three systems referred to above by two expert liver pathologists (JV & SH) in tandem using a multihead microscope, with the intention to reach consensus.

Grading

Grading was scored according to the Ishak system, evaluating degree of interface hepatitis (score 0-4), confluent necrosis (score 0-6), lobular inflammation (score 0-4) and portal inflammation (score 0-4).¹⁹ Furthermore slides were scored according to the Nakanuma system, encompassing degree of cholangitis activity (CA) (score 0-3) and hepatitis activity (HA)(score 0-3).²¹ (Supplementary Table 1.)

Staging

Staging was performed according the Ishak, as well as the Nakanuma and Ludwig systems.¹⁹⁻²¹ With Ishak staging system degree of fibrosis is evaluated (0-6). The Nakanuma staging system is based on semi-quantitative scoring of three histological features- fibrosis (score 0-3), bile duct loss (score 0-3) and deposition of orcein positive granules (score 0-3). The final Nakanuma stage is obtained from the total score of these three features. Stage I (no or minimal progression) is a score of 0, stage II (mild progression) a score of 1-3, stage III (moderate progression) a score of 4-6 and stage IV (advanced progression) a score of 7-9.²⁴ Ludwig staging system consists of 4 stages; stage I, cholangitis or portal hepatitis; stage II, periportal fibrosis or hepatitis; stage III, septal fibrosis, bridging necrosis or both; and stage IV, biliary cirrhosis. (Supplementary Table 2.)

Endpoints

For analyses of association with endpoints, three different endpoints were chosen. The first endpoint was transplant-free survival, defined as time to PSC-related death (death from end-stage liver disease, liver surgery, and colorectal carcinoma), LTx and presentation with cholangiocarcinoma. Since the occurrence of cholangiocarcinoma may not be predictable by liver histology at time of diagnosis, second endpoint was time to LTx alone. The third endpoint was the occurrence of liver cirrhosis related symptoms (CRS) at follow-up.

Statistical Analysis

Patient characteristics were expressed as mean $\pm$ standard deviation and median and interquartile range where ever appropriate.

Association of histologic grade and stage with transplant-free survival, time to LTx and development of CRS was estimated using Kaplan Meier survival curve and Wilcoxon log-rank test. Due to relatively small sample size, and the resulting small amount of patients per grade and stage, survival analyses were performed in grouped subcategories. In this reclassification the order of severity of grade/stage was maintained; those subgroups with very few or no patients were grouped together with the grade/stage of most similar severity. Ishak grading components interface hepatitis, lobular inflammation and portal inflammation were reclassified in score 0, 1, ≥ 2 (original score: 0-4), component confluent necrosis was reclassified in 0, ≥ 1 (original score: 0-6), and the total Ishak grade was reclassified in score 1-4 (original score: 0-1 = 1; 2,3 = 2; 4 = 3; ≥ 5 = 4). Nakanuma grading component cholangitis activity was reclassified in 0, 1, ≥ 2 (original score: 0-3), component hepatitis activity in 0, ≥ 1 (original score: 0-3), and the total Nakanuma grade score in 0, 1, ≥ 2 (original score: 0-6). Ishak stage was reclassified in score 1-3 (original score : 0,1 = 1; 2,3 = 2; ≥ 4 = 3). For the Nakanuma staging system the original scoring system was maintained. Ludwig staging system was scored 1-4 in which a Ludwig score of 0 was classified as score 1.

Cox proportional hazard analysis was performed and associated hazard ratio (HR) was calculated for clinical, biochemical and histopathological variables. Exploratory analyses for correlations between liver biochemistry and histologic grade and stage to assess if liver biochemistry could reflect the degree of liver injury measured by liver histology, were performed by Spearman's ranking test. Statistical analyses were performed using SPSS version 20.0 software (SPSS, Inc., Chicago, IL). $p < 0.05$ was considered statistically significant.

RESULTS

Patient and biopsy characteristics

Sixty-four patients were included, 40 male and 24 female, with a median follow-up of 112 months (IQR 70-172). Patient characteristics, serum liver tests, and characteristic histologic features are summarized in Table 1. Laboratory results are expressed in times upper limit of normal (xULN), as assays varied between hospitals and over time. Biopsies had a median length of 14 mm (IQR 11-19). The median number of portal tracts was 13 (IQR 9-19). Median disease duration at time of biopsy was 0 months (range 0-20).

Table 1. Patient characteristics.

N	64	
Male [n (%)]	40	(63)
Age follow-up (years) [mean (SD)]	49	(±15)
Age at diagnosis PSC (years) [mean (SD)]	38	(±14)
Large duct PSC [n (%)]	54	(84)
Inflammatory bowel disease [n (%)]	43	(67)
Ulcerative colitis [n (%)]	32	(50)
Crohn's disease [n (%)]	8	(12)
Unspecified [n (%)]	3	(5)
Portal tracts [median (IQR)]	13	(9-19)
Biopsy length (mm) [median (IQR)]	14	(11-19)
Disease duration at time of biopsy (months) [median (range)]	0	(0-20)
Follow up time (months) [median (IQR)]	112	(70-172)
AST xULN [median (IQR)]*	1.40	(1.04-2.64)
ALT xULN [median (IQR)]*	2.04	(1.40-4.43)
ALP xULN [median (IQR)]*	1.65	(1.24-3.39)
γGT xULN [median (IQR)]*	5.80	(3.02-11.10)
Total bilirubin xULN [median (IQR)]*	0.82	(0.52-1.21)
MRS*	-0.28	(-0.77-0.78)

SD = standard deviation; PSC = primary sclerosing cholangitis; IQR = inter quartile range; mm = millimeter; AST = aspartate aminotransferase; ALT = alanine aminotransferase; ALP = alkaline phosphatase; γGT = gamma-glutamyl transferase; xULN = times upper limit of normal; MRS = Mayo risk score

* Data available for: AST n=56, ALT n= 51, ALP n=52, γGT n=46 total bilirubin n= 49, MRS n=33.

Distribution grade and stage

Consensus on the evaluation of grade and stage was reached in 100% of cases.

Grading

Ishak grading component interface hepatitis showed a median grade of 1 (range 1-2). Confluent necrosis was a relative infrequent finding in PSC liver biopsies with a median grade of 0 (range 0-1). The median grade of lobular inflammation and portal inflammation was 1 (range 0-2) and 1 (range 0-3), respectively. The majority of liver biopsies did not show Nakanuma grading components CA with a median of 0 (range 0-3) nor HA, median 0 (range 0-2). Distribution of total Ishak and Nakanuma grade is shown in Fig. 1A,B.

Staging

Fig. 1C represents the Ishak stage, with a median of 2 (range 0-6). In Fig.1D the Nakanuma stage (median 2 (range1-4)) is shown, derived from the components fibrosis, bile duct loss and deposition of orcein positive granules score. Ludwig stage is displayed in Fig. 1E, showing a normal distribution (mean 2, SD 1).

The distribution of Ishak and Nakanuma grading components, and the Nakanuma staging components is illustrated in Supplementary Fig. 1.

Fig.1. Distribution of grading and staging systems.
(A) Ishak grading system. (B) Nakanuma grading system.
(C) Ishak staging system. (D) Nakanuma staging system. (E) Ludwig staging system.

Patient outcome

PSC related death, liver transplantation and cholangiocarcinoma

Long-term follow up data were present for all patients. A total of 11 (17%) patients developed an endpoint; 3 patients were diagnosed with cholangiocarcinoma of whom 2 patients died and eight patients underwent LTx. The median time to endpoint was 86 months (IQR 43-142).

Liver cirrhosis related symptoms

Data on the occurrence of CRS were available for 63 patients. A total of 16 (25%) patients developed CRS, of which nine developed more than one CRS. Nine (14%) patients developed varices, of whom one had a variceal bleeding. Ascites was found in six (10%) patients, splenomegaly in twelve (19%) and three (5%) patients presented with hepatic encephalopathy. The median time until presentation of the first CRS was 53 months (range 0-280), one patient presented with CRS at time of liver biopsy.

Association of histologic grade and stage with patient outcome

The prognostic significances of clinical, biochemical and histopathological parameters in predicting transplant-free survival, time to LTx and CRS, calculated by univariable Cox proportional hazard analyses are summarized in Table 2.

Grading

No significant association of the Ishak and Nakanuma grading systems and outcome measures transplant-free survival, time to LTx and CRS were shown.

Staging

A significant association was shown for Ishak staging system and transplant-free survival ($p=0.04$) as well as Ishak stage and time to LTx ($p=0.005$) (Fig. 2A,B). Univariable Cox proportional hazard analysis revealed a hazard ratio of 2.56; 95% CI 1.11-5.89 ($p=0.03$) and HR 4.18; 95% CI 1.51-11.56 ($p=0.006$) for transplant-free survival and time to LTx, respectively (Table 2). Nakanuma staging components fibrosis and CBP deposition were significantly associated with transplant-free survival and time to LTx, in which degree of CBP deposition was most pronounced ($p < 0.001$ vs $p=0.005$ respectively, Supplementary Table 3). This resulted in a strong significant association of Nakanuma staging system as a whole with both transplant-free survival ($p < 0.001$) and time to LTx ($p < 0.001$) (Fig. 2 C,D). The accompanying hazard ratio was 6.53; 95%CI 2.01-21.22 ($p=0.002$) and HR 7.05; 95% CI 1.77-28.11 ($p=0.006$) for transplant-free survival and time to LTx, respectively. (Table 2). The Ludwig staging system also showed a significant association

with outcome measurements transplant-free survival and time to LTx (log-rank $p=0.027$ and 0.002 , respectively), HR 1.94; 95% CI 1.00-3.79 ($p=0.05$) and HR 3.13; 95% CI 1.42-6.87 ($p=0.005$), respectively, (Fig. 2 E,F; Table 2).

The endpoint cirrhosis related symptoms showed a significant association with the Ludwig staging system (log rank $p=0.044$). There was no association between the Nakanuma staging system as a whole and the development of CRS, though the Nakanuma staging component fibrosis did show a significant association (log rank $p=0.04$). The Ishak staging system was not associated with the development of cirrhosis related symptoms (Supplementary Table 3).

Univariable analyses of the other parameters showed that bilirubin and ALP were predictive factors for time to LTx. Though a trend for similar associations was seen for the endpoint transplant-free survival, this was not significant. However, MRS was a predictive factor for transplant-free survival (Table 2).

An overview of median survival time per grade and stage, as well as the associations of histologic stage and grade with the three outcome measurements, calculated by log-rank test is given in the Supplementary Table 3.

Table 2. Prognostic significances of clinical, biochemical and histopathological parameters in predicting transplant-free survival, time to LTx and cirrhosis related symptoms, calculated by univariable Cox proportional hazard analyses.

	Transplant-free survival		Time to LTx		Cirrhosis related symptoms*	
	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value
Sex	0.97 (0.28 – 3.32)	0.96	0.57 (0.12-2.85)	0.50	0.37 (0.11-1.30)	0.12
Age at PSC diagnosis	1.01 (0.96 – 1.05)	0.63	1.00 (0.95-1.05)	0.97	1.01 (0.97-1.04)	0.70
PSC type	0.67 (0.09-5.28)	0.71	0.04 (0.00-301.42)	0.48	0.37 (0.05-2.78)	0.33
Coexisting IBD	0.43 (0.13-1.43)	0.17	0.66 (0.16-2.77)	0.57	1.95 (0.56-6.84)	0.30
Cirrhosis related symptoms*	1.19 (0.31-4.62)	0.80	2.07 (0.46-9.25)	0.34	NA	NA
AST*	1.18 (1.00-1.40)	0.05	1.11 (0.87-1.43)	0.41	1.07 (0.91-1.26)	0.42
ALT*	1.05 (0.91-1.21)	0.53	1.02 (0.84-1.24)	0.86	1.07 (0.97-1.19)	0.18
ALP*	1.27 (0.96-1.68)	0.10	1.41 (1.03-1.93)	0.03	1.02 (0.77-1.34)	0.90
γGT*	0.95 (0.82-1.10)	0.51	0.98 (0.84-1.14)	0.78	1.05 (0.97-1.14)	0.20
Total bilirubin*	1.53 (0.99-2.35)	0.05	1.73 (1.07 – 2.79)	0.03	1.17 (0.80-1.70)	0.42
MRS*	3.41 (1.15-10.09)	0.03	3.19 (0.97-10.47)	0.06	1.31 (0.72-2.38)	0.38
Ishak grade (1-4)	1.50 (0.86-2.61)	0.15	1.56 (0.82-2.99)	0.18	1.11 (0.69-1.79)	0.68
Nakanuma grade (0, 1, ≥2)	1.35 (0.65-2.82)	0.43	1.42 (0.60-3.31)	0.43	1.26 (0.69-2.30)	0.45
Ishak stage 1-3	2.56 (1.11-5.89)	0.03	4.18 (1.51-11.56)	0.006	1.09 (0.95-3.81)	0.07
Nakanuma stage (1-4)*	6.53 (2.01-21.22)	0.002	7.05 (1.77-28.11)	0.006	1.57 (0.70-3.51)	0.28
Ludwig stage (1-4)	1.94 (1.00-3.79)	0.05	3.13 (1.42-6.87)	0.005	1.16 (0.66-2.00)	0.62

HR = Hazard Ratio; CI = confidence interval; LTx = liver transplantation; PSC = primary sclerosing cholangitis; IBD = inflammatory bowel disease; AST = aspartate aminotransferase; ALT = alanine aminotransferase; ALP = alkaline phosphatase; γGT = gamma-glutamyl transferase; MRS = Mayo risk score; NA = not assessed.

* Data available for: AST n=56, ALT n=51, ALP n=52, γGT n=46, total bilirubin n=49, MRS n=33, Nakanuma stage n=58; cirrhosis related symptoms n=63.

Fig. 2. Kaplan Meier survival curves. Endpoints transplant-free survival and time to liver transplantation, shown for: (A,B) Ishak staging system. (C,D) Nakanuma staging system. (E,F) Ludwig staging system.

Correlation between histologic grade and stage and liver biochemistry values

In Tables 3a and 3b correlation between liver biochemistry and histologic grade and stage are shown, calculated using Spearman's rank correlation test. For some patients serum liver tests at time of liver biopsy were not available. Liver biochemistry could be retrieved for AST n=56, ALT n= 51, ALP n=52, γ GT n=46, total bilirubin n= 49.

For grading, significant positive correlations between serum ALP level and the Ishak grading component portal inflammation ($r=0.31$, $p=0.027$), Ishak grading system as a whole ($r=0.31$, $p=0.023$) and the Nakanuma grading component HA ($r=0.28$, $p=0.047$) were shown.

For staging, a positive correlation between total bilirubin level and Ishak staging system ($r=0.32$, $p=0.027$), as well as Nakanuma staging component fibrosis were demonstrated ($r=0.29$, $p=0.042$).

Correlation coefficients did not exceed 0.33.

Table 3a. Correlation between grading and liver biochemistry values.

	Inter- face activity (0, 1, ≥2)	Con- fluent necrosis (0, ≥1)	Lobular inflam- mation (0, 1, ≥2)	Portal inflam- mation (0, 1, ≥2)	Ishak total (1-4)	CA (0, 1, ≥2)	HA (0, ≥1)	Na- kanuma total (0, 1, ≥2)
AST*	0.062 p=0.650	-0.018 p=0.896	0.094 p=0.490	0.066 p=0.627	0.171 p=0.207	-0.038 p=0.783	0.236 p=0.080	0.100 p=0.462
ALT*	-0.122 p=0.393	-0.041 p=0.774	0.055 p=0.704	-0.025 p=0.863	-0.009 p=0.951	-0.201 p=0.158	0.052 p=0.716	-0.146 p=0.307
ALP*	0.256 p=0.067	-0.247 p=0.078	0.163 p=0.248	0.306 p=0.027	0.314 p=0.023	0.143 p=0.311	0.277 p=0.047	-0.132 p=0.360
γGT*	0.146 p=0.333	0.062 p=0.683	-0.093 p=0.540	0.001 p=0.996	0.115 p=0.446	-0.018 p=0.905	0.239 p=0.110	0.126 p=0.402
Total* bilirubin	0.141 p=0.334	-0.011 p=0.940	-0.072 p=0.621	0.268 p=0.063	0.254 p=0.090	-0.241 p=0.095	0.093 p=0.526	-0.050 p=0.731

CA = cholangitis activity; HA = hepatitis activity; AST = aspartate aminotransferase; ALT = alanine aminotransferase; ALP = alkaline phosphatase; γGT = gamma-glutamyl transferase.

* Data available for: AST n=56, ALT n= 51, ALP n=52, γGT n=46, total bilirubin n= 49.

Table 3b. Correlation between staging and liver biochemistry value, calculated by Spearman's rank correlation test.

	Ishak (1-3)	Fibrosis (0-3)	Bile duct loss (0-3)	Orcein positive granules** (0-3)	Nakanuma total** (1-4)	Ludwig (1-4)
AST*	0.050 p=0.717	-0.028 p=0.840	0.039 p=0.778	-0.028 p=0.843	0.074 p=0.606	-0.011 p=0.938
ALT*	-0.167 p=0.241	-0.191 p=0.179	-0.040 p=0.780	-0.109 p=0.465	-0.085 p=0.569	-0.162 p=0.256
ALP*	0.180 p=0.201	0.088 p=0.535	0.001 p=0.992	0.024 p=0.870	0.146 p=0.322	0.131 p=0.356
γGT*	0.096 p=0.528	0.049 p=0.748	0.110 p=0.468	0.188 p=0.233	0.114 p=0.473	0.045 p=0.769
Total* bilirubin	0.317 p=0.027	0.292 p=0.042	0.245 p=0.090	0.121 p=0.429	0.255 p=0.091	0.090 p=0.537

AST = aspartate aminotransferase; ALT = alanine aminotransferase; ALP = alkaline phosphatase; γGT = gamma glutamyl transferase.

* Data available for: AST n=56, ALT n= 51, ALP n=52, γGT n=46, total bilirubin n= 49.

** Data available for: n=58.

DISCUSSION

With this study we have demonstrated that the Ishak, Nakanuma, and Ludwig histological **staging** systems are predictive indicators of transplant-free survival and time to LTx in PSC.

The Nakanuma staging system appeared to have the strongest predictive power, given its highest incremental hazard of long-term outcome corresponding to the stage progression. This may be explained by the fact that the Nakanuma staging system includes features that could be considered PSC appropriate. When focusing on the prognostic value of these individual components, degree of fibrosis and CBP deposition both predicted transplant-free survival and time to LTx, in which degree of CBP deposition was most discriminative ($p < 0.001$, $p = 0.005$ respectively). The Nakanuma staging component bile duct loss was not associated with outcome. This is in contrast with findings in PBC, where bile duct loss has been reported to be a predictor of disease progression.²⁵ Moreover, in PSC a peripheral liver biopsy might not be the best way to evaluate the extent of bile duct loss, since the disease affects the entire biliary tree and obliteration or obstruction of the larger central bile ducts may greatly affect progression to liver failure.¹⁵

The prognostic value of the Ishak and Ludwig staging system, as well as Nakanuma component fibrosis implicate that degree of fibrosis is of important prognostic value in PSC. Similar findings have recently been demonstrated by Ruiz *et al.*, who analyzed radiologic disease course in PSC by 3-dimensional MRC, and found that risk of progression was mainly dependent on the effects of biliary disease on liver parenchyma, rather than the effects of severe stricturing.²⁶

Necroinflammatory activity may predispose to development of CRS, however, no associations of histologic **grading** systems with CRS were shown. In PBC, associations between histology and CRS seem to be more pronounced. Kakuda *et al.* showed a significant prognostic value of Nakanuma components fibrosis and CBP deposition, as well as Nakanuma and Ludwig staging systems as a whole for the development of cirrhosis related conditions before ursodeoxycholic acid (UDCA) treatment.²⁴ In a retrospective cohort of 58 PBC patients Chan *et al.* confirmed the Nakanuma system as a prognostic factor for liver related events, while the Ludwig system was not.²⁷ It must be stated that, despite that PSC and PBC are both cholangiopathies, sharing biochemical, clinical and some morphological features, considerable differences exist in both disease entities¹⁵. The patchy distribution of affected bile ducts throughout the liver, with resulting sampling variability may be greater in PSC than in PBC.¹⁸ Furthermore, the differences in statistical analyses as well as in definition of CRS between studies may contribute to these conflicting results.

The extent to which histologic features, scored by the different scoring systems were present in PSC biopsies may indicate their applicability in PSC. The Nakanuma grading component cholangitis activity – though well recognized in PBC – is not typically seen in PSC. This was reflected in the presence of cholangitis activity in only 38% of cases in our cohort and the associated lack of predictive value. Hepatitis activity was detected in a minority of 27% of cases and had no predictive value. This may partly be explained by the exclusion of patients with auto-immune hepatitis overlap syndrome, in whom more severe hepatitis activity would be expected. PSC appropriate Nakanuma staging components bile duct loss and orcein positive granules were present in only 50% of cases. This could be due to the inclusion of biopsies taken at time of diagnosis, creating a cohort of patients with relatively early-stage disease.

Exploratory analyses for correlations between liver biochemistry and histologic grade and stage were performed to assess if liver biochemistry could reflect the degree of liver injury measured by liver histology. No strong correlations could be found, since correlation coefficients did not exceed 0.33. This suggests that liver biochemistry does not adequately reflect the degree of histologic injury. This finding is supported by a recent study of Queen *et al.* who evaluated the clinical course, endoscopic and pathologic findings of patients with normal liver biochemistry and showed that PSC patients can have cirrhosis and significant ductal disease, in the setting of normal liver biochemistry.²⁸

In 1995, Olsson *et al.* demonstrated a sampling variability of at least 20% for blind needle biopsy in PSC, which is often thought to refute its use as a prognostic measure.²⁹ The results of our study show that sampling variability is apparently not a major confounder. This is in line with results from a study by Angulo *et al.*, who assessed the time course of histologic progression in PSC.³⁰ They observed progression in Ludwig stage in liver biopsies over two years' time in 53% of PSC patients with an initial Ludwig stage of I-III.³⁰ In a Markov model they estimated that the risk of progression from Ludwig stage II to stage III or IV was 66% after two years. These results indicate that despite the issue of sampling variability, progression of liver disease can be evaluated using liver biopsy.³⁰ In addition, several clinical trials have included histologic change in grade-/stage as clinical endpoint.⁷⁻¹³ Angulo *et al.* noted a significant improvement in the degree of portal inflammation – grade – after one year of budesonide treatment.¹³ Degree of fibrosis and disease stage progressed despite treatment.¹³ In a 2-year double-blind placebo-controlled study of high-dose UDCA, Mitchell *et al.* showed a significant reduction in progression of disease stage in the treatment group.⁹ Fifty percent of the placebo group showed no change in disease stage, while the other fifty percent showed progression by a single stage.⁹ These studies indicate that

both progression and improvement of liver disease grade-/stage can adequately be assessed by liver biopsy.^{9,13,30} In addition, with the present study the prognostic value of liver biopsy on outcome was demonstrated. Therefore, liver histology appears to be a useful biomarker for progression, and its use as outcome measure in clinical trials deserves consideration. Moreover, for regulatory bodies, histology is still the gold standard as outcome parameter in many liver diseases.

A limitation of this study is that multivariable Cox proportional hazard regression analysis, to analyze if the prognostic value of the histology is independent of clinical and biochemical variables, could not be applied for methodological reasons. When performing a multivariable Cox proportional hazard regression analysis, the total numbers of events – not the total number of included cases – is important, to retain the validity of the outcome.³¹ The most appropriate number of events per variable (EPV) to use in a multivariable analysis is 10.^{32,33} For the present study a total of 140 events – an increase of the sample size up to 10 times – would be necessary to generate stable and interpretable results by multivariable analysis, which was unfortunately not feasible.

The retrospective collection of serum liver tests and data on occurrence of liver CRS may be incomplete and hence may have led to a relative underestimation of the occurrence of liver CRS. However, follow-up data on solid clinical endpoints were complete in all 64 patients. Its retrospective character is also an asset in the sense that it yielded a median follow-up time of 112 months. Guido *et al.* described liver biopsy samples of 22 mm or more in length, with at least 11 portal tracks to be most reliable for grading and staging at least in the setting of viral hepatitis.³⁴ Although the quality of liver biopsies of the present study was high with a median amount of portal tracks of 13 (IQR 9-19) and a median biopsy length of 14 mm (IQR 11-19), these optimal requirements were not met in all cases. The biopsies included in this study were all taken at time of diagnosis, which may result in spectrum bias. However, in Fig. 1 a normal distribution of grades and stages for most of the classification systems is shown, demonstrating that PSC patients diagnosed in both early and late disease stage were included in this study. Lastly the percentage of small duct PSC in this study (16%) is relatively high in comparison with the 9% small duct PSC patients reported in a large population based epidemiology study.¹⁷ Small duct PSC patients are known to have a better prognosis, which could influence results of outcome. For this reason analyses for associations between histological grading and staging and outcome were repeated in only the large duct patients of this cohort. Similar results were demonstrated (data not shown).

In conclusion, applying the Nakanuma, Ishak, and Ludwig histopathological staging systems in PSC is feasible and clinically relevant. Our results support

the notion that histopathological scoring systems are likely candidates for surrogate endpoints and stratification in clinical trials in PSC, and may be used to assess noninvasive biomarkers for future trials. Validation of the applicability of these staging systems as well as determination of inter-observer variability to test the robustness of these scoring systems in PSC in a large multi-center cohort is warranted.

REFERENCES

1. Hirschfield, G. M., Karlsen, T. H., Lindor, K. D. & Adams, D. H. Primary sclerosing cholangitis. *Lancet* 382, 1587–99 (2013).
2. Maneesh, D., Elmunzer, B. J., Dwame-na, B. & Higgins, P. Primary Sclerosing Cholangitis: Meta-Analysis of Diagnostic Performance of MR Cholangiopancreatography. *Radiology* 256, 387–396 (2010).
3. Weber, C. et al. Magnetic resonance cholangiopancreatography in the diagnosis of primary sclerosing cholangitis. *Endoscopy* 40, 739–745 (2008).
4. Maccarty, R., Larusso, N., Wiesner, R. & Ludwig, J. Primary Sclerosing Cholangitis: Findings on Cholangiography and Pancreatography. *Radiology* 149, 39–44 (1983).
5. Chapman, R. et al. Diagnosis and management of primary sclerosing cholangitis. *AASLD Practice Guidelines. Hepatology* 51, 660–78 (2010).
6. European Association for the Study of the Liver. EASL Clinical Practice Guidelines: management of cholestatic liver diseases. *J. Hepatol.* 51, 237–67 (2009).
7. Lindor, K. D. Ursodiol for primary sclerosing cholangitis. *NEJM* 336, 691–695 (1997).
8. Lindor, K. D. et al. High-dose ursodeoxycholic acid for the treatment of primary sclerosing cholangitis. *Hepatology* 50, 808–14 (2009).
9. Mitchell, S. a. et al. A preliminary trial of high-dose ursodeoxycholic acid in primary sclerosing cholangitis. *Gastroenterology* 121, 900–907 (2001).
10. Färkkilä, M. et al. Metronidazole and ursodeoxycholic acid for primary sclerosing cholangitis: a randomized placebo-controlled trial. *Hepatology* 40, 1379–86 (2004).
11. Hommes, D. W. et al. A double-blind, placebo-controlled, randomized study of infliximab in primary sclerosing cholangitis. *J. Clin. Gastroenterol.* 42, 522–6 (2008).
12. Thiel, D. H. Van et al. Tacrolimus (FK), a Treatment for Primary Sclerosing Cholangitis; Results of an Open-Label Preliminary Trial. *Am J Gastroenterol* 90, 455–459 (2010).
13. Angulo, P., Batts, K. P., Jorgensen, R. a, LaRusso, N. a & Lindor, K. D. Oral budesonide in the treatment of primary sclerosing cholangitis. *Am. J. Gastroenterol.* 95, 2333–7 (2000).
14. Barbatis, C. et al. Histological features of sclerosing cholangitis in patients with chronic ulcerative colitis. *J Clin Pathol* 38, 778–783 (1985).
15. Scheuer, P. & Path, F. Pathologic Features and Evolution of Primary Biliary Cirrhosis and Primary Sclerosing Cholangitis. *Ludwig Symposium on Biliary Disorders-PartII. Mayo Clin Proc* 73, 179–183 (1998).
16. Goodman, Z. D. Grading and staging systems for inflammation and fibrosis in chronic liver diseases. *J. Hepatol.* 47, 598–607 (2007).
17. Boonstra, K. et al. Population-based epidemiology, malignancy risk, and outcome of primary sclerosing cholangitis. *Hepatology* 58, 2045–55 (2013).
18. Portmann, B. & Zen, Y. Inflammatory disease of the bile ducts-cholangiopathies: liver biopsy challenge and clinico-pathological correlation. *Histopathology* 60, 236–48 (2012).
19. Ishak, K. et al. Histological grading and staging of chronic hepatitis. *J. Hepatol.* 22, 696–9 (1995).
20. Ludwig, J., Dickson, E. & McDonald, G. Staging of chronic nonsuppurative destructive cholangitis (syndrome of primary biliary cirrhosis). *Virchows Arch A Pathol Anat Histol* 22, 103–12 (1978).

21. Nakanuma, Y., Zen, Y., Harada, K. & Sasaki, M. Application of a new histological staging and grading system for primary biliary cirrhosis to liver biopsy specimens: Interobserver agreement. *Pathol. Int.* 60, 167–174 (2010).
22. Hiramatsu, K., Aoyama, H., Zen, Y. & Aishima, S. Proposal of a new staging and grading system of the liver for primary biliary cirrhosis. *Histopathology* 49, 466–478 (2006).
23. Hennes, E. M. et al. Simplified criteria for the diagnosis of autoimmune hepatitis. *Hepatology* 48, 169–76 (2008).
24. Kakuda, Y., Harada, K., Sawada-Kitamura, S., Ikeda, H. & Nakanuma, Y. Evaluation of a new histologic staging and grading system for primary biliary cirrhosis in comparison with classical systems. *Hum. Pathol.* 44, 1109–1117 (2013).
25. Kumagi, T. et al. Baseline Ductopenia and Treatment Response Predict Long-Term Histological Progression in Primary Biliary Cirrhosis. *Am. J. Gastroenterol.* 105, 2186–2194 (2010).
26. Ruiz, A. et al. Radiologic course of primary sclerosing cholangitis: assessment by three-dimensional magnetic resonance cholangiography and predictive features of progression. *Hepatology* 59, 242–50 (2014).
27. Chan, A. W. H. et al. Evaluation of histological staging systems for primary biliary cirrhosis: correlation with clinical and biochemical factors and significance of pathological parameters in prognostication. *Histopathology* 65, 174–86 (2014).
28. Queen, T., Cox, K. & Adler, D. G. Primary Sclerosing Cholangitis in the Setting of Normal Liver Chemistries Can Be Associated with Severe Ductal Disease and Dominant Strictures. *Dig Dis Sci* 229–230 (2014).
29. Olsson, R. et al. Sampling variability of percutaneous liver biopsy in primary sclerosing cholangitis. *J. Clin. Pathol.* 48, 933–5 (1995).
30. Angulo, P. et al. Time course of histological progression in primary sclerosing cholangitis. *Am. J. Gastroenterol.* 94, 3310–3 (1999).
31. Kawada, T. The appropriate number of endpoints to keep validity for Cox proportional hazard analysis. *Int. J. Cardiol.* 153, 110–1 (2011).
32. Concato, J., Peduzzi, P., Holford, T. R. & Feinstein, A. R. Importance of events per independent variable in proportional hazards analysis I. Background, goals, and general strategy. *J. Clin. Epidemiol.* 48, 1495–1501 (1995).
33. Peduzzi, P., Concato, J. & Feinstein, A. R. Importance of events per independent variable in proportional hazard regression analysis II. Accuracy and precision of regression estimates. *J. Clin. Epidemiol.* 48, 1503–1510 (1995).
34. Guido, M. & Rugge, M. Liver biopsy sampling in chronic viral hepatitis. *Semin Liver Dis* 24, 89–97 (2004).

Supplementary Fig. 1. Distribution of grading and staging components. (A) Ishak grading component; interface hepatitis. (B) Ishak grading component; confluent necrosis. (C) Ishak grading component; lobular inflammation. (D) Ishak grading component; portal inflammation. (E) Nakanuma grading component; cholangitis activity. (F) Nakanuma grading component; hepatitis activity. (G) Nakanuma staging component; fibrosis. (H) Nakanuma staging component; bile duct loss. (I) Nakanuma staging component; orcein positive granules.

Supplementary Table 1. Grading according Ishak and Nakanuma.^{19,21}

Score	Criteria
<i>Ishak</i>	
Periportal or periseptal interface hepatitis (piecemeal necrosis)	
0	None
1	Mild (focal, few, portal areas)
2	Mild/moderate (focal, most portal areas)
3	Moderate (continuous around <50% of tracts or septa)
4	Severe (continuous around >50% of tracts or septa)
Confluent necrosis	
0	None
1	Focal confluent necrosis
2	Zone 3 necrosis in some areas
Confluent necrosis	
3	Zone 3 necrosis in most areas
4	Zone 3 necrosis + occasional portal-central bridging
5	Zone 3 necrosis + multiple portal-central bridging
6	Panacinar or multiacinar necrosis
Focal (spotty) lytic necrosis, apoptosis and focal inflammation	
0	None
1	One focus or less per 10x objective
2	Two to four foci per 10x objective
3	Five to ten foci per 10x objective
4	More than 10 foci per 10x objective
Portal inflammation	
0	None
1	Mild, some or all portal areas
2	Moderate, some or all portal areas
3	Moderate/marked, all portal areas
4	Marked, all portal areas
<i>Nakanuma</i>	
A. Cholangitis activity	
0 (no activity)	No cholangitis but mild damage to the epithelium of the duct may be present
1 (mild activity)	1 evident chronic cholangitis in the specimen
2 (moderate activity)	≥ 2 bile ducts with evident chronic cholangitis
3 (marked activity)	≥ 1 CNSDC in the specimen
B. Hepatitis activity	
0 (no activity)	No interface hepatitis and no or minimal lobular hepatitis
1 (mild activity)	Interface hepatitis affecting 10 continuous hepatocytes at limiting plate in 1 portal tract or fibrous septa and mild-to-moderate lobular hepatitis
2 (moderate activity)	Interface hepatitis affecting 10 continuous hepatocytes at limiting plate in ≥ 2 portal tract or fibrous septa and mild-to-moderate lobular hepatitis
3 (marked activity)	Interface hepatitis affecting 20 continuous hepatocytes at limiting plates in more than half of the portal tracts and moderate lobular hepatitis or bridging/zonal necrosis

Supplementary Table 2. Staging according Ishak, Nakanuma, and Ludwig.^{19–21}

Score	Criteria
<i>Ishak</i>	
0	No fibrosis
1	Fibrous expansion of some portal areas, with or without short fibrous septa
2	Fibrous expansion of most portal areas, with or without short fibrous septa
3	Fibrous expansion of most portal areas with occasional portal to portal bridging
4	Fibrous expansion of portal areas with marked bridging (portal to portal as well as portal to central)
5	Marked bridging (portal-portal and/or portal-central) with occasional nodules (incomplete cirrhosis)
6	Cirrhosis, probable or definite
<i>Nakanuma</i>	
A. Fibrosis	
0	No portal fibrosis or fibrosis limited to portal tracts
1	Portal fibrosis with periportal fibrosis or incomplete septal fibrosis
2	Bridging fibrosis with variable lobular disarray
3	Liver cirrhosis with regenerative nodules and extensive fibrosis
B. Bile duct loss	
0	No bile duct loss
1	Bile duct loss in less than one third of portal tracts
2	Bile duct loss in one third to two thirds of portal tracts
3	Bile duct loss in more than two-thirds of portal tracts
C. Deposition of orcein positive granules	
0	No deposition of granules
1	Deposition of granules in a couple of zone 1 hepatocytes at less than one-third of portal tracts
2	Deposition of granules in a variable number of zone 1 hepatocytes at one-third to two-thirds of portal tracts
3	Deposition of granules in most zone 1 hepatocytes at more than two-thirds of portal tracts
<i>Ludwig</i>	
0	No fibrosis
1	Cholangitis or portal hepatitis
2	Periportal fibrosis or hepatitis
3	Septal fibrosis, bridging necrosis or both
4	Biliary cirrhosis

Supplementary Table 3. Association of histological grade and stage with outcome measurements transplant-free survival, time to LTx and cirrhosis related symptoms, by log-rank test.

		Transplant-free survival			Time to LTx			Cirrhosis related symp- toms*				
GRADE		Median survival months (IQR)		p-value (log- rank)	Median time to LTx months (IQR)		p-value (log- rank)	Median time to CRS months (IQR)		p-value (log- rank)		
Ishak												
Interface	0:	164	(108-194)	0.172	164	(108-194)	0.294	161	(111-174)	0.997		
hepatitis	1:	93	(65-197)			(65-197)			(55-179)			
(0 vs. 1 vs. ≥2)	≥2:	111	(66-140)		93	(66-140)		75	(72-175)			
					111			112				
Confluent	0:	111	(69-175)	0.520	111	(69-175)	0.620	111	(65-177)	0.507		
necrosis	1:	135	-			135		-			108	-
(0 vs. ≥1)	≥2:											
Lobular	0:	125	(92-201)	0.569	125	(92-201)	0.619	111	(53-171)	0.533		
inflammation	1:	107	(65-169)			107		(65-169)			111	(66-180)
(0 vs. 1 vs. ≥2)	≥2:	120	-			120		-			111	-
Portal	0:	146	(89-169)	0.098	146	(89-169)	0.057	135	(99-170)	0.522		
inflammation	1:	122	(66-205)			122		(66-205)			111	(65-203)
(0 vs. 1 vs. ≥2)	≥2:	87	(66-132)			89		(66-132)			85	(63-132)
Total (1-4)	1 (0-1):	163	(105-189)	0.115	163	(105-189)	0.059	163	(112-189)	0.970		
	2 (2,3):	111	(65-204)			111		(65-204)			93	(60-197)
	3 (4):	81	(61-120)			81		(61-120)			90	(67-152)
	4 (≥5):	111	(66-156)			111		(66-156)			111	(66-169)
Nakanuma												
Cholangitis	0:	111	(67-201)	0.099	111	(67-201)	0.238	111	(65-203)	0.891		
activity	1:	124	(65-169)			124		(65-169)			111	(26-169)
(0 vs. 1 vs. ≥2)	≥2:	103	(76-149)			103		(76-160)			95	(70-142)
Hepatitis	0:	120	(70-179)	0.220	120	(70-179)	0.059	117	(66-175)	0.170		
activity	≥1:	111	(57-142)			111		(57-142)			78	(34-150)
(0 vs. ≥1)												
Total (0-6)				0.855			0.769			0.765		
Total	0:	118	(72-201)	0.686	118	(72-201)	0.668	117	(65-203)	0.734		
(0 vs. 1 vs. ≥2)	1:	115	(42-181)			115		(42-181)			114	(37-186)
	≥2:	108	(75-131)			108		(75-134)			77	(63-119)

Supplementary Table 3. Continued Association of histological grade and stage with outcome measurements transplant-free survival, time to LTx and cirrhosis related symptoms, by log-rank test.

STAGE		Median survival months (IQR)		p-value (log- rank)	Median time to LTx months (IQR)		p-value (log- rank)	Median time to CRS months (IQR)		p-value (log- rank)
<i>Ishak (1-3)</i>	1 (0,1):	161	(91-208)	0.041	163	(91-208)	0.005	126	(75-203)	0.119
	2 (2,3):	100	(66-148)		100	(66-148)		89	(61-154)	
	3 (≥4):	66	(21-156)		66	(21-156)		119	(14-180)	
<i>Nakanuma</i>										
Fibrosis (0-3)	0:	147	(115-198)	0.022	147	(115-198)	0.001	147	(115-199)	0.040
	1:	119	(78-203)		119	(78-203)		111	(65-185)	
	2:	81	(63-134)		81	(63-134)		77	(47-118)	
	3:	21	-		21	-		180	-	
Bile duct loss (0-3)	0:	126	(72-199)	0.614	126	(72-199)	0.848	126	(68-204)	0.514
	1:	111	(63-146)		111	(63-158)		95	(55-154)	
	2:	-			-			-		
	3:	111	(111-111)		111	(111-111)		111	(111-111)	
Orcein positive granules (0- 3)**	0:	135	(96-204)	<0.001	142	(96-204)	0.041	117	(62-181)	0.427
	1:	112	(65-192)		112	(65-192)		112	(74-192)	
	2:	58	-		58	-		58	-	
	3:	68	(23-82)		68	(23-84)		80	(28-157)	
Orcein positive granules (1,2)**	1(0-1):	127	(88-203)	<0.001	127	(88-203)	0.005	117	(66-179)	0.602
	2 (2-3):	66	(21-86)		66	(21-89)		71	(14-180)	
Total (1-4)**	1:	188	-	<0.001	188	-	<0.001	189	-	0.355
	2:	127	(92-203)		127	(92-203)		117	(65-179)	
	3:	66	(51-115)		66	(51-115)		76	(25-151)	
	4:	9	(9-9)		9	(9-9)		180	(180-180)	
<i>Ludwig (1-4)</i>	1 (0,1):	164	(115-189)	0.027	165	(115-189)	0.002	152	(111-173)	0.044
	2:	103	(70-205)		103	(70-205)		90	(62-204)	
	3:	81	(63-134)		81	(63-134)		77	(47-118)	
	4:	21	-		21	-		180	-	

CRS = cirrhosis related symptoms; IQR = inter quartile rang; LTx = liver transplantation.

* data available of 63 patients

** Data available for: n=58

- Unable to calculate IQR due to small numbers

Validation of the prognostic value of histologic scoring systems in primary sclerosing cholangitis; an international cohort study

Elisabeth M.G. de Vries, Joanne Verheij, Martti Farkkila, Johanna Arola, Peter Schirmacher, Daniel Gotthardt, Benjamin Goeppert, Palak J. Trivedi, Gideon M. Hirschfield, Henriette Ytting, Ben Vainer, Henk R. van Buuren, Katharina Biermann, Maren H. Harms, Olivier Chazouillers, Dominique Wendum, Astrid D. Kemgang, Roger W. Chapman, Lai Mun Wang, Kate D. Williamson, Ulrich H. Beuers, Stefan G. Hubscher, Cyriel Y. Ponsioen

Submitted

ABSTRACT

Introduction

Histologic scoring systems specific for primary sclerosing cholangitis (PSC) are not validated. We recently determined the applicability and prognostic value of three histological scoring systems in a single cohort of PSC patients. The aim of this study was to validate their prognostic utility, using a well-characterized multicenter PSC cohort.

Material and Methods

Liver biopsies were collected from PSC patients across 7 European centers. Histologic scoring was performed using the Nakanuma, Ishak, and Ludwig scoring systems. The prognostic value of clinical, biochemical, and all three histologic scoring systems on predicting composite endpoint 1: PSC related death and liver transplantation, endpoint 2: liver transplantation, endpoint 3: liver related events, was assessed using uni- and multivariable Cox proportional hazards modelling.

Results

119 PSC patients who had undergone liver biopsy were identified, the median follow-up was 142 months. During follow-up 31 patients died (20 PSC related deaths), 31 underwent liver transplantation, 35 experienced one or more liver related events. All three staging systems were independent predictors of endpoints 2 and 3; Nakanuma HR3.16 (95%CI 1.49-6.68), HR 2.05 (95%CI 1.17-3.57); Ishak: HR 1.55 (95%CI 1.10-2.18), HR 1.43 (95%CI 1.10-1.85); Ludwig: HR 2.62 (95%CI 1.19-5.80), HR 2.06 (95%CI 1.09-3.89), respectively. Only the Nakanuma staging system was independently associated with endpoint 1: HR2.14 (95%CI 1.22-3.77)

Conclusion

We substantiate the independent prognostic value of staging disease progression in PSC by the Nakanuma, Ishak and Ludwig histologic staging systems. The Nakanuma staging system – incorporating features of chronic biliary disease – was again shown to hold strongest predictive value, and may be preferable for scoring PSC liver biopsies. Histologic stage can be useful for risk stratification, and warrants further study as surrogate endpoint in clinical trials.

INTRODUCTION

Primary sclerosing cholangitis (PSC) is an idiopathic cholestatic liver disease, affecting the intra- and/or extrahepatic bile ducts.¹ Chronic inflammation induces formation of fibrotic strictures, interspersed with dilatations, resulting in a typical 'beaded' appearance when visualizing the PSC bile ducts on a cholangiogram.

Liver biopsy is no longer routinely performed to establish a diagnosis of PSC.^{2,3} However, there are still some circumstances in which the assessment of liver histology is important diagnostically, such as the suspicion of small duct PSC or PSC/auto-immune hepatitis (AIH) overlap syndrome. Beside its diagnostic purpose, histological assessment of disease severity may also be useful to stratify clinical outcomes in autoimmune and cholestatic liver diseases; particularly when determining entry into clinical trials and evaluating the efficacy of prospective therapeutic interventions.⁴

Histologically, the characteristic fibrosing duct lesions of PSC mainly involve medium-sized (septal) ducts and are thus infrequently seen in needle biopsy specimens.⁵ Varying degrees of portal and periportal inflammation are commonly present, particularly during the early stages of PSC, and may be associated with bile duct inflammation. As the disease progresses, there is loss of small and medium-sized ducts (ductopenia) and the development of secondary changes related to chronic cholestasis including ductular reaction and the accumulation of copper binding protein in periportal hepatocytes.⁶ These changes are associated with the development of progressive periportal fibrosis, ultimately leading to the development of biliary cirrhosis.⁶

In the absence of a validated histologic scoring system specifically designed to assess disease severity in PSC, we recently assessed the applicability and prognostic value of three histologic scoring systems, designed primarily to assess disease severity in chronic hepatitis (Ishak system) and primary biliary cholangitis (PBC) (Ludwig and Nakanuma systems), across a group of patients with PSC.⁷ Staging by all of these three scoring systems was shown to be strongly associated with disease progression in PSC.⁷ However, this study was performed in a single center, with a relatively small sample size, and included only biopsies taken at time of diagnosis.

With the present study we aimed to validate the prognostic value of liver histology, scored by the three previously evaluated scoring systems, using several well-characterized external cohorts of PSC patients.

PATIENTS AND METHODS

Patients and tissue preparation

For this multicenter cohort study, PSC patients from 7 centers in 6 European countries were included: Helsinki University Hospital, Finland; University Hospital Heidelberg, Germany; Queen Elizabeth Hospital, United Kingdom; Rigshospitalet, Denmark; Erasmus University Medical Center, the Netherlands; Hôpital Saint Antoine, France; John Radcliffe Hospital, United Kingdom. PSC diagnosis was established according to the European Association for the Study of the Liver guidelines.² A diagnosis of PSC/AIH-overlap syndrome was made in keeping with expertise of the contributing center.

To ensure adequate follow-up time, PSC liver biopsies taken at least 10 years ago were requested from the archives by liver pathologists in the participating centers. Biopsies taken between 1980-2006 were included. The original haematoxylin and eosin, connective tissue, and orcein stained sections were collected and the quality of the staining was checked. In those cases where staining had faded over time, additional stainings were carried out. In the event that the pathology laboratory of the participating center did not perform orcein histochemistry, additional orcein staining was performed by the laboratory of SH in Birmingham. Retrospectively, clinical data were collected from patient files considering patient demographics, medication use, liver biochemistry at time of diagnosis ($\pm$ 1 month), liver related events (variceal bleeding, ascites, splenomegaly or hepatic encephalopathy), and the occurrence of cholangiocarcinoma, liver transplantation, or death.

Medical Research Involving Human Subjects Act (WMO) did not apply to this study. Participation by individual centers was approved at a local level. Patient data were treated anonymously.

Histologic evaluation

Two expert liver pathologists (JV, SH) scored the liver biopsies in tandem using a multihead microscope, with the intention to reach consensus. Liver biopsy sections containing less than 6 evaluable portal tracts were excluded from the study. The criteria for scoring according to the Nakanuma, Ishak and Ludwig scoring systems have been described previously.⁸⁻¹⁰ (Supplementary Table 1,2) To summarize, both Nakanuma and Ishak scoring systems evaluate disease severity in terms of the histologic grade. The Nakanuma grading system includes degree of cholangitis activity (score 0-3), and hepatitis activity (score 0-3),⁹ while the Ishak grading system evaluates degree of interface hepatitis (score 0-4), confluent necrosis (score 0-6), lobular inflammation (score 0-4) and portal inflammation (score 0-4),⁸ (Supplementary Table 1)

All three histologic scoring systems evaluate disease progression in terms of stage, by scoring the degree of fibrosis; Nakanuma score 0-3, Ishak score 0-6; Ludwig score 0-4.⁸⁻¹⁰ The Nakanuma scoring system in addition scores the degree of bile duct loss (0-3) and deposition of orcein positive granules (representing copper binding protein) (0-3). The final Nakanuma stage is obtained from the total score of these three features: stage I (no or minimal progression) corresponds to a score of 0, stage II (mild progression) corresponds to a score of 1-3, stage III (moderate progression) corresponds to a score of 4-6 and stage IV (advanced progression) corresponds to a score of 7-9.⁹ (Supplementary Table 2.)

Endpoints

Several endpoints were assessed for the time-to-event analyses of association with endpoints. Endpoint 1 was transplant-free survival, defined as a composite of PSC-related death (death from end-stage liver disease, death from liver surgery, death from cholangiosepsis and death from cholangiocarcinoma (CCA)), and liver transplantation. Although the risk of developing CCA in PSC is increased up to 20%¹¹, the pathophysiological mechanism leading to CCA in PSC is still to be clarified, and will probably partly depend on the duration and intensity of the chronic inflammation, and in part to genetic predisposition of the host. It is uncertain if death from CCA lies within a biological pathway similar to disease progression in PSC, and hence death from CCA may not be heralded by liver histology at time of diagnosis. To accommodate for this uncertainty, endpoint 2: liver transplantation, was introduced, which focusses purely on disease progression. Endpoint 3 was the occurrence of liver related events at follow-up, which was defined as the occurrence of gastro-esophageal variceal bleeding, development of ascites, splenomegaly¹² or hepatic encephalopathy. Ascites and splenomegaly were assessed during yearly surveillance imaging- either abdominal ultrasonography or magnetic resonance cholangiopancreatography. Splenomegaly was defined as a spleen length of more than 120 mm.¹²

Statistical Analysis

Patient characteristics and laboratory values were expressed as mean $\pm$ standard deviation, and median and interquartile range when having a symmetrical and skewed distribution, respectively. Dichotomous variables were expressed as percentage (%) of the cohort. Reference values of biochemical variables may vary between hospitals and over time, therefore biochemical variables were expressed as ratio of upper limit of normal, or lower limit of normal.

Associations of grading and staging with the different clinical endpoints were estimated using Kaplan Meier curve and log-rank test. The potential prognostic value on predicting outcome of clinical, biochemical, and histological vari-

ables was assessed using univariable Cox proportional hazards model. Potential multicollinearity between the different staging systems, as well as Mayo risk score and its laboratory components aspartate aminotransferase, albumin, and bilirubin were checked by Spearman's rank correlation coefficient. Using stepwise multivariable Cox proportional hazards regression, the independent predictive value of the histologic scoring systems – entered as continuous variables – was evaluated, from which the hazard ratio (HR), and corresponding 95% confidence interval (CI) could be calculated. The criterion for retaining predictors was a p-value <0.05. Statistical analyses were performed using SPSS version 22.0 software (SPSS, Chicago, IL). A p-value of <0.05 was considered statistically significant.

RESULTS

Patient and biopsy characteristics

We received a total of 154 liver biopsies, of which 17 biopsies were excluded due to insufficient amount of portal tracts, poor quality of the staining or an accidental case of sending the wrong tissue biopsy. Of a total of 12 PSC cases, one or more consecutive biopsies were received, of which the first taken biopsy was included in the final cohort.

Table 1 summarized the patient characteristics. A total 119 patients (corresponding to 119 liver biopsies) were included in the final cohort, of which 81 (68%) were male, 38 (32%) female. The majority of 101 (85%) patients were diagnosed with large duct PSC, 2 (2%) small duct PSC, and 16 (13%) patients had PSC/AIH-overlap syndrome. Concomitant inflammatory bowel disease (IBD) was present in 80 (67%) patients, the majority (68 (57%)) suffered from ulcerative colitis. There was a median follow-up from PSC diagnosis until reaching an endpoint or date of last follow up, of 142 months (inter quartile range (IQR) 98-188).

Biopsies had a median of 17 (IQR 11-24) portal tracts, and a median biopsy length of 15 mm (IQR 11-20). The median disease duration at time of biopsy was 0 (range 0-38) months. Sixteen (13%) biopsies showed concentric periductal fibrosis, characteristic for PSC diagnosis.

Table 1. Patient Characteristics.

	Validation cohort	
N	119	
Male [n (%)]	81	(68)
Age at PSC diagnosis (years) [median (IQR)]	38	(27-50)
PSC type		
Large duct PSC [n (%)]	101	(85)
Small duct PSC [n (%)]	2	(2)
AIH-overlap [n (%)]	16	(13)
Inflammatory bowel disease [n (%)]	80	(67)
Ulcerative colitis [n (%)]	68	(57)
Crohn's disease [n (%)]	10	(8)
Unspecified [n (%)]	2	(2)
Portal tracts per biopsy [median (IQR)]	17	(11-24)
Biopsy length (mm) [median (IQR)]	15	(11-20)
Disease duration at time of biopsy (months) [median (range)]	0	(0-38)
Follow-up time (months) [median (IQR)]	142	(98-188)
Ursodeoxycholic acid at time of biopsy [n (%)]	48	(45)
Aminosalicylates at time of biopsy [n (%)]	34	(33)
Corticosteroid at time of biopsy [n (%)]	18	(18)
Thiopurine therapy at time of biopsy [n (%)]	9	(9)
Methotrexate therapy at time of biopsy [n (%)]	0	(0)
Anti-TNF therapy at time of biopsy [n (%)]	1	(1)
Biochemical parameters at PSC diagnosis		
AST xULN [median (IQR)]	1.64	(0.89-2.90)
ALT xULN [median (IQR)]	2.28	(1.03-4.00)
ALP xULN [median (IQR)]	2.56	(1.25-4.18)
γGT xULN [median (IQR)]	4.53	(1.98-10.75)
Total bilirubin xULN [median (IQR)]	1.00	(0.58-2.00)
Albumin xLLN [median (IQR)]	1.12	(1.06-1.30)
INR xULN [median (IQR)]	1.00	(0.83-1.00)
Platelet count xLLN [median (IQR)]	1.61	(1.03-2.06)
Creatinine xULN [median (IQR)]*	0.79	(0.62-1.00)
Mayo Risk Score [median (IQR)]	-0.02	(-0.77-0.96)

PSC = primary sclerosing cholangitis; IQR = inter quartile range; AIH = autoimmune hepatitis; mm = millimeter; AST = aspartate aminotransferase; ALT = alanine aminotransferase; ALP = alkaline phosphatase; γGT = gamma-glutamyl transferase; xULN = times upper limit of normal; xLLN= times upper limit of normal INR =International normalized ratio

* Data available for: Ursodeoxycholic acid n = 106; Aminosalicylates n = 104;Thiopurines n = 104; Corticosteroids n=103; Methotrexate n = 104; Anti-TNF n = 104; AST n= 98, ALT n= 91, ALP n=109, γGT n=70 total bilirubin n=106, albumin n=95 INR n = 65, platelet count n=101, creatinine n = 93 Mayo risk score n=.85

Patient outcome

During follow-up, 31 (26%) patients died, 20/31 (65%) of deaths were PSC related. In that subgroup, the median time from PSC diagnosis until PSC related death was 152 months (IQR 92-223). 31 (26%) patients underwent liver transplantation, which they underwent within a median time of 79 months (IQR 17-112) from PSC diagnosis. A total of 11 (9%) patients were diagnosed with CCA, for those patients the median time from diagnosis until developing CCA was 146 months (IQR 62-224); 8 (73%) patients died.

A total of 35 (29%) patients experienced one of more liver related events within a median of 85 months (IQR 14-132) after PSC diagnosis. 7 (6%) patients had a variceal bleed. Ascites occurred in 13 (11%) patients, 18 (15%) patients had an enlarged spleen, and 5 (4%) patients developed hepatic encephalopathy.

Distribution of grade and stage

Evaluation of grade and stage according to the three histologic scoring systems was reached in consensus in 100% of cases by the two in tandem scoring pathologists.

Nakanuma grading components cholangitis activity and hepatitis activity were absent in the vast majority of cases, resulting in a median grade of 0 (range 0-3). The Ishak grading component interface hepatitis was distributed with a median of 1 (range 0-3), while confluent necrosis was only present in 4 cases, resulting in a median score of 0 (range 0-6). Degree of lobular inflammation and portal inflammation as scored by Ishak showed a median of 1 (range 0-4), and 1 (range 0-3), respectively. Figures 1 A and B illustrate the total Nakanuma and Ishak grades. The distribution of Nakanuma and Ishak grading components is illustrated in Supplementary Figure 1.

All three histologic components (fibrosis, bile duct loss, and orcein positive granules) staged by Nakanuma were distributed with a median of 1 (range 0-3), resulting in a median total Nakanuma stage of 2 (range 1-4). (Figure 1C; Supplementary Figure. 1) The distribution of degree of fibrosis, as scored according Ishak and Ludwig are shown in Figure 1D, E: median 2 (range 0-6) and 2 (range 0-4), respectively.

Fig.1. Distribution of grading and staging systems.
 (A) Nakanuma grading system. (B) Ishak grading system.
 (C) Nakanuma staging system. (D) Ishak staging system. (E) Ludwig staging system.

Association between histologic grade and stage with clinical outcome.

Kaplan Meier survivorship estimates indicated a significant association of the Nakanuma stage, Ishak stage and Ludwig stage with endpoints 1 and 2. The Ishak and Ludwig staging systems were significantly associated with endpoint 3 (Figure 2A-C; supplementary Figure A-F). Table 2 shows the results of the univariable Cox proportional hazard analyses, which were used to analyze potential associations between clinical, biochemical and histological variables and outcome in PSC- as defined by the three different endpoints.

The Ishak and Ludwig staging systems were highly correlated, with a correlation coefficient of 0.93, $p < 0.001$. This multicollinearity precluded the analyses of all three staging system in one multivariable cox model. Therefore, using step-wise regression, the independence of each individual staging system was assessed by entering each staging system separately. In addition, there was a significant correlation between Mayo risk score and aspartate aminotransferase (correlation coefficient 0.665, $p < 0.001$), Mayo risk score and bilirubin (correlation coefficient 0.685, $p < 0.001$), and Mayo risk score and albumin (correlation coefficient -0.647, $p < 0.001$). Therefore, Mayo risk score was not included in multivariable regression analyses.

Fig. 2. Kaplan Meier survival curves of the Nakanuma staging system. (A) Endpoint 1; Transplant-free survival (B) Endpoint 2; Liver transplantation, (C) Endpoint 3; Liver related events.

Endpoint 1. Composite PSC related death and liver transplantation

Nakanuma grade, Nakanuma stage, Ishak stage, Ludwig stage, center of inclusion, total bilirubin, albumin, platelet count, and Mayo risk score were predictors of endpoint 1 in univariable analysis. (Table 2A). Separate multivariable analyses for the three different staging systems showed that the Nakanuma grading system was inversely associated with endpoint 1 (corrected for Ishak stage: HR 0.55 (95%CI 0.31-0.99), p=0.046; corrected for Ludwig stage: HR 0.54 (95%CI 0.30-0.98), p=0.041). Each incremental grade corresponds to a 55% and 54% decreased risk, respectively, of reaching an endpoint. By contrast, the Nakanuma staging system was positively associated with endpoint 1 (HR 2.14 (95%CI 1.22-3.77), p=0.008). (Table 3A)

Table 2A. Univariable analysis composite endpoint 1. PSC related death and liver transplantation; endpoint 2. Liver transplantation.

	Endpoint 1		Endpoint 2	
	HR (95% CI)	p-value	HR (95% CI)	p-value
Sex	0.88 (0.45-1.70)	0.698	0.68 (0.29-1.59)	0.375
Age at diagnosis PSC/10Yrs	1.16 (0.93-1.44)	0.183	1.03 (0.79-1.33)	0.834
PSC type	0.74 (0.44-1.26)	0.270	0.61 (0.29-1.27)	0.187
Co-existing IBD	0.63 (0.33-1.18)	0.148	0.63 (0.30-1.33)	0.222
Center of inclusion	1.28 (1.12-1.47)	<0.001	1.28 (1.08-1.51)	0.004
AST	0.99 (0.90-1.08)	0.825	1.01 (0.92-1.11)	0.804
ALT	0.99 (0.92-1.06)	0.813	0.99 (0.90-1.09)	0.819
ALP	1.05 (0.98-1.14)	0.176	1.07 (0.99-1.17)	0.109
γGT	1.03 (0.99-1.07)	0.144	1.04 (0.98-1.09)	0.185
Total bilirubin	1.18 (1.10-1.25)	<0.001	1.18 (1.10-1.26)	<0.001
Albumin	0.02 (0.00-0.18)	<0.001	0.00 (0.00-0.04)	<0.001
INR	11.25 (0.71-177.33)	0.085	12.37 (0.52-294.06)	0.120
Platelet count	0.39 (0.21-0.71)	0.002	0.25 (0.11-0.59)	0.002
Creatinine	1.81 (0.76-4.34)	0.183	2.74 (0.99-7.60)	0.053
Mayo risk score	1.62 (1.23-2.13)	0.001	1.83 (1.31-2.57)	<0.001
Ursodeoxycholic acid at biopsy	0.56 (0.28-1.11)	0.097	0.65 (0.27-1.61)	0.355
Aminosalicylates at biopsy	0.80 (0.38-1.67)	0.546	1.15 (0.46-2.88)	0.773
Thiopurines at biopsy	0.34 (0.05-2.53)	0.295	0.52 (0.07-3.92)	0.528
Corticosteroids at biopsy	0.91 (0.38-2.21)	0.838	0.77 (0.23-2.66)	0.682
Methotrexate at biopsy	NA		NA	
Anti-TNF at biopsy	NA		NA	
Nakanuma cholangitis activity (0-3)	0.53 (0.30-0.93)	0.028	0.33 (0.12-0.87)	0.025
Nakanuma hepatitis activity (0-3)	0.67 (0.37-1.23)	0.199	0.74 (0.38-1.43)	0.364
Nakanuma grade (0-6)	0.60 (0.39-0.90)	0.013	0.51 (0.29-0.90)	0.020
Ishak grade (0-18)	1.00 (0.88-1.13)	0.993	1.01 (0.88-1.17)	0.873
Nakanuma orcein deposition (0-3)	1.79 (1.35-2.37)	<0.001	2.12 (1.51-2.98)	<0.001
Nakanuma fibrosis (0-3)	2.34 (1.36-4.00)	0.002	3.78 (1.94-7.35)	<0.001
Nakanuma bile duct loss (0-3)	1.50 (1.06-2.14)	0.023	1.67 (1.11-2.51)	0.013
Nakanuma stage (1-4)	2.58 (1.66-4.03)	<0.001	3.38 (2.01-5.69)	<0.001
Ishak stage (0-6)	1.37 (1.13-1.65)	0.001	1.58 (1.26-1.98)	<0.001
Ludwig stage (0-4)	2.16 (1.28-3.65)	0.004	3.34 (1.74-6.41)	<0.001

PSC = primary sclerosing cholangitis; HR = Hazard Ratio; CI = Confidence Interval; IBD = inflammatory bowel disease; AST = aspartate aminotransferase; ALT = alanine aminotransferase; ALP = alkaline phosphatase; γGT = gamma-glutamyl transferase; INR = International normalized ratio; NA = not assessable.

* Data available for: Ursodeoxycholic acid n = 106; Aminosalicylates n = 104; Thiopurines n = 104; Corticosteroids n=103; Methotrexate n = 104; Anti-TNF n = 104; AST n= 98, ALT n= 91, ALP n=109, γGT n=70 total bilirubin n=106, albumin n=95 Creatinine n = 93 INR n = 65, platelet count n=101, Mayo risk score; n=.85; Nakanuma orcein deposition n=107; Nakanuma fibrosis n = 118; Nakanuma stage n = 107; Ishak stage n = 118; Ludwig stage n = 118

Table 2B. Univariable analysis endpoint 3. Liver related events.

	Endpoint 3	
	HR (95% CI)	p-value
Sex	1.57 (0.80-3.08)	0.194
Age at diagnosis PSC/10yrs	1.33 (1.04-1.70)	0.024
PSC type	0.80 (0.46-1.40)	0.432
Co-existing IBD	0.41 (0.21-0.82)	0.011
Center of inclusion	1.20 (1.02-1.40)	0.029
AST	1.01 (0.93-1.09)	0.791
ALT	0.99 (0.92-1.07)	0.866
ALP	1.05 (0.96-1.15)	0.272
γGT	1.03 (1.00-1.07)	0.087
Total bilirubin	1.12 (1.04-1.21)	0.005
Albumin	0.24 (0.03-1.71)	0.154
INR	0.15 (0.01-4.79)	0.286
Platelet count	0.65 (0.36-1.18)	0.154
Creatinine	0.96 (0.36-2.58)	0.958
Mayo risk score	1.68 (1.27-2.22)	0.000
Ursodeoxycholic acid at biopsy	1.36 (0.65-2.84)	0.409
Aminosaliclates at biopsy	0.46 (0.18-1.21)	0.117
Thiopurines at biopsy	0.38 (0.05-2.82)	0.345
Corticosteroids at biopsy	0.46 (0.14-1.53)	0.205
Methotrexate at biopsy	NA	
Anti-TNF at biopsy	NA	
Nakanuma cholangitis activity (0-3)	0.88 (0.55-1.41)	0.586
Nakanuma hepatitis activity (0-3)	0.98 (0.61-1.56)	0.915
Nakanuma grade (0-6)	0.92 (0.66-1.29)	0.640
Ishak grade (0-18)	1.01 (0.89-1.15)	0.888
Nakanuma fibrosis (0-3)	1.99 (1.09-3.65)	0.026
Nakanuma orcein deposition (0-3)	1.54 (1.12-2.12)	0.008
Nakanuma bile duct loss (0-3)	1.41(0.94-2.10)	0.094
Nakanuma stage (1-4)	1.67 (1.04-2.68)	0.035
Ishak stage (0-6)	1.37 (1.07-1.66)	0.009
Ludwig stage (0-4)	2.82 (1.02-3.27)	0.043

PSC = primary sclerosing cholangitis; HR = Hazard Ratio; CI = Confidence Interval; IBD = inflammatory bowel disease; AST = aspartate aminotransferase; ALT = alanine aminotransferase; ALP = alkaline phosphatase; γGT =gamma-glutamyl transferase; xULN = times upper limit of normal; INR = International normalized ratio; NA = not assessable

* Data available for: Ursodeoxycholic acid n = 106; Aminosaliclates n = 104; Thiopurines n = 104; Corticosteroids n=103; Methotrexate n = 104; Anti-TNF n = 104; AST n= 98, ALT n= 91, ALP n=109, γGT n=70 total bilirubin n=106, albumin n=95 Creatinine n = 93 INR n = 65, platelet count n=101, Mayo risk score; n=.85; Nakanuma orcein deposition n=107; Nakanuma fibrosis n = 118; Nakanuma stage n = 107; Ishak stage n = 118; Ludwig stage n = 118

Endpoint 2. Liver transplantation

The same predictors of endpoint 1 were shown to be associated with endpoint 2 in univariable Cox regression analysis (Table 2A). When correcting for potential interacting variables in multivariable analyses, the Nakanuma, Ishak, and Ludwig staging systems were all independently associated with liver transplantation, HR 3.16 (95% CI 1.49-6.68), $p=0.003$; HR 1.55 (95% CI 1.10-2.18), $p=0.013$; HR 2.62 (95% CI 1.19-5.80), $p=0.017$, respectively. The Nakanuma grading system was inversely associated with liver transplantation in the multivariable analysis correcting for the Ishak and Ludwig staging systems, HR 0.25 (95% CI 0.07-0.85), $p=0.027$; HR 0.25 (95% CI 0.07-0.84), $p=0.025$, respectively. (Table 3B)

Endpoint 3. Liver related events

The variables Nakanuma stage, Ishak stage and Ludwig stage, age at diagnosis PSC, co-existing IBD, center of inclusion, total bilirubin, and Mayo risk score were significant predictors for the occurrence of liver related events in univariable analysis (Table 2B). The Nakanuma, Ishak and Ludwig staging systems were all shown to be independently associated with the occurrence of liver related events in multivariable analyses: HR 2.05 (95% CI 1.17-3.57), $p=0.012$; HR 1.43 (95% CI 1.10-1.85) $p=0.007$; HR 2.06 (95% CI 1.09-3.89), $p=0.027$. (Table 3C)

Table 3A. Multivariable analysis endpoint 1. PSC related death and liver transplantation.

	Univariable		Nakanuma Stage		Ishak Stage		Ludwig Stage	
	HR (95% CI)	p-value	Multivariable HR (95% CI)	p-value	Multivariable HR (95% CI)	p-value	Multivariable HR (95% CI)	p-value
Center of inclusion	1.28 (1.12-1.47)	<0.001	1.13 (0.90-1.43)	0.287	1.11 (0.91-1.35)	0.301	1.23 (0.93-1.37)	0.221
Total bilirubin	1.18 (1.10-1.25)	<0.001	1.10 (1.00-1.22)	0.061	1.23 (1.02-1.24)	0.015	1.12 (1.02-1.23)	0.017
Albumin	0.02 (0.00-0.18)	<0.001	0.19 (0.01-3.65)	0.271	0.19 (0.02-1.91)	0.158	0.19 (0.02-1.99)	0.165
Platelet count	0.39 (0.21-0.71)	0.002	0.60 (0.33-1.08)	0.090	0.58 (0.32-1.06)	0.074	0.57 (0.31-1.02)	0.060
Nakanuma grade (0-6)	0.60 (0.39-0.90)	0.013	0.58 (0.32-1.04)	0.069	0.55 (0.31-0.99)	0.046	0.54 (0.30-0.98)	0.041
Nakanuma stage (1-4)	2.58 (1.66-4.03)	<0.001	2.14 (1.22-3.77)	0.008				
Ishak stage (0-6)	1.37 (1.13-1.65)	0.001			1.19 (0.93-1.52)	0.169		
Ludwig stage (0-4)	2.16 (1.28-3.65)	0.004					1.53 (0.86-2.70)	0.147

The independent prognostic value of the individual staging systems was assessed by stepwise regression.

Table 3B. Multivariable analysis endpoint 2. Liver transplantation.

	Univariable		Nakanuma Stage		Ishak Stage		Ludwig Stage	
	HR (95% CI)	p-value	HR (95%CI)	p-value	HR (95%CI)	p-value	HR (95%CI)	p-value
Center of inclusion	1.28 (1.08-1.51)	0.004	0.95 (0.70-1.31)	0.768	0.90 (0.67-1.21)	0.470	0.92 (0.68-1.25)	0.602
Total bilirubin	1.18 (1.10-1.26)	<0.001	1.10 (0.95-1.26)	0.207	1.13 (0.99-1.03)	0.078	1.13 (0.98-1.29)	0.088
Albumin	0.00 (0.00-0.04)	<0.001	0.02 (0.00-1.91)	0.092	0.02 (0.00-0.80)	0.037	0.02 (0.00-0.73)	0.034
Platelet count	0.25 (0.11-0.59)	0.002	0.68 (0.31-1.49)	0.335	0.71 (0.33-1.53)	0.379	0.67 (0.31-1.43)	0.298
Nakanuma grade (0-6)	0.51 (0.29-0.90)	0.020	0.33 (0.11-1.03)	0.056	0.25 (0.07-0.85)	0.027	0.25 (0.07-0.84)	0.025
Nakanuma stage (1-4)	3.38 (2.01-5.69)	<0.001	3.16 (1.49-6.68)	0.003				
Ishak stage (0-6)	1.58 (1.26-1.98)	<0.001			1.55 (1.10-2.18)	0.013		
Ludwig stage (0-4)	3.34 (1.74-6.41)	<0.001					2.62 (1.19-5.80)	0.017

The independent prognostic value of the individual staging systems was assessed by stepwise regression.

Table 3C. Multivariable analysis endpoint 3. Liver related events.

	Univariate		Nakanuma Stage		Ishak Stage		Ludwig Stage	
	HR (95% CI)	p-value	HR (95%CI)	p-value	HR (95%CI)	p-value	HR (95%CI)	p-value
Co-existing IBD	0.41 (0.21-0.82)	0.011	0.80 (0.33-1.94)	0.621	0.79 (0.35-1.80)	0.5744	0.65 (0.29-1.44)	0.288
Age at diagnosis	1.33 (1.04-1.70)	0.024	1.36 (1.01-1.84)	0.044	1.37 (1.02-1.83)	0.036	1.29 (0.97-1.71)	0.076
PSC/10yrs								
Center of inclusion	1.20 (1.02-1.40)	0.029	1.15 (0.93-1.42)	0.2019	1.02 (0.84-1.23)	0.8650	1.08 (0.90-1.29)	0.431
Total bilirubin	1.12 (1.04-1.21)	0.005	1.06 (0.97-1.15)	0.212	1.06 (0.97-1.15)	0.197	1.06 (0.97-1.15)	0.217
Nakanuma stage (1-4)	1.67 (1.04-2.68)	0.035	2.05 (1.17-3.57)	0.012				
Ishak stage (0-6)	1.37 (1.07-1.66)	0.009			1.43 (1.10-1.85)	0.007		
Ludwig stage (0-4)	2.82 (1.02-3.27)	0.043					2.06 (1.09-3.89)	0.027

The independent prognostic value of the individual staging systems was assessed by stepwise regression.

Multivariable analysis Nakanuma staging components

When analyzing the prognostic value of the individual Nakanuma staging components to predict endpoints 1, 2 and 3, orcein deposition was shown to be the strongest independent predictor: HR 1.59 (95% CI 1.16-2.18), $p=0.004$; HR 1.76 (95% CI 1.19-2.61), $p=0.005$; HR 1.40 (95% CI 0.95-2.06), $p=0.088$, respectively. The Nakanuma component liver fibrosis was an independent predictor of endpoint 2 only (HR 2.25 (95% CI 1.08-4.68), $p=0.030$), while the Nakanuma component bile duct loss was never associated with outcome. (Table 4)

Table 4. Multivariable analysis Nakanuma staging components, endpoint 1. PSC related death and liver transplantation; endpoint 2. Liver transplantation, endpoint 3 liver related events.

	Endpoint 1		Endpoint 2		Endpoint 3	
	HR (95% CI)	<i>p</i> -value	HR (95%CI)	<i>p</i> -value	HR (95%CI)	<i>p</i> -value
Orcein deposition	1.59 (1.16-2.18)	0.004	1.76 (1.19-2.61)	0.005	1.40 (0.95-2.06)	0.088
Fibrosis	1.73 (0.94-3.21)	0.081	2.25 (1.08-4.68)	0.030	1.56 (0.77-3.17)	0.222
Bile duct loss	1.05 (0.71-1.56)	0.795	1.07 (0.68-1.67)	0.785	1.02 (0.63-1.65)	0.943

DISCUSSION

This study confirms the prognostic value of staging disease progression using the Nakanuma, Ishak, and Ludwig staging systems.

We demonstrated that all three staging systems were independent predictors of liver transplantation and development of liver related events. In contrast, the Ishak and Ludwig scoring systems were not independent predictors of endpoint 1 (composite PSC related death and liver transplantation). The inclusion of death from CCA in endpoint 1— which may not be predictable by histology — could have contributed to this difference. Additional univariable analysis showed that, indeed, neither grading nor staging liver histology were of prognostic value for the development of CCA. (Supplementary Table 3.)

In accordance with the previous study on the prognostic value of histology in PSC by de Vries et al., the Nakanuma staging system was shown to be the strongest histological predictor of long-term outcome.⁷ Also in PBC the Nakanuma staging system was repeatedly shown to be superior to the classical Ludwig and Scheuer staging systems in predicting clinical outcome.^{13,14} These observations can be explained by the inclusion of three staging components that reflect histopathological progression of liver injury in vanishing bile duct diseases.¹⁴ Of the individual staging components, orcein deposition was the most valuable histologic predictor. In addition, the degree of orcein deposition was strongly correlated with alkaline phosphatase – an important prognostic biochemical marker in PSC.¹⁵ (Supplementary Table 4.) These results are in concordance with the recent findings in PBC of both Kakuda et al. and Chan et al., which demonstrated that the copper binding protein score was the most powerful histologic prognostic Nakanuma staging component.^{13,14} The deposition of orcein positive granules is a sensitive marker of chronic cholestasis in the liver¹⁶, most likely related to the inability to excrete copper associated protein in bile due to bile duct loss. The accumulation of copper-associated protein in periportal hepatocytes is frequently associated with other morphological changes related to the accumulation of toxic bile acids, such as hepatocyte ballooning and Mallory-Denk body formation (“cholate stasis”), and may thus be an important marker of disease progression in PSC.⁵ Although bile duct loss might theoretically be a more sensitive marker of disease progression in ductopenic diseases such as PSC, loss of bile ducts is often patchy in distribution and may thus be subject to problems with sampling variability in liver biopsy specimens. The absence of an association between degree of bile duct loss and clinical outcomes was therefore not surprising, and similar results were shown in a recent study by the Vries et al.⁷ Unfortunately, our study was insufficiently powered to assess the independent prognostic value of the in-

dividual Nakanuma staging components in a multivariable model, which also included clinical and biochemical parameters.

Similar to what has been demonstrated in PBC, we found no association between Nakanuma grade and liver related events.^{13,14} There was, however, an unexpected negative association between Nakanuma grade and the composite endpoint 1 and endpoint 2 - a higher Nakanuma grade being associated with a decreased chance on reaching these endpoints. This difference could not be explained by the inclusion of PSC/AIH-patients, who would be expected to have relatively high scores for the degree of hepatitis activity, since the prognostic value of Nakanuma grade was driven by the cholangitis activity component. (Table 2.) We checked if Nakanuma grade could reflect an early (inflammatory) disease stage and a corresponding better prognosis, with less fibrosis by assessing the association between cholangitis activity and degree of fibrosis, but no such association was found. (supplementary Table 5.) Overall, cholangitis activity and hepatitis activity are not prominent features in PSC, and were present in only 27% versus 14% of cases, respectively. The resulting limited sample size per grade is important to take into consideration, and the negative association between Nakanuma grade and outcome should therefore be interpreted with caution.

The proportion of PSC/AIH-overlap patients (13%) included in the present study was relatively high when compared with the 4-7% reported in large epidemiologic studies.^{11,17} PSC/AIH-overlap may be regarded as a distinct disease entity, with an altered prognosis.^{18,19} A sensitivity analysis in which PSC/AIH-overlap patients were excluded was performed to assess the potential impact of PSC/AIH-overlap on the prognostic value of the histologic scoring systems. No differences were found. (data not shown) Moreover, the independent association of all three histologic staging systems with outcome confirms the applicability of liver histology as a risk stratifier in PSC. Carbone et al. recently reported a similar independent association between degree of fibrosis and long-term outcome in PBC.²⁰ Through risk stratification (into low/high risk for poor outcome) patient subgroups that potentially have the greatest benefit from inclusion and treatment in clinical trials may be identified using their initial liver biopsy.⁴

The patchy distribution of the PSC liver carries the risk of sampling variability in PSC biopsies.^{21,22} In addition, the use of liver biopsy is hampered by its invasive nature, and interpretative inconsistencies. The correlation of histology with clinical outcomes in the present study suggests that sampling variability when it comes to staging is not a significant issue in the context of using liver biopsy findings as surrogate endpoints in clinical trials. Repeated biopsies were present in a small subset of PSC cases; n= 12, of which orcein staining was missing in 1 case. A sub analysis of the remaining 11 cases showed progression of Na-

kanuma stage in 5 cases, steady status in 5 cases, and regression in 1 case (data not shown). This result supports the ability of the Nakanuma staging system to study histologic change over time – despite the risk of sampling variability.

The development of liver fibrosis and the end stage of liver cirrhosis were long thought to be unidirectional and irreversible. However, increased understanding of the pathophysiology of hepatic fibrogenesis – especially in the field of chronic viral hepatitis and auto-immune hepatitis – has revealed that this is a dynamic process, which can be prevented or reversed by either eliminating the etiologic agent, or intervening with the pathogenic mechanism.^{23–25} Further research is necessary to assess if treatment effects could be measured by a change in histologic stage. If this would be the case, this would warrant further study as a potential surrogate endpoint in clinical trials.

Promising non-invasive modalities such as the enhanced liver fibrosis score and transient elastography may be a future substitute for the measure of fibrosis.^{26,27} However, the value of liver histology in PSC is not restricted to staging liver fibrosis, and the important additional prognostic value of staging the degree of orcein deposition cannot currently be replaced by non-invasive tools. An additional advantage of liver histology is that it measures disease progression within the affected organ itself and has the potential advantage of construct validity. Furthermore, it allows for mechanistic studies to investigate the effect of an investigational drug.^{4,28}

This study has several limitations. Firstly, since not all PSC patients routinely undergo liver biopsy, and since this was a retrospective accrual of cases, selection bias of patients cannot be excluded. Secondly, the development of liver related events has not been systematically assessed, but relied on clinician reported incidence that may yield an underestimation. Lastly, since the recently presented novel prognostic model for PSC (de Vries et al. unpublished data) – including biochemistry values at time of diagnosis – showed adequate performance in predicting prognosis in PSC, we decided to use available biochemistry values at time of diagnosis for the multivariable regression analyses. However, biochemistry results at time of diagnosis may differ from the results at time of biopsy since not all biopsies were taken at the time of diagnosis. A potential limitation of the liver histologic staging systems overall is that they may have no prognostic value for development of CCA, which is a significant cause of morbidity and death of PSC patients.

In conclusion, the histologic staging systems according to Nakanuma, Ishak, and Ludwig were confirmed to be independent predictors of long-term outcome in PSC, in this multicenter cohort. The Nakanuma staging system –incorporating features of chronic biliary disease – was shown to hold strongest predic-

tive value, and may be considered the preferred system for scoring PSC liver biopsies. Histologic scoring may be a useful tool for risk stratification. in clinical trials of PSC, as well as clinical practice. Our findings warrant testing of the robustness in clinical practice by evaluating inter-observer agreement.

REFERENCES

1. Hirschfield, G. M., Karlsen, T. H., Lindor, K. D. & Adams, D. H. Primary sclerosing cholangitis. *Lancet* 382, 1587–99 (2013).
2. European Association for the Study of the Liver. EASL Clinical Practice Guidelines: management of cholestatic liver diseases. *J. Hepatol.* 51, 237–67 (2009).
3. Chapman, R. et al. Diagnosis and management of primary sclerosing cholangitis. *AASLD Practice Guidelines. Hepatology* 51, 660–78 (2010).
4. Trivedi, P. J., Corpechot, C., Pares, A. & Hirschfield, G. M. Risk Stratification in autoimmune cholestatic liver diseases: Opportunities for clinicians and trialists. *Hepatology* 63, 644–59 (2016).
5. Nakanuma, Y., Zen, Y. & Portmann, B. Diseases of the bile ducts. In: Burt AD, Portmann BC, Ferrell LD, eds. *MacSween's pathology of the liver*. 6th ed. Edinburgh: Churchill Livingstone. (2012).
6. Portmann, B. & Zen, Y. Inflammatory disease of the bile ducts-cholangiopathies: liver biopsy challenge and clinicopathological correlation. *Histopathology* 60, 236–48 (2012).
7. De Vries, E. M. G. et al. Applicability and prognostic value of histologic scoring systems in primary sclerosing cholangitis. *J. Hepatol.* 63, 1212–1219 (2015).
8. Ishak, K. et al. Histological grading and staging of chronic hepatitis. *J. Hepatol.* 22, 696–9 (1995).
9. Nakanuma, Y., Zen, Y., Harada, K. & Sasaki, M. Application of a new histological staging and grading system for primary biliary cirrhosis to liver biopsy specimens: Interobserver agreement. *Pathol. Int.* 60, 167–174 (2010).
10. Ludwig, J., Dickson, E. & McDonald, G. Staging of chronic nonsuppurative destructive cholangitis (syndrome of primary biliary cirrhosis). *Virchows Arch A Pathol Anat Histol* 22, 103–12 (1978).
11. Boonstra, K. et al. Population-based epidemiology, malignancy risk, and outcome of primary sclerosing cholangitis. *Hepatology* 58, 2045–55 (2013).
12. Ehlken, H. et al. Spleen size for the prediction of clinical outcome in patients with primary sclerosing cholangitis. *Gut* feb 26, n/a–n/a (2016).
13. Kakuda, Y., Harada, K., Sawada-Kitamura, S., Ikeda, H. & Nakanuma, Y. Evaluation of a new histologic staging and grading system for primary biliary cirrhosis in comparison with classical systems. *Hum. Pathol.* 44, 1109–1117 (2013).
14. Chan, A. W. H. et al. Evaluation of histological staging systems for primary biliary cirrhosis: correlation with clinical and biochemical factors and significance of pathological parameters in prognostication. *Histopathology* 65, 174–86 (2014).
15. Williamson, K. & Chapman, R. Editorial: further evidence for the role of serum alkaline phosphatase as a useful surrogate marker of prognosis in PSC. *Aliment. Pharmacol. Ther.* 41, 149–151 (2015).
16. Nakanuma, Y., Karino, T. & Otha, G. Orcein positive granules in the hepatocytes in chronic intrahepatic cholestasis. *Virchows Arch A Pathol Anat Histol* 382, 21–30 (1979).
17. Tischendorf, J. J. W., Hecker, H., Krüger, M., Manns, M. P. & Meier, P. N. Characterization, outcome, and prognosis in 273 patients with primary sclerosing cholangitis: A single center study. *Am. J. Gastroenterol.* 102, 107–14 (2007).
18. Floreani, A. et al. Clinical course and outcome of autoimmune hepatitis/primary sclerosing cholangitis overlap syndrome. *Am. J. Gastroenterol.* 100, 1516–1522 (2005).

19. Lüth, S. et al. Characteristics and long-term prognosis of the autoimmune hepatitis/primary sclerosing cholangitis overlap syndrome. *J. Clin. Gastroenterol.* 43, 75–80 (2009).
20. Carbone, M. et al. P1198 : Histological stage is relevant for risk-stratification in primary biliary cirrhosis. *J. Hepatol.* 62, S805 (2015).
21. Scheuer, P. & Path, F. Pathologic Features and Evolution of Primary Biliary Cirrhosis and Primary Sclerosing Cholangitis. Ludwig Symposium on Biliary Disorders-PartII. *Mayo Clin Proc* 73, 179–183 (1998).
22. Olsson, R. et al. Sampling variability of percutaneous liver biopsy in primary sclerosing cholangitis. *J. Clin. Pathol.* 48, 933–5 (1995).
23. Sohrabpour, a. a., Mohamadnejad, M. & Malekzadeh, R. Review article: The reversibility of cirrhosis. *Aliment. Pharmacol. Ther.* 36, 824–832 (2012).
24. Czaja, A. J. Hepatic inflammation and progressive liver fibrosis in chronic liver disease. *World J. Gastroenterol.* 20, 2515–32 (2014).
25. Hammel, P. et al. Regression of liver fibrosis after biliary drainage in patients with chronic pancreatitis and stenosis of the common bile duct. *N. Engl. J. Med.* 344, 418–423 (2001).
26. Vesterhus, M. et al. Enhanced Liver Fibrosis score predicts transplant-free survival in PSC independently of the Mayo risk score. *J. Hepatol. Conf.* 49th Annu. Meet. Eur. Assoc. Study Liver, Int. Liver Congr. London United Kingdom 188 (2014).
27. Corpechot, C. 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* 146, 970–9; quiz e15–6 (2014).
28. Ponsioen, C. Y. 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 publication.*

Validation of the prognostic value of histologic scoring systems in PSC

Supplementary Figure 1. Distribution of grading and staging components.(A) Nakanuma grading component; cholangitis activity. (B) Nakanuma grading component; hepatitis activity. (C) Ishak grading component; interface hepatitis. (D) Ishak grading component; confluent necrosis. (E) Ishak grading component; lobular inflammation. (F) Ishak grading component; portal inflammation. (G) Nakanuma staging component; fibrosis. (H) Nakanuma staging component; bile duct loss. (I) Nakanuma staging component; orcein positive granules.

Supplementary Figure 2. Kaplan Meier curves. (A-C) Ishak stage, endpoint 1 (A), endpoint 2 (B), endpoint 3 (C)(D-F)Ludwig stage, endpoint 1 (D), endpoint 2 (E), endpoint 3 (F)

Supplementary Table 1. Grading according Ishak and Nakanuma.^{8,9}

Score	Criteria
<i>Ishak</i>	
Periportal or periseptal interface hepatitis (piecemeal necrosis)	
0	None
1	Mild (focal, few, portal areas)
2	Mild/moderate (focal, most portal areas)
3	Moderate (continuous around <50% of tracts or septa)
4	Severe (continuous around >50% of tracts or septa)
Confluent necrosis	
0	None
1	Focal confluent necrosis
2	Zone 3 necrosis in some areas
3	Zone 3 necrosis in most areas
4	Zone 3 necrosis + occasional portal-central bridging
5	Zone 3 necrosis + multiple portal-central bridging
6	Panacinar or multiacinar necrosis
Focal (spotty) lytic necrosis, apoptosis and focal inflammation	
0	None
1	One focus or less per 10x objective
2	Two to four foci per 10x objective
3	Five to ten foci per 10x objective
4	More than 10 foci per 10x objective
Portal inflammation	
0	None
1	Mild, some or all portal areas
2	Moderate, some or all portal areas
3	Moderate/marked, all portal areas
4	Marked, all portal areas
<i>Nakanuma</i>	
A. Cholangitis activity	
0 (no activity)	No cholangitis but mild damage to the epithelium of the duct may be present
1 (mild activity)	1 evident chronic cholangitis in the specimen
2 (moderate activity)	>/= 2 bile ducts with evident chronic cholangitis
3 (marked activity)	>/= 1 CNSDC in the specimen
B. Hepatitis activity	
0 (no activity)	No interface hepatitis and no or minimal lobular hepatitis
1 (mild activity)	Interface hepatitis affecting 10 continuous hepatocytes at limiting plate in 1 portal tract or fibrous septa and mild-to-moderate lobular hepatitis
2 (moderate activity)	Interface hepatitis affecting 10 continuous hepatocytes at limiting plate in >/=2 portal tract or fibrous septa and mild-to-moderate lobular hepatitis
3 (marked activity)	Interface hepatitis affecting 20 continuous hepatocytes at limiting plates in more than half of the portal tracts and moderate lobular hepatitis or bridging/zonal necrosis

Supplementary Table 2. Staging according Ishak, Nakanuma and Ludwig.^{8–10}

Score	Criteria
<i>Ishak</i>	
0	No fibrosis
1	Fibrous expansion of some portal areas, with or without short fibrous septa
2	Fibrous expansion of most portal areas, with or without short fibrous septa
3	Fibrous expansion of most portal areas with occasional portal to portal bridging
4	Fibrous expansion of portal areas with marked bridging (portal to portal as well as portal to central)
5	Marked bridging (portal-portal and/or portal-central) with occasional nodules (incomplete cirrhosis)
6	Cirrhosis, probable or definite
<i>Nakanuma</i>	
A. Fibrosis	
0	No portal fibrosis or fibrosis limited to portal tracts
1	Portal fibrosis with periportal fibrosis or incomplete septal fibrosis
2	Bridging fibrosis with variable lobular disarray
3	Liver cirrhosis with regenerative nodules and extensive fibrosis
B. Bile duct loss	
0	No bile duct loss
1	Bile duct loss in less than one third of portal tracts
2	Bile duct loss in one third to two thirds of portal tracts
3	Bile duct loss in more than two-thirds of portal tracts
C. Deposition of orcein positive granules	
0	No deposition of granules
1	Deposition of granules in a couple of zone 1 hepatocytes at less than one-third of portal tracts
2	Deposition of granules in a variable number of zone 1 hepatocytes at one-third to two-thirds of portal tracts
3	Deposition of granules in most zone 1 hepatocytes at more than two-thirds of portal tracts
<i>Ludwig</i>	
0	No fibrosis
1	Cholangitis or portal hepatitis
2	Periportal fibrosis or hepatitis
3	Septal fibrosis, bridging necrosis or both
4	Biliary cirrhosis

Supplementary Table 3. Univariable analysis using the endpoint diagnosis of cholangiocarcinoma

	Diagnosis cholangiocarcinoma	
	HR (95% CI)	p-value
Sex	1.61 (0.47-5.56)	0.449
Age at diagnosis PSC/10Yrs	1.16 (0.70-1.92)	0.571
PSC type	0.88 (0.30-2.59)	0.822
Co-existing IBD	1.16 (0.25-5.54)	0.848
Center of inclusion	1.20 (0.91-1.60)	0.196
AST	0.93 (0.66-1.32)	0.689
ALT	0.96 (0.80-1.17)	0.704
ALP	0.99 (0.80-1.24)	0.945
γGT	0.98 (0.87-1.11)	0.744
Total bilirubin	1.16 (0.89-1.51)	0.288
Albumin	0.65 (0.01-49.20)	0.843
INR	NA	
Platelet count	0.46 (0.13-1.61)	0.222
Creatinine	0.78 (0.09-6.99)	0.827
Mayo risk score	0.64 (0.23-1.77)	0.392
Ursodeoxycholic acid at biopsy	0.53 (0.15-1.81)	0.309
Aminosalicylates at biopsy	1.10 (0.32-3.79)	0.882
Thiopurines at biopsy	1.76 (0.21-14.81)	0.602
Corticosteroids at biopsy	3.65 (0.97-13.73)	0.055
Methotrexate at biopsy	NA	
Anti-TNF at biopsy	NA	
Nakanuma cholangitis activity (0-3)	0.54 (0.15-1.87)	0.329
Nakanuma hepatitis activity (0-3)	1.12 (0.50-2.51)	0.775
Nakanuma grade (0-6)	0.82 (0.40-1.69)	0.595
Ishak grade (0-18)	1.11 (0.91-1.36)	0.293
Ishak stage (0-6)	0.93 (0.61-1.41)	0.723
Nakanuma orcein deposition (0-3)	0.68 (0.34-1.36)	0.271
Nakanuma fibrosis (0-3)	1.00 (0.32-3.19)	0.995
Nakanuma bile duct loss (0-3)	0.53 (0.21-1.33)	0.175
Nakanuma stage (1-4)	0.44 (0.13-1.51)	0.193
Ludwig stage (0-4)	1.08 (0.37-3.14)	0.885

PSC = primary sclerosing cholangitis; HR = Hazard Ratio; CI = Confidence Interval; IBD = inflammatory bowel disease; AST = aspartate aminotransferase; ALT = alanine aminotransferase; ALP = alkaline phosphatase; γGT = gamma-glutamyl transferase; INR = International normalized ratio; NA = not assessable

* Data available for: Ursodeoxycholic acid n = 106; Aminosalicylates n = 104; Thiopurines n = 104; Corticosteroids n=103; Methotrexate n = 104; Anti-TNF n = 104; AST n= 98, ALT n= 91, ALP n=109, γGT n=70 total bilirubin n=106 , albumin n=95 INR n = 65, creatinine n = 93; platelet count n=101, Mayo risk score; n=.85; Nakanuma orcein deposition n=107; Nakanuma fibrosis n = 118; Nakanuma stage n = 107; Ishak stage = 118; Ludwig stage n= 118

Supplementary Table 4. Correlation between histologic stage and liver biochemistry

	Fibrosis (0-3)	Bile duct loss (0-3)	Orcein positive granules (0-3)	Nakanuma total (1-4)	Ishak (0-6)	Ludwig (0-4)
AST	0.259 p=0.010	0.309 p=0.002	0.368 p=0.001	0.373 p=0.000	0.336 p=0.001	0.240 p=0.018
ALT	0.178 p=0.092	0.240 p=0.022	0.265 p=0.018	0.357 p=0.001	0.232 p=0.027	0.158 p=0.134
ALP	0.214 p=0.026	0.357 p=0.000	0.390 p=0.000	0.420 p=0.000	0.261 p=0.006	0.203 p=0.05
γGT	0.133 p=0.272	0.523 p=0.000	0.419 p=0.001	0.427 p=0.001	0.128 p=0.292	0.113 p=0.351
Total bilirubin	0.242 p=0.013	0.186 p=0.056	0.241 p=0.020	0.238 p=0.021	0.314 p=0.001	0.235 p=0.016
INR	0.103 p=0.420	-0.354 p=0.004	-0.202 p=0.136	-0.203 p=0.134	0.107 p=0.400	0.116 p=0.362

AST = aspartate aminotransferase; ALT = alanine aminotransferase; ALP = alkaline phosphatase; γGT = gamma-glutamyl transferase; INR = International normalized ratio

* Data available for: AST n= 98; ALT n= 91; ALP n=109; γGT n=70; total bilirubin n=106; INR n = 65; Nakanuma orcein deposition n=107; Nakanuma fibrosis n = 118; Nakanuma stage n = 107; Ishak stage = 118; Ludwig stage n= 118

Supplementary Table 5. Association between Nakanuma grade and fibrosis

Grading system	Staging system	OR (95% CI) *	p-value *
Nakanuma grade (0-1) 0= 0; 1= >0	Nakanuma fibrosis stage (0-3)	1.472 (0.82-2.64)	0.196
Nakanuma grade (0-1) 0= 0; 1= >0	Ishak stage (0-6)	1.11 (0.87-1.40)	0.408
Nakanuma grade (0-1) 0= 0; 1= >0	Ludwig stage (0-4)	1.15 (0.87-2.64)	0.143

* Calculated by Univariable logistic regression

A novel prognostic model for transplant-free survival in primary sclerosing cholangitis

Elisabeth M.G. de Vries*, Junfeng Wang*, Kate D. Williamson, Mariska M.G. Leeflang, Kirsten Boonstra, Rinse. K. Weersma, Ulrich H. Beuers, Roger W. Chapman, Ronald B. Geskus#, Cyriel Y. Ponsioen#

* These authors contributed equally to this work

Shared senior authors

Journal of Hepatology 2015;63:1212-9

ABSTRACT

Introduction

Most prognostic models for primary sclerosing cholangitis (PSC) are based on patients referred to tertiary care and may not be applicable for the majority of PSC patients. The aim of this study was to construct and externally validate a novel, broadly applicable prognostic model for transplant-free survival of PSC, based on a large, predominantly population-based cohort using readily available variables.

Methods

The derivation cohort consisted of 692 PSC patients from the Netherlands, the validation cohort of 264 PSC patients from the UK. Retrospectively, clinical and biochemical variables were collected. We derived the prognostic index (PI) from a multivariable Cox regression model in which predictors were selected and parameters were estimated using the least absolute shrinkage and selection operator. The composite endpoint of PSC-related death and liver transplantation was used. To validate the model, we calculated the c-statistic as discrimination index, and compared predicted curves with Kaplan-Meier estimates to establish calibration accuracy.

Results

The final model included the variables: PSC subtype, age at PSC diagnosis, albumin, platelets, aspartate aminotransferase, alkaline phosphatase and bilirubin. The model yielded a C-statistic of 0.706 (95%CI 0.617-0.796), and 0.679 (95%CI 0.508-0.849), in the derivation and validation cohorts, respectively. Calibration was satisfactory. The discriminative power of the PI remained stable when prediction was based on biochemical variables collected at follow-up.

Conclusion

This novel prognostic PSC model showed adequate performance in estimating PSC-related death and/or liver transplant in a predominantly population based setting. The transplant-free survival probability can be recalculated contemporaneously when updated biochemical values are available.

INTRODUCTION

Primary sclerosing cholangitis (PSC) is an insidious, progressive cholestatic liver disease in which inflammation of the intra- and extra hepatic bile ducts leads to sclerosis, obstruction and destruction of the biliary tract. This results in chronic cholestasis, biliary fibrosis and eventually liver cirrhosis. PSC affects predominantly men (male to female ratio 2:1) at a mean age of 40 years, and up to 70% of patients suffer from concomitant inflammatory bowel disease (IBD).¹ Disease course is highly variable, with a reported median transplant-free survival from diagnosis onwards, ranging from 13 years in patients seen at tertiary referral centres to 21 years in a population-based cohort.¹ Various drugs have been studied in PSC, but none have been shown to be efficacious in halting disease progression.² For patients suffering from end-stage liver disease or severe complications of cholestasis, the only curative option is liver transplantation.³

In the absence of medical treatment options for PSC, estimating prognosis is important to improve patient counselling and management, for instance with regard to optimal patient selection and timing of listing for liver transplantation (LTx). The latter is one of the most expensive treatments currently available, hence accurate prediction of future expenditure is of great importance to health care officials. Furthermore, an accurate prognostic model may serve as a tool for risk stratification in clinical trials and its prognostic index (PI) may even be explored as a candidate surrogate endpoint.

Several prognostic models combining markers for disease progression in PSC patients have been developed (Table 1).^{4–11} Most of these models were exclusively based on liver transplant referral centre cohorts, and their predictive value may be restricted to specific patient groups. In addition, definition of endpoints differed between studies, and only two of these models were externally validated. (Table 1).^{8,12}

At present, the most widely used prognostic model in PSC is the Mayo risk score, based on the variables age, bilirubin, aspartate aminotransferase (AST), variceal bleeding and albumin.⁸ The time of origin in this study was set at date of referral instead of date at diagnosis, and this model's prediction is limited to four years, and performs best in end-stage PSC patients.⁸ In addition, the model only estimates time to (all cause) death, and not time to LTx.⁸ Death was projected for those patients that underwent liver transplantation, by making an assumption of how long they would have lived had they not undergone liver transplantation. Lastly, 25% of patients were recruited from the placebo and treatment arm of an ursodeoxycholic acid trial, which was justified by virtue of the negative outcome of the trial.⁸ However, ursodeoxycholic acid has an inherent effect on liver biochemistry, hence it may have influenced the prognostic value of the variables in part of the derivation patients.^{13–15}

The aim of this study was to construct and externally validate a novel and broadly applicable prognostic model for transplant-free survival of PSC, consisting of readily available disease characteristics and biochemical variables, based on a large, predominantly population-based cohort.

Table 1. Previous prognostic models for primary sclerosing cholangitis

First Auteur	Year	Patient cohort: FU, survival time	Endpoint	Variables included in -	
				Age	Histology
Wiesner et al.	1989	N=174. Mayo clinic's cohort,: referral center US FU: mean 6yr (range 2.7-15.5) Survival: median 11.9 yrs	-Death, all cause - LTx	X	X
Farrant et al.	1991	N=126. Kings College; referral center UK FU: median 5.8 yrs, Survival: medium 12 yrs	-Death, from liver related disease -LTx	X	X
Dickson et al.	1992	N=426. Five medical referral centers UK and US FU: median 3 yrs (range, 0.01-16.6 years) Survival: not described	Death, all cause	X	X
Broome et al.	1996	N=305. Nine 9 Swedish centers, all served as primary and referral centers FU: median 63 months (range 1-194) Survival: medium 12 yrs	-Death, from liver related disease -LTx	X	X
Kim et al.	2000	N=529. Patients included in Dickson trial + patient from UDCA study Mayo, US. N=405 model building, N=124 validation (patients from Kings College) FU: median 3 3 yrs (range 0.01-25,1) Survival: not described	-Death, all cause	X	
Boberg et al.	2002	N=330. Five European centers, all served as primary and referral centers FU: median 8.4 yrs Survival: median 11.7 yrs	-Death, from PSC -LTx	X	
Ponsioen et al.	2002	N=130. Four Dutch centers,1 referral, 3 primary centers FU: median 4 yrs (0.2-18.4) Survival: median 18 yrs	-Death, from liver disease -LTx	X	
	2010	N= 111. One center Olso, referral center FU: median 6.3 yrs (range 0.1-26.2)	-Death, from liver disease -LTx		
Tischendorf et al.	2007	N=273.medical school of Hannover, Germany, referral center. FU: median 6.3 yrs (range 0.1-23.3) Survival: median 9.3 yrs	- Death, all cause - LTx	X	

- the prognostic model									
<i>Hepatomegaly</i>	<i>Splenomegaly</i>	<i>Variceal bleeding</i>	<i>IBD</i>	<i>ALP</i>	<i>Bilirubin</i>	<i>AST</i>	<i>Albumin</i>	<i>Hb</i>	<i>Intra-extrahep cholangiogram</i>
			X		X			X	
X	X			X					
	X				X				
					X				
		X			X	X	X		
					X		X		
									X
									X
X	X				X		X		X

METHODS

Study design and patients

Derivation cohort

The derivation cohort consisted of all patients seen in 44 hospitals that were located in a geographically defined area of 6 adjacent provinces, comprising 50% of the Dutch population (2007: 7,758,980 inhabitants) (40 basic care centres and 4 academic centres without transplant facilities) and that were alive at January 2000.

Additionally we included 44 patients from a referral center for liver transplantation outside the geographically defined area. This way we provided a derivation cohort with a case-mix that is as representative as possible for most PSC patient series.

The protocol was approved by the central Committee for Research Ethics in Utrecht and all local ethics committees of the participating hospitals in the Netherlands (trialregister.nl number, NTR2813).

Validation cohort

The validation cohort included all PSC patients that visited the John Radcliff hospital, UK, from 1981 onwards. During time of patient inclusion, this centre served as non-transplant tertiary care liver centre. Patients that did not have follow-up in this hospital after PSC diagnosis were excluded from the validation cohort.

PSC diagnosis was established according to the European Association for the Study of the Liver guidelines.¹⁶ Both small and large duct patients were included. A diagnosis of autoimmune hepatitis (AIH) overlap syndrome (PSC-AIH) was made in patients with a characteristic cholangiogram who, in addition, met the simplified AIH criteria.¹⁷ Inflammatory bowel disease (IBD) diagnosis was based on the Lennard-Jones criteria.¹⁸

Data collection

Data of clinical and biochemical variables were retrospectively retrieved from patient records. Biochemical parameters that were collected at time of diagnosis (+/- 3 months) included: AST, alanine aminotransferase (ALT), alkaline phosphatase (ALP), gamma-glutamyl transpeptidase (γ GT), total bilirubin, albumin and platelets. Follow-up liver biochemistry values were collected for AST, ALT, ALP and total bilirubin during the first 3 years after diagnosis in patients included in the derivation cohort. For patients included in the validation cohort, follow-up liver biochemistry values were collected for all biochemistry variables during the entire length of follow-up.

Statistical analysis

Patient characteristics were summarized either as median and interquartile range (IQR), or in percentages. As assays may vary between hospitals and over time, biochemical variables were expressed as ratio of upper limit of normal (xULN) or lower limit of normal (xLLN). Since these ratios can only have positive values and were positively skewed, we applied a base 10 logarithmic transformation to all biochemical variables. Missing values were imputed with multivariate imputation by chained equations (MICE)¹⁹, and 20 imputed datasets were created. (Methods described in supplementary appendix).

PSC diagnosis, defined by the first pathological magnetic resonance cholangiography, endoscopic retrograde cholangiography or liver biopsy, was used as start of follow-up. A composite clinical endpoint was defined, composed of LTx or PSC related death (death from end-stage liver failure, death from liver surgery, death from cholangiocarcinoma (CCA) or death from colorectal carcinoma). Data on liver transplantation and death were retrieved from the cohort database.¹ In case no endpoint was reached, individuals were censored at the date of last follow-up at the outpatient clinic, or - for the derivation cohort - end of the study data collection (January 2012), whichever came first.

All clinical and biochemical variables that were regarded potentially relevant by expert opinion were assessed as potential predictors. This included: age at PSC diagnosis, PSC subtype, sex, IBD type, AIH-overlap syndrome, ursodeoxycholic acid use, AST, ALT, ALP, bilirubin, γ GT, albumin, platelets.

First, the functional form of the relation between the biochemical variables and the risk of the composite endpoint was investigated in univariate Cox models by means of restricted cubic splines.²⁰ Additional variable transformation was performed if strong non-linear effects were shown.

Next, a multivariable Cox regression model was fitted via the least absolute shrinkage and selection operator (Lasso). This penalized likelihood approach creates a model in which several variables are set to zero, whereas others are shrunk to zero to avoid overfitting.²¹ (Methods described in supplementary appendix). Since Lasso also served for variable selection purpose, the parameters may shrink too much. For this reason an additional adjustment factor was needed to compensate for over-shrinkage.

Based on the fitted model, we combined the parameter estimates and values of the variables to calculate a 'raw' prognostic index (PI). The 'raw' PI was then used as a single predictor to re-fit the Cox model, and the coefficient of this Cox model served as adjustment factor. The final PI was obtained by multiplying the 'raw' PI by the adjustment factor. The PI only reflects relative hazards, but does not give absolute risks, which is more informative in a clinical setting and is

needed for calibration. The transplant-free survival probability of individual patients can be estimated at any follow-up time if we have an estimate of the baseline “survival”.²² (Methods described in supplementary appendix).

Validation of the model was based on the discriminative power and on calibration accuracy. The discriminative power is the ability of the model to distinguish high-risk patients from low-risk patients. This was assessed in both the derivation and validation cohort via Harrell’s C-statistic.²³ The C-statistic was calculated in each of the 20 imputed datasets, and then averaged with Rubin’s rule.^{24,25} For assessing calibration accuracy, we divided patients into four risk groups, based on their PI’s, using threshold points at 16th, 50th and 84th percentiles in derivation data.²² Next, we compared the averaged predicted survival probability with the observed Kaplan-Meier survival probability in the four risk groups. For validation cohort, we recalibrated the baseline survival before estimating the probability of survival for each individual.

We also assessed whether the discriminative power of the model remained stable over time, by calculating the C-statistic using updated data at 1, 2 and 3 years after diagnosis.

Statistical analyses were performed using R version 3.1.2²⁶, and packages MICE, rms, glmnet and ggplot2.

RESULTS

Patient characteristics

The derivation cohort included 692 patients of which 447 (65%) were male, and 630 (91%) were diagnosed with large duct PSC. The median age at PSC diagnosis was 37 years (IQR 27-49). The median follow-up time was 110 months (IQR 69-184). (Table 2)

The validation cohort consisted of 264 patients, of which 165 (63%) were male. A total of 231 (88%) patients had large duct PSC and the median age at PSC diagnosis was 45 (IQR 33-57) years. The median follow-up time was 103 months (IQR 53-153). (Table 2)

The transplant-free survival probability of patients included in the derivation cohort, was slightly lower compared to the validation cohort. (Figure 1)

Prognostic model development

Variable transformation based on univariable analyses

The relationship between platelets and outcome followed a U-shaped pattern, with a turning point around 3.16xLLN. Therefore, the values for this variable were transformed by taking the absolute distance between log(platelets) (in unit xLLN), and the turning point $\log(3.16) \approx 0.5$. No strong non-linear effect was observed in other biochemical variables.

Table 2. Patient Characteristics.

	Derivation cohort (n=692)	Missing values [n (%)]	Validation cohort (n=264)	Missing values [n (%)]
Male [n (%)]	447 (65)		165 (63)	
Large duct PSC [n (%)]	630 (91)		231 (88)	
Age at diagnosis PSC (years) [median (IQR)]	37 (27-49)		45 (33-57)	
AIH overlap [n (%)]	37 (5)		5 (2)	
Inflammatory bowel disease [n (%)]	480 (70)		195 (74)	
Ulcerative colitis [n (%)]	373 (54)		143 (54)	
Crohn's disease [n (%)]	89 (13)		39 (15)	
Unspecified [n (%)]	18 (3)		13 (5)	
Ursodeoxycholic acid [n (%)]*	326 (80)	281 (41)		
Follow up time (months) [medi- an (IQR)]	110 (69-184)		103 (53-153)	
PSC related death [n (%)]	71 (10)		37 (14)	
Liver transplantation [n (%)]	121 (18)		18 (7)	
Alkaline phosphatase (xULN) [median (IQR)]	1,97 (1.43-3.28)	489 (71)	2.27 (1.27-3.83)	177 (67)
Aspartate aminotransferase (xULN) [median (IQR)]	1.55 (1.0-3.03)	490 (71)	1.21 (0.79-2.05)	198 (75)
Alanine aminotransferase (xULN) [median (IQR)]	2.29 (1.34-4.66)	496 (72)	1.40 (0.88-2.58)	238 (90)
Bilirubin (xULN) [median (IQR)]	0.82 (0.53-1.71)	503 (73)	0.71 (0.47-1.29)	178 (67)
Gamma-glutamyl transpeptidase (xULN) (median (IQR))	6.11 (3.36-11.88)	513 (74)	7.15 (1,98- 14.30)	210 (80)
Albumin (xLLN) [median (IQR)]	1.14 (1.02-1.23)	549 (79)	1.23 (1.11-1.31)	182 (69)
Platelets (xLLN) [median (IQR)]	1.79 (1.36-2.29)	547 (79)	1.89 (1.53-2.75)	189 (71)

Missing values: Missing laboratory value available at time of diagnosis (+/- 3 months)

Laboratory values taken at time of diagnosis (+/- 3 months)

Multivariable analysis and model selection

The model with the highest C-statistic included a total of 13 variables. A model including 13 variables would not be easy to use in clinical practice. When allowing for a performance of 10% below the optimal C-statistic, seven variables remained in the model: PSC subtype, age at PSC diagnosis, ALP, AST, bilirubin, platelets and albumin. (Supplementary Figure 1)

Figure 1. Kaplan-Meier Curves for derivation and validation cohorts.
Solid line: derivation cohort (NL); Dashed line: validation cohort (UK).

Calculation Prognostic index (PI)

The PI was calculated as 1.890 times the ‘raw’ PI which is a linear combination of selected variables:

$$PI = 0.323 \times \text{PSC subtype} + 0.018 \times \text{Age at diagnosis} - 2.485 \times \text{Albumin} + 2.451 \times \text{Platelets} + 0.347 \times \text{AST} + 0.393 \times \text{ALP} + 0.337 \times \text{Total Bilirubin}$$
 (Formula 1)

- PSC subtype: large duct PSC = 1; small duct PSC = 0.
- AST, ALP and Total Bilirubin are expressed in xULN, transformed to (log)
- Albumin is expressed in xLLN, transformed to (log)
- Platelets are expressed in xLLN, transformed to abs(log-0.5)

The distribution of the final PI is plotted in Figure 2. In Figure 3 the relation between the PI and 5-year, 10-year and 15-year survival probabilities is illustrated. To calculate the survival probability for an individual patient at any year after diagnosis, the baseline survival probabilities are provided in Supplementary Table 1. An online calculator will be provided for this calculation: www.epipscpbc.nl/prognosis.

Table 3. Thresholds and corresponding proportion of patients included in risk groups 1-4 .

Risk Group	Range in PI	Proportion	
		Derivation	Validation
1. Low risk	<1.032	16.0%	16.7%
2. Low- Intermediate risk	1.032- 1.578	34.0%	28.8%
3. Moderate risk	1.578- 2.266	34.0%	33.5%
4. High risk	>=2.266	16.0%	21.0%

Figure 2. Distribution of prognostic index in the derivation and validation cohorts. Upper part: derivation cohort; lower part: validation cohort.

Figure 3. Prediction of 5, 10 and 15 years survival probability versus final prognostic index. Solid line: 5 years; Dashed line: 10 years; Dotted line: 15 years.

Model performance in derivation cohort

The discriminative power of the model in the derivation cohort yielded a C-statistic of 0.706 (95% CI, 0.622, 0.791).

The Kaplan Meier survival curves of the four risk groups that we defined were well separated, confirming the discriminative power of the PI. (Table 3 and Figure 4A) The mean predicted survival curves were comparable with the observed Kaplan-Meier curves per risk group, which indicated good calibration. (Figure 4A).

Figure 4. Predicted versus observed survival probability per risk group (1-4) (A) Derivation cohort (B) Validation cohort (baseline hazard recalibrated). Solid line: Observed Kaplan-Meier Curve; Dashed line: Mean predicted survival curves.

Model performance in validation cohort

The PI of each individual included in the validation cohort was calculated using formula 1. The distribution of the final PI in the validation cohort is plotted in Figure 2.

The discriminative power of the model in the validation cohort was 0.663 (95% CI, 0.604, 0.722). This is slightly lower than in the derivation cohort.

To assess the calibration accuracy of the validation cohort, we classified the patients into the same four risk groups as used in the derivation cohort. The proportion of patients included in each risk group is shown in Table 3. The Kaplan-Meier survival curves of the risk groups were well separated, confirming the high discriminative power of the PI in the validation cohort.(Figure 4B)

Performance of the model over time

The C-statistics at 1, 2 and 3 years follow-up, were slightly lower than the C-statistic based on data at time of diagnosis in the derivation cohort, but were very comparable in the validation cohort. (Supplementary table 2)

A patient example of how the model can be used in clinical practice can be found in the supplementary appendix.

DISCUSSION

This study provides a novel prognostic model for transplant-free survival of PSC patients, based on an unique, well phenotyped, predominantly population-based PSC cohort. The model is based on seven objectively measured, and readily available variables: PSC subtype, age at PSC diagnosis, ALP, AST, total bilirubin, albumin, and platelets. Validation of the model in an external PSC cohort showed its robustness and adequate performance.

The usefulness and applicability of previously composed prognostic models has been hampered by the use of tertiary referral only cohorts, and variation in definitions of the time of origin in the studies (date of PSC diagnosis versus date of referral), PSC diagnosis, and clinical endpoints (Table 1). By using a predominantly population-based cohort as the derivation cohort, the present study largely mitigated this limitation. In addition, established definitions for PSC diagnosis and clinical endpoints were used, creating a prognostic model that represents a heterogeneous population of PSC patients, and is broadly applicable.¹⁶

The optimal way to validate a prognostic model is to assess its performance in an independent, but similar patient cohort.²² We used a PSC patient cohort from the United Kingdom. A notable difference in frequency of LTx between the derivation and validation cohort was observed. This may in part have resulted from the non-population based 44 patients included in the derivation cohort, that may have led to a relative over-representation of more severely diseased PSC patients. Thus the model may overestimate the individual's risk of reaching the end point when applied to patients in other settings. However, including a case mix of both population based and referral cohort PSC patients in the derivation cohort, we believe gives the optimal representation of most PSC patient series. The biochemistry parameters that are included in the model will generally change throughout the disease course. Early biochemical markers include an elevated serum ALP- the hallmark of PSC-, and serum AST and ALT levels that are often elevated up to 2 to 3 xULN at time of diagnosis.¹⁶ In contrast, serum total bilirubin is often normal at PSC diagnosis, and will increase in case of advanced disease stage, a dominant stricture or suppurative cholangitis. Serum albumin will only decrease once hepatic synthetic function is affected in advanced disease. The same holds true for serum platelets, which will only deviate from normal in advanced disease stage. The liver biochemistry variables implemented in the model are measured at every routine patient check, thus are frequently updated, objective, and readily available. This greatly benefits the applicability of the present model in clinical practice.

A recent study by our group aiming to assess the prognostic value of ALP, demonstrated that ALP levels one year after diagnosis (T1) had a better prognostic value compared with ALP values at diagnosis, or percentage change between

diagnosis and T1.²⁷ Given this result, we also made an attempt to develop another model based on T1 data, following the same modelling strategy. This T1 model included 2 more variables, and yielded a lower C-statistic of 0.692 (0.675 after adjustment for optimism).

To assess if the current model – based on data at time of diagnosis- was also applicable at later time points, additional C-statistics were calculated when using laboratory values at a follow-up of 1, 2 and 3 years. Similar C-statistics were found, indicating that the model has good performance also when it is used for prediction at the first years after diagnosis.

The calibration accuracy in the validation cohort was acceptable after recalibrating the baseline survival probability, represented by the difference in mean predicted transplant-free survival curves, when compared with the observed Kaplan-Meier curves per risk group. This indicates that recalibration should be considered when the model is applied to a different cohort. A prognostic model developed based on patients from one country may not always be valid for patients from other parts of the world.²⁸ With this study, we have reported all the information that is necessary for further recalibration, or revision of the current model, to fit a specific external patient cohort.

Current epidemiologic data on the natural history of PSC – including our Epi PSC PBC cohort- are predominantly based on Western, mostly Caucasian populations, and cohorts including children are lacking. However, there seems to be a geographical variation in incidence and prevalence rates, with lower rates in populations from Southern European and Asian descent.²⁹ Unfortunately, true population-based studies are scarce, and none have been performed in Asia and Africa.²⁹ Factors that are suggested to play a role in the variable global distribution of PSC, are differences in frequency of IBD, and in HLA susceptibility between ethnic populations.^{30–33} To be able to confirm the applicability of the present prognostic model in all populations, a better understanding of possible differences in incidence, prevalence and natural history between various ethnical and racial populations, and age groups is warranted.

Because variceal bleeding was not recorded for most patients, we were not able to compare the performance of the present model with the Mayo risk score. Contrary to the Mayo risk score, which has a horizon of only 4 years, the present model includes a more than 3 times longer follow-up time, and was based on patients retrieved from a predominantly population-based cohort. Hence, we believe our model has a broader applicability as compared with the Mayo risk score. Retrospective data collection always has the inherent drawback of incomplete data collection. Consequently, biochemical values were not available for all patients at all time-points. However, using multiple imputation, we could study

the prognostic value of these biochemical values using data from all individuals, thereby improving the validity of the results.³⁴

In conclusion, with this novel prognostic model for PSC consisting of seven clinical and biochemical variables, long-term transplant-free survival probabilities of patients with PSC can accurately be predicted. It may prove a useful tool in patient counseling, health care budget planning, as well as for risk stratification in clinical trials.

REFERENCES

1. Boonstra, K. et al. Population-based epidemiology, malignancy risk, and outcome of primary sclerosing cholangitis. *Hepatology* 58, 2045–55 (2013).
2. Karlsen, T. H., Vesterhus, M. & Boberg, K. M. Review article: controversies in the management of primary biliary cirrhosis and primary sclerosing cholangitis. *Aliment. Pharmacol. Ther.* 39, 282–301 (2014).
3. Karlsen, T. H. & Boberg, K. M. Update on primary sclerosing cholangitis. *J. Hepatol.* 59, 571–82 (2013).
4. Wiesner, R. et al. Primary sclerosing cholangitis: natural history, prognostic factors and survival analysis. *Hepatology* 10, 430–6 (1989).
5. Farrant, J. et al. Natural history and prognostic variables in primary sclerosing cholangitis. *Gastroenterology* 100, 1710–7 (1991).
6. Dickson, E. et al. Primary sclerosing cholangitis: refinement and validation of survival models. *Gastroenterology* 103, 1893–901 (1992).
7. Broomé, U. et al. Natural history and prognostic factors in 305 Swedish patients with primary sclerosing cholangitis. *Gut* 38, 610–615 (1996).
8. Kim, W. et al. A revised natural history model for primary sclerosing cholangitis. *Mayo Clin Proc.* 75, 688–94 (2000).
9. Boberg, K. M. et al. Time-dependent Cox regression model is superior in prediction of prognosis in primary sclerosing cholangitis. *Hepatology* 35, 652–7 (2002).
10. Ponsioen, C. Y. et al. Natural history of primary sclerosing cholangitis and prognostic value of cholangiography in a Dutch population. *Gut* 51, 562–6 (2002).
11. Tischendorf, J. J. W., Hecker, H., Krüger, M., Manns, M. P. & Meier, P. N. Characterization, outcome, and prognosis in 273 patients with primary sclerosing cholangitis: A single center study. *Am. J. Gastroenterol.* 102, 107–14 (2007).
12. Ponsioen, C. Y., Reitsma, J. B., Boberg, K. M., Aabakken, L. & Schrumpf, E. Validation of a cholangiographic prognostic model in primary sclerosing cholangitis. *Endoscopy* 742–747 (2010).
13. Beuers, U. et al. Ursodeoxycholic acid for treatment of primary sclerosing cholangitis: a placebo-controlled trial. *Hepatology* 1992 Sep;1, 707–14 (1992).
14. Lindor, K. D. Ursodiol for primary sclerosing cholangitis. *NEJM* 336, 691–695 (1997).
15. Mitchell, S. a. et al. A preliminary trial of high-dose ursodeoxycholic acid in primary sclerosing cholangitis. *Gastroenterology* 121, 900–907 (2001).
16. European Association for the Study of the Liver. EASL Clinical Practice Guidelines: management of cholestatic liver diseases. *J. Hepatol.* 51, 237–67 (2009).
17. Hennes, E. M. et al. Simplified criteria for the diagnosis of autoimmune hepatitis. *Hepatology* 48, 169–76 (2008).
18. Lennard-Jones, J. Classification of inflammatory bowel disease. *Scand J Gastroenterol* 170, 16–9 (1989).
19. Van Buuren, S. & Groothuis-Oudshoorn, K. Multivariate Imputation by Chained Equations. *J. Stat. Softw.* 45, 1–67 (2011).
20. Harrell, F. E., J. Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. (2001).
21. Simon, N., Friedman, J., Hastie, T. & Tibshirani, R. Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent. *J. Stat. Softw.* 39, 1–13 (2011).

22. Royston, P. & Altman, D. G. External validation of a Cox prognostic model: principles and methods. *BMC Med. Res. Methodol.* 13, 33 (2013).
23. Harrell, F. J., Lee, K. & Mark, D. Multi-variable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat. Med.* 15, 361–87 (1996).
24. Rubin, D. *Multiple Imputation for Non-response in Surveys.* (John Wiley and Sons, 2004).
25. Marshall, a, Altman, D. G., Holder, R. L. & Royston, P. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. *BMC Med. Res. Methodol.* 9, 57 (2009).
26. R Core Team (2014). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
27. de Vries, E. M. et al. Alkaline Phosphatase at Diagnosis of Primary Sclerosing Cholangitis and One Year Later: Evaluation of Prognostic Value. *Liver Int.* n/a–n/a (2016).
28. Van Houwelingen, H. C. Validation, calibration, revision and combination of prognostic survival models. *Stat. Med.* 19, 3401–3415 (2000).
29. Boonstra, K., Beuers, U. & Ponsioen, C. Y. Epidemiology of primary sclerosing cholangitis and primary biliary cirrhosis: a systematic review. *J. Hepatol.* 56, 1181–8 (2012).
30. Economou, M. & Pappas, G. New global map of Crohn's disease: Genetic, environmental, and socioeconomic correlations. *Inflamm. Bowel Dis.* 14, 709–720 (2008).
31. Neri, T. et al. Novel association of HLA-haplotypes with primary sclerosing cholangitis (PSC) in a southern European population. *Dig Liver Dis* 35, 571–6 (2003).
32. Spurkland, a et al. HLA class II haplotypes in primary sclerosing cholangitis patients from five European populations. *Tissue Antigens* 53, 459–469 (1999).
33. Bittencourt, P. . et al. Susceptibility to primary sclerosing cholangitis in Brazil is associated with HLA-DRB1*13 but not with tumour necrosis factor α -308 promoter polymorphism. *Gut* 51, 609–616 (2002).
34. Sterne, J. a C. et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *Bmj* 338, b2393–b2393 (2009).

Supplementary Table 1. Baseline survival probability of the prognostic model.

Year	Baseline Survival	Year	Baseline Survival
0	1	13	0.718900081
1	0.976003276	14	0.695637963
2	0.958190391	15	0.674046812
3	0.944872387	16	0.649928168
4	0.921801174	17	0.632527844
5	0.911738594	18	0.626239076
6	0.886368701	19	0.613255379
7	0.870414839	20	0.599392599
8	0.844120125	21	0.564462416
9	0.812447482	22	0.527004006
10	0.788752707	23	0.503629598
11	0.781935324	24	0.456791571
12	0.745666457	25	0.425503951

Supplementary Table 2. C-statistics of PI calculated at n years of follow-up.

Derivation Cohort			Validation Cohort	
Follow-up	C-statistic	95% CI	C-statistic	95% CI
1 year	0.681	(0.588, 0.775)	0.664	(0.604, 0.724)
2 years	0.686	(0.590, 0.781)	0.657	(0.594, 0.721)
3 years	0.665	(0.564, 0.767)	0.664	(0.598, 0.729)

Supplementary Figure 1. Model selection procedure by tuning λ .

X axis: Lasso's penalty parameter ' λ ', ranges from 0- 0.1.

Left Y axis: Harrell's C-statistic after adjustment for optimism.

Right Y axis: Number of variables in the model.

The blue line corresponds to the left Y axis, which represents the trend of C-statistic as λ increases. The horizontal red line represents the minimal performance requirement, which is 10% below the optimal performance for our model. The combination of the red horizontal and the red vertical line represent where the model performance crosses the minimal performance requirement – at that λ value 7 variables were selected. The yellow vertical line represents the optimal performance we could achieve with these 7 variables, and the final λ that was chosen (0.041).

STATISTICAL METHODS IN MODEL DERIVATION AND VALIDATION

Multiple imputation

Data of biochemical variables (aspartate aminotransferase (AST), alanine aminotransferase (ALT), alkaline phosphatase (ALP), gamma-glutamyl transpeptidase (γ GT), total bilirubin, albumin and platelets) were collected at time of diagnosis and during follow-up. These values were not all available at time of diagnosis, even when time interval of ± 3 months was allowed. The percentages of missing values of each variable can be found in Table 2.

Missing values were imputed with multivariate imputation by chained equations (MICE), and 20 imputed datasets were created. In each imputation, time of iterations was set as 25. For continuous variables the predictive mean matching (pmm) method was used; for binary variables logistic regression (logreg) was used; for longitudinal variables (AST, ALT, ALP and total bilirubin) two-level linear model (2l.norm) was used.

After 20 imputed datasets were created, we fused all datasets and created one combined dataset. The further analyses were based on this combined dataset.

Variable selection by Lasso

Lasso's penalty parameter ' λ ' was chosen based on the discriminative power of the model, using Harrell's C-statistic²³. The C-statistic was corrected for optimism using the bootstrap procedure (1000 resamples).²⁰ We chose the λ based on the criterion that the resulting model has as few predictors as possible while yielding a C-statistic that is no more than 10% below the optimal one. Within a range of λ 's with the same number of predictors, we chose the one that had the highest C-statistic. With this λ , fewer variables were selected, but the parameters may shrink too much. For this reason an additional adjustment factor was needed to compensate for over-shrinkage.

Adjustment factor through recalibration

The "raw" was calculated with the formula derived by the Lasso model in 1.2. Then in this recalibration step, the 'raw' PI was used as a single predictor to re-fit the Cox model, and the coefficient of this Cox model served as the adjustment factor. In other words, the final PI will be the "raw" PI multiplied by the adjustment factor.

Actually, this adjustment factor will inflate all the parameters in the final model which is aimed to compensate the over shrinkage. This method may introduce an over-fitting issue, but since the relation between all parameters cannot change (they were all multiplied by the same adjustment factor), the prevention of over-fitting in the Lasso step is still effective.

Estimate Baseline Survival and Individual Survival

Another product in 1.3 is the baseline survival estimates. After recalibrating the model with the final PI, baseline survival estimates which are corresponding to the patients with the average PI were also calculated for each year of follow-up. Then the survival probability of other patients can be estimated by the formula: $S_i(t) = S_0(t)^{\exp(P_i - \bar{P})}$, where $S_0(t)$ is the baseline survival at time t , P_i is the prognostic index of patient i and $\bar{P}$ is the average PI in derivation cohort.

Model use in clinical practice

The following patient example explains how the model can be used in clinical practice: *A forty year old male patient visits the outpatient clinic. He was recently diagnosed with large duct PSC, and his blood tests show the following results:*

- Albumin (g/L) 38 (reference value: > 35)
- Platelets: 250 x 109/L (reference range: 150-35 x 109/L 0)
- AST (U/L): 88 (reference value: <35)
- Alkaline phosphatase (U/L): 360 (reference value: <120)
- Bilirubin (μmol/L): 26 (reference value: <17)

When we transform clinical and laboratory data according to the methods described in the formula, we get the following numbers to be entered in formula 1 (Step 1: expression in xULN/xLLN; Step 2: variable transformation):

- PSC subtype large duct: 1
- Age at diagnosis: 40
- Albumin (g/L)
 - Step 1: $38/35 = 1.086 \times \text{LLN}$; Step 2: $\log(1.086) = 0.041$
- Platelets
 - Step 1: $250/150 = 1.667 \times \text{LLN}$; Step 2: $\text{abs}(\log(1.667) - 0.5) = 0.270$
- AST (U/L)
 - Step 1: $88/35 = 2.514 \times \text{ULN}$; Step 2: $\log(2.514) = 0.398$
- Alkaline phosphatase (U/L):
 - Step 1: $360/120 = 3.0 \times \text{ULN}$; Step 2: $\log(3.0) = 0.477$
- Bilirubin (Umol/L)
 - Step 1: $26/17 = 1.529 \times \text{ULN}$; Step 2: $\log(1.529) = 0.176$

When entering these transformed data in the formula one can calculate the prognostic index:

$$\text{PI} = 0.323 \times 1 + 0.018 \times 40 - 2.485 \times 0.041 + 2.451 \times 0.270 + 0.347 \times 0.398 + 0.393 \times 0.477 + 0.337 \times 0.176 = 1.986$$

To calculate the 5, 10 and 15 years survival probability, the user needs the baseline survival for each year - provided in Supplementary table 1 - , as well as the mean of the final PI in the derivation cohort which is 1.646.

5-year survival probability : $\hat{S}_i(5) = \hat{S}_0(5)^{\exp(Pi - \bar{Pi})} = 0.912^{\exp(1.986 - 1.646)} = 87.86\%$

10-year survival probability : $\hat{S}_i(10) = \hat{S}_0(5)^{\exp(Pi - \bar{Pi})} = 0.789^{\exp(1.986 - 1.646)} = 71.68\%$

15-year survival probability : $\hat{S}_i(15) = \hat{S}_0(5)^{\exp(Pi - \bar{Pi})} = 0.674^{\exp(1.986 - 1.646)} = 57.45\%$

The relation between final PI and 5, 10 and 15 years survival probability is illustrated in Figure 3.

10

Summary, discussion, and future perspectives

SUMMARY

Biomarkers for disease progression in primary sclerosing cholangitis

Primary sclerosing cholangitis (PSC) is a dreadful cholestatic liver disease affecting the intra- and/or extrahepatic bile ducts. Inflammation of the biliary tree leads to the destruction of bile ducts, biliary fibrosis and eventually results in liver cirrhosis. PSC is strongly associated with inflammatory bowel disease, and patients are at increased risk of developing hepatobiliary malignancies.^{1–3} The aetiopathogenesis is still largely unknown, however, PSC is thought to be an immune mediated disorder that arises in a genetically susceptible host. With the upswing of the recent genome wide association (GWAS) studies in PSC, multiple genetic risk factors have been discovered.^{4–9} Notwithstanding, only a small fraction (7.3%) of PSC liability can be explained by these risk genes, suggesting that environmental risk factors are of major importance in triggering the disease development.¹⁰ Risk factors that were previously suggested to be associated with PSC include smoking behaviour and appendectomy. These risk factors were never studied on a population level and previous studies included only a limited number of PSC patients.^{11–14}

Chapter 2 describes a case control study in which potential risk factors for developing PSC – smoking, appendectomy, family history and geographical distribution – were assessed in a population-based cohort of 343 PSC patients, as compared to 370 IBD control patients, and 232 healthy controls. Smoking was shown to be associated with a lower risk of developing PSC. Appendectomy was not associated with the risk of developing PSC, neither was family history of IBD and auto-immune liver disease, nor residential environment.

Despite two decades of research in which treatment by various drugs (immunosuppressives, antibiotics, bile acids) has extensively been studied, so far there is no medical treatment option that has been proven to alter the progression of PSC.¹⁵ To improve patient management and come up with a rationally targeted medical therapy, a better understanding of pathophysiology is vital. The identification of genetic and environmental risk factors provides biological and molecular insight that is helpful in unravelling PSC pathogenesis, and may give clues to future therapeutic approaches.¹⁰

To find a treatment for PSC, one has to be able to measure its effect on outcome, and a robust design of upcoming therapeutic trials is important to be able to demonstrate treatment effects. The most reliable way to assess the clinical impact of a therapeutic intervention is through its effect on a well-defined *clinically meaningful endpoint*, such as survival (PSC related death) and liver transplantation.^{16,17} However, these clinical endpoints may be impossible to reach in clinical trials of very rare, slowly progressive diseases such as PSC, and selection of an appropriate indirect measure - a biomarker - to substitute for clinical

responses, could reduce the size, duration and costs of clinical trials. In addition, biomarkers can provide a basis for the identification of suitable candidates for clinical trials, contribute to the understanding of the pharmacology of candidates, and for characterization of the subtypes of disease for which a therapeutic intervention is most appropriate.^{16,18} If we can identify biomarkers that have prognostic value of clinical outcome in PSC, these may function as risk stratifiers and/or potential surrogate endpoints in clinical trials. Furthermore, in a heterogeneous disease such as PSC, stratification of patients into those with low versus high risk of poor outcome by means of prognostic biomarkers, can result in a more patient tailored counselling and management.¹⁸

Chapter 3 provides a comprehensive overview of biomarkers for progression of PSC that have been investigated to date. Categories that can be distinguished include clinical and biochemical biomarkers, histology, imaging, prognostic modelling and genetics. So far, clinical symptoms have been poorly associated with progression of PSC and therefore appear not useful as biomarkers of disease progression.^{20–23} Biochemical biomarkers such as alkaline phosphatase and enhanced liver fibrosis (ELF) test were found to be of strong prognostic value - at least in stratifying patients in favourable and unfavourable disease course -, and may be useful biomarkers of progression in PSC.^{24–29}

An increased serum alkaline phosphatase is the diagnostic hallmark in PSC, and is found in 95% of patients.³⁰ Several publications have reported that a reduction of alkaline phosphatase in a certain time frame is of prognostic value for PSC disease course.^{24–27} These studies were, however, limited by the absence of a limit for time to alkaline phosphatase reduction, thereby affecting the validity of the alkaline phosphatase measurement to predict time to event.^{24,25} In addition, all studies were undertaken in relatively small cohorts, and time until reduction, as well as thresholds for reduction of alkaline phosphatase differed between studies.^{24–27}

Chapter 4 describes a retrospective cohort study in which the prognostic value of alkaline phosphatase at diagnosis, at one year follow up, and percentage change between both time points is evaluated. Alkaline phosphatase levels at diagnosis, at one year follow up, and percentage change were found to be of prognostic value for endpoint-free survival in PSC, with an optimal threshold of 1.3 times the upper limit of normal at one year follow up. This is the first study in which the performance of alkaline phosphatase to predict prognosis in PSC is evaluated by means of a robust measure of prognostic value, the c-statistic. The c-statistic ranged between 0.5 and 0.7 and the wide corresponding confidence intervals indicated that the predictive accuracy of alkaline phosphatase at diagnosis and one year later is relatively low, but statistically significant. Therefore, alkaline

phosphatase is an important candidate surrogate endpoint / stratifier in PSC. The ELF test is a novel, non-invasive method to measure liver fibrosis, incorporating three 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). Liver fibrosis is a well-established predictor of outcome in PSC – exemplified by the implementation of liver histology in several prognostic models for PSC.^{31–34} ELF was previously shown to be predictive of transplant-free survival in PSC independently of other risk factors or risk score.³⁵

Chapter 5 describes a multi-center cohort study in which the prognostic utility of the ELF test was validated. In a well characterized cohort of 534 patients with large duct PSC, ELF test levels were shown to be significantly higher in patients reaching an endpoint than in those censored. By subdividing ELF test into three groups based of fibrosis severity, based on cut-off levels provided by the manufacturer, patients with PSC could be stratified into low, intermediate and high risk of poor outcome. One unit increase in the ELF test was shown to be independently associated with a 1.31-fold increased risk of death or liver transplantation. ELF test enabled a good discrimination between PSC patients that do, and do not reach an endpoint. With this study, the prognostic value of ELF test and its ability to stratify risk of poor outcome in PSC was confirmed.

Little is known about the effect of genetic variants on the course and progression of complex diseases in general, and on PSC in particular. The large variability of disease severity in PSC may be explained by the existence of modifier effects in PSC.

Chapter 6 describes the largest genotype-phenotype study in PSC to date, including 3402 PSC patients from 11 countries throughout Europe, the US and Canada. The SNP rs853974 was identified to be genome-wide significantly associated with transplant-free survival in PSC patients. The *R-spondin 3* gene was identified as candidate gene for this association, and expression of *R-spondin 3* in key effector cells, such as human and murine cholangiocytes and human hepatic stellate cells, was demonstrated.

Despite its invasive nature and risk of sampling variability, there still seems to be a place for histology as biomarker for disease progression in PSC. The applicability and prognostic value of liver histology in PSC is described in **Chapter 7**. In the absence of a validated histologic scoring system specifically designed to assess disease severity in PSC, the prognostic value of evaluating both degree of necroinflammatory activity (grade) and fibrosis (stage) was assessed using three histologic scoring systems that were originally designed for different liver diseases: the Nakanuma, Ishak and Ludwig scoring systems. In a Dutch PSC cohort, including 64 PSC patients, staging by all three histologic scoring systems

was strongly associated with the long term clinical outcome. In **Chapter 8**, the association between scoring liver histology and clinical outcome was validated in a large multi-center cohort, including PSC patients from 7 centers throughout 6 European countries. Histologic grading and staging was again performed using the Nakanuma, Ishak and Ludwig scoring systems. This study confirmed the independent prognostic utility of staging disease progression by all three scoring systems. The Nakanuma staging system – incorporating features of chronic biliary disease – was shown to hold strongest predictive value, and may be preferable for scoring PSC liver biopsies.

Combining biomarkers of disease progression in a prognostic model may improve estimation of survival. **Chapter 9** describes a novel prognostic model for transplant-free survival of PSC patients. This model was developed using a predominantly population based PSC cohort, including 692 patients, and validated in an external PSC cohort. The model is based on seven objectively measured, and readily available variables: PSC subtype (small or large duct), age at PSC diagnosis, aspartate aminotransferase, alkaline phosphatase, bilirubin, albumin and platelets, and can be used to accurately predict 5, 10 and 15 years transplant-free survival probabilities. The performance of the model is stable over time. Therefore, transplant-free survival probability can be recalculated when updated biochemical values are available.

DISCUSSION AND FUTURE PERSPECTIVES

What is a biomarker

In 2001, the Biomarkers Definitions Working Group defined a biomarker to be *‘A characteristic that is objectively measured as an indicator of normal biological process, pathogenic process, or pharmacological response to a therapeutic intervention’*.¹⁶ In general, three aspects of disease pathophysiology can be distinguished in which biomarkers may play a role: markers of trait, state and rate.³⁶ Firstly, biomarker of trait can be used to predict the likelihood of developing a disease, and is typically considered to be a risk factor.³⁶ These biomarkers may provide clues to the underlying pathogenesis of the disease. Secondly, a biomarker of disease state can be used to detect abnormalities that are associated with a certain disease entity, and is typically considered to be a diagnostic marker.³⁶ In a heterogeneous disease as PSC it may be difficult to identify a biomarker of disease state with a high sensitivity and specificity. Lastly, a biomarker of rate can be used as a clinical measurement tool to determine disease progression, or to accurately detect effects of therapeutic intervention.³⁶

A valid biomarker should measure what we want it to measure, and should be ‘fit-for-purpose’ - described by the face validity of a test. Overall, the

evaluation of the validity of a biomarker is complex, and the strength of the evidence for a biomarker can be evaluated in a three step approach – content, construct and criterion validity.^{37,38} The first step, content validity shows the degree of which the biomarker reflects the concept of interest (i.e the biological phenomenon studied), and is a largely qualitative assessment. Secondly, construct validity assesses if the measurements of the biomarker are consistent with the hypotheses- do the theory based predictions hold? The third step, criterion validity has two components; the ability of the biomarker to predict outcomes, and validation against a referent biomarker – is the measurement an adequate reflection of the gold standard? Criterion validity is usually measured by sensitivity, specificity and predictive power.³⁹

A surrogate endpoint is defined as: ‘A biomarker that is intended to substitute for a clinical endpoint. A surrogate endpoint is expected to predict clinical benefit (or harm or lack of benefit or harm) based on epidemiologic, therapeutic, pathophysiology, or other scientific evidence.’ Even though all surrogate endpoints are a subset of biomarkers, not all biomarkers are qualified to be surrogate endpoints; as nicely quoted by Fleming and DeMets. “A correlate does not a surrogate make”.^{40,41} For a biomarker to be qualified as a surrogate endpoint it should ideally lie directly in the causal pathway linking disease to outcome, and evidence should indicate that the biomarker predicts clinical risk or benefit.^{16,17} However, these causal mechanisms are often multidimensional, and in most circumstances not all treatment effects are fully accounted for by a single biomarker.¹⁷ In addition, potential ‘off-target’ drug effects may not be captured by the biomarker, further affecting its reliability.¹⁷ An important source of confounding originates from the failure to identify factors that may influence the measurement of the biomarker.³⁹

In the absence of profound understanding of the causal pathways linking the onset of PSC to its outcome, the theoretic approach of identifying the ‘ideal’ surrogate endpoint is not feasible in PSC. The focus of this thesis was mainly to identify potential biomarkers of outcome, establish construct validity, and to explore their association with clinically meaningful endpoints, in retrospective cohorts. The next step in establishing the validity of these biomarkers to serve as surrogate endpoint will be to include them as exploratory endpoints in randomized controlled clinical trials.⁴¹ In these trials the *accuracy* of the biomarker can be assessed by prospectively demonstrating a correlation between the biomarker and the clinically meaningful endpoint and its *reliability* can be derived from the reproducibility of the biomarker.¹⁷ Other important aspects that should be

considered when assessing the validity of a biomarker include the *interpretability* – degree of measurement errors, risk of bias or confounding – and its *feasibility* – taking into consideration the costs and acceptability of the biomarker.³⁹

In summary, validation of the biomarkers for disease progression of PSC that have been established within this thesis to function as surrogate endpoints in clinical trials should focus on the following three criteria:⁴²

The surrogate endpoint must be correlated with the true clinically meaningful endpoint.⁴³

The surrogate endpoint should capture a reliable and sufficiently large proportion of the treatment effect on the true clinical endpoint.

The surrogate endpoint should be able to predict the treatment effect on the true clinical endpoint.

Future biomarkers for disease progression in PSC

There are several important potential biomarkers for disease progression in PSC that deserve further exploration.

Patient-reported outcome

Patient-reported outcome is defined as any report on the status of a patient's health condition that comes directly from the patient, without interpretation by clinicians or others.⁴⁴ They are designed to measure an explicit concept, and can range from documentation of pruritus severity by means of a visual analogue scale, assessment of functioning by a diary, or measurement of patients health related quality of life using validated questionnaires.⁴⁵ Emphasis on patient-centered research is increasing, illustrated by the important increase in the use of patient reported outcome in the current clinical research.⁴⁶

Currently, there are no validated patient reported outcome measures in PSC. Symptoms, such as pruritus have been associated with a reduced health related quality of life,⁴⁷ but an association between symptoms or HRQL and disease severity has never been demonstrated.^{22,23,47,48} It is important to develop a valid and reliable instrument to measure patient reported outcome in PSC, and use this instrument to gain insight in their symptoms, sense of wellbeing, functional status, health related quality of life, and patients' adherence to treatment. The first step in developing such an instrument would be to establish its content validity.⁴⁴ Next, additional patient reported outcome properties should be studied; reliability, construct validity, if appropriate criterion validity, and the ability to detect change.⁴⁴ Once proved to be reliable and valid, the patient reported outcome measure may be used for measuring treatment effect, by implementation as a primary or secondary endpoint in clinical trials in PSC.⁴⁵

Microbiota

The distinct association between PSC and IBD, and increasing knowledge on the role of microbiota in the pathogenesis of IBD⁴⁹ has led to the hypothesis that microbiota may play a role in initiating and determining PSC phenotype. Two recent studies assessed the faecal microbiota – sequenced by 16SrDNA – of PSC patients compared with healthy controls and IBD patients, and showed dysbiosis in PSC patients with a reduced microbiota diversity.^{50,51} This dysbiosis was comparable in PSC patients with and without IBD, independent of treatment with ursodeoxycholic acid, and distinct from IBD patients.^{50,51} Sabino et al. described a three genera microbial signature in PSC, including a significant overrepresentation of *Enterococcus*, *Fusobacterium* and *Lactobacillus*,⁵¹ while Kummen et al demonstrated a marked enrichment in the *Veilonella* genus.⁵⁰ The mucosa-associated microbiota of patients with PSC has been evaluated by Rossen et al. who showed a slightly altered gut microbiota, with a reduced diversity and an under representation of uncultured Clostridialis II.⁵²

Altogether, these studies have shown that patients with PSC have a distinct microbiota profile, thereby establishing the construct validity of using microbiota as a biomarker for PSC. Modification of intestinal microbiota could potentially influence the disease process, and may therefore be a future biomarker to measure treatment effect in PSC.

Imaging modalities

Several high quality and non-invasive imaging techniques have the potential to be used as non-invasive biomarkers for disease progression in PSC. The most promising is the vibration-controlled transient elastography (FibroScan), which uses measurement of liver stiffness to establish degree of liver fibrosis.⁵³ Corpechot et al. have analysed the usefulness and prognostic value of FibroScan in PSC, and confirmed an independent association between liver stiffness measurement and histologically staged degree of fibrosis.⁵⁴ In addition, the progression rate of liver stiffness measurement was independently associated with clinical outcome.⁵⁴ A large international cohort study (FICUS trial) aiming to validate the use of FibroScan as biomarker of disease progression in PSC is currently being performed.

REFERENCES

1. Boonstra, K. et al. Population-based epidemiology, malignancy risk, and outcome of primary sclerosing cholangitis. *Hepatology* 58, 2045–55 (2013).
2. Burak, K. et al. Incidence and risk factors for cholangiocarcinoma in primary sclerosing cholangitis. *Am J Gastroenterol* 99, 523–6 (2004).
3. Claessen, M. M. H., Vleggaar, F. P., Tytgat, K. M. a J., Siersema, P. D. & van Buuren, H. R. High lifetime risk of cancer in primary sclerosing cholangitis. *J Hepatol* 50, 158–64 (2009).
4. Karlsen, T. H. et al. Genome-wide association analysis in primary sclerosing cholangitis. *Gastroenterology* 138, 1102–11 (2010).
5. Melum, E. et al. Genome-wide association analysis in primary sclerosing cholangitis identifies two non-HLA susceptibility loci. *Nat. Genet.* 43, 17–9 (2011).
6. Srivastava, B. et al. Fine mapping and replication of genetic risk loci in primary sclerosing cholangitis. *Scand. J. Gastroenterol.* 47, 820–6 (2012).
7. Folseraas, T. et al. Extended analysis of a genome-wide association study in primary sclerosing cholangitis detects multiple novel risk loci. *J. Hepatol.* 57, 366–75 (2012).
8. Ellinghaus, D. et al. Genome-wide association analysis in primary sclerosing cholangitis and ulcerative colitis identifies risk loci at GPR35 and TCF4. *Hepatology* 58, 1074–1083 (2013).
9. Liu, J. Z. et al. Dense genotyping of immune-related disease regions identifies nine new risk loci for primary sclerosing cholangitis. *Nat. Genet.* 45, 670–5 (2013).
10. Henriksen, E., Meluma, E. & Karlsen, T. Update on primary sclerosing cholangitis genetics. *Curr. Opin. Gastroenterol.* 30, 310–9 (2014).
11. van Erpecum, K. J. et al. Risk of primary sclerosing cholangitis is associated with nonsmoking behavior. *Gastroenterology* 110, 1503–6 (1996).
12. Loftus, E. V et al. Primary sclerosing cholangitis is associated with nonsmoking; a case-control study. *Gastroenterology* 110, 1496–1502 (1996).
13. Mitchell, S. et al. Cigarette smoking, appendectomy, and tonsillectomy as risk factors for the development of primary sclerosing cholangitis: a case control study. *Gut* 51, 567–73 (2002).
14. Florin, T., Pandeya, N. & Radford-Smith, G. Epidemiology of appendectomy in primary sclerosing cholangitis and ulcerative colitis: its influence on the clinical behaviour of these diseases. *Gut* 53, 973–979 (2004).
15. Karlsen, T. H., Vesterhus, M. & Boberg, K. M. Review article: controversies in the management of primary biliary cirrhosis and primary sclerosing cholangitis. *Aliment. Pharmacol. Ther.* 39, 282–301 (2014).
16. Atkinson A.J., J. et al. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. *Clin. Pharmacol. Ther.* 69, 89–95 (2001).
17. Fleming, T. R. & Powers, J. H. Biomarkers and surrogate endpoints in clinical trials. *Stat. Med.* 31, 2973–2984 (2012).
18. Trivedi, P. J., Corpechot, C., Pares, A. & Hirschfield, G. M. Risk Stratification in autoimmune cholestatic liver diseases: Opportunities for clinicians and trialists. *Hepatology* 63, 644–59 (2016).
19. Ponsioen, C. Y. et al. Surrogate endpoints for clinical trials in primary sclerosing cholangitis ; review and results from an International PSC Study Group consensus process. *Hepatology* 64, epub ahead of print (2016),

20. ter Borg, P. C. J., Fekkes, D., Vrolijk, J. M. & van Buuren, H. R. The relation between plasma tyrosine concentration and fatigue in primary biliary cirrhosis and primary sclerosing cholangitis. *BMC Gastroenterol.* 5, 11 (2005).
21. Björnsson, E., Simren, M., Olsson, R. & Chapman, R. W. Fatigue in patients with primary sclerosing cholangitis. *Scand. J. Gastroenterol.* 39, 961–8 (2004).
22. Benito de Valle, M. et al. Factors that reduce health-related quality of life in patients with primary sclerosing cholangitis. *Clin. Gastroenterol. Hepatol.* 10, 769–775.e2 (2012).
23. Gotthardt, D. N. et al. Pruritus is associated with severely impaired quality of life in patients with primary sclerosing cholangitis. *Eur. J. Gastroenterol. Hepatol.* 1–6 (2014).
24. Stanich, P. P. et al. Alkaline phosphatase normalization is associated with better prognosis in primary sclerosing cholangitis. *Dig. Liver Dis.* 43, 309–13 (2011).
25. Al Mamari, S., Djordjevic, J., Halliday, J. S. & Chapman, R. W. Improvement of serum alkaline phosphatase to <1.5 upper limit of normal predicts better outcome and reduced risk of cholangiocarcinoma in primary sclerosing cholangitis. *J. Hepatol.* 58, 329–34 (2013).
26. Lindström, L., Hultcrantz, R., Boberg, K. M., Friis-Liby, I. & Bergquist, A. Association between reduced levels of alkaline phosphatase and survival times of patients with primary sclerosing cholangitis. *Clin. Gastroenterol. Hepatol.* 11, 841–6 (2013).
27. Rupp, C. et al. Reduction in alkaline phosphatase is associated with longer survival in primary sclerosing cholangitis, independent of dominant stenosis. *Aliment. Pharmacol. Ther.* 1–10 (2014).
28. de Vries, E., Boonstra, K., Leeflang, M., Beuers, U. & Ponsioen, C. Normalization of alkaline phosphatase within the first year after diagnosis is of prognostic value in primary sclerosing cholangitis. *Gastroenterol. Conf. Dig. Dis. Week 2014, DDW 2014 Chicago, United States* 907 (2014).
29. Vesterhus, M. et al. Enhanced Liver Fibrosis Score Predicts Transplant-Free Survival in Primary Sclerosing Cholangitis. *Hepatology* 62, 188–197 (2015).
30. Singh, S. & Talwalkar, J. a. Primary sclerosing cholangitis: diagnosis, prognosis, and management. *Clin. Gastroenterol. Hepatol.* 11, 898–907 (2013).
31. Wiesner, R. et al. Primary sclerosing cholangitis: natural history, prognostic factors and survival analysis. *Hepatology* 10, 430–6 (1989).
32. Farrant, J. et al. Natural history and prognostic variables in primary sclerosing cholangitis. *Gastroenterology* 100, 1710–7 (1991).
33. Dickson, E. et al. Primary sclerosing cholangitis: refinement and validation of survival models. *Gastroenterology* 103, 1893–901 (1992).
34. Broomé, U. et al. Natural history and prognostic factors in 305 Swedish patients with primary sclerosing cholangitis. *Gut* 38, 610–615 (1996).
35. Vesterhus, M. et al. Enhanced Liver Fibrosis score predicts transplant-free survival in PSC independently of the Mayo risk score. *J. Hepatol. Conf. 49th Annu. Meet. Eur. Assoc. Study Liver, Int. Liver Congr. London United Kingdom* 188 (2014).
36. Fox, N. & Growdon, J. H. Biomarkers and Surrogates. *NeuroRx* 1, 181 (2004).
37. Streiner, D. L. & Norman, G. R. 'Precision' and 'accuracy': Two terms that are neither. *J. Clin. Epidemiol.* 59, 327–330 (2006).
38. Cronbach, L. & Meehl, P. Construct validity in psychological tests. *Psychol. Bull.* 129, 3–9 (1955).

39. Mayeux, R. Biomarkers: potential uses and limitations. *NeuroRx* 1, 182–8 (2004).
40. Weir, C. J. & Walley, R. J. Statistical evaluation of biomarkers as surrogate endpoints: A literature review. *Stat. Med.* 25, 183–203 (2006).
41. Fleming, T. R. & DeMets, D. L. Surrogate end points in clinical trials: are we being misled? *Ann. Intern. Med.* 125, 605–13 (1996).
42. Hunter, D. J. et al. A pathway and approach to biomarker validation and qualification for osteoarthritis clinical trials. *Curr. Drug Targets* 11, 536–45 (2010).
43. Prentice, R. Surrogate endpoints in clinical trials: definition and operational criteria. *Stat. Med.* 8, 431–440 (1989).
44. FDA & HHS. Guidance for Industry Use in Medical Product Development to Support Labeling Claims Guidance for Industry. *Clin. Fed. Regist.* 1–39 (2009).
45. Gnanasakthy, A., Lewis, S., Clark, M., Mordin, M. & DeMuro, C. Potential of patient-reported outcomes as nonprimary endpoints in clinical trials. *Health Qual. Life Outcomes* 11, 83 (2013).
46. Vodicka, E. et al. Inclusion of patient-reported outcome measures in registered clinical trials: Evidence from Clinical-Trials.gov (2007-2013). *Contemp. Clin. Trials* 43, 1–9 (2015).
47. Kremer, A. E., Bolier, R., van Dijk, R., Oude Elferink, R. P. J. & Beuers, U. Advances in pathogenesis and management of pruritus in cholestasis. *Dig. Dis.* 32, 637–45 (2014).
48. Björnsson, E., Simren, M., Olsson, R. & Chapman, R. W. Fatigue in patients with primary sclerosing cholangitis. *Scand. J. Gastroenterol.* 39, 961–968 (2004).
49. Kostic, A. D., Xavier, R. J. & Gevers, D. The microbiome in inflammatory bowel disease: Current status and the future ahead. *Gastroenterology* 146, 1489–1499 (2014).
50. Kummen, M. et al. The gut microbial profile in patients with primary sclerosing cholangitis is distinct from patients with ulcerative colitis without biliary disease and healthy controls. *Gut* [Epub ahead of print] (2016).
51. Sabino, J. et al. Primary sclerosing cholangitis is characterised by intestinal dysbiosis independent from IBD. *Gut* 1–9 (2016).
52. Rossen, N. G. et al. The mucosa-associated microbiota of PSC patients is characterized by low diversity and low abundance of uncultured Clostridiales II. *J. Crohns. Colitis* 9, 342–348 (2015).
53. Tapper, E. B., Castera, L. & Afdhal, N. H. FibroScan (Vibration-Controlled Transient Elastography): Where Does It Stand in the United States Practice. *Clin. Gastroenterol. Hepatol.* (2014).
54. Corpechot, C. 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* 146, 970–9; quiz e15–6 (2014).

NEDERLANDSE SAMENVATTING

Biomarkers voor ziekteprogressie in primaire scleroserende cholangitis

Primaire scleroserende cholangitis is een cholestatische leverziekte waarbij zowel de intra als de extra hepatische galwegen aangetast kunnen zijn. Ontsteking van de galboom leidt tot destructie van de galwegen, biliare fibrose en zal uiteindelijk resulteren in lever cirrose. PSC is sterk geassocieerd met inflammatoire darmziekten (IBD) en PSC patiënten hebben daarnaast een verhoogd risico op het ontwikkelen van hepatobiliaire maligniteiten. De ontstaanswijze van PSC is nog grotendeels onbekend, al wordt gedacht dat het een immuun gemedieerde ziekte is, die ontstaat in patiënten die daarvoor genetische aanleg hebben. Met de opkomst van de ‘genome wide association studies’ in PSC zijn verschillende genetische risicofactoren ontdekt. Desondanks wordt slechts een klein deel (7.3%) van de vatbaarheid voor PSC verklaard door dergelijke risicogenen, en spelen omgeving gerelateerde risicofactoren een belangrijke rol in het ontstaan van de ziekte. Risicofactoren waarvan reeds een associatie met PSC is aangetoond zijn rookgedrag en appendectomie. Deze risicofactoren zijn echter nooit onderzocht op een populatie gebaseerd niveau, en eerdere studies bevatten slechts een beperkt aantal PSC patiënten.

Hoofdstuk 2 beschrijft een case-controle studie waarin potentiële risicofactoren voor het ontstaan van PSC – roken, appendectomie, familie ziektegeschiedenis, en geografische spreiding – onderzocht zijn in een populatie-gebaseerd cohort bestaande uit 343 PSC patiënten, 370 IBD controle patiënten en 232 gezonde controles. Roken bleek geassocieerd te zijn met een verminderd risico op het ontwikkelen van PSC. Appendectomie, familiegeschiedenis van IBD of auto-immuun leverziekten en geografische spreiding waren niet geassocieerd met het risico op PSC.

Ondanks 20 jaren van onderzoek naar een behandeling voor PSC, is er nog altijd geen medicamenteuze behandeloptie gevonden die de progressie van de ziekte kan beïnvloeden. Een beter begrip van de pathogenese is van groot belang om een gerichte therapie te kunnen ontwikkelen. Door genetische en omgeving gerelateerde risicofactoren voor PSC te identificeren, vergroot men het biologische en moleculaire inzicht dat noodzakelijk is om de onderliggende PSC pathologie te kunnen ontrafelen en richting te kunnen geven aan toekomstige therapeutische benaderingen.

Om een effectieve behandeling voor PSC te kunnen vinden, is een meetbaar effect van de behandeling op de uiteindelijke uitkomst noodzakelijk. Hier moet rekening mee worden gehouden bij het ontwerpen van nieuwe therapeutische studies in PSC. De meest betrouwbare manier om het effect van een therapeutische interventie te meten, is door middel van klinische – harde –

eindpunten, zoals overleving (PSC-gerelateerd overlijden) en levertransplantatie. Het gebruik van dergelijk klinische eindpunten kan echter inefficiënt zijn en een geschikte indirecte maat – een biomarker – voor klinische respons kan de grootte, duur en kosten van een klinische studie reduceren. Dergelijk biomarkers zouden daarnaast gebruikt kunnen worden om geschikte kandidaten voor klinische studies te identificeren en om ziektefenotypen waarin therapeutische interventies de meeste waarde hebben te onderscheiden. Wanneer we biomarkers kunnen identificeren die van prognostische waarde zijn voor de klinische uitkomst van PSC, zouden deze gebruikt kunnen worden voor risico stratificatie en/of als surrogaat eindpunt in klinische studies. Het stratificeren van patiënten in een laag versus hoog risico op slechte uitkomst zou kunnen resulteren in een meer patiënt specifieke counseling en behandeling.

In **Hoofdstuk 3** wordt een overzicht gegeven van alle biomarkers voor ziekteprogressie in PSC die op dit moment onderzocht zijn. Categorieën die kunnen worden onderscheiden zijn klinische en biochemische biomarkers, histologie, beeldvorming, prognostische modellen en genetica. Doordat er tot dusver geen associatie tussen klinische symptomen en ziekteprogressie in PSC is aangetoond, lijken klinische symptomen geen nuttige biomarkers voor ziekteprogressie. Daarentegen is de sterke prognostische waarde van biochemische biomarkers zoals het alkalische fosfatase en de ‘enhance liver fibrosis’ (ELF) test wel reeds aangetoond, en dit zouden potentiële biomarkers voor ziekteprogressie van PSC kunnen zijn.

Een verhoogd alkalisch fosfatase is het belangrijkste diagnostische kenmerk van PSC en wordt gevonden in 95% van de patiënten. Verschillende publicaties hebben beschreven dat een daling van het alkalische fosfatase in een bepaalde tijdspanne van prognostische waarde zou zijn voor het ziektebeloop van PSC. Deze studies hadden echter geen tijdslimiet waarin deze daling moest plaatsvinden en vonden daarbij plaats in relatief kleine cohorten. Verder verschilden zowel de tijd tot afname van het alkalische fosfatase als de afkapwaarde tussen de verschillende studies. **Hoofdstuk 4** beschrijft een retrospectieve cohort studie waarin de prognostische waarde van het alkalische fosfatase is bestudeerd ten tijde van diagnose, één jaar later en het percentuele verschil tussen beiden tijdspunten. De levels van het alkalische fosfatase ten tijde van diagnose, één jaar later en het percentuele verschil bleken van voorspellende waarde voor eindpuntvrije overleving in PSC, met een optimale afkapwaarde van 1.3 keer de bovenwaarde van normaal na één jaar follow-up. Dit is de eerste studie die de prognostische waarde van het alkalische fosfatase meet middels een robuuste statistische maat, de c-statistic. De c-statistic had een range tussen de 0.5 en 0.7, met wijde confidentie intervallen, wat betekende dat de voorspellende

waarde van het alkalische fosfatase ten tijde van diagnose en na één jaar relatief laag, maar wel statistisch significant was. Het alkalische fosfatase is dan ook een belangrijk potentieel surrogaat eindpunt en risico stratificator in PSC.

De ELF test is een nieuwe, niet invasieve methode om leverfibrose te meten. Het bestaat uit een algoritme van drie serum markers van fibrose: hyaluronisch zuur, tissue inhibitor of metalloproteinases-1 (TIMP-1) en amino-terminal pro-peptide of type III pro-collagen (PIIINP). Lever fibrose is een erkende voorspeller van uitkomst in PSC, wat blijkt uit de implementatie van leverhistologie in verschillende prognostische modellen voor PSC. In een eerdere studie is aangetoond dat ELF een voorspeller van transplantatievrije overleving in PSC is, onafhankelijk van andere risicofactoren of risicoscores. **Hoofdstuk 5** beschrijft een multicenter cohort studie waarin de prognostische toepasbaarheid van ELF is gevalideerd. In een cohort van 534 large-duct PSC patiënten was de ELF test significant hoger in patiënten die een eindpunt bereikten dan in patiënten die gecensureerd werden. Door de ELF test onder te verdelen in drie groepen van lever fibrose, gebaseerd op de afkapwaarden die voorgesteld zijn door de fabrikant, konden PSC patiënten worden gestratificeerd in laag, middelmatig en hoge kans op een slechter uitkomst. Een punt toename in ELF was onafhankelijk geassocieerd met een 1.31 keer zo grote kans op dood of levertransplantatie. Door middel van de ELF test kon een goed onderscheid worden gemaakt tussen patiënten die wel en geen eindpunt behaalden. Met deze studie werd de prognostisch waarde van ELF en de mogelijkheid van ELF om te stratificeren naar risico op slechte uitkomst bevestigd.

Er is maar weinig bekend over de effecten van genetische varianten op het ziekte beloop en progressie van complexe ziekten in het algemeen, en PSC in het bijzonder. De grote variabiliteit van ziekte ernst in PSC zou verklaard kunnen worden door het bestaan van zogenaamde 'modifier effects' in PSC. In **Hoofdstuk 6** wordt de grootste genotype-fenotype studie in PSC beschreven, waarin 3402 patiënten vanuit 11 verschillende landen zijn geïnccludeerd. De SNP rs853974 was genome-wide significant geassocieerd met transplantatievrije overleving van PSC patiënten. Het R-spondin 3 gen bleek verantwoordelijk voor deze associatie, en expressie van r-spondin 3 in cellen die een belangrijke rol lijken te spelen in het ontstaan PSC- cholangiocyten van muizen en mensen, menselijke stamcellen – werd aangetoond.

Ondanks dat een leverbiopt een invasief onderzoek is met risico op sampling variabiliteit lijkt histologie nog altijd een belangrijke biomarker van ziekteprogressie in PSC te zijn. De toepasbaarheid en prognostisch waarde van leverhistologie in PSC werd beschreven in **Hoofdstuk 7**. Aangezien er geen gevalideerd histologisch scoring systeem bestaat dat speciaal is ontworpen

om de ziekte ernst van PSC te evalueren, werd in deze studie gebruik gemaakt van histologische scoringssystemen die oorspronkelijk werden ontworpen voor andere leverziekten; de Nakanuma, Ishak en Ludwig scoringssystemen. Hiermee kon de mate van necro-inflammatoire activiteit (grade) en fibrose (stage) worden geëvalueerd. In een Nederlands cohort bestaande uit 64 PSC patiënten bleek dat staging middels alle drie de scoringssystemen sterk geassocieerd was met lange termijn klinische uitkomst. In **Hoofdstuk 8** werd de associatie tussen het scoren van leverhistologie en klinische uitkomst gevalideerd middels een grote multicenter studie, waarin PSC patiënten vanuit 7 centra in 6 Europese landen geïnccludeerd werden. Opnieuw werd het scoren van de histologische grade en stage verricht middels de Nakanuma, Ishak en Ludwig scoringssystemen. Met deze studie werd de onafhankelijke voorspellende waarde van het stageren van ziekteprogressie middels alle drie de scoringssystemen bevestigd. Het Nakamuma scoringsysteem, welke kenmerken van chronische biliaire ziekte bevat, had de grootste voorspellende waarde en zou dan ook het systeem van voorkeur kunnen zijn in PSC.

Door verschillende biomarkers van ziekteprogressie te combineren in een prognostisch model, zou de nauwkeurigheid van het voorspellen van overleving kunnen verbeteren. In **Hoofdstuk 9** werd een nieuw prognostisch model voor transplantatievrije overleving van PSC patiënten beschreven. Dit model werd ontwikkeld aan de hand van een zo goed als populatie-gebaseerd cohort van 692 PSC patiënten, en gevalideerd in een extern PSC cohort. Het model was gebaseerd op 7 objectieve en gemakkelijk te meten variabelen: PSC type, leeftijd ten tijde van PSC diagnose, aspartate aminotransferase, alkalisch fosfatase, bilirubine, albumine en het trombocyten getal. Met dit prognostische model kon een nauwkeurige voorspelling van de 5, 10 en 15 jaars overleving van PSC patiënten gemaakt worden.

LIST OF PUBLICATIONS

Alkaline Phosphatase at Diagnosis of Primary Sclerosing Cholangitis and One Year Later: Evaluation of Prognostic Value.

Elisabeth M.G. de Vries, Junfeng Wang, Mariska M.G. Leeflang, Kirsten Boonstra, Rinse K. Weersma, Ulrich Beuers, Ronald B. Geskus, Cyriel Y. Ponsioen.

Liver International 2016; epub ahead of print

Risk factors for primary sclerosing cholangitis.

Kirsten Boonstra, **Elisabeth M.G. de Vries**, Nan van Geloven, Karel J. van Erpecum, Marcel Spanier, Alexander C. Poen, Carin M. van Nieuwkerk, Ben J. Witteman, Hans A. Tuijnman, Anton H. Naber, Paul J. Kingma, Ulrich Beuers, Cyriel Y. Ponsioen, on behalf of the Epi PSC PBC Study Group.

Liver International 2016;36:84-91

Applicability and prognostic value of histologic scoring systems in primary sclerosing cholangitis.

Elisabeth M.G. de Vries, Joanne Verheij, Stefan G. Hubscher, Mariska M.G. Leeflang, Kirsten Boonstra, Ulrich Beuers, Cyriel Y. Ponsioen.

Journal of Hepatology 2015;63:1212-9

Fecal microbiota transplantation as novel therapy in gastroenterology: A systematic review.

Noortje G. Rossen, J K. MacDonald, **Elisabeth M.G. de Vries**, Geert R. D'Haens, Willem M. de Vos, Erwin G. Zoetendal, Cyriel Y. Ponsioen.

World Journal of Gastroenterology 2015;21:5359-71

Biomarkers for disease progression of primary sclerosing cholangitis.

Elisabeth M.G. de Vries, Ulrich Beuers, Cyriel Y. Ponsioen.

Current Opinion in Gastroenterology 2015;31:239-46

Biodegradable stent placement before neoadjuvant chemoradiotherapy as a bridge to surgery in patients with locally advanced esophageal cancer.

Maarten W. van den Berg, Didi Walter, **Elisabeth M.G. de Vries**, Frank P. Vleggaar, Mark I. van Berge Henegouwen, Richard van Hillegersberg, Peter D. Siersema, Paul Fockens, Jeanin E. van Hooft.

Gastrointestinal Endoscopy 2014;80:908-13

Primaire scleroserende cholangitis.

Elisabeth M.G. de Vries, Cyriel Y. Ponsioen.

Critical care 2013;2:12-14

CONTRIBUTING AUTHORS

Rudi Alberts

Department of Gastroenterology and
Hepatology
University of Groningen and University Medical
Centre Groningen
Groningen, the Netherlands

Graeme Alexander

Department of Medicine, Division of
Hepatology
University of Cambridge
Cambridge, United Kingdom

Domenico Alvaro

Department of Clinical Medicine Division of
Gastroenterology
Sapienza University of Rome
Rome, Italy

Carl A. Anderson

Wellcome Trust Sanger Institute
Wellcome Trust Genome Campus, Hinxton
Cambridge, United Kingdom

Johanna Arola

Department of Pathology
Helsinki University and Helsinki University
Hospital
Helsinki, Finland

Annika Bergquist

Center for Digestive Diseases
Karolinska University Hospital, Karolinska
Institutet
Stockholm, Sweden

Francesca Bernuzzi

Center for Autoimmune Liver Diseases
Humanitas Clinical and Research Center
Rozzano, Italy

Ulrich Beuers

Department of Gastroenterology and
Hepatology
Academic Medical Center
Amsterdam, the Netherlands

Katharina Biermann

Department of Pathology
Erasmus University Medical Center
Rotterdam, the Netherlands

Niklas K. Björkström

Center for Infectious Medicine, Department of
Medicine Huddinge
Karolinska Institutet, Karolinska University
Hospital
Stockholm, Sweden

Einar Björnsson

Department of Internal Medicine, Division of
Gastroenterology and Hepatology
Landspítali University Hospital
Reykjavik, Iceland

Kirsten Muri Boberg

Norwegian PSC Research Center, Division of
Cancer Medicine, Surgery and Transplantation
Oslo University Hospital, Rikshospitalet
Oslo, Norway

Kirsten Boonstra

Department of Gastroenterology and
Hepatology
Academic Medical Center
Amsterdam, the Netherlands

Christopher L. Bowlus

Division of Gastroenterology and
Hepatology
University of California Davis
Davis, United States

Maria C. Bragazzi

Sapienza University of Rome
Rome, Italy

Henk van Buuren

Department of Gastroenterology and
Hepatology
Erasmus University Medical Center
Rotterdam, the Netherlands

Marco Carbone

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

Roger W. Chapman

Translational Gastroenterology Unit
John Radcliffe Hospital
Oxford, United Kingdom

Olivier Chazouillères

Department of Hepatology
Hôpital Saint Antoine, Sorbonne Universités
Paris, France

Angela Cheung

Gen. Internal Medicine
University Health Network- Toronto General
Hospital
Toronto, Canada

Georgios Dalekos

Department of Clinical Medicine, Division of
Gastroenterology
Sapienza University of Rome
Rome, Italy

John Eaton

Division of Gastroenterology and Hepatology
Mayo Clinic, Rochester, Minnesota
United States

Bertus Eksteen

Snyder Institute of Chronic Diseases
Department of Medicine, University of Calgary
Calgary, Canada

David Ellinghaus

Institute of Clinical Molecular Biology
Christian-Albrechts-University of Kiel
Kiel, Germany

Karel J. van Erpecum

Department of Gastroenterology and
Hepatology
University Medical Center Utrecht
Utrecht, the Netherlands

Martti Färkkilä

Department of Gastroenterology and
Hepatology
Helsinki University and Helsinki University
Hospital
Helsinki, Finland

Eleonora A.M. Festen

Department of Gastroenterology and
Hepatology
University of Groningen and University Medical
Centre Groningen
Groningen, the Netherlands

Annarosa Floreani

Dept. of Surgical, Oncological and
Gastroenterological Sciences
University of Padova
Padova, Italy

Trine Folseraas

Norwegian PSC Research Center, Division of
Cancer Medicine, Surgery and Transplantation
Oslo University Hospital Rikshospitalet
Oslo, Norway

Irene Franceschet

Department of Surgery, Oncology and
Gastroenterology
University of Padova
Padova, Italy

Andre Franke

Institute of Clinical Molecular Biology
Christian-Albrechts-University of Kiel
Kiel, Germany

Nan van Geloven

Clinical Research Unit
Academic Medical Center
Amsterdam, the Netherlands

Ronald B Geskus

Department of Clinical Epidemiology,
Biostatistics and Bioinformatics
Academic Medical Center
Amsterdam, the Netherlands

Odd Helge Gilja

National Centre for Ultrasound in
Gastroenterology
Department of Medicine, Haukeland University
Hospital
Bergen, Norway

Benjamin Goeppert

Institute of Pathology
University Hospital Heidelberg
Heidelberg, Germany

Elizabeth C. Goode

Norwich Medical School, Faculty of Medicine
and Health Sciences
University of East Anglia
Norwich, United Kingdom

Daniel Gotthardt

Department of Gastroenterology and
Hepatology
University Hospital Heidelberg
Heidelberg, Germany

Maren H Harms

Department of Gastroenterology and
Hepatology
Erasmus University Medical Center
Rotterdam, the Netherlands

Gideon M Hirschfield

National Institute for Health Research (NIHR)
Birmingham Liver Biomedical Research Unit
(BRU)
Institute of Immunology and Immunotherapy
University of Birmingham, United Kingdom

Bart van Hoek

Department of Gastroenterology and
Hepatology
Leiden University Medical Centre
Leiden, the Netherlands

Kristian Holm

Norwegian PSC Research Center, Division of
Cancer Medicine, Surgery and Transplantation

Oslo University Hospital, Rikshospitalet
Oslo, Norway

Simon Hohenester

Liver Center Munich, Department of Medicine
University of Munich
Munich, Germany

Johannes R. Hov

Norwegian PSC Research Center, Division of
Cancer Medicine, Surgery and Transplantation
Oslo University Hospital, Rikshospitalet
Oslo, Norway

Stefan G Hubscher

School of Cancer Sciences, University of Bir-
mingham
Department of Cellular Pathology
Queen Elizabeth Hospital
Birmingham, United Kingdom

Floris Imhann

Department of Gastroenterology and
Hepatology
University of Groningen and University Medical
Centre Groningen
Groningen, the Netherlands

Pietro Invernizzi

Center for Autoimmune Liver Diseases
Humanitas Clinical and Research Center
Rozzano, Italy

Xiaojun Jiang

Norwegian PSC Research Center, Division of
Cancer Medicine, Surgery and Transplantation
Oslo University Hospital, Rikshospitalet
Oslo, Norway

Brian D. Juran

Division of Gastroenterology and Hepatology
Mayo Clinic, Rochester
Minnesota, United States

Tom H. Karlsen

Norwegian PSC Research Center, Division of
Cancer Medicine, Surgery and Transplantation
Oslo University Hospital, Rikshospitalet
Oslo, Norway

Astrid Donald Kemgang

Department of Hepatology
Hôpital Saint Antoine, Sorbonne Universités
Paris, France

Paul J. Kingma

Department of Gastroenterology and
Hepatology
Tergooziekenhuizen
Hilversum/Blaricum, the Netherlands

Konstantinos Lazaridis

Division of Gastroenterology and Hepatology
Mayo Clinic, Rochester
Minnesota United States

Mariska MG Leeflang

Department of Clinical Epidemiology, Biostatistics
and Bioinformatics
Academic Medical Center
Amsterdam, the Netherlands

Henrike Lenzen

Department of Gastroenterology, Hepatology
and Endocrinology
Hannover Medical School
Hannover, Germany

Wolfgang Lieb

Popgen Biobank, University Hospital
Schleswig-Holstein
Christian-Albrechts-University
Kiel, Germany

Jimmy Z. Liu

Wellcome Trust Sanger Institute
Wellcome Trust Genome Campus, Hinxton
Cambridge, UK

Anton H. Naber

Department of Gastroenterology and
Hepatology
Tergooziekenhuizen
Hilversum/Blaricum, the Netherlands

Carin M. van Nieuwkerk

Department of Gastroenterology and
Hepatology
VU Medical Center
Amsterdam, the Netherlands

Michael P. Manns

Department of Gastroenterology, Hepatology
and Endocrinology
Hannover Medical School
Hannover, Germany

Hanns-Ulrich Marschall

Department of Molecular and Clinical Medicine
Institute of Medicine, Sahlgrenska Academy
Gothenburg, Sweden

Marco Marzioni

Department of Gastroenterology, Università
Politecnica delle Marche
Ospedali Riuniti University Hospital
Ancona, Italy

Andrew Mason

Division of Gastroenterology and Hepatology
University of Alberta
Edmonton, Alberta, Canada

Espen Melum

Norwegian PSC Research Center, Division of
Cancer Medicine, Surgery and Transplantation
Oslo University Hospital, Rikshospitalet
Oslo, Norway

Piotr Milkiewicz

Liver and Internal Medicine Unit
Medical University of Warsaw
Warsaw, Poland

Tobias Müller

Department of Internal Medicine, Hepatology
and Gastroenterology
Charité Universitätsmedizin Berlin, Campus
Virchow Klinikum
Berlin, Germany

Fotis Sampaziotis

Wellcome Trust-Medical Research Council Stem
Cell Institute, Anne McLaren Laboratory
Department of Surgery, University of Cam-
bridge
Cambridge, United Kingdom

Richard N. Sandford

Academic Department of Medical Genetics
University of Cambridge
Cambridge, UK

Christoph Schramm

1st Department of Medicine
University Medical Center Hamburg-Eppendorf
Hamburg, Germany

Stefan Schreiber

Institute of Clinical Molecular Biology
Christian-Albrechts-University of Kiel
Kiel, Germany

Erik Schrumpf

Norwegian PSC Research Center, Division of
Cancer Medicine, Surgery and Transplantation
Oslo University Hospital, Rikshospitalet
Oslo, Norway

Mark Silverberg

Inflammatory Bowel Disease (IBD) Group, Zane
Cohen Centre for Digestive Diseases
Mount Sinai Hospital Toronto
Ontario, Canada

Brijesh Srivastava

Academic Department of Medical Genetics
University of Cambridge
Cambridge, UK

Martina Sterneck

Department of Hepatobiliary Surgery and
Transplantation
University Medical Center Hamburg-Eppendorf
Hamburg, Germany

Albert Pares

Liver Unit, Hospital Clínic, IDIBAPS, CIBERehd
University of Barcelona
Barcelona, Spain

Massimo Pinzani

Institute for Liver and Digestive Health
University College London
Royal Free Hospital
London, UK

Christian Rupp

Department of Internal Medicine IV
University Hospital of Heidelberg
Heidelberg, Germany

Simon M. Rushbrook

Norwich Medical School, Faculty of Medicine
and Health Sciences
University of East Anglia
Norwich, United Kingdom

Christian Rust

Department of Medicine I,
Krankenhaus Barmherzige Brüder
Munich, Germany

Peter Schirmacher

Institute of Pathology
University Hospital Heidelberg
Heidelberg, Germany

Marcel Spanier

Department of Gastroenterology and
Hepatology
Rijnstate Hospital
Arnhem, the Netherlands

Albert Pares

Liver Unit, Hospital Clínic, IDIBAPS, CIBERehd,
University of Barcelona
Barcelona, Spain

Alexander C. Poen

Department of Gastroenterology and
Hepatology
Isala Clinics
Zwolle, the Netherlands

Cyriel Y Ponsioen

Department of Gastroenterology and
Hepatology
Academic Medical Center
Amsterdam, the Netherlands

Krista Rombouts

Integrated Research and Treatment
Center-Transplantation
Hannover Medical School
Hannover, Germany

William Rosenberg

UCL Institute for Liver and Digestive Health,
Division of Medicine
University College London & Royal Free Lon-
don, NHS Foundation Trust
London, United Kingdom

Helge Røsjø

K.G. Jebsen Cardiac Research Centre and
Center for Heart Failure Research
University of Oslo
Oslo, Norway

Katrin Schölzel

Institute for Liver and Digestive Health
University College London, Royal Free Hospital
London, UK

Andreas Teufel

1st Department of Medicine
University of Mainz
Mainz, Germany

Douglas Thorburn

UCL Institute for Liver and Digestive Health,
Division of Medicine
University College London & Royal Free Lon-
don, NHS Foundation Trust
London, UK

Hans A. Tuynman

Department of Gastroenterology and
Hepatology
Medical Center Alkmaar
Alkmaar, the Netherlands

Palak J. Trivedi

National Institute for Health Research (NIHR)
Birmingham Liver Biomedical Research Unit
(BRU)
Institute of Immunology and Immunotherapy
University of Birmingham, United Kingdom

Ben Vainer

Department of Pathology
Rigshospitalet, University of Copenhagen
Copenhagen, Denmark

Ludovic Vallier

Wellcome Trust-Medical Research Council
Stem Cell Institute, Anne McLaren Laboratory,
Department of Surgery
University of Cambridge
Cambridge, UK

Joanne Verheij

Department of Pathology
Academic Medical Center
Amsterdam, the Netherlands

Mette Vesterhus

Norwegian PSC Research Center, Division of
Cancer Medicine, Surgery and Transplantation
Oslo University Hospital, Rikshospitalet
Oslo, Norway

Arnau Vich Vila

Department of Gastroenterology and
Hepatology
University of Groningen and University Medical
Centre Groningen
Groningen, the Netherlands

Boudewijn de Vries

Department of Gastroenterology and
Hepatology
University Medical Center Groningen
Groningen, the Netherlands

Junfeng Wang

Department of Clinical Epidemiology,
Biostatistics and Bioinformatics
Academic Medical Center
Amsterdam, the Netherlands

Lai Mun Wang

Translational Gastroenterology Unit
John Radcliffe Hospital
Oxford, United Kingdom

Weiwei Wang

Division of Gastroenterology and Hepatology
University of Alberta
Edmonton, Alberta, Canada

Rinse K Weersma

Department of Gastroenterology and
Hepatology
University Medical Center Groningen
Groningen, the Netherlands

Tobias J. Weismüller

Department of Gastroenterology, Hepatology
and Endocrinology
Hannover Medical School
Hannover, Germany

Dominique Wendum

Department of Pathology
Hôpital Saint Antoine, Sorbonne Universités
Paris, France

Kate D. Williamson

Translational Gastroenterology Unit
John Radcliffe Hospital
Oxford, United Kingdom

Ben J. Witteman

Department of Gastroenterology and
Hepatology
Gelderse Vallei Hospital
Ede, the Netherlands

Ewa Wunsch

Department of Clinical and Molecular Bio-
chemistry
Pomeranian Medical University
Szczecin, Poland

Henriette Ytting

Department of Hepatology
Rigshospitalet, University of Copenhagen
Copenhagen, Denmark

Kalliopi Zachou

Department of Internal Medicine
University of Thessaly
Larissa, Greece

PHD PORTFOLIO

Courses	Year	ECTS
Basic course Legalisation and Organisation for Clinical Reasearchers (BROK)	2012	0.9
Practical Biostatistics	2013	1.1
Clinical Epidemiology	2014	0.6
Scientific Writing in English	2014	1.5
Course Infectious Disease	2015	1.0
Specific courses and seminars		
Immunohistochemistry Course Birmingham	2012	0.2
Biweekly seminar in gastroenterology	2012-2016	1.5
Biweekly progress reports in IBD	2012-2016	1.5
Gutclub meetings	2012-2016	1.5
Oral presentations		
Normalization of alkaline phosphatase within the first year after diagnosis is of prognostic value in primary sclerosing cholangitis DDW, Chicago, United States	2014	0.5
A novel prognostic model for primary sclerosing cholangitis	2015	0.5
Najaarscongres NVGE, Veldhoven, the Netherlands	2015	0.5
UEGW, Barcelona, Spain	2015	0.5
AASLD, San Fransisco, United States		
Validation of the prognostic value of histologic scoring systems in primary sclerosing cholangitis; a multicenter cohort study EASL, Barcelona, Spain	2016	0.5
Poster presentations		
Novel prognostic model for primary sclerosing cholangitis UEGW, Berlin, Germany	2013	0.5
Risk factors for the onset of primary sclerosing cholangitis; A case control study UEGW, Berlin, Germany	2013	0.5
Normalization of alkaline phosphatase within the first year after diagnosis is of prognostic value in primary sclerosing cholangitis. Falk Symposium, Paris, France	2014	0.5
Applicability and prognostic value of histologic scoring systems in primary sclerosing cholangitis. AASLD, Boston, United States	2014	0.5

(Inter)national conferences		
NVGE Voor-en Najaarsvergaderingen	2013-2015	1.5
European Crohns and Colitis Organisation (ECCO)	2013	0.5
European Association for the study of the liver (EASL)	2013, 2014, 2016	1.5
Falk Symposium, Paris, France	2014	0.5
Digestive Disease Week (DDW)	2014	0.5
American Association for the Study of Liver Diseases (AASLD)	2014, 2015	1.0
Teaching		
Lecture for 2nd year medical students (science course)	2013-2016	1.5
Lecture for pharmaceutical representatives	2013,2014	1.0
Lecture for the PSC patient organization Nieuwegein, the Netherlands	2015	1.0
Student coaching		
Robin Bokelaar	2013-2014	1.0
Roos Sablerolles, post ERCP complications in PSC	2015	0.5
Parameters of esteem		
NVGE travel grant	2015	
UEGW travel grant	2015	
Oral poster prize, open session cholestatic liver disease AASDL	2015	
EASL travel grant	2016	

DANKWOORD

Dank voor..

PSC Patiënten

Het vermogen vooruit te kijken in de zoektocht naar een behandeling voor PSC, jullie essentiële bijdrage aan dit proefschrift.

Opponenten - Prof. dr. Jansen, Prof. dr. Roskams, Prof. dr. van Gulik, dr. Rauws, dr. van Buuren

Het beoordelen van mijn proefschrift, jullie zitting in mijn commissie.

dr. Ponsioen, Cyriel

Jouw geduld en gedrevenheid om langlopende, grootse, tijdrovende, maar daarmee baanbrekende projecten te leiden. Het voorbeeld dat je bent als dokter, als onderzoeker, en als mentor – betrokken, doelgericht, en alles met de nodige ongedwongen humor en luchtigheid. Jouw vertrouwen in mij doet het beste naar boven komen!

Prof. dr. Beuers, Ulrich

Jouw toewijding en zorgvuldigheid waarmee je kritisch elk stuk doornam, en garant staat voor kwaliteit.

dr. Verheij, Joanne

Jouw energie en aanstekelijke enthousiasme voor het vak, dat ons bracht tot mooie ideeën, projecten en nieuwe inzichten. Het pragmatisme en de efficiëntie die ik af kon kijken, jouw betrokkenheid op alle fronten.

Junfeng Wang

Your endless patience in trying to make me understand statistics. Your brightness, your hard working spirit, our valuable collaboration.

dr. Leeflang, Mariska

Jouw gave om epidemiologische en statistische resultaten te vertalen naar klinische relevantie.

dr. Geskus, Ronald

Jouw bijzonder kritisch werkwijze, waarmee je onze projecten bracht tot op het punt waar alles klopt.

Members of the Internationale PSC study Group

Allowing me the participate in PSC research projects within this international collaboration of inspiring scientists.

Kirsten

De glimmende Epi PSC PBC trein die jij zo zorgvuldig beheerde en voor mij achterliet, jouw interesse en betrokkenheid bij alles wat ik deed.

Noortje

De houvast in de wereld van onderzoek, van SPSS naar presentaties, van IBD meetings naar ieder congres, van Spotify naar Xavier Rudd.

AMC collega's - C2.310. Pim, Maarten, Margriet, Jessica, Selma, Emo, Froukje

Het ritme en de regelmaat. De zuurstof en de zon!

Margriet: Jouw- ogenschijnlijk onbewogen- Drentse mentaliteit; brengt rust maar ook eindeloos droog vermaak.

Sel: De talloze life events met jou als klankbord op rechts achter – samen van de reservebank naar de basis..!

Emo: jouw mannelijkheid, dat ik een van jouw bitches mocht zijn.

Kim: jouw nuchtere en zo gemotiveerde aanpak, die mij het volste vertrouwen geeft in het vervolg.

Manon: Jij doet in het lab wat ik nooit deed, met passie en met doorzettingsvermogen!

Paranimfen: Geen pieken zonder dalen – jullie altijd aan mijn zij.

Heleen: the twin mystique, nooit aflatende steun en toeverlaat!

Suzanne: van Groningen via Deventer naar Amsterdam, met jou langs elk station.

Vrienden - Link met omgekeerde n

Jullie, áltijd daar! Zonder jullie geen boek!

Renske

Jouw guidance tussen de Lange Leidse en het AMC - wilde fases - dilemma's - liefdes- dips-

Willemijn

De morele steun.. Onvoorwaardelijke vriendschap.

Suzan

Mijn grote zus zijn; jij doet mij relativeren, draait wat negatief is positief, leeft enthousiasme!

Papa en mama

Het maakt jullie niet uit wat ik doe – of jullie het onderwerp nou begrijpend uit kunnen spreken of niet – altijd geïnteresseerd, altijd enthousiast, altijd motiverend.

~Jaap

Zonder promotie niet met jou, jij stimuleerde mij tot dit punt. Nu is jouw later mijn later!

ABOUT THE AUTHOR

Liesbeth de Vries was born in Leiderdorp on the sixth of February 1986, where she grew up with her parents, her older sister, and her twin sister. After finishing her secondary school at the Rijnlands Lyceum in Oestgeest, she started medical school at the Rijksuniversiteit, Groningen. During her studies she travelled to India where she participated in a developmental internship within the Rajastahan Network for people living with HIV/AIDS, and did an internship focusing on social and tropical medicine in Wasso Hospital, Tanzania. In 2011 she moved to Amsterdam where she finished the last part of her studies. After graduation she gained clinical experience working as a hospital doctor on the internal medicine ward of the Kennemer Gasthuis, Haarlem. Subsequently she got the opportunity to start a PhD project focusing on the identification of biomarkers for disease progression in PSC, under supervision of Cyriel Ponsioen and Ulrich Beuers. The first of May 2016, she started her training in gastroenterology and hepatology at the Academic Medical Center, Amsterdam, with a residency at the department of internal medicine of the OLVG West, Amsterdam.