RISK STRATIFICATION IN PRIMARY BILIARY CHOLANGITIS AND PRIMARY SCLEROSING CHOLANGITIS

Jorn Cornelis Goet

Colophon

ISBN 978-94-6423-210-3

Layout and printing: Proefschriftmaken.nl, Vianen, The Netherlands

Cover design: Kristina S. Fluitman

The work presented in this thesis was conducted at the department of Gastroenterology and Hepatology, Erasmus University Medical Center, Rotterdam, The Netherlands.

Financial support for printing this thesis was kindly provided by: Nederlandse Vereniging voor Hepatologie, afdeling Maag-, Darm- en Leverziekten van het Erasmus MC, Erasmus Universiteit Rotterdam, SBOH.

RISK STRATIFICATION IN PRIMARY BILIARY CHOLANGITIS AND PRIMARY SCLEROSING CHOLANGITIS

Risicostratificatie van patiënten met primaire biliaire cholangitis
en primaire scleroserende cholangitis

Proefschrift

ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam
op gezag van de rector magnificus
Prof.dr. F.A. van der Duijn Schouten

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
woensdag 12 mei 2021 om 10:30 uur

door

Jorn Cornelis Goet

geboren op dinsdag 15 mei 1990 te Voorburg

Erasmus University Rotterdam

PROMOTIECOMMISSIE

Promotor	Prof.dr. H.J. Metselaar
Overige leden	Prof.dr. C.J. van der Woude
	Prof.dr. U.H.W. Beuers
	Prof.dr. J.P.H. Drenth
Co-promotoren	Dr. H.R. van Buuren
	Dr. B.E. Hansen

TABLE OF CONTENTS

CHAPTER 1

GENERAL INTRODUCTION

Part A - Primary Biliary Cholangitis

The general introduction is based on and adapted from:

Risk Stratification and Prognostic Modeling in Primary Biliary Cholangitis
Jorn C Goet, Maren H Harms, Marco Carbone, Bettina E Hansen
Best Practice & Research Clinical Gastroenterology. 2018; 34-35: 95-106.

Guideline review: British Society of Gastroenterology/UK-PBC Primary Biliary Cholangitis treatment and management guidelines
Jorn C Goet, Gideon M Hirschfield.
Frontline Gastroenterology 2019;10(3):316-319

Primary biliary cholangitis (PBC) is a chronic cholestatic liver disease that predominantly affects middle-aged women[1,2]. The disease is rare and classified as a so-called orphan disease. Thus far, its pathogenesis has not been elucidated but is thought to result from an interplay of multiple genetic factors with superimposed environmental triggers[1-3]. PBC is a usually slowly progressive disorder that may lead to hepatic damage, fibrosis, cirrhosis and eventually liver failure requiring a liver transplant (LT), or death[1,2].

FROM CIRRHOSIS TO CHOLANGITIS: A MOST WELCOME CHANGE OF NOMENCLATURE

Primary Biliary Cholangitis as a disease entity was already described as early as 1851 by Addison and Gull[4] and by *Hanot et al*. in 1876[5]. In 1938 the term 'xanthomatous biliary cirrhosis' was proposed because of the observed periorbital xanthoma formation, together with destruction of intrahepatic bile ducts leading to cirrhosis[6]. The name ''Primary Biliary Cirrhosis" was introduced by Ahrens et al. in 1950[7] and was gradually adopted and used worldwide for more than half a century. In 1965, the term 'chronic suppurative destructive cholangitis' was proposed[7], but there was a lack of general acceptance of this more satisfactory terminology. During the last decades, however, the name Primary Biliary Cholangitis was increasingly regarded an unfortunate misnomer as most patients present without cirrhosis and a substantial proportion will never develop cirrhosis or the complications of end-stage liver disease. Moreover, the inclusion of 'cirrhosis' in the nomenclature has potential negative implications for patients including a stigma of alcohol abuse. In 2015 an authorative international platform, with strong support of patient organizations, proposed a new name: 'Primary Biliary Cholangitis'[8] . This name had the advantage that the acronym PBC could be maintained. Primary Biliary Cholangitis was implemented in the relevant 2017 European[9] and 2018 American[10] clinical Practice Guidelines, and since then has become the undisputed new name of the disease.

CLINICAL PRESENTATION AND DIAGNOSIS

Nowadays, many PBC patients lack specific disease-related symptoms at presentation[11-14] and are diagnosed prompted by abnormal results of routine liver biochemical testing. Diagnostic studies for fatigue and pruritus, the most common symptoms of the disease, lead to the diagnosis in a subset of patients, whereas a small subgroup presents with cirrhosis-related complications such as ascites or, very rarely, variceal bleeding. Although many patients may present at early symptom-free stages of disease, the large majority (~95%) will develop PBC-specific symptoms over time, fatigue and pruritus being the most prevalent[15-19]. In addition to fatigue and itch, patients may suffer from social isolation, sicca complex (dry eyes and mouth), abdominal pain, xanthelasma, jaundice and arthralgia. These symptoms significantly impact patients' quality of life and adequate therapeutic interventions are limited[20,21].

A diagnosis of Primary Biliary Cholangitis is usually made, in the correct context, in patients with repeated unexplained elevation of serum levels of alkaline phosphatase

(ALP) (> 6 months) in combination with the presence of antimitochondrial antibodies (AMA) (titre >1:40; immunofluorescence test) or highly PBC-specific antinuclear antibodies. Approximately 10% of patients test negative for AMA: AMA-negative PBC. AMA against pyruvate dehydrogenase complex E2 in serum, as determined by ELISA, is a major serological hallmark and present in ~95% of patients[22]. In this setting, liver biopsy is no longer required to establish a definite diagnosis[9], and is nowadays not performed routinely. Histological hallmarks of PBC are the presence of chronic non-suppurative destructive inflammation (cholangitis) of portal and small interlobular and septal bile ducts, frequently associated with portal triad granulomas (florid bile duct lesions). In more advanced stages progressive fibrosis, bile duct loss (ductopenia) and copper accumulation can be observed[23,24]. A liver biopsy to confirm diagnosis is indicated in the absence of diagnostic serological findings, in case of otherwise diagnostic uncertainty, or to clarify diagnosis when clinical suspicion of co-existing liver disease (e.g. features of autoimmune hepatitis; non-alcoholic fatty liver disease) may impact on patient management. Imaging should be used to exclude alternative diagnoses, such as obstructive biliary and infiltrative disease, and to assess the stage of the disease. In most cases abdominal ultrasound is sufficient. This should preferentially be accompanied by transient elastography as a non-invasive tool to determine the degree of liver fibrosis. In cases with diagnostic uncertainty magnetic resonance cholangiopancreatography (MRCP) can be indicated to detect disorders such as primary sclerosing cholangitis (PSC). Magnetic resonance imaging (MRI) is typically normal in patients with PBC, and is not indicated routinely.

EPIDEMIOLOGY

PBC predominantly affects middle-aged women, with a female to male ratio of 10:1[1,2], although recent studies suggest an increasing male prevalence[25,26]. Although most patients are diagnosed at an average age of 50 years, patients may present in their twenties but may also be considerably older. The incidence and prevalence of PBC is increasing[27]. In a recent systematic review, incidence rates ranged from 0.33 to 5.8 per 100.000 inhabitants/year and prevalence from 1.91 to 40.2 per 100.000[27-30]. Incidence and prevalence vary greatly between different geographical areas. To date, these findings remain unexplained.

AETIOLOGY AND PATHOPHYSIOLOGY

Thus far, the pathogenesis of PBC has not been elucidated, and is thought to result from an interplay of genetic susceptibility and environmental triggers[3]. Sibling studies in PBC have shown that first-degree relatives carry higher risk of developing PBC, indicating a genetic component in the development of the disease[31]. In addition, Genome-Wide Association Studies (GWAS) have identified different risk loci[32-34]. Suggested environmental triggers in PBC include hair dyes, nail polish, cigarette smoking, and infectious agents including Escheria coli, Mycobacterium gordonae, and retroviruses[35,36].

PBC has long been considered an autoimmune liver disease. Supporting this classification is the association between PBC and other autoimmune disorders including CREST,

Sjögren, autoimmune hypothyroidism and rheumatoid arthritis[37] as well as the female preponderance and presence of anti-mitochondrial antibodies[22]. A strong argument against PBC as a classical autoimmune disease is the absence of a significant or clear beneficial therapeutic effect of a wide range of immune-based therapies. It may well be that the features suggesting a possible autoimmune aetiology are secondary to destructive cholangitis due to another pathophysiological mechanism.

Key in the pathogenesis of PBC are immunoregulatory changes (immunomodulatory pathway) and associated selective destruction of intrahepatic cholangiocytes (biliary pathway)[31,38,39]. The immune-mediated pathway is directed against mitochondrial and nuclear autoantigens of the biliary epithelial cells lining the interlobular bile ducts in the liver[1,2,40]. This process is directed against the pyruvate dehydrogenase complex E2[22]. Why this immunological intolerance develops is unknown. However, this intolerance triggers a cascade of progressive bile duct destruction and cholestasis[41]. Histopathologically, there is portal inflammation and destruction of the intrahepatic bile ducts, which causes a decrease in bile secretion with retention of toxic substances in the liver, ultimately leading to hepatic damage, fibrosis, cirrhosis and eventually liver failure[1,2].

In this regard, an important advance in our understanding of the possible pathogenesis of PBC was the development of the 'biliary umbrella' concept. Central to this theory is an intact Cl-/HCO3- exchanger (AE2; anion exchanger 2) and an intact biliary glycocalyx, thereby maintaining a protective biliary 'umbrella' against invasion of hydrophobic bile acids monomers normally present in bile. In PBC, AE2 is downregulated resulting in less bicarbonate secretion with subsequently more toxic bile. This toxic bile causes the destruction of the cholangiocyte cell membrane and mitochondria, resulting in apoptosis of cholangiocytes. Within this cascade, AE2 expression is further downregulated by bile acids by inducing reactive oxygen species in the biliary epithelial cells, thereby inducing cell senescence leading to bile duct inflammation[42].

Further understanding of the aetiology of PBC can be derived from the gut-liver axis in which nuclear hormone receptors play a key role. These nuclear receptors regulate genes that are involved in bile acid homeostasis[43]. One key receptor is the farnesoid X receptor (FXR) which is predominantly expressed in the liver and small intestine. Inhibition of FXR reduces bile acid uptake, bile acid synthesis, hepatic inflammation and fibrinogenesis. In addition, pathways leading to more bile acid export and hepatic regeneration are increased[44]. In animal models mimicking PBC, activation of the FXR pathway has been shown to be protective.

THERAPIES

The search for therapies for PBC has not only been hampered by the unidentified etiology but also by to the slowly progressive nature of PBC, thereby making it virtually impossible to address end-points such as death or LT in the time-span of a clinical trial. Over the past decades, many potential therapies have been studied, including immunosuppressive, immunomodulating, chelating and anti-inflammatory agents, such as prednisone,

budesonide, D-penicillamine, colchicine, cyclosporine, azathioprine, methotrexate, chlorambucil, malotilate, and rituximab. However, none of these therapies have been proven effective and none are currently licensed for PBC. Currently licenced agents in PBC are UDCA and obeticholic acid (OCA).

First-line therapies

Ursodeoxycholic acid (UDCA) at a dose of 13-15 mg/kg/d is the established first-line therapy for PBC, usually continued for live[9,45]. UDCA accounts for approximately 1-3% of endogenous bile acids. With the use of exogenous UDCA as a pharmacotherapy, UDCA becomes the predominant bile acid in serum and bile. UDCA has been shown to improve liver biochemistry[46], can delay histological progression[47-49] and has been suggested to impact on long-term outcome[46,50,51]. Recently, *Harms et al.* studied the long-term effects of UDCA in a cohort of 3902 UDCA-treated and 373 untreated patients using an inverse probability-weighing analyses (IPTW)[52]. Using IPTW, and thereby mimicking a clinical trial, they showed UDCA therapy significantly reduced the risk of death or LT (hazard ratio [HR] 0.46, 95% confidence interval [CI] 0.40-0.52). Importantly, this protective effect remained regardless of disease stage or biochemical response. Clinical benefit of UDCA as expressed by absolute measures has also been studied recently and has shown that the number needed to treat to prevent one death or LT within 5 years in PBC is on average 11[53]. Recently, UDCA given after liver transplantation for PBC was shown to be associated with a lower risk for recurrence of the disease, graft loss, and death[54]. These findings further support the use of UDCA as mainstay therapy and its prompt initiation following a PBC diagnosis.

Second-line therapies

At present, most patients are diagnosed at an early stage of disease and are usually promptly treated with an adequate dose of UDCA[14]. However, despite adequate treatment, patients can remain at risk of developing cirrhosis and associated complications. In addition, about one third of patients has inadequate biochemical response to UDCA treatment which is associated with reduced transplant-free survival[55]. For these patients second-line add-on therapies are needed to delay disease progression.

Obeticholic acid is the only presently licenced add-on agent for patients with insufficient response to UDCA. OCA is prescribed at an initial dose of 5 mg/day, titrating to 10 mg/day at 6 months if tolerated. The effects of OCA were evaluated in a 12-month, double-blind, placebo-controlled, phase 3 trial in which 217 patients with PBC who had inadequate response to UDCA were randomly assigned to UDCA+OCA at a dose of 10 mg, OCA at a dosage of 5 mg with titration to 10mg if tolerated, or a placebo[56]. This study found positive effects of OCA on important surrogates of disease progression among a significant proportion of patients, with ~46% of patients reaching the primary endpoint of ALP-level less than 1.67 times the upper limit of normal (ULN), with a reduction of at least 15% from baseline, and a normal bilirubin. Pruritus is a side effect of OCA that needs pro-active recognition.

Although not currently licensed for PBC, *bezafibrate*, has recently been evaluated as add-on therapy in a randomised controlled trial. Bezafibrate acts on the peroxisome proliferator-activated receptors (PPAR)[57] resulting in downregulation of bile acid uptake and synthesis in

hepatocytes[58]. *Corpechot et al.* showed that approximately 67% of patients with inadequate response to UDCA have improvement in biochemical parameters after 2 years of bezafibrate add-on therapy[59]. Importantly, a decrease in pruritus was reported in the treatment arm. These finding confirm the results of a large number of small, uncontrolled studies, many originating from Japan, reporting beneficial effects of fibrates, including bezafibrate and fenofibrate and used either as monotherapy or in combination with UDCA, on serum liver tests and pruritus[58,60-67]. The currently available data suggest that the effects of beza- and fenofibrate are comparable and that these drugs are safe[68].

A recent randomized, placebo-controlled trial confirmed the anti-pruretic potential of bezafibrate in treating cholestatic itch[69,70]. Although not officially licensed, fibrates are now used on a world-wide scale as second line treatment for PBC. It seems doubtful whether there will be future initiatives to further document the efficacy of fibrates as add-on therapy on hard endpoint in a randomized controlled fashion. In contrast to obeticholic acid, fibrate therapy is cheap. In the Netherlands, the price per day is approximately Euro 0.36.

DISEASE COURSE AND PROGNOSIS

Although PBC is a slowly progressive disease in the majority of patients, the clinical course may differ greatly between patients and prognosis is largely dependent on development of cirrhotic complication[71]. The introduction of UDCA as mainstay treatment has greatly affected the clinical disease course[45-49,72]. At present, most patients are diagnosed at earlier stages of disease and are usually promptly treated with an adequate dose of UDCA[14]. However, despite adequate treatment, patients can remain at risk of developing cirrhosis and associated complications.

An early study showed that approximately 20% of UDCA-treated patients progress to cirrhosis within 4 years after start of treatment[73]. Later, *Corpechot et al.* showed that 7% of UDCA-treated patients progress to extensive fibrosis or cirrhosis anually[48]. A study with 254 paired biopsies over 655 patient-years showed that the rate of histological progression to cirrhosis under UDCA treatment from fibrosis stage I, II, and III was 4%, 12% and 59% respectively[74]. Median time to cirrhosis in these stages was 25, 20, and 4 years, respectively[74]. A recent cohort study by *Trivedi et al.* in 511 patients showed a cumulative incidence of cirrhosis development of 40% over a 10-year follow-up period, indicating that nowadays a substantial percentage of patients still develops cirrhosis[75,76]. Despite this reported histological progression under UDCA, a large international cohort study involving 3224 UDCA treated patients, recently reported by Harms et al, found that 85% of patients remained free of non-neoplastic cirrhotic complications (ascites; variceal bleeding; hepatic encephalopathy) after 15 years of follow-up[77]. An incomplete response to UDCA and a high AST/platelet ratio were independent risk factors for future complications. Another important finding was that the cumulative incidence of complications has decreased over the last decades. The ten-year transplant-free survival of UDCA-treated PBC patients is considered good and is approximately 80%[11,50,51,78-83]. In addition, absolute numbers of liver transplantations for PBC have decreased over the last three decades in North America and Europe and remained stable at a low level in the last 10 years[84-86].

RISK STRATIFICATION

Often the objective of studying a new marker is limited to the assessment of its association with a future clinical event. In the setting of risk stratification, the aim is to estimate the likelihood of a clinical event taking place. Assessment of the risks and risk parameters allow identification of patients or patient groups with mild or more progressive disease pathways, and thereby allow for the targeting of care. For decades, UDCA was the only available drug to treat PBC. Currently, the clinical scenario is changing as new agents (have) become available, all with different profiles of efficacy and tolerability. In this scenario, risk stratification is a critical tool to distinguish patients in need of second-line therapies from patients who can be safely continued on UDCA monotherapy. Over the past decade, our ability to identify subgroups of patients with a higher chance to have a progressive disease, and to estimate the risk of future adverse events based on readily available clinical parameters, has improved considerably. Newly developed prediction models now enable us to quantify the risk of future events for individual patients and provide important tools in clinical practice for patient counselling, timing of diagnostic procedures and therapeutic interventions, and selection of patients for clinical trials.

Liver biochemistry

Shapiro et al. (1979) were the first to start a long history of studying factors associated with disease progression in PBC, recognizing the association between serum bilirubin levels and survival (**Table 1**)[87]. They found that patients with bilirubin levels >2 mg/dl in two subsequent measurements within 6 months had an average survival of 4.1 years, whereas the average survival was 2.1 and 1.4 years when two subsequent measurements were above 6 mg/dl or 10 mg/dl, respectively. Furthermore, they showed that the behavioural pattern of bilirubin is characterized by two distinct phases: a phase in which bilirubin remains relatively stable for many years followed by an 'acceleration' phase with rapidly increasing values cumulating in death within a few years[87]. Similar patterns are observed in other end-stage liver diseases[88]. Confirming these phases, *Harms et al.* (2016) showed (n=3529) that the curve breaking point of bilirubin – meaning the point where, after a gradual increase in bilirubin, bilirubin increases rapidly towards occurrence of a clinical endpoint – was found at a bilirubin level 1.6 times the ULN. From this breaking point onward there was a median of 19 months before a clinical endpoint occurred (death or LT)[89]. This may suggest that bilirubin is a "late" biomarker, i.e. increasing only shortly before a clinical event, and thereby less applicable for early detection of disease progression. Interestingly, a recent study by the Global PBC study group showed that bilirubin values within the normal range (stratified into quartiles) both at baseline and after one year of UDCA therapy were also predictive of transplant-free survival[90]. Five-year transplant-free survival rates for Q1-Q4 of bilirubin were for 97%, 95%, 96%, and 91%, respectively ($p<0.001$). In addition, higher bilirubin (per 0.1xULN increase) was associated with an increased chance of death or transplantation (HR 1.14). Alkaline phosphatase (ALP) and gamma-glutamyl transpeptidase (GGT) are considered to be early markers of disease[72]. Although elevation of GGT may precede that of ALP in early stages of PBC, baseline GGT levels itself have not been associated with disease progression in PBC[108].

Table 1. Prognostic factors in primary biliary cholangitis

Prognostic factor	Prognostic significance	Strengths ✓ and limitations ✕
Demographics		
Male sex vs. female sex[91-93]	Associated with UDCA non-response (72% vs. 80%, p<0.005)[91] (Attributable to more advanced disease)[92] Increased HCC risk in male non-responders and male cirrhotic patients[93]	✕ Lack of external validation
Younger age at diagnosis[91,92]	Response rate in patients < 50 years old < 68% vs. 86% in aged > 60	✕ Lack of external validation
Symptomatic patients[15-19,94]	Fatigue: may be associated with worse transplant-free survival[94] Pruritus: conflicting data[15-19].	✕ Lack of standardization in symptom definitions
Serological markers		
Anti-gp210[95,96]	*Hepatic decompensation subtype* (associated with 6-fold increase in LT/death)	✕ Unclear whether prognostic impact is independent to UDCA response; needs replication in large-scale cohort
Anti-centromere[95,97]	Portal hypertension subtype	
Biochemical variables		
Bilirubin[78]	Surrogate marker of disease activity and stage	✓ Externally validated and considered the most robust markers of disease activity in PBC
Alkaline phosphatase[78]	Surrogate marker of disease activity	
Albumin[51,74,78]	Low levels are independent predictor of liver-related mortality	✕ May not be a realistic marker for the risk stratification of early-stage populations
AST and ALT[82,98,99]	Elevated levels (after 12 months UDCA) associated with worse survival/liver-related events	
APRI[76]	>0.54 associated with death/LT (HR 2.75)	✓ Externally validated ✓ Prognostic value additive to biochemical response
Markers of disease stage		
ELF	Each point increase associated with 3-fold increase of adverse events[100]	✕ Unclear whether marker of disease activity or stage ✕ Calibration not assessed ✕ Unclear cost-effectiveness compared to LFTs and TE
Liver stiffness measurement [101,102]	LSM progression associated with prognosis[101] kPa > 9.6 associated with decompensation, LT or death (HR 5.1; 95% CI 1.5-15.9)[101] Significantly improves newly introduced risk stratification models[102]	✕ Large inter-operator variability ✕ Poor discrimination in subtle changes in liver fibrosis ✕ Overestimates of fibrosis stage in cholestasis [103]
Histological parameters[24,82,104-107]	Advanced histological stages: associated with poor prognosis[24,104] Interface hepatitis: associated with the development of cirrhosis and LT or liver-related death and[74,107]	✕ Biopsy is an invasive and costly procedure ✓ Prognostic value additive prognostic to prognostic models
	Premature ductopenic variant associated with progressive disease[105] Baseline ductopenia (>50% loss) predicts histological progression[106]	✕ Lack of external validation

HCC, hepatocellular carcinoma; LT, liver transplantation; AST, aspartate aminotransferase; ALT, alanine aminotransferase; APRI, aspartate aminotransferase to platelet ratio; CI, confidence interval; HR, hazard ratio; ELF, enhanced liver fibrosis; LFTs, liver function tests; TE, transient elastography; LSM, liver stiffness measurement; kPa, kilopascal.

However, a decrease in GGT of 70% or normal levels within 6 months after start of UDCA treatment is associated with better prognosis[109,110]. In meta-analyses of 4845 patients *Lammers et al.* showed that both alkaline phosphatase (>2.0 x ULN) and bilirubin (> 1.0 x ULN) are independent predictors of liver transplantation and death[78]. Although outcomes were best predicted by biochemistry measured one year after initiation of UDCA, ALP and bilirubin measured at other time points remained strongly associated with clinical outcomes. Importantly, ALP levels held additive prognostic value to bilirubin and this effect was independent of sex, follow-up time, presenting age, UDCA treatment and disease stage. Thus, this landmark paper showed that ALP and bilirubin levels are strongly associated with long-term outcomes in PBC. Both are considered the most robustly validated markers of disease activity (ALP and bilirubin) and disease stage (bilirubin) in PBC. ALP and bilirubin are accepted to be "reasonably likely to predict clinical benefit" in PBC and are used as an endpoint in clinical trials[111].

In addition to bilirubin, albumin is an important predictor of transplant-free survival in PBC. Both low serum albumin and high bilirubin values are independent predictors of the development of cirrhosis and mortality[51,74,78,112,113]. *Ter Borg et al.* showed that stratifying patients according to biochemical disease stage based on the combination of bilirubin and albumin allows us to identify UDCA-treated PBC patients at baseline that are at low (both normal bilirubin and albumin), median (abnormal bilirubin or albumin) and high risk (abnormal bilirubin and albumin) of future clinical events[51]. The low-risk group had a liver-transplantation free survival comparable with general population. In contrast, the intermediate- and high-risk groups had a significantly worse survival with a 5-year transplant-free survival of ~82% and ~33%, respectively.

Elevation of aspartate aminotransferase (AST) and alanine aminotransferase (ALT) are associated with the level of lobular inflammation and necrosis, especially in a setting of elevated IgG[104]. Both elevated AST and ALT are strongly associated with long-term outcome[82,98,99]. The UK-PBC study group showed that elevated transaminases after 12 months of UDCA therapy were associated with an increased risk of a liver-related event[98].

The concept of biochemical response

Angulo et al. were the first to recognize that changes in biochemical parameters during UDCA treatment were associated with clinical outcome[114]. In a cohort of 180 UDCA-treated patients, they showed that patients with serum ALP >2 times the ULN after six months of therapy were more likely to encounter severe disease progression (11% vs. 33%, p<0.04). In 2006, a Spanish study found (n=192) that a 40% reduction of ALP after one year of treatment was associated with a similar survival as that of matched controls from a general population (*Barcelona criteria*). In contrast, the prognosis of those who did not meet these criteria was worse than that of a general population (relative risk for liver transplantation or death 7.47 (95% CI 1.87-29.78)). Since then, several response criteria have been proposed, all with different combinations of biochemical variables to capture incomplete response to UDCA and thereby identifying patients that are at risk of events: *Paris I, Rotterdam, Ehime, Toronto, Paris II and Momah/Lindor criteria*[80-82,99,106,109,110] (**Table 2**). Percentages of incomplete response vary between 24 and 52%, underlining that there is still a significant proportion of patients that may benefit from additional treatment. Most

criteria evaluate biochemical response after one year of UDCA. However, the optimal time point for biochemical evaluation has yet to be determined and it may already be possible to assess response to therapy after 6 months[115]. *Leuschner et al.* showed that approximately 80% of decrease of alkaline under UDCA treatment occurs within 6 months of UDCA therapy, suggesting that most criteria are best applied after at least 6 months of therapy[116]. The Paris-I criteria are accurate and thoroughly validated dichotomous criteria for stratifying patients into low- and high-risk categories for events[55,76,80,117]. However, the optimal response criteria may differ between patients and study populations. For example, Paris II criteria were designed for early stage disease patients[99]. Combined analyses of various proposed criteria showed they have independent prognostic significance, suggesting that none of these criteria is an optimal measure of response[118]. Furthermore, some criteria are mainly focused at the assessment of response to treatment and do not incorporate markers of disease severity or stage (e.g. albumin and/or bilirubin). These criteria may not sufficiently capture the baseline difference in survival that is associated with difference stages of disease[51]. Nonetheless, biochemical response criteria provide a readily available way to identify patients that are likely to benefit from additional therapies or clinical trials.

Age and gender

In PBC, young age at diagnosis and male sex have both been associated with a reduced chance of biochemical response to UDCA therapy. In a large cohort study the UK-PBC study group showed (n=2353) male patients had significantly lower rates of biochemical response than females (72% vs. 80%, $p<0.005$)[91]. In addition, the study group showed that there is an inverse correlation between age at diagnosis and attainment of biochemical response at 1 year of UDCA therapy. Attainment rates of response were < 30%, 68%, 86% and 90% in patients aged younger than 30, younger than 50, older than 60 and older than 70 at PBC diagnosis, respectively. In accordance with the latter study, another large, multicentre long-term follow up study (n=4355) found that young PBC patients (aged < 45) had significantly lower response rates to UDCA than their older counterparts (aged >65) (odds ratio 5.48; 95% CI, 3.92-7.67; P<.0001)), regardless of sex[120]. In addition, risk of liver transplant or death decreased significantly with age: HR for patients aged ≤ 35, 14.59 (95% CI, 9.66-22.02) versus HR for patients > 65 years, 1.39 (95% CI, 1.23-1.57) (P<.0001), also regardless of sex[120]. More recently, a large population-based study from Italy and Denmark showed that males had worse survival than females (HR for all-cause mortality 2.36 in the Italian and 3.04 in the Danish population)[25]. However, this study used an administrative registry which is often limited by availability of treatment data and underreporting of outcomes. In addition, the study data did not allow for a multivariable analysis with inclusion of liver biochemistry or disease severity. In summary, in clinical practice, male patients are likely to be at a higher risk of an unfavourable outcome, although it is uncertain whether sex is an independent risk factor or that this finding is explained by differences in age and disease stage at time of diagnosis. Young-presenting PBC patients may be an important subgroup with more aggressive disease.

Table 2. Proposed criteria for the assessment of biochemical response to UDCA in PBC

Criteria	Formula	Time point (months UDCA)	Clinical endpoint / Non-response, %	PBC group / n	Follow-up time	Remarks (strengths ✓ and limitations ✕)
Rochester/ Mayo, 1999[114]	ALP <2.0 × ULN	6	Liver transplant-free survival / Not reported	All / 180	Range 0.4-7.0 years	✕ Short FU time ✕ Small sample size ✕ Does not incorporate disease severity
Barcelona, 2006 [81]	>40% decrease of ALP or normalization	12	Liver transplant-free survival / 39	All /192	Range 1.5-14 years	✓ Matched survival with general population ✕ Does not incorporate disease severity ✕ 17 endpoints (9 died, 8 fulfilled criteria for liver transplantation (treatment failure))
Paris I, 2008[82]	ALP <3.0 × ULN and AST <2.0 × ULN and normalization of bilirubin	12	Liver transplant-free survival / 39	Advanced / 292	Median 5.3 (range 1.0-21.5) years	✓ Externally validated [55,76,80,98,99] ✓ Combination of multiple variables ✓ Systematic determination of cut-offs ✕ Limited to advanced disease stage
Rotterdam, 2009[80]	Normalization of abnormal bilirubin and/or albumin	12	Liver transplant-free survival / 24	All / 375	Median 9.7 (range 1.0-17.3) years	✓ Long-term FU, prospective large sample ✓ Incorporation of disease stage ✓ Included patients from general- (39) and university (7) hospitals
Ehime, 2009/2011[109,110]	≥70% decrease of γ-GT	6	Liver transplant-free survival / 52 [109] and 79 [110]	Asymptomatic / 83 109 and 134 110	Mean 5.2±4.4 years [109] and median 4.6 (range 0.8-24.3) years	✓ Systematic determination of cut-off ✕ Small sample size [109] ✕ Short FU time ✕ Does not incorporate disease severity
Toronto, 2010[106]	ALP ≤ 1.67 × ULN	24	Histological progression and liver transplant-free Survival / 43	All / 69	Mean 9.4 years	✓ Systematic determination of cut-off and optimal timing of assessment (1 or 2 years of therapy) ✓ Paired liver biopsy ✕ Small sample size ✕ Does not incorporate disease severity
Paris II, 2011[99]	ALP and AST ≤ 1.5 × and normalization of bilirubin	12	Liver transplant-free survival, ascites, variceal bleeding, encephalopathy, HCC / 52	Early PBC (Ludwig I and II) / 165	Median 7 (range 1.6-20.3) years	✓ Applicable in early disease stage presenting patients ✓ Systematic determination of cut-off ✓ Inclusion of multiple endpoints significant to prognostication in PBC ✕ Only 11 endpoints
Momah/ Lindor, 2012[119]	ALP ≤ 1.67 × and bilirubin ≤ 1mg/ dL	12	Development of varices, ascites, encephalopathy, liver transplantation or death /48	All / 73	Mean of 3 years	✓ Use of multiple endpoints significant to prognostication in PBC ✓ Systematic determination of cut-off ✕ Small sample size

UDCA, ursodeoxycholic acid; ALP, alkaline phosphatase; ULN, upper limit of normal; AST, aspartate aminotransferase; γ-GT, gamma-glutamyl transpeptidase; HCC, hepatocellular carcinoma.

Symptom profile

Nowadays, most PBC patients present without specific disease-related symptoms[11-14]. However, the large majority of patients (~95%) will develop these symptoms over time, of which fatigue and pruritus are the most prevalent ones[15-19]. Although pruritus and fatigue may significantly impact quality of life[20], studies regarding the impact of such symptoms on life expectancy yield conflicting results[15-19]. Some studies suggest that a symptomatic presentation may predict inadequate UDCA response and poor prognosis. *Quarnati et al.* showed (n=216, 141 asymptomatic and 75 with pruritus and/or fatigue), that symptomatic patients were younger, more often female, and were significantly less likely to respond to UDCA therapy (63% vs. 81%, p<0.01)[121]. In addition, symptomatic patients were more likely to develop cirrhosis and associated complications (31 vs. 13%, p<0.01). *Jones et al.* (n=136) reported that fatigued patients had significantly worse survival than matched controls, independent of UDCA response and disease severity (56% vs. 74%, p<0.001)[94]. However, other large studies have reported that absence of symptoms merely represents an earlier histological stage and better biochemical profile (lower bilirubin- and higher albumin levels)[12,17,122]. Furthermore, two large studies suggest that once symptoms develop, survival in asymptomatic-presenting patients is comparable to those initially presenting with symptoms[17,122]. Therefore, development or presence of symptoms may rather be a surrogate of disease stage than an independent factor impacting prognosis. The potential additive prognostic value of symptoms to existing risk stratification models is currently unknown and this may require further exploration in a prospective setting.

Serological markers

Anti-mitochondrial antibodies (AMA) are the serological hallmark of PBC and are present in approximately 95% of patients[22]. A large proportion of individuals positive for AMA in the absence of abnormal liver function tests (n=29) will eventually developed cholestatic liver function tests (83%) or symptoms of liver disease (75%)[18]. AMA positivity or AMA titre are however not associated with prognosis in PBC[123,124]. Antinuclear antibodies (ANA) are present in approximately 30% of patients with PBC. There are multiple staining patterns, but two distinct patterns are considered specific for PBC: a multinuclear dots pattern (Sp100) and perinuclear rims (anti-gp210). Antibodies against gp-210 have been associated with an unfavourable disease course with a 6-fold increase in the risk of disease progression to transplantation or death (*hepatic failure subtype*)[95,96]. The presence of anti-centromere antibodies has been associated with a *portal hypertension type* progression[95,97]. As such, ANAs may serve as early markers for a high risk of progressive disease[125]. Further exploration in prospective studies is warranted to assess whether ANA patterns are reliable markers of prognosis and whether the prognostic impact of ANAs is independent to UDCA response.

Serum markers of fibrosis

Serum markers of fibrosis provide an outcome on a continuous outcome scale, potentially providing more information than the categorized histological disease stages. The first marker suggested to be sensitive for progressive liver damage in PBC was serum hyaluronate[126]. Later, *Mayo et al.* showed that the ELF (enhanced liver fibrosis) score, a combination of serum concentrations of hyaluronic acid, procollagen III peptide and tissue inhibitor of metalloproteinase 1, is a fairly accurate non-invasive measure of disease

severity and a prognostic marker for clinical progression in PBC (Area Under the Receiver Operating Characteristic Curves (AUROC) of 0.68-0.78). In 161 patients with serial biannual liver biopsies and ELF measurements for a median of 7.3 years, they showed each point increase in ELF was associated with a 3-fold increase of cirrhotic complications, LT, or liver-related death[100]. Although ELF is easily obtainable and applicable in clinical practice, it may be less sensitive in patients with early disease. Moreover, whether the association between ELF and outcome is independent of biochemical response remains to be determined.

An AST to ALT ratio greater than 1.1 may be a marker of ongoing liver fibrosis. *Nyblom et al.* (n=121) found that the AST/ALT ratio was significantly higher in PBC patients with cirrhosis than in those without cirrhosis (0.8 versus 1.5, P<0.0001)[127]. However, the positive predictive value for cirrhosis of the different cut-offs of this ratio varied between 45-61%. The use of these cut-offs may thus result in unjust identification of cirrhosis. The performance of AST/ALT ratio has been evaluated in two independent studies. In an Italian study (n=120), the performance of AST/ALT ratio as a measure for the different stages of fibrosis was reported to be poor with AUROC scores of 0.53, 0.57 and 0.58 for fibrosis stage II, III and IV, respectively[128]. A large multicentre study in 2488 UDCA-treated patients from the Global PBC study group showed that AST/ALT ratio was associated with survival in univariable but not in multivariable analyses indicating that AST/ALT ratio is not strongly associated with outcome[98].

Finally, an aspartate aminotransferase-to platelet ratio index (APRI) > 0.54, as a surrogate for liver fibrosis and portal hypertension, is an important non-invasive marker and prognostic factor associated with cirrhotic complications, death and liver transplantation in PBC patients[71,76]. APRI is associated with outcome independent of response to treatment with UDCA and thus imparts additional prognostic value to existing biochemical response criteria[76]. An APRI of > 0.54 after one year of UDCA therapy is associated with an almost 3-fold increase in risk of death or liver transplantation.

Liver stiffness measurement

Liver stiffness measurement (LSM) with vibration-controlled transient elastography provides a simple measure of liver fibrosis stage, especially in severe fibrosis and cirrhosis[101,128,129]. In an Italian cohort study (n=120), LSM by transient elastography was better in identifying any grade of fibrosis and cirrhosis (AUROC 0.89, 0.92 and 0.99 for fibrosis stage II, III and IV, respectively) than non-invasive surrogate markers of fibrosis such as APRI (AUROC 0.66,0.67 and 0.84 for fibrosis stage II, III and IV, respectively) and the AST/ALT ratio (AUROC 0.53,0.57 and 0.58 for fibrosis stage II, III and IV, respectively). LSM also positively correlated with the Mayo risk score, indicating that it may be suitable as a prognostic measure in PBC[128]. Subsequently, *Corpechot et al* showed LSM values above 9.6 kPA carry a hazard of 5 for adverse outcomes (decompensation, liver transplantation or death)[101]. In addition, progression of liver stiffness at a cut-off of 2.1 kPa/year is associated with an increased risk of adverse outcomes[101]. Recent studies suggest that poor biochemical response is associated with higher rates of LSM progression and that LSM progression is able to predict clinical outcomes in PBC independently of UDCA response[101,130]. Moreover, preliminary data suggests that LSM significantly improves risk stratification of newly established prognostic scores[102]. However, transient elastography is

not uniformly available in all clinics, requires experience and may be unreliable in obese patients. Moreover, cholestasis can falsely increase LSM values resulting in inaccurate estimates of fibrosis severity[103]. These factors currently limit the possibilities of including LSM into prognostic tools for the general clinician.

Histological parameters

In the current clinical setting, in which liver biopsy is no longer recommended for standard diagnostic or management purposes[45,72], the significance of liver biopsy in prognostication is limited. In addition, the role of liver biopsy in risk stratification for the individual patient may be limited by sample variation. Widely divergent fibrosis scores have been observed when samples are taken simultaneously from different areas in of the liver in PBC[131]. Still, in the setting of risk stratification, liver biopsy may have significance in the determination of underlying features that explain persistent incomplete biochemical response. Advanced histological stages (extensive fibrosis or cirrhosis) of disease predict poor prognosis[24,104]. In addition, baseline ductopenia predicts histological stage progression[106]. Another important histological feature is interface hepatitis which is independently associated with the development of cirrhosis and liver transplantation or liver-related death and imparts additive prognostic impact on accuracy of established prognostic models in PBC[74,107]. The rare premature ductopenic variant of PBC – clinically characterized by severe pruritus, progressive icteric cholestasis, incomplete response to UDCA, and histologically by bile duct loss without significant fibrosis or cirrhosis – is associated with rapid progression to end-stage liver disease requiring liver transplantation[105].

PREDICTION MODELS

Previously proposed biochemical response criteria in PBC provide an easy-to-use tool for clinical practice to identify patients at risk of adverse outcome. However, they provide a crude dichotomization of risk into high or low risk, resulting in loss of predictive information. Especially patients with biochemical values that are close to the criteria's threshold(s) may be unjustly characterized as low risk instead of high risk and vice versa. Moreover, dichotomous criteria do not allow clinicians to provide their patients with individualized prediction of prognosis[55,118]. Prediction models partly encompass the aforementioned shortcomings of dichotomous criteria.

The purpose of prediction models is to predict future events. The quality of these models depends on performance, validation, and generalizability, as well as their ease of use, and their applicability at various stages of the disease. These models may aid physicians in identification of patients at risk of future event and in clinical decision making. Three levels of application are recognized: the individual patient level, the group level, and trial settings[132]. For an individual patient, a prediction model may provide important information on what to expect from their disease regarding complications or survival. On a group level, prediction models measure epidemiological differences between groups. The prediction, a value between 0 and 100%, allows us to stratify patients into risk groups. Such strata may include low, moderate, and high-risk groups, which may be used to recommend changes in clinical management. In a trial setting, risk stratification may be

helpful in selecting the in- and exclusion criteria as well as in determining an appropriate sample size.

Models predicting transplant-free survival

Before the introduction of UDCA as the mainstay treatment, *Christensen et al.* were the first to fit a multivariable model with factors associated with clinical outcome (n=248). They reported that increased bilirubin, older age and cirrhosis were independently associated with a poor prognosis[133]. Over the subsequent decades, several prediction models have been proposed, all comprising both strengths and limitations (**Table 3**). In 1989, *Dickson et al.* introduced the Mayo Risk Score (MRS) for predicting prognosis in PBC patients based on readily available clinical variables[134]. The MRS is the most frequently used model to predict survival in PBC and was extensively cross-validated in other cohorts[134-137]. In a cohort of 312 untreated patients, bilirubin, albumin, prothrombin time, age and severity of oedema were identified as independent predictors of prognosis. The model predicts survival up to 7 years of follow-up and was originally developed for the assessment of the ideal timing of liver transplantation. Later, it was adjusted to a model that could be used at any time point during follow-up to predict prognosis 2 years after a patient visit[135]. As the MRS does not include histological parameters, a major advantage as compared to previously proposed models was its use of readily available variables, although oedema may be considered a subjective parameter. Data regarding the prognostic performance of the MRS in UDCA-treated patients is conflicting, with studies showing that the MRS continues to stratify patients into high- and low risk groups[114,138] and studies showing that survival of patients using UDCA is significantly better than predicted by the Mayo Risk Score[51,79,81]. More general models used in PBC are the albumin-bilirubin score (ALBI), Child-Turcotte-Pugh (CTP) score, and the Model for End-stage Liver Disease (MELD) (**Table 3**)[139-141]. The MELD score is currently predominantly used to allocate patients for liver transplantation and as an endpoint in clinical trials. Recent studies suggest that allocation based on MELD score results in higher waitlist mortality in PBC compared to most other aetiologies of end-stage liver disease[142,143].

Continuous models predicting transplant-free survival

The aforementioned risk prediction models were mostly developed for end-stage PBC. They primarily focus on short-term survival, and do not incorporate biochemical response or disease activity (e.g. ALP). Therefore, in this era with mostly UDCA-treated and early-disease stage presenting patients, these models may not be sufficient. Recently, two new models were proposed that address these shortcomings. In 2015, the GLOBE score was introduced (www.globalpbc.com). Recently, a mobile application was launched to further facilitate the use of the GLOBE score in clinical practice (**Figure 1**). This model was constructed using a derivation cohort of 2488- and a validation cohort of 1634 UDCA-treated patients, and comprises age, bilirubin, albumin, alkaline phosphatase, and platelet count after 1 year of UDCA treatment as independent predictors of liver transplantation or death in UDCA-treated patients[55]. Also introduced in 2015, the UK-PBC risk score (www.uk-pbc.com) was developed in a nation- wide cohort of 1916 patients (derivation cohort) and validated in a cohort of 1249 UDCA-treated PBC patients, this score predicts the risk of liver-related death or liver transplantation with a model comprising baseline albumin and platelet count, as well as bilirubin, transaminases, and alkaline phosphatase after 1 year

of UDCA therapy[98]. With C-statistics of >0.8, both these models have superior predictive performances for incomplete response to UDCA compared with previously proposed dichotomous criteria[55,98].

Both scores use variables on a continuous scale, which has the benefit of preserving a greater amount of predictive information. The outcomes of the scores have a continuous scale too and thus provide gradual individualized estimates of survival, rather than crude differentiation into high- and low-risk groups. Importantly, they take into account biochemical response to UDCA by incorporating biochemistry after one year of therapy, thus combining predictive information of both disease severity and response to treatment. Therefore, these models are better able to accurately predict survival than previously introduced models and biochemical response criteria.

The GLOBE score was initially constructed to estimate the risk of death or liver transplantation after 1 year of UDCA therapy. However, recent analyses indicate that the score can also be used to risk stratify UDCA-treated patients at later points in time[144-146]. An advantage of the GLOBE score is its use of age-specific thresholds beyond which survival significantly deviates from a sex and age-matched general population. The score presents the median survival of this matched population at 3, 5, 10 and 15 years. This provides clinicians and patients with a clear sense of the prognosis.

Integration of continuous models into clinical practice

Both the GLOBE and UK-PBC risk scores have been proposed in the recent EASL Clinical Practice Guidelines as a 'tool for the selection of patients for second-line therapies, either in routine care or in therapeutic research; and for the stratification of risks for clinical trials in order to account for the prognostic disparity between patients at inclusion'[9]. While both scores were validated in a Western population, recent studies from China and Japan indicate that both scores provide reliable estimates in populations of other ethnicities as well[147-149].

Table 3. Proposed prediction models for estimating prognosis in primary biliary cholangitis in the pre-UDCA era

Model	Formula	Strengths ✓	Limitations[a] ✕
Yale Model[24]	0.037*1.037^{age}+0.74*2.1*(hepatomegaly; 0=no, 1=yes)−1.34*0.26*(portal fibrosis; 0=no, 1=yes)+0.82*2.26*(BILI≥5 [mg/dl]; 0=no, 1=yes)−0.73*0.48*(BILI<1.5 [mg/dl]; 0=no, 1=yes)		✕ Invasive, biopsy needed
European model, 1985[150]	R=2.51*log (BILI (µmol/L))+0.0069*exp(age(years)-20)/10)+0.88*(cirrhosis; 0=no, 1=yes)−0.05*(ALB[g/L])+0.68*(central cholestasis; 0=no, 1=yes)+0.52*(azathioprine; 0=no, 1=yes)		✕ Invasive, biopsy needed ✕ Short-term survival estimates only
Mayo Model, 1989 [134]	R=0.039*(age(years))+0.871*log (BILI[mg/dL])+2.38*log (PT (sec))+0.859*(edema)−2.53*log (ALB[g/dL])	✓ Readily available variables ✓ Thoroughly validated ✓ Provides individual survival estimates up to 7 years of FU	✕ Oedema as a subjective variable [a] ✕ Inclusion of patients from tertiary centres only [a] ✕ Uncertain reliability in UDCA treated patients [51,79,81] [a]
European model, 1993[151]	R=2.53*log (BILI [µmol/L]−1.53)+1.39*(ascites)−0.085*(ALB [g/L]−34.3)+0.040*(age(years)−55)+0.65*(gastro intestinal bleeding)	✓ Readily available variables ✓ Can be used at any time point during follow-up	
Mayo Model, 1994 [135]	R=0.051*(age(years))+1.209*log (BILI[mg/dL])+2.754*log (PT(sec))+0.675*(edema)−3.304*log (ALB[g/dL])	✓ Readily available variables ✓ Thoroughly validated ✓ Applicable any time point during follow-up	✕[a] ✕ Short term survival estimates (up to 2 years)
Mayo model, 2000 [136]	*Points* / *Age (years)* / *Bilirubin (mg/dL)* / *Albumin (g/dL)* / *Prothrombin time (s)* / *Oedema* 0 / <38 / <1 / > 4.1 1 / 38-62 / 1-1.7 / 2.8-4.1 / ≤ ULN / No 2 / >62 / 1.7-6.4 / <2.8 / >ULN / Yes 3 / / >6.4 / / /	✓ Readily available variables ✓ Applicable at any time point during follow-up ✓ Easy to use, quick estimate	✕[a] ✕ Use of risk-categories resulting in loss of predictive information/no individualized survival estimates
MELD score, 2001 [141,152]	R=3.8*log (BILI [mg/dL])+11.2*log (INR)+9.6*log (creatinin[mg/dL])	✓ May not be suitable in early stages of disease	✕ Not PBC-specific ✕ Unclear role in early risk stratification ✕ May underestimate risk of death [142,143]
Newcastle model, 2002[12]	0.0742*age+0.2610*ln(ALP/ULN)−2.53*(ALB/LLN)+0.195*ln(BILI/ULN)	✓ Inclusion of ALP as a marker of disease activity	
ALBI score[b] [139,140]	-0.085*(ALB[g/L])+0.66*ln(BILI*[µmol/L])	✓ Easy to use, quick estimate	✕ Use of risk-categories resulting in loss of predictive information/no individualized survival estimates [b] ✕ Only validated in Chinese cohort with small sample size ✕ Low number of patients in high-risk group (n=7)

BILI, bilirubin; PT, prothrombin time; ALB, albumin; ULN, upper limit of normal; LLN, lower limit of normal; INR, international normalized ratio; ALP, alkaline phosphatase; ALBI, Albumin-bilirubin; LT, liver transplantation.

[a] These limitations also apply to the other versions of the Mayo risk model in this table.

[b] The score divides patients into 3 groups of risk according to cut-offs of grade 1 (≤ −2.60), grade 2 (> −2.60 to −1.39), and grade 3 (> −1.39).[139,140]

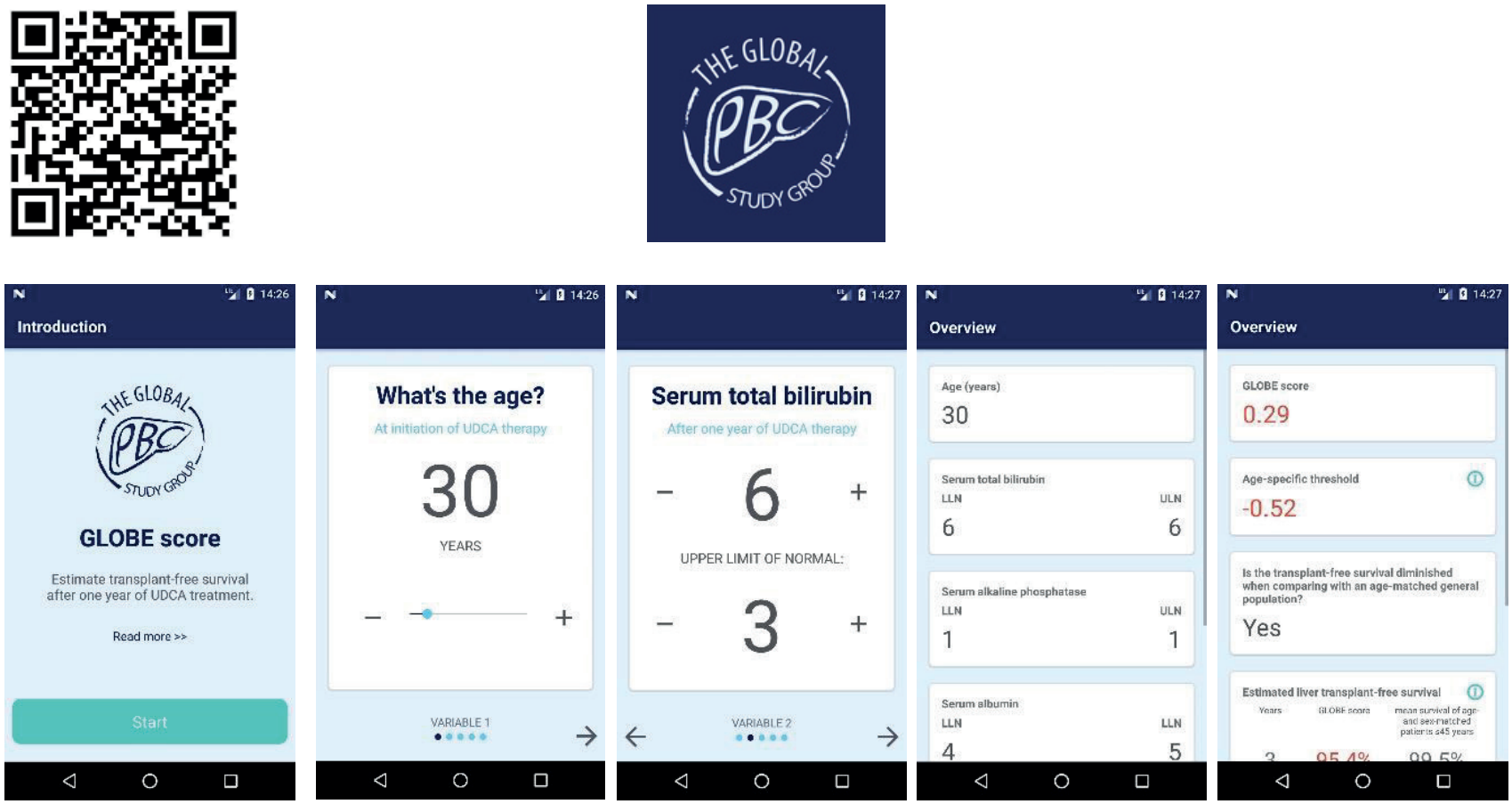


Figure 1 - The GLOBE score app. On 26th of April 2018 the GLOBE score app became available for Android and iOS devices. This application facilitates clinicians to implement the GLOBE score in daily clinical practice. The application was developed in collaboration with *Synappz medical apps.*

How should one interpret the overall survival (GLOBE) or the risk of end-stage liver disease (UK-PBC) estimated by these scores? No clear-cut thresholds have been suggested to determine the need for second-line therapies or de-escalation of follow-up back to primary care. The identification of specific risk thresholds would oversimplify clinical decision-making. Thresholds such as the ones outlined above are patient-specific and can even change over time within one and the same patient, influenced by the patient profile (e.g. age, fibrosis stage, pruritus severity), side effects, and the efficacy of a specific agent. 'Risk' must therefore be contextualized. For example, a 10-year risk of end-stage liver disease of 20% would be undesirable for a 40-years old patient without comorbidities, but it might be acceptable for a 75-years old patient with other life-shortening diseases. In other words, treatment targets should be determined by the effectiveness of the treatment, its side-effect profile, and the extent to which the individual patient would benefit from the risk reduction. An additional approach to interpret the risk estimated by the scores is a comparison with survival in a population not affected by the disease, such as provided by the GLOBE score. Ideally, the survival estimate in the general population might serve as a benchmark for the identification of the level of risk acceptable for the disease population.

Part B - Primary Sclerosing Cholangitis

Primary sclerosing cholangitis (PSC) is an infrequent, chronic, slowly progressive cholestatic liver disease. The disease is characterized by intrahepatic and extrahepatic fibro-inflammatory biliary strictures, chronic cholestasis and development of cirrhosis[153-155]. PSC mainly affects relatively young males and is strongly associated with inflammatory bowel disease (IBD), in particular ulcerative colitis (UC). The aetiology of PSC remains obscure. The clinical course of PSC is highly variable and can be complicated by recurrent cholangitis and development of dominant benign and malignant biliary strictures. In addition, there is an increased risk for other malignancies including hepatocellular, gallbladder and colorectal carcinoma. For patients that progress to end-stage liver disease, the only curative treatment is liver transplantation, with an excellent survival rate of approximately 80% at 5 years[156-158]. However, post liver transplantation patients are at a high risk of recurrence, which may require retransplantation[159-163].

CLINICAL PRESENTATION AND DIAGNOSIS

In contrast to PBC, approximately 50% of PSC patients may be symptomatic at disease presentation[156]. Symptoms include fatigue, pruritus and abdominal pain. In addition, patients may present with weight loss, jaundice and episodic fever and chills[164]. A small subgroup presents with cirrhosis-related complications such as ascites or variceal bleeding, or even more rare, with hepatobiliary malignancy.

A diagnosis of PSC is usually made in patients with repeated unexplained elevation of alkaline phosphatase and/or GGT (> 6 months) in combination with characteristic findings of multifocal bile duct strictures and segmental dilatations on biliary imaging (**Figure 1**). These strictures and dilatations lead to a characteristic beaded appearance on magnetic resonance cholangiopancreatography (MRCP) or endoscopic retrograde cholangio-pancreatography (ERCP)[156]. Other frequent cholangiography findings are intra and/or extrahepatic biliary stones, atrophy and hypertrophy, usually of the right and left liver lobe, respectively, and an increased gallbladder volume. Secondary sclerosing cholangitis and other cholestatic liver disease should be excluded[165]. Different phenotype variants of PSC can be recognized. Classic large-duct PSC is characterized by multifocal intra- and extrahepatic strictures with ductal dilatation, while small-duct PSC has typical histological changes in absence of radiographic abnormalities[156]. Typical histological changes are nonsuppurative paucicellular cholangitis, periductal fibrosis, ductular reaction and ductopenia. The autoimmune hepatitis component in PSC forms a continuous spectrum with higher autoimmune activity in younger patients[166]. PSC and autoimmune hepatitis (AIH) can occur simultaneously or concurrently and this so-called PSC with features of auto-immune hepatitis/PSC-AIH overlap syndrome should always be considered when confronted with cases with an atypical course or insufficient response to treatment[167].

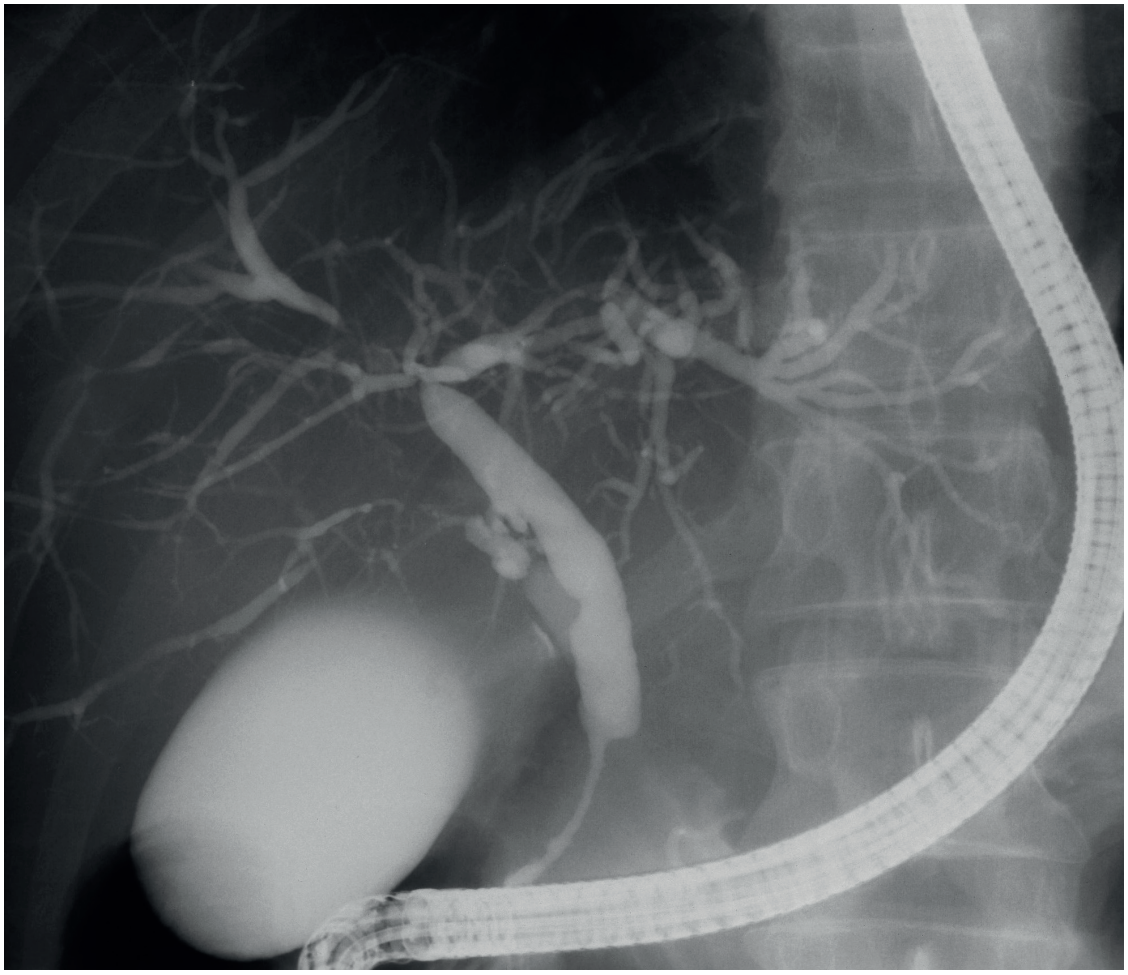


Figure 1. ERCP showing an irregular contour and a marked, long distal stenosis of the common bile duct, and strictures and caliber changes of intrahepatic bile ducts, particularly in the right liver lobe.

EPIDEMIOLOGY

In PSC, the male-to-female ratio is 2 to 1 and patients typically present around the age of 40. A systemic review by *Molodecky et al*. including 8 studies from North America and European countries, estimated a PSC incidence rate of 0.77 (0.45-1.09) per 100,000 person-years. However, when excluding the two non-population based studies, the incidence rate increased to 1.00 (0.82-1.17)[168]. Interestingly, all studies reported an increase in the incidence rate over time. Similar results were reported by *Boonstra et al.* who in a systematic review of 11 studies from North America and Europe, which include the 8 studies covered by *Moledecky et al*, reported an incidence rate ranging from 0 to 1.3 per 100,000 inhabitants/year[27]. The reported prevalence rate was 0–16.2 per 100,000 inhabitants. This study also observed an increase in prevalence and incidence rates over time. An in-depth population-based study from the Netherlands, covering approximately 50% of the Dutch population, reported an incidence rate of 0.5 (0.4 for women and 0.6 for

men) and a point prevalence of 6.0 per 100.000[169]. This implies an estimated 1000-1200 PSC patients are living in the Netherlands.

PSC AND INFLAMMATORY BOWEL DISEASE

An interesting aspect of PSC is its strong association with inflammatory bowel disease (IBD). IBD is present in 60-80% of PSC patients in Western countries[153,164]. Conversely, PSC occurs in 2.5-8% of IBD cases[170,171], while a recent report shows rates as high as 14%[172]. While IBD can be diagnosed during any stage of PSC[153,173-175], in a majority of cases IBD precedes the development of PSC. The vast majority of PSC patients suffering from IBD have ulcerative colitis (UC), accounting for up to 80% of IBD cases in PSC[175,176]. Approximately 10% of IBD in PSC is defined as Crohn's disease (CD), and the remaining 10% is indeterminate[175,177]. An IBD diagnosis in PSC can be made years before a PSC diagnosis, at the time of PSC diagnosis, and throughout the course of PSC[178]. Patients can develop IBD after LT for PSC[179]. Conversely, PSC can occur years after colectomy for IBD[170,180,181]. De novo IBD after LT develops in about 18% of PSC patients[182]. PSC-related UC phenotype differs from the usual course of UC. It follows a clinically mild course and is characterized by a higher prevalence of pancolitis with backwash ileitis, and rectal sparing[175,182]. Therefore, IBD in PSC patients can be considered a distinct clinical entity as compared to classical UC or Crohn's disease.

AETIOLOGY AND PATHOPHYSIOLOGY

Akin to PBC, PSC's pathogenesis has not been elucidated, but it is thought to result from an interplay of genetic susceptibility and environmental triggers. This may lead to an aberrant immune response resulting in cholangitis, fibro-inflammatory biliary strictures, and finally obliteration, destruction and loss of bile ducts. This process is associated with variable portal and periportal inflammation, and progressive fibrosis culminating in cirrhosis. Histologically, a characteristic but not constant finding is obliterative concentric periductal biliary fibrosis, so-called onionskinning (**Figure 2**)[183].

Linked to PSC's close association with UC, smoking appears protective[184]. Coffee consumptions has been suggested to be protective as well[184]. Changes in microbiome with reduced bacterial diversity and altered taxonomy appear a possible pathogenetic factor in PSC[185]. Some data suggests use of hormonal contraception may be associated with development of PSC[184]. Sibling studies in PBC have shown that first-degree relatives are at higher risk of developing PSC, indicating a genetic component in the development of the disease[186]. In addition, GWAS have identified different risk loci for PSC[187-190]. Interestingly, first-degree relatives of PSC patients without IBD are also at an increased risk of ulcerative colitis, indicating possible shared genetic susceptibility factors for PSC and UC. The close association between PSC and inflammatory bowel disease support the hypothesis of an underlying autoimmune process. Further evidence for an auto-immune origin comes from a comparison of genetic risk between PSC and other auto-immune diseases such as coeliac disease, rheumatoid arthritis or type 1 diabetes, which has shown

a 50% genetic overlap[189]. Several other observations support immune-mediated bile duct injury as a key mechanism in the pathogenesis of PBC. Immunoregulatory changes reported include: elevation of IgM level[191]; presence of auto-antibodies with titres similar to those observed in autoimmune hepatitis, including anti-smooth muscle and antinuclear antibodies in approximately 75% of patients[192]; an increase in the number of CD4-positive T cells in the liver[193] and aberrant HLA class II expression on cholangiocytes[194]. Interestingly, cholangiocytes in PSC express antigens that cross-react with epithelial cells in the colon[195]. In addition to immunoregulatory changes, an interesting observation is the similarity between PSC and other pathways leading to biliary injury (e.g. ischemia, infection or toxin-related) on ERCP or MRCP as well as on histological assessment [165,154]. This finding may suggest a common pathway in pathogenesis of bile duct injury.

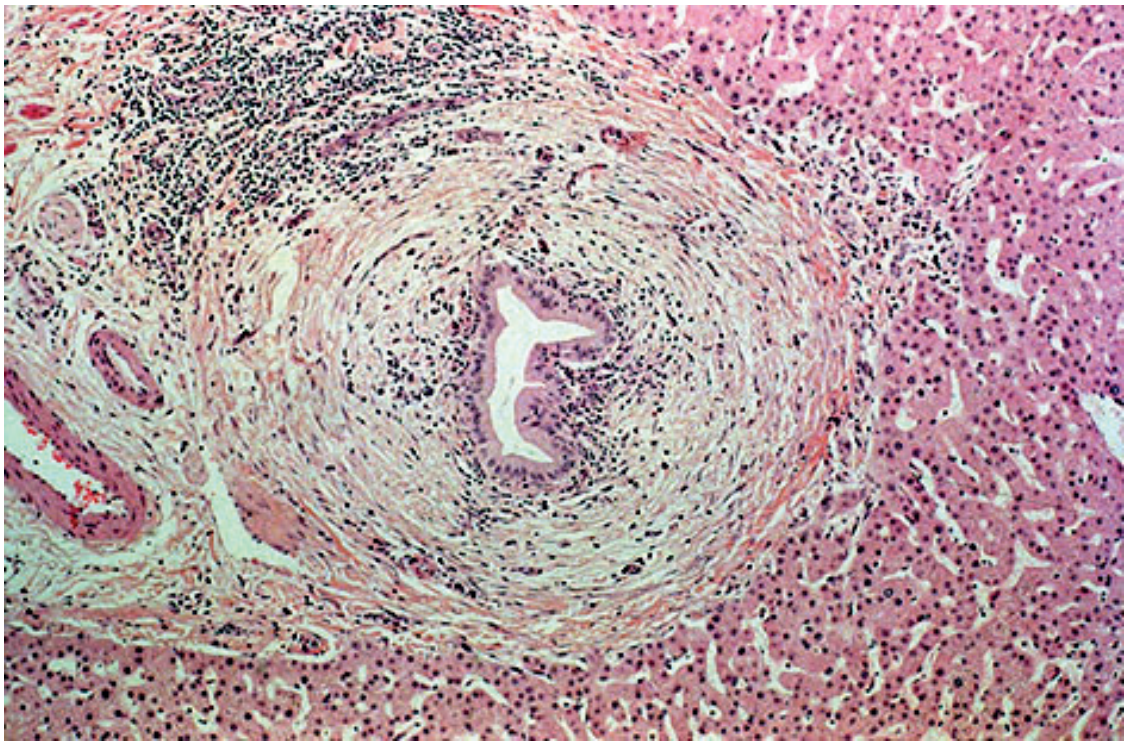


Figure 2 showing portal tract of the liver with lymphocytic inflammatory infiltrate and a bile duct with extensive periductular "onion-skin" fibrosis, and lymphocytes focally invading the biliary epithelium.

Several hypothesis specific to PSC's pathogenesis have been proposed among which the toxic bile acid hypothesis, and the hypothesis that microbiota translocated from the gut to the liver may trigger an aberrant response[196,197]. Another hypothesis is derived from the close association between PSC and IBD and considers PSC as the hepatobiliary manifestation of inflammatory bowel disease[154]. This theory assumes that IBD is a likely omnipresent underlying or co-existing, but often asymptomatic disorder. This is supported by the observation that approximately 89% of PSC patients may have a

colon biopsy compatible with IBD, while only 47% may have endoscopic signs of colonic inflammation[198] Central to this so-called 'dual homing hypothesis' is an immunological interaction between the gut and liver in which memory T-cells, expressing gut-homing markers, may migrate to the liver via aberrantly expressed mucosal adhesion molecules such as MAdCAM-1[199,200].

THERAPIES

In contrast to PBC, no medical treatment with the potential to change the natural course of the disease has been identified for PSC. The search for new therapies has been hampered by uncertainty regarding the pathogenesis of the disease as well as the disease pathways that affect the disease progression. These gaps in our knowledge of PSC has frustrated the development of disease-specific agents. In addition, the slow-progressive nature of PSC and the low-event rate make it almost impossible to study the impact of new drugs on hard clinical endpoints within the timespan of a clinical trial. Finally, the heterogenous nature of PSC, characterized by a highly variable rate of progression, the varying and unpredictable occurrence of hepatobiliary malignancies and of episodes of bacterial cholangitis and variability in the degree of cholestasis associated with development of dominant benign bile duct strictures, and concurrent inflammatory bowel disease in itself requiring anti-inflammatory and/or immune base treatment, markedly complicate the search for an effective drug treatment for PSC.

Ursodeoxycholic acid

Over the past decades the potential benefits of ursodeoxycholic acid (UDCA) in PSC have been extensively studied. UDCA has a number of potential beneficial mechanisms in protecting the liver from further biliary injury. UDCA is a hydrophilic bile acid and may therefore render bile more hydrophilic resulting in cytoprotective effects[201,202]. A more direct effect may be by inducing choleresis by stimulating hepato-biliary secretion[203-206]. UDCA may also have an immunomodulatory effect as well as a protective effect against bile acid-induced apoptosis of the cholangiocytes[206]. Finally, UDCA has been suggested to have chemoprotective properties on several tumour cell lines, including colon cancer[207]. Despite unequivocal biochemical improvements (i.e. reduction in ALP and other liver enzymes) reported with UDCA, a parallel improvement in survival or other hard clinical end-points has not been documented[208-213]. Moreover, high dose UDCA (28-30 mg/kg/day) has been associated with poorer clinical outcomes compared to placebo[209]. Nonetheless, patients treated with UDCA whose ALP levels decrease have a longer survival than patients without a decrease[214]. In a systematic review of 8 randomized UDCA trials (592 patients) no significant reduction of risk of death, treatment failure, or liver histology was observed[215]. However, most trials conducted lacked the statistical power to evaluate hard clinical endpoints. In fact, two trials on UDCA with hard clinical endpoints showed event rates of 3-4% within 5 years of follow-up[209,213]. In addition, duration of UDCA therapy in most trials is short (<3 years). This might not be appropriate to detect a significant effect in a slow-progressive disease like PSC. Finally, most patients in these trials had advanced histological stage with substantial fibrosis. As such we do not know the effect of UDCA on early- or moderate stage PSC. Still, despite a lack of clinical proof of its ability to alter

the course of the disease, UDCA is a widely used anti-cholestatic treatment in PSC. The recommended dosage is 13-15 mg/kg/day.

Potential new therapies

In addition to UDCA, many other immunosuppressive and anti-inflammatory drug therapies have been studied including glucocorticoids, cyclosporine, methotrexate, vancomycin, azathioprine and 6-mercaptopurine, tacrolimus, D-penicillamine, etanercept and OCA. There is no evidence that these drugs modify the disease course in PSC. Vedolizumab, a monoclonal antibody, which has been shown to be effective in the treatment of PSC-IBD[216,217], has been studied in PSC as well. These studies reported conflicting results of changes in biochemical markers of liver disease activity[216,218,219]. Dose dependant improvement in serum levels of ALP, AST, and other liver enzymes has been reported in patients treated with the UDCA homologue nor-UDCA, a drug with promising anti-fibrotic, anti-inflammatory and anti-cholestatic effects in a PSC mouse model[220]. Nor-UDCA is now being tested in a phase 3 clinical randomized trial. Recently, the results of a randomized, placebo-controlled, phase II study of obeticholic acid were published, reporting serum ALP was significantly reduced with OCA 5-10 mg versus placebo after 24 weeks treatment[221]. However, 60-67% of patients treated with OCA versus 46% in the placebo group reported pruritus.

DISEASE COURSE AND PROGNOSIS

The clinical course of PSC is highly variable, but in general the prognosis is grim with a reported median time of 10-15 years from diagnosis to transplantation or death. These results, however, were published by tertiary centers, and therefore were likely subject to selection bias[154,164,222,223]. Indeed, a much more reliable estimate was derived from a large population-based study covering 50% of the Dutch population. This key study reported a much more favourable median transplantation-free survival of approximately 21 years[169].

Malignancy risk

Aside from progression of liver disease and associated complications, the higher incidence of hepatobiliary malignancies may significantly impact disease course and prognosis[169,223,224]. The most common hepatobiliary malignancy in PSC is cholangiocarcinoma (CCA), which develops in 6-36% of PSC patients[164,224-226] and greatly affects life expectancy. Patients with PSC and UC have a higher risk of colorectal carcinoma (CRC) than patients with UC alone and may develop CRC at a younger age[169,227-229]. Results from studies that assess the risk of CRC in PSC-CD patients is conflicting[223,228,230-234]. CRC-risk is increased after liver transplantation; however, this risk may be attributable to IBD duration[235-237].

RISK STRATIFICATION

Risk stratification plays an important role in PSC. This is due the aforementioned problems in the development of disease-specific agents, the slow-progressive and highly variable nature of the disease, the low-event rate, and unfavourable survival rates in patients

who develop hepatobiliary malignancy[169,223,224]. The definition of specific risk profiles is of importance for the identification of PSC subgroups that may have an unfavourable disease course, or for selection of patients for clinical trials. In addition, risk stratification helps the clinician in patient counselling. There is a great need for surrogate endpoints for disease progression in PSC akin to those in PBCs[78,111]. While personalized risk assessment in PBC has improved greatly in recent years, the heterogeneous disease course of PSC and the lack of a therapeutic drug have made the development of prediction models for PSC difficult. Still, over the last few years large multicentre cohort studies, such as performed by the International PSC Study Group, have improved our knowledge regarding factors associated with progressive disease. Previously proposed risk prediction models in PSC were designed in late-stage disease patients and were not discriminatory in early disease stage patients. Conversely, the introduction of newly developed prognostic models for PSC, namely the Amsterdam-Oxford score, UK-PSC score and Primary Sclerosing Cholangitis Risk Estimate Tool (PREsTo) enable to quantify the risk of future events in individual patients, each with their own strengths and limitations[238-240]. These newly introduced risk prediction models permit earlier long-term and short-term risk estimates. Consequently, they allow clinicians to identify patients that are likely to suffer a more progressive course at an early stage, and thereby target diagnostic procedures and interventions. The prognostic index of these risk prediction models, if strongly associated with disease progression and/or hard clinical endpoints, could serve as a surrogate endpoint[241]. The use of the newly developed risk prediction models as surrogate endpoints, however, requires validation.

Phenotype variants of PSC

Approximately 6-16% of PSC patients have the small-duct phenotype variant[242]. Of these patients, eventually 20% will progress to large-duct PSC within a median of 7.4 years[243]. Small-duct PSC has been associated with a lower risk of CCA and better transplant-free survival[243,244]. These results have recently been robustly validated by the international PSC Study group in a study of 7121 patients. This study found that small-duct PSC retained a better transplant-free survival in multivariable analyses, independent of sex, IBD type or timing of PSC diagnosis[244]. Data regarding PSC with features of autoimmune hepatitis are conflicting. *Floreani et al.* reported PSC-AIH patients are younger, have more pronounced elevations in transaminases and IgG, and may benefit from UDCA and immunosuppression, and may have better transplant-free survival rates[245]. These results are contradicted by *Weismuller et al.* who found, in a cohort of 7121 PSC patients (470 PSC-AIH variant), that there were no differences in survival between PSC-AIH and classical PSC. However, a lower risk of CCA was observed[244]. Finally, IgG4 elevation is observed in 9-15% of PSC patients[246-250]. Although these patients present with more significant elevations in liver biochemistry and Mayo risk score, and a high prevalence of cirrhosis, the impact on clinical outcome is unclear[246-249]. One study in 127 patients (12 with elevated IgG4) found that the transplant-free survival for individuals with elevated IgG4 was significantly shorter than that for individuals with normal IgG4 levels (1.7 vs 6.5 years). However, these results could not be replicated in another study that showed increased IgG4 was not associated with an increased risk of CCA, liver transplantation or liver-related death[250]. Further studies that assess IgG4 as a risk stratifier are therefore required.

Inflammatory Bowel Disease in PSC

The presence of IBD has been robustly validated as a risk stratifier in PSC[251]. Progressive PSC requiring liver transplantation is associated with a milder course of UC (reduced disease activity, less use of steroids, azathioprine, and surgery) whereas patients with a less severe course of PSC often had a more severe course of UC[251,252]. In accord with these findings, *Naveneethan et al*. reported PSC patients with severe colitis had lower rates of liver transplantation[253]. Furthermore, colectomy prior to PSC diagnosis is associated with a decreased risk of liver transplantation or death[254]. Finally, patients with severe PSC requiring liver transplantation that have concomitant UC have less severe UC and lower rates of colectomy. These findings indicate that there might be a reciprocal influence on disease course between PSC and IBD. Probably the most robust assessment of the impact of IBD on disease course has been performed by *Weismuller et al*. who evaluated the impact of IBD on LT-free survival as a time-dependent covariate. In a cohort of 7121 patients, they showed that patients with Crohn's disease (CD) had a better transplant-free survival and lower risk of hepatobiliary malignancy development relative to UC, irrespective of sex and PSC phenotype[244]. No differences in terms of survival or hepatobiliary malignancy risk were observed between PSC patients with CD versus those without IBD. Ulcerative colitis and development of UC following PSC diagnosis were associated with an increased risk of liver transplantation or death and of hepatobiliary malignancy as compared to PSC-CD and PSC without IBD. The presence of IBD, in particular UC, should therefore be considered an important prognostic stratifier.

Age and sex

Advanced age at diagnosis has invariably been associated with worse survival. Age as a component of most risk prediction models in PSC[164,222,255-260] and was recently confirmed as an important prognostic factor in a large multicentre study of > 7000 PSC patients[261]. Studies suggest female PSC patients have a more favourable disease course as compared to male patients, with lower rates of liver transplantation or death and a lower risk of malignancies. Interestingly, a recent study in IBD patients showed that approximately 2/3 of newly diagnosed asymptomatic PSC patients are female. The better prognosis for females may be explained by a diagnosis of PSC in earlier, asymptomatic stages[171]. In addition, females more often present with non-classical PSC sub-phenotypes than male patients; these phenotypes are associated with a better prognosis[244]. Moreover, *Weismuller et al* showed[244], in a large multicenter study, that IBD phenotype overrides the prognostic impact of patient sex, i.e. differences in transplant-free survival and malignancy risk disappear when males and females are matched for IBD phenotype.

Symptom profile

The most commonly reported symptoms associated with PSC are pruritus and fatigue. Other symptoms include abdominal pain, chills, fever, jaundice and weight loss[262,263]. These symptoms may severely impact on quality of life and increased attention is being given to accurately capture the impact of these symptoms on patients' well-being[264,265]. Symptoms are not incorporated in any of the risk-prediction models for PSC. However, symptomatic presentations have been associated with worse transplant-free survival and an increased risk of malignancies[222,223].

Liver biochemistry

In recent years increased attention arose for risk stratification based on readily available, simple biochemical variables, either separately or combined in a prognostic model. A wide range of biochemical variables has been studied, including alkaline phosphatase (ALP), aspartate aminotransaminase (AST), alanine aminotransferase (ALT), albumin, gamma-glutamyl transpeptidase (GGT), and platelet count.

Traditionally, bilirubin and albumin have been considered important markers of liver-related disease progression in PSC. Bilirubin serum levels reflect the severity of biliary obstruction and of liver function impairment, and are part of most prognostic scores in PSC, including traditional scores like the Child-Pugh-Turcott score and the Model for End-Stage Liver disease (MELD) score. Bilirubin is usually normal in early stages of the disease and may be a more appropriate marker for late-stage disease. Increasing and persistently elevated levels of bilirubin are associated with hepatobiliary malignancy[222]. Bilirubin may increase and fluctuate due to bacterial cholangitis or development and treatment of dominant strictures or biliary stones. Likewise, albumin levels alter late in disease course and may not be appropriate for risk stratification in early stages of PSC[222,238,258,259]. Bilirubin, albumin and AST are components of the PSC Mayo risk model, which predicts survival up to 4 years[258].

In contrast, alkaline phosphatase (ALP) is considered a useful early marker of the disease. Elevations in ALP, in the setting of hepatobiliary disease, results from the damaged liver regurgitating hepatic alkaline phosphatase back into serum[266] or , in the setting of obstruction, due to leakage of hepatic ALP into the serum[267]. Chronic elevation in ALP is a diagnostic hallmark of PSC and the most frequently observed biochemical abnormality[268,269]. In PSC, ALP has often been used in trials as a primary endpoint and several studies report that a reduction or an increase of ALP levels stratifies patients into low- and high risk categories for progressive disease[214,241,270-273] (**Table 1**). These studies, however, are limited by the use of random criteria and thresholds for ALP (reduction). In addition, some studies[270,271] assessed ALP reduction without a set time limit, as such it is likely that the differences found are due to the fact that patients that are close to a clinical endpoint will have an increased ALP. In the largest study so far (N=336, mostly UDCA-treated), *De Vries et al*.[214] evaluated the prognostic value of ALP levels in a systematic approach with assessment of the C-statistic as a measure of optimal predictive discrimination. The prognostic value of ALP levels at diagnosis and 1 year after diagnosis, as well as percentage change between both time points were studied. The decrease in ALP levels between T0 and T1, and a threshold of 1.3xULN at T1 were found to discriminate patients at risk of an event versus those without an event.

RISK PREDICTION MODELS

Several natural history models for PSC have been derived over the past decades (**Table 2**). The first PSC specific model for PSC was introduced by the Mayo Clinic in 1989[255]. This model included the variables age, bilirubin, haemoglobin, histological stage and inflammatory bowel disease. The Mayo risk score (MRS) was later adjusted[258] and now

includes the following variables to estimate the probability of survival for a period of 4 years: age, bilirubin, aspartate aminotransferase, history of variceal bleeding, and albumin. To date, the MRS remains a frequently used score to assess the short-term mortality risk of PSC patients.

Recently, three new natural history models for PSC have been introduced (**Table 3**). First, the Amsterdam-Oxford model (AOM) was developed to predict the long-term risk of PSC-related death and/or liver transplantation in a population-based cohort of 692 patients with PSC from the Netherlands and a validation cohort of 264 patients[238]. PSC-related death was defined as death from end-stage liver failure, death from liver surgery, death from cholangiocarcinoma, or death from colorectal carcinoma. The AOM incorporates PSC subtype, age at PSC diagnosis, aspartate aminotransferase (AST), alkaline phosphatase (ALP), total bilirubin, albumin and platelet count. The c-statistic of this model was a modest 0.68 (95% CI 0.51 to 0.85) and had satisfactory calibration. Second, Goode et al. included 1001 patients from the UK-PSC research consortium and used survival analyses to derive a short-term (2-year window) and long-term (10-year survival) natural history model [239]. The endpoint used was all-cause mortality or liver transplantation. Variables included in the models are age, values of bilirubin, alkaline phosphatase, albumin, platelets, presence of extrahepatic biliary disease, and variceal haemorrhage. Both models were validated in a cohort 451 patients and yielded c-statistics of 0.81 and 0.80, respectively). Finally, Eaton et al. introduced the PREsTo using a different statistical approach with machine learning that showed an impressive performance (C-statistic 0.90 [95% CI 0.84-0.95]) for the prediction of hepatic decompensation[240]. This model was derived in a cohort of 509 patients and validated in a cohort of 278 patients and included the following nine predictors: bilirubin, albumin, serum alkaline phosphatase, platelets, aspartate aminotransferase, hemoglobin, sodium, patient age, and number of years since PSC was diagnosed. This score is, however, limited by the prediction of hepatic decompensation rather than solid clinical endpoints such as liver transplantation, death, or CCA.

Table 1. Proposed ALP thresholds in PSC

Study	Year of publication	Patient cohort, *n*	Follow-up time (years)	ALP threshold	Composite endpoint	Strengths ✓and limitations ✕
Stanich et al. [270]	2011	87	7 (IQR 5-9)	Normalization (1 x ULN): Normalization: one/more ALP values of 1 x ULN or less	CCA LT Death (all cause)	✓ Inclusion of hepatobiliary malignancy and liver decompensation as an outcome
Al Mamari et al. [271]	2013	139	10 (IQR 6-15)	1.5 x ULN: At least 3 ALP values < 1.5 ULN, at least 6 months apart, continued to be < 1.5 9 ULN for this study follow-up period	Liver decompensation LT Death (liver-related, including death from CCA)	✓ Use of ALP values at different timepoints ✕ Censoring of other causes of death
Lindstrom et al. [272]	2012	195	11 (IQR 0-13)	Normalization (1 x ULN) or reduced ALP by >40%: after 1 year in the trial	CCA LT Death (all cause)	✓ Inclusion of hepatobiliary malignancy as an outcome
Rupp et al. [273]	2014	215	10	1. ALP normalization, 50% reduction compared with baseline values, or reduction < 1.5 x ULN: within 6 months. 2. ALP reduction < 1.5 9 ULN: within 12 months	LT Death (all cause)	✓ use of ALP values at different timepoints
Goode et al. [274]	2015	1000	5.7	ALP < 1.5 x ULN at 1 year OR ALP <2.0 x ULN at 2 years	LT	✕ Only LT as a clinical endpoint ✓ Large cohort
De Vries et al. [214]	2016	366	8.3 (IQR 5.6-12.5)	1.3 x ULN at 12 months UDCA therapy.	PSC-related death LT	✓ systematic determination of ALP threshold ✕ Censoring of other causes of death

IQR, interquartile range; ULN, upper limit of normal; CCA, cholangiocarcinoma; LT, liver transplantation; ALP, alkaline phosphatase.

Table 2. Proposed prediction models for estimating prognosis in primary sclerosing cholangitis

Model	Variables included
Mayo, *Wiesner et al.* (1989)[255] N=174	Age, bilirubin, haemoglobin, histological stage, inflammatory bowel disease
King's College Hospital, *Farrant et al.* (1991)[256] N=126	presence of hepatomegaly, presence of splenomegaly, alkaline phosphatase, histological stage, age in years
Multicentre model, *Dickson et al.* (1992) [257] N= 426	Age, bilirubin, histological stage, presence of splenomegaly
Scandinavian model, *Broomé et al.* (1996)[164] N= 305	Age, bilirubin, histological stage
New Mayo Model, *Kim et al.* (2000)[258] N= 405	Age, bilirubin, aspartate aminotransferase, history of variceal bleeding, albumin
Child Pugh[275], *Pugh et al.* (1973) N = 38	Bilirubin, albumin, INR, encephalopathy, ascites
MELD[152], *Kamath et al.* (2001) N = 282	Bilirubin, creatinine, INR
Time-dependent score, *Boberg et al.* (2002) [259] N = 330	Age at diagnosis, bilirubin, albumin
PSC score[222], *Tischendorf et al.* (2007) N = 273	Age, albumin, bilirubin elevation >3 months, hepatomegaly, splenomegaly, dominant bile duct stenosis, intra- and extrahepatic bile duct changes

Table 3. Newly introduced prediction models for estimating prognosis in primary sclerosing cholangitis

Model	Variables included	Cohort	Methods	Follow-up time/median survival	(Composite) endpoint	Remarks (strengths ✓ and limitations ✕)
Amsterdam Oxford Model, *De Vries et al.* (2016) [238]	PSC subtype, age at diagnosis, albumin, platelet count, aspartate aminotransferase, bilirubin	Derivation N=692 Validation N=264	Cox -regression/ survival analyses	110 (IQR 69–184) months	PSC-related death or liver transplantation	✓ Population based cohort ✕ Censoring of non-PSC-related death
UK-PSC risk scores, *Goode et al*. 2018[239]	**Short-term** = bilirubin, albumin, platelets, haemoglobin **Long-term** = age, bilirubin, ALP, albumin, haemoglobin, platelets, and presence of extrahepatic biliary disease	Derivation N=1001 Validation N=451	Cox -regression/ survival analyses	7904 patient-years	All-cause mortality or liver transplantation	✓ Mix of transplant- and non-transplant centres (57%) ✓ Large cohort
PREsTO, *Eaton et al*. 2018[240]	Bilirubin, albumin, serum alkaline phosphatase, platelets, aspartate aminotransferase, hemoglobin, sodium, patient age, and number of years since PSC was diagnosed	Derivation N=509 Validation N=278	Machine learning	6.09 (IQR 2.82-13.10) years	Hepatic decompensation (ascites, variceal haemorrhage, or encephalopathy)	✓ Application of novel promising technique ✕ Reproducibility; no formula available ✕ Endpoint

IQR, interquartile range; ALP, alkaline phosphatase; PREsTO, Primary Sclerosing Cholangitis Risk Estimate Tool.

Scope and aims of this thesis

This thesis focuses on risk stratification in primary biliary cholangitis and primary sclerosing cholangitis.

The papers regarding PBC in this thesis are based on data collected by the Global PBC Study Group. This multicentre international collaboration between medical centres involved in PBC research in Europe, Canada, United States of America, South America, China, and Japan currently includes long-term follow-up data from > 6000 PBC patients. This unique database allowed us to conduct multiple studies, including the assessment of changes in clinical presentation of PBC over the years. Whereas in earlier decades most patients presented with an advanced histological stage and with symptoms such as fatigue and itch, nowadays most patients may present during an earlier, and often asymptomatic, stage. Therefore, the underlying assumption that PBC, as a disease, is a static entity may not be accurate. **Chapter 2** describes disease characteristics over a 44-year period in patients diagnosed between 1970 and 2014, from 17 centres across Europe and North America. **Chapter 3** further focuses on the subpopulation of patients that present with an early biochemical disease stage, defined by normal albumin and normal bilirubin levels. This study describes the proportions of these patients who progress to moderate or advanced PBC and factors associated with progression and patient survival in this important subgroup.

Key to the clinical management of PBC is the assessment of prognosis at one year of UDCA therapy. At this timepoint clinicians may want to identify not only individuals who may benefit from second-line therapies, but also to identify cases in which UDCA monotherapy can be continued safely. The GLOBE score can be used to adequately risk stratify PBC patients at this timepoint. However, in daily clinical practice data with respect to the initial treatment response or the initial estimate of transplant-free survival may not, or no longer, be available. Moreover, (repeated) risk stratification after more prolonged treatment can be considered at least of equal key relevance in patient management. **Chapter 4** addresses these issues and assesses the application and predictive performance of the GLOBE score during prolonged UDCA therapy. Subsequently, **chapter 5** compares the predictive performance of the GLOBE score with other newly introduced and previously introduced prediction models.

Chapter 6 assesses the predictive performance and utility of the Amsterdam-Oxford model in correctly estimating the risk of liver transplantation or death in patients with primary sclerosing cholangitis in a large cohort of patients from three tertiary centres in Europe. In addition, we compare the Amsterdam-Oxford model with the Mayo risk score.

As patients with PSC progress towards cirrhosis and end-stage liver disease with associated complications, the only curative treatment is liver transplantation. Prioritization for liver transplantation on the liver transplantation waiting list in the Netherlands is based on the MELD score, which aims to transplant patients at highest short-term mortality risk based on objective parameters. As some complications in PSC, such as recurrent cholangitis or hepatobiliary malignancy, may not be adequately captured by the MELD score, PSC patients may receive MELD exception points. **Chapter 7** studies the influence of MELD exception points on waiting list survival and waiting list mortality as well as post-transplant outcomes between PSC and non-PSC patients by current waiting list policy in the Netherlands.

REFERENCES

1. Kaplan MM, Gershwin ME. Primary biliary cirrhosis. *N Engl J Med.* 2005;353(12):1261-1273.
2. Poupon R. Primary biliary cirrhosis: a 2010 update. *J Hepatol.* 2010;52(5):745-758.
3. Selmi C, Mayo MJ, Bach N, et al. Primary biliary cirrhosis in monozygotic and dizygotic twins: genetics, epigenetics, and environment. *Gastroenterology.* 2004;127(2):485-492.
4. Addison T, Gull W. On a certain affection of the skin vitiligoidea- α plana, β tuberosa. . *Guy's Hosp Rep.* 1851;7:265-276.
5. Hanot V. Étude sur une forme de cirrhose hypertrophique du foie [cirrhose hypertrophique avec ictere chronique. *JBBaillère, Paris.* 1876.
6. Thannhauser S, Magendantz H. The different clinical groups of xanthomatous disease: a clinical physiological study of 22 cases. *Ann Intern Med.* 1938;11:1662–1746.
7. Rubin E, Schaffner F, Popper H. Primary Biliary Cirrhosis. Chronic non-suppurative destrictive cholangitis. *Am J Pathol.* 1965;46:387-407.
8. Beuers U, Gershwin ME, Gish RG, et al. Changing nomenclature for PBC: From 'cirrhosis' to 'cholangitis'. *J Hepatol.* 2015;63(5):1285-1287.
9. EASL Clinical Practice Guidelines: The diagnosis and management of patients with primary biliary cholangitis. *J Hepatol.* 2017;67(1):145-172.
10. Lindor KD, Bowlus CL, Boyer J, Levy C, Mayo M. Primary Biliary Cholangitis: 2018 Practice Guidance from the American Association for the Study of Liver Diseases. *Hepatology.* 2019;69(1):394-419.
11. Floreani A, Caroli D, Variola A, et al. A 35-year follow-up of a large cohort of patients with primary biliary cirrhosis seen at a single centre. *Liver Int.* 2011;31(3):361-368.
12. Prince MI, Chetwynd A, Craig WL, Metcalf JV, James OF. Asymptomatic primary biliary cirrhosis: clinical features, prognosis, and symptom progression in a large population based cohort. *Gut.* 2004;53(6):865-870.
13. Prince MI, James OF. The epidemiology of primary biliary cirrhosis. *Clin Liver Dis.* 2003;7(4):795-819.
14. Murillo Perez F, Goet JC, Lammers WJ, et al. Milder disease stage in patients with primary biliary cholangitis over a 44-year period: A changing natural history. *Hepatology.* 2017; doi: 10.1002/hep.29717. [Epub ahead of print].
15. Balasubramaniam K, Grambsch PM, Wiesner RH, Lindor KD, Dickson ER. Diminished survival in asymptomatic primary biliary cirrhosis. A prospective study. *Gastroenterology.* 1990;98(6):1567-1571.
16. Prince M, Chetwynd A, Newman W, Metcalf JV, James OF. Survival and symptom progression in a geographically based cohort of patients with primary biliary cirrhosis: follow-up for up to 28 years. *Gastroenterology.* 2002;123(4):1044-1051.
17. Mitchison HC, Lucey MR, Kelly PJ, Neuberger JM, Williams R, James OF. Symptom development and prognosis in primary biliary cirrhosis: a study in two centers. *Gastroenterology.* 1990;99(3):778-784.
18. Metcalf JV, Mitchison HC, Palmer JM, Jones DE, Bassendine MF, James OF. Natural history of early primary biliary cirrhosis. *Lancet.* 1996;348(9039):1399-1402.
19. Springer J, Cauch-Dudek K, O'Rourke K, Wanless IR, Heathcote EJ. Asymptomatic primary biliary cirrhosis: a study of its natural history and prognosis. *Am J Gastroenterol.* 1999;94(1):47-53.
20. Mells GF, Pells G, Newton JL, et al. Impact of primary biliary cirrhosis on perceived quality of life: the UK-PBC national study. *Hepatology.* 2013;58(1):273-283.

21. Jacoby A, Rannard A, Buck D, et al. Development, validation, and evaluation of the PBC-40, a disease specific health related quality of life measure for primary biliary cirrhosis. *Gut.* 2005;54(11):1622-1629.
22. Oertelt S, Rieger R, Selmi C, et al. A sensitive bead assay for antimitochondrial antibodies: Chipping away at AMA-negative primary biliary cirrhosis. *Hepatology.* 2007;45(3):659-665.
23. Ludwig J, Dickson ER, McDonald GS. Staging of chronic nonsuppurative destructive cholangitis (syndrome of primary biliary cirrhosis). *Virchows Arch A Pathol Anat Histol.* 1978;379(2):103-112.
24. Roll J, Boyer JL, Barry D, Klatskin G. The prognostic importance of clinical and histologic features in asymptomatic and symptomatic primary biliary cirrhosis. *N Engl J Med.* 1983;308(1):1-7.
25. Lleo A, Jepsen P, Morenghi E, et al. Evolving Trends in Female to Male Incidence and Male Mortality of Primary Biliary Cholangitis. *Sci Rep.* 2016;6:25906.
26. Tanaka A, Mori M, Matsumoto K, Ohira H, Tazuma S, Takikawa H. Increase trend in the prevalence and male-to-female ratio of primary biliary cholangitis, autoimmune hepatitis, and primary sclerosing cholangitis in Japan. *Hepatol Res.* 2019;49(8):881-889.
27. Boonstra K, Beuers U, Ponsioen CY. Epidemiology of primary sclerosing cholangitis and primary biliary cirrhosis: a systematic review. *J Hepatol.* 2012;56(5):1181-1188.
28. Boonstra K, Kunst AE, Stadhouders PH, et al. Rising incidence and prevalence of primary biliary cirrhosis: a large population-based study. *Liver Int.* 2014;34(6):e31-38.
29. Griffiths L, Dyson JK, Jones DE. The new epidemiology of primary biliary cirrhosis. *Semin Liver Dis.* 2014;34(3):318-328.
30. Dahlqvist G, Gaouar F, Carrat F, et al. Large-scale characterization study of patients with antimitochondrial antibodies but nonestablished primary biliary cholangitis. *Hepatology.* 2017;65(1):152-163.
31. Gershwin ME, Mackay IR. The causes of primary biliary cirrhosis: Convenient and inconvenient truths. *Hepatology.* 2008;47(2):737-745.
32. Poupon R, Ping C, Chretien Y, et al. Genetic factors of susceptibility and of severity in primary biliary cirrhosis. *J Hepatol.* 2008;49(6):1038-1045.
33. Hirschfield GM, Liu X, Xu C, et al. Primary biliary cirrhosis associated with HLA, IL12A, and IL12RB2 variants. *N Engl J Med.* 2009;360(24):2544-2555.
34. Liu X, Invernizzi P, Lu Y, et al. Genome-wide meta-analyses identify three loci associated with primary biliary cirrhosis. *Nat Genet.* 2010;42(8):658-660.
35. Triger DR. Primary biliary cirrhosis: is there an environmental contribution? *J Gastroenterol Hepatol.* 1991;6(6):568-569.
36. O'Donohue J, McFarlane B, Bomford A, Yates M, Williams R. Antibodies to atypical mycobacteria in primary biliary cirrhosis. *J Hepatol.* 1994;21(5):887-889.
37. Floreani A, Franceschet I, Cazzagon N. Primary biliary cirrhosis: overlaps with other autoimmune disorders. *Semin Liver Dis.* 2014;34(3):352-360.
38. McNally RJ, James PW, Ducker S, James OF. Seasonal variation in the patient diagnosis of primary biliary cirrhosis: further evidence for an environmental component to etiology. *Hepatology.* 2011;54(6):2099-2103.
39. McNally RJ, Ducker S, James OF. Are transient environmental agents involved in the cause of primary biliary cirrhosis? Evidence from space-time clustering analysis. *Hepatology.* 2009;50(4):1169-1174.
40. Jones DE. Pathogenesis of primary biliary cirrhosis. *Gut.* 2007;56(11):1615-1624.

41. Beuers U. Drug insight: Mechanisms and sites of action of ursodeoxycholic acid in cholestasis. *Nat Clin Pract Gastroenterol Hepatol.* 2006;3(6):318-328.
42. Hohenester S, Wenniger LM, Paulusma CC, et al. A biliary HCO3- umbrella constitutes a protective mechanism against bile acid-induced injury in human cholangiocytes. *Hepatology.* 2012;55(1):173-183.
43. Karpen SJ, Trauner M. The new therapeutic frontier--nuclear receptors and the liver. *J Hepatol.* 2010;52(3):455-462.
44. Boyer JL, Trauner M, Mennone A, et al. Upregulation of a basolateral FXR-dependent bile acid efflux transporter OSTalpha-OSTbeta in cholestasis in humans and rodents. *Am J Physiol Gastrointest Liver Physiol.* 2006;290(6):G1124-1130.
45. Lindor KD, Gershwin ME, Poupon R, et al. Primary biliary cirrhosis. *Hepatology.* 2009;50(1):291-308.
46. Shi J, Wu C, Lin Y, Chen YX, Zhu L, Xie WF. Long-term effects of mid-dose ursodeoxycholic acid in primary biliary cirrhosis: a meta-analysis of randomized controlled trials. *Am J Gastroenterol.* 2006;101(7):1529-1538.
47. Angulo P, Batts KP, Therneau TM, Jorgensen RA, Dickson ER, Lindor KD. Long-term ursodeoxycholic acid delays histological progression in primary biliary cirrhosis. *Hepatology.* 1999;29(3):644-647.
48. Corpechot C, Carrat F, Bonnand AM, Poupon RE, Poupon R. The effect of ursodeoxycholic acid therapy on liver fibrosis progression in primary biliary cirrhosis. *Hepatology.* 2000;32(6):1196-1199.
49. Poupon RE, Lindor KD, Pares A, Chazouilleres O, Poupon R, Heathcote EJ. Combined analysis of the effect of treatment with ursodeoxycholic acid on histologic progression in primary biliary cirrhosis. *J Hepatol.* 2003;39(1):12-16.
50. Corpechot C, Carrat F, Bahr A, Chretien Y, Poupon RE, Poupon R. The effect of ursodeoxycholic acid therapy on the natural course of primary biliary cirrhosis. *Gastroenterology.* 2005;128(2):297-303.
51. ter Borg PC, Schalm SW, Hansen BE, van Buuren HR, Dutch PBCSG. Prognosis of ursodeoxycholic Acid-treated patients with primary biliary cirrhosis. Results of a 10-yr cohort study involving 297 patients. *Am J Gastroenterol.* 2006;101(9):2044-2050.
52. Harms MH, van Buuren HR, Corpechot C, et al. Ursodeoxycholic acid therapy and liver transplant-free survival in patients with primary biliary cholangitis. *J Hepatol.* 2019;71(2):357-365.
53. Harms MH, de Veer RC, Lammers WJ, et al. Number needed to treat with ursodeoxycholic acid therapy to prevent liver transplantation or death in primary biliary cholangitis. *Gut.* 2019.
54. Corpechot C, Chazouilleres O, Belnou P, et al. Long-term impact of preventive UDCA therapy after transplantation for primary biliary cholangitis. *J Hepatol.* 2020.
55. Lammers WJ, Hirschfield GM, Corpechot C, et al. Development and Validation of a Scoring System to Predict Outcomes of Patients With Primary Biliary Cirrhosis Receiving Ursodeoxycholic Acid Therapy. *Gastroenterology.* 2015;149(7):1804-1812 e1804.
56. Nevens F, Andreone P, Mazzella G, et al. A Placebo-Controlled Trial of Obeticholic Acid in Primary Biliary Cholangitis. *N Engl J Med.* 2016;375(7):631-643.
57. Kota BP, Huang TH, Roufogalis BD. An overview on biological mechanisms of PPARs. *Pharmacol Res.* 2005;51(2):85-94.

58. Honda A, Ikegami T, Nakamuta M, et al. Anticholestatic effects of bezafibrate in patients with primary biliary cirrhosis treated with ursodeoxycholic acid. *Hepatology.* 2013;57(5):1931-1941.
59. Corpechot C, Chazouilleres O, Rousseau A, et al. A Placebo-Controlled Trial of Bezafibrate in Primary Biliary Cholangitis. *N Engl J Med.* 2018;378(23):2171-2181.
60. Nakai S, Masaki T, Kurokohchi K, Deguchi A, Nishioka M. Combination therapy of bezafibrate and ursodeoxycholic acid in primary biliary cirrhosis: a preliminary study. *Am J Gastroenterol.* 2000;95(1):326-327.
61. Akbar SM, Furukawa S, Nakanishi S, Abe M, Horiike N, Onji M. Therapeutic efficacy of decreased nitrite production by bezafibrate in patients with primary biliary cirrhosis. *J Gastroenterol.* 2005;40(2):157-163.
62. Kanda T, Yokosuka O, Imazeki F, Saisho H. Bezafibrate treatment: a new medical approach for PBC patients? *J Gastroenterol.* 2003;38(6):573-578.
63. Kita R, Takamatsu S, Kimura T, Kokuryu H, Osaki Y, Tomono N. Bezafibrate may attenuate biliary damage associated with chronic liver diseases accompanied by high serum biliary enzyme levels. *J Gastroenterol.* 2006;41(7):686-692.
64. Hazzan R, Tur-Kaspa R. Bezafibrate treatment of primary biliary cirrhosis following incomplete response to ursodeoxycholic acid. *J Clin Gastroenterol.* 2010;44(5):371-373.
65. Dohmen K, Tanaka H, Haruno M. Effectiveness of fenofibrate in comparison to bezafibrate for patients with asymptomatic primary biliary cirrhosis. *Fukuoka Igaku Zasshi.* 2013;104(10):350-361.
66. Ohira H, Sato Y, Ueno T, Sata M. Fenofibrate treatment in patients with primary biliary cirrhosis. *Am J Gastroenterol.* 2002;97(8):2147-2149.
67. Han XF, Wang QX, Liu Y, et al. Efficacy of fenofibrate in Chinese patients with primary biliary cirrhosis partially responding to ursodeoxycholic acid therapy. *J Dig Dis.* 2012;13(4):219-224.
68. Harms M, Cheung A, van Buuren H, et al. Comparable Beneficial Effects of Beza- and Fenofibrate in Primary Biliary Cholangitis. An International Study in UDCA-treated individuals - Plenary and Parallel Sessions (Abstracts 1–258). *Hepatology.* 2016;64(S1):1-136.
69. Bolier R, de Vries ES, Pares A, et al. Fibrates for the treatment of cholestatic itch (FITCH): study protocol for a randomized controlled trial. *Trials.* 2017;18(1):230.
70. De Vries E, Bolier R, Goet J, et al. Bezafibrate is More Effective than Placebo in Pruritus of Chronic Cholestasis: The Fitch Trial - Oral abstract presentation at the AASLD The Liver Meeting, Boston. *Hepatology.* 2019;70(S1):1-187.
71. Harms MH, Lammers WJ, Thorburn D, et al. Major Hepatic Complications in Ursodeoxycholic Acid-Treated Patients With Primary Biliary Cholangitis: Risk Factors and Time Trends in Incidence and Outcome. *Am J Gastroenterol.* 2018;113(2):254-264.
72. EASL Clinical Practice Guidelines: The diagnosis and management of patients with primary biliary cholangitis. *J Hepatol.* 2017.
73. Degott C, Zafrani ES, Callard P, Balkau B, Poupon RE, Poupon R. Histopathological study of primary biliary cirrhosis and the effect of ursodeoxycholic acid treatment on histology progression. *Hepatology.* 1999;29(4):1007-1012.
74. Corpechot C, Carrat F, Poupon R, Poupon RE. Primary biliary cirrhosis: incidence and predictive factors of cirrhosis development in ursodiol-treated patients. *Gastroenterology.* 2002;122(3):652-658.
75. Trivedi PJ, Hirschfield GM. Primary biliary cirrhosis: Renaming primary biliary cirrhosis-clarity or confusion? *Nat Rev Gastroenterol Hepatol.* 2015;12(12):678-679.

76. Trivedi PJ, Bruns T, Cheung A, et al. Optimising risk stratification in primary biliary cirrhosis: AST/platelet ratio index predicts outcome independent of ursodeoxycholic acid response. *Journal of Hepatology*. 2014;60(6):1249-1258.
77. Harms MH, Lammers WJ, Thorburn D, et al. Major Hepatic Complications in Ursodeoxycholic Acid-Treated Patients With Primary Biliary Cholangitis: Risk Factors and Time Trends in Incidence and Outcome. *Am J Gastroenterol*. 2018;113(2):254-264.
78. Lammers WJ, van Buuren HR, Hirschfield GM, 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*. 2014;147(6):1338-1349.
79. Poupon RE, Bonnand AM, Chretien Y, Poupon R. Ten-year survival in ursodeoxycholic acid-treated patients with primary biliary cirrhosis. The UDCA-PBC Study Group. *Hepatology*. 1999;29(6):1668-1671.
80. Kuiper EM, Hansen BE, de Vries RA, et al. Improved prognosis of patients with primary biliary cirrhosis that have a biochemical response to ursodeoxycholic acid. *Gastroenterology*. 2009;136(4):1281-1287.
81. Pares A, Caballeria L, Rodes J. Excellent long-term survival in patients with primary biliary cirrhosis and biochemical response to ursodeoxycholic Acid. *Gastroenterology*. 2006;130(3):715-720.
82. Corpechot C, Abenavoli L, Rabahi N, et al. Biochemical response to ursodeoxycholic acid and long-term prognosis in primary biliary cirrhosis. *Hepatology*. 2008;48(3):871-877.
83. Myers RP, Shaheen AA, Fong A, et al. Epidemiology and natural history of primary biliary cirrhosis in a Canadian health region: a population-based study. *Hepatology*. 2009;50(6):1884-1892.
84. Lee J, Belanger A, Doucette JT, Stanca C, Friedman S, Bach N. Transplantation trends in primary biliary cirrhosis. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association*. 2007;5(11):1313-1315.
85. Kuiper EM, Hansen BE, Metselaar HJ, et al. Trends in liver transplantation for primary biliary cirrhosis in the Netherlands 1988-2008. *BMC Gastroenterol*. 2010;10:144.
86. Harms MH, Janssen QP, Adam R, et al. Trends in liver transplantation for primary biliary cholangitis in Europe over the past three decades. *Aliment Pharmacol Ther*. 2019;49(3):285-295.
87. Shapiro JM, Smith H, Schaffner F. Serum bilirubin: a prognostic factor in primary biliary cirrhosis. *Gut*. 1979;20(2):137-140.
88. Christensen E, Schlichting P, Fauerholdt L, Juhl E, Poulsen H, Tygstrup N. Changes of laboratory variables with time in cirrhosis: prognostic and therapeutic significance. *Hepatology*. 1985;5(5):843-853.
89. Harms MH, Pares A, Mason AL, et al. Behavioral Patterns of Total Serum Bilirubin Prior to Major Clinical Endpoints in 3529 Patients with Primary Biliary Cholangitis. 2016;64:S633–S634
90. Murillo Perez CF, Gulamhusein A, van Buuren H, et al. Bilirubin is Predictive of Transplant-free Survival Even Within the Normal Range in Primary Biliary Cholangitis Patients. *Hepatology*. 2017;66(1):S41A.
91. Carbone M, Mells GF, Pells G, et al. Sex and age are determinants of the clinical phenotype of primary biliary cirrhosis and response to ursodeoxycholic acid. *Gastroenterology*. 2013;144(3):560-569 e567; quiz e513-564.

92. Cheung A LW, Hirschfield GM, Invernizzi P, Mason A, Ponsioen CY, Floreanim, A, Corpechot C, Mayo MJ, Pares A, Battezzati PM, Nevens F, Thorburn D, Kowdley KV, Trivedi PJ, Kumagi T, Lleo A, LaRusso NF, Boonstra K, Cazzagon N, Franceschet I, Poupon R, Caballeria J, Lindor KD, Hansen BE, Janssen HLA, Van Buuren HR. Age, bilirubin and albumin, regardless of sex, are the strongest independent predictors of biochemical response and transplantation-free survival in patients with primary biliary cirrhosis. *J Hepatol.* 2015;62:S798.
93. Trivedi PJ, Lammers WJ, van Buuren HR, et al. Stratification of hepatocellular carcinoma risk in primary biliary cirrhosis: a multicentre international study. *Gut.* 2015.
94. Jones DE, Al-Rifai A, Frith J, Patanwala I, Newton JL. The independent effects of fatigue and UDCA therapy on mortality in primary biliary cirrhosis: results of a 9 year follow-up. *J Hepatol.* 2010;53(5):911-917.
95. Nakamura M, Kondo H, Mori T, et al. Anti-gp210 and anti-centromere antibodies are different risk factors for the progression of primary biliary cirrhosis. *Hepatology.* 2007;45(1):118-127.
96. Muratori P, Muratori L, Ferrari R, et al. Characterization and clinical impact of antinuclear antibodies in primary biliary cirrhosis. *Am J Gastroenterol.* 2003;98(2):431-437.
97. Nakamura M, Kondo H, Tanaka A, et al. Autoantibody status and histological variables influence biochemical response to treatment and long-term outcomes in Japanese patients with primary biliary cirrhosis. *Hepatol Res.* 2015;45(8):846-855.
98. Carbone M, Sharp SJ, Flack S, et al. The UK-PBC risk scores: Derivation and validation of a scoring system for long-term prediction of end-stage liver disease in primary biliary cirrhosis. *Hepatology.* 2015:doi: 10.1002/hep.28017.
99. Corpechot C, Chazouilleres O, Poupon R. Early primary biliary cirrhosis: biochemical response to treatment and prediction of long-term outcome. *J Hepatol.* 2011;55(6):1361-1367.
100. 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(5):1549-1557.
101. Corpechot C, Carrat F, Poujol-Robert A, et al. Noninvasive elastography-based assessment of liver fibrosis progression and prognosis in primary biliary cirrhosis. *Hepatology.* 2012;56(1):198-208.
102. Corpechot C GF, Lemoinne S, Kemgang Fankem A, Poupon R, Chretien Y, et al. Fibroscan improves the ability of the new prognostic scoring systems to predict outcomes of PBC. *J Hepatol.* 2016;64:S177–S178.
103. Chang PE, Goh GB, Ngu JH, Tan HK, Tan CK. Clinical applications, limitations and future role of transient elastography in the management of liver disease. *World J Gastrointest Pharmacol Ther.* 2016;7(1):91-106.
104. Poupon R, Chazouillères O, Balkau B, Poupon RE. Clinical and biochemical expression of the histopathological lesions of primary biliary cirrhosis. *Journal of Hepatology.* 1999;30(3):408-412.
105. Vleggaar FP, van Buuren HR, Zondervan PE, ten Kate FJ, Hop WC, Dutch Multicentre PBCsg. Jaundice in non-cirrhotic primary biliary cirrhosis: the premature ductopenic variant. *Gut.* 2001;49(2):276-281.
106. Kumagi T, Guindi M, Fischer SE, et al. Baseline ductopenia and treatment response predict long-term histological progression in primary biliary cirrhosis. *Am J Gastroenterol.* 2010;105(10):2186-2194.
107. Carbone M, Sharp SJ, Heneghan MA, et al. P1198 : Histological stage is relevant for risk-stratification in primary biliary cirrhosis. *Journal of Hepatology.* 2015;62:S805.

108. Matsuo I, Omagari K, Kinoshita H, Hata T, Tagawa M, Kohno S. Elevation of serum gamma-glutamyl transpeptidase precedes that of alkaline phosphatase in the early stages of primary biliary cirrhosis. *Hepatology Research.* 1999;14(3):223-232.
109. Azemoto N, Abe M, Murata Y, et al. Early biochemical response to ursodeoxycholic acid predicts symptom development in patients with asymptomatic primary biliary cirrhosis. *J Gastroenterol.* 2009;44(6):630-634.
110. Azemoto N, Kumagi T, Abe M, et al. Biochemical response to ursodeoxycholic acid predicts long-term outcome in Japanese patients with primary biliary cirrhosis. *Hepatol Res.* 2011;41(4):310-317.
111. Fleming TR, Powers JH. Biomarkers and surrogate endpoints in clinical trials. *Stat Med.* 2012;31(25):2973-2984.
112. van Hoogstraten HJ, Hansen BE, van Buuren HR, ten Kate FJ, van Berge-Henegouwen GP, Schalm SW. Prognostic factors and long-term effects of ursodeoxycholic acid on liver biochemical parameters in patients with primary biliary cirrhosis. Dutch Multi-Centre PBC Study Group. *J Hepatol.* 1999;31(2):256-262.
113. Bonnand AM, Heathcote EJ, Lindor KD, Poupon RE. Clinical significance of serum bilirubin levels under ursodeoxycholic acid therapy in patients with primary biliary cirrhosis. *Hepatology.* 1999;29(1):39-43.
114. Angulo P, Lindor KD, Therneau TM, et al. Utilization of the Mayo risk score in patients with primary biliary cirrhosis receiving ursodeoxycholic acid. *Liver.* 1999;19(2):115-121.
115. Zhang LN, Shi TY, Shi XH, et al. Early biochemical response to ursodeoxycholic acid and long-term prognosis of primary biliary cirrhosis: results of a 14-year cohort study. *Hepatology.* 2013;58(1):264-272.
116. Leuschner M, Maier KP, Schlichting J, et al. Oral budesonide and ursodeoxycholic acid for treatment of primary biliary cirrhosis: results of a prospective double-blind trial. *Gastroenterology.* 1999;117(4):918-925.
117. Carbone M, Sharp SJ, Flack S, et al. The UK-PBC risk scores: Derivation and validation of a scoring system for long-term prediction of end-stage liver disease in primary biliary cholangitis. *Hepatology.* 2016;63(3):930-950.
118. Lammers WJ, van Buuren H.R., Pares, A., et al. Defining optimal laboratory response criteria in UDCA treated primary biliary cirrhosis. Results of an internationalmulticenter long-term follow-up study. *Hepatology.* 2013;58:249A.
119. Momah N, Silveira MG, Jorgensen R, Sinakos E, Lindor KD. Optimizing biochemical markers as endpoints for clinical trials in primary biliary cirrhosis. *Liver Int.* 2012;32(5):790-795.
120. Cheung AC, Lammers WJ, Murillo Perez CF, et al. Effects of Age and Sex of Response to Ursodeoxycholic Acid and Transplant-free Survival in Patients With Primary Biliary Cholangitis. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association.* 2019;17(10):2076-2084 e2072.
121. Quarneti C, Muratori P, Lalanne C, et al. Fatigue and pruritus at onset identify a more aggressive subset of primary biliary cirrhosis. *Liver Int.* 2015;35(2):636-641.
122. Mahl TC, Shockcor W, Boyer JL. Primary biliary cirrhosis: survival of a large cohort of symptomatic and asymptomatic patients followed for 24 years. *J Hepatol.* 1994;20(6):707-713.
123. Vleggaar FP, van Buuren HR. No prognostic significance of antimitochondrial antibody profile testing in primary biliary cirrhosis. *Hepatogastroenterology.* 2004;51(58):937-940.
124. Palmer JM, Yeaman SJ, Bassendine MF, James OF. M4 and M9 autoantigens in primary biliary cirrhosis--a negative study. *J Hepatol.* 1993;18(2):251-254.

125. Trivedi PJ, Corpechot C, Pares A, Hirschfield GM. Risk stratification in autoimmune cholestatic liver diseases: Opportunities for clinicians and trialists. *Hepatology.* 2016;63(2):644-659.
126. Nyberg A, Engstrom-Laurent A, Loof L. Serum hyaluronate in primary biliary cirrhosis--a biochemical marker for progressive liver damage. *Hepatology.* 1988;8(1):142-146.
127. Nyblom H, Bjornsson E, Simren M, Aldenborg F, Almer S, Olsson R. The AST/ALT ratio as an indicator of cirrhosis in patients with PBC. *Liver Int.* 2006;26(7):840-845.
128. Floreani A, Cazzagon N, Martines D, Cavalletto L, Baldo V, Chemello L. Performance and utility of transient elastography and noninvasive markers of liver fibrosis in primary biliary cirrhosis. *Dig Liver Dis.* 2011;43(11):887-892.
129. Gomez-Dominguez E, Mendoza J, Garcia-Buey L, et al. Transient elastography to assess hepatic fibrosis in primary biliary cirrhosis. *Aliment Pharmacol Ther.* 2008;27(5):441-447.
130. Corpechot C GF, Lemoinne S, Ben Belkacem K, Kemgang Fankem A, Chretien Y, et al. Relationships between biochemical response to UDCA and progression of liver stiffness as determined by Fibroscan in patients with PBC. *Hepatology.* 2016;64:194A–195A.
131. Garrido MC, Hubscher SG. Accuracy of staging in primary biliary cirrhosis. *J Clin Pathol.* 1996;49(7):556-559.
132. Moons KG, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? *BMJ.* 2009;338:b375.
133. Christensen E, Crowe J, Doniach D, et al. Clinical pattern and course of disease in primary biliary cirrhosis based on an analysis of 236 patients. *Gastroenterology.* 1980;78(2):236-246.
134. Dickson ER, Grambsch PM, Fleming TR, Fisher LD, Langworthy A. Prognosis in primary biliary cirrhosis: model for decision making. *Hepatology.* 1989;10(1):1-7.
135. Murtaugh PA, Dickson ER, Van Dam GM, et al. Primary biliary cirrhosis: prediction of short-term survival based on repeated patient visits. *Hepatology.* 1994;20(1 Pt 1):126-134.
136. Kim WR, Wiesner RH, Poterucha JJ, et al. Adaptation of the Mayo primary biliary cirrhosis natural history model for application in liver transplant candidates. *Liver Transpl.* 2000;6(4):489-494.
137. Grambsch PM, Dickson ER, Kaplan M, LeSage G, Fleming TR, Langworthy AL. Extramural cross-validation of the Mayo primary biliary cirrhosis survival model establishes its generalizability. *Hepatology.* 1989;10(5):846-850.
138. Kilmurry MR, Heathcote EJ, Cauch-Dudek K, et al. Is the Mayo model for predicting survival useful after the introduction of ursodeoxycholic acid treatment for primary biliary cirrhosis? *Hepatology.* 1996;23(5):1148-1153.
139. Johnson PJ, Berhane S, Kagebayashi C, et al. Assessment of liver function in patients with hepatocellular carcinoma: a new evidence-based approach-the ALBI grade. *J Clin Oncol.* 2015;33(6):550-558.
140. Chan AW, Chan RC, Wong GL, et al. New simple prognostic score for primary biliary cirrhosis: Albumin-bilirubin score. *J Gastroenterol Hepatol.* 2015;30(9):1391-1396.
141. Malinchoc M, Kamath PS, Gordon FD, Peine CJ, Rank J, ter Borg PC. A model to predict poor survival in patients undergoing transjugular intrahepatic portosystemic shunts. *Hepatology.* 2000;31(4):864-871.
142. Singal AK, Fang X, Kaif M, et al. Primary biliary cirrhosis has high wait-list mortality among patients listed for liver transplantation. *Transpl Int.* 2017;30(5):454-462.

143. Younossi Z.M. SM, de Avila L., Ahmed A., Wong R.J., Saab S., Kowdley K.V., Brown K.A., Kugelmas M., Racila A. Abstract #312: Liver Transplant (LT) Candidates with Primary Biliary Cholangitis (PBC) Experience Higher Wait List Mortality: Data from the Scientific Registry of Transplant Recipients (SRTR). *Hepatology.*66(1):176A-177A.
144. Goet JC, Lammers WJ, Floreani A, et al. The GLOBE score identifies PBC patients at increased risk of liver transplantation or death in different age-categories over time. *Journal of Hepatology.* 2017;66(1):S543 - S544.
145. Goet JC LW, Janssen HL, Hirschfield GM, Corpechot C, Lindor KD, et al. The GLOBE score identifies PBC patients at increased risk of liver transplantation or death during follow-up. *Hepatology.* 2016;189A–190A.
146. Lammers WJ, Goet JC, Mayo MJ, et al. Time-dependent factors associated with diminished transplant-free survival in patients with primary biliary cholangitis and an optimal response to ursodeoxycholic acid after one year of treatment. *Hepatology.* 2017;66(1):Suppl 157A-158A.
147. Yang F, Yang Y, Wang Q, et al. The risk predictive values of UK-PBC and GLOBE scoring system in Chinese patients with primary biliary cholangitis: the additional effect of anti-gp210. *Aliment Pharmacol Ther.* 2017.
148. Huang A. SY, Wang C., Wang Y., Chen J., Wu V., Zou Z., Lau G. . Abstract # 318: The risk predictive values of UK-PBC and GLOBE scoring system and performance of biochemical response criteria in Chinese patients with primary biliary cholangitis (PBC) on ursodeoxycholic acid. *Hepatology.* 2017;66(1):suppl 180A.
149. Cheung KS, Seto WK, Fung J, Lai CL, Yuen MF. Prognostic Factors for Transplant-Free Survival and Validation of Prognostic Models in Chinese Patients with Primary Biliary Cholangitis Receiving Ursodeoxycholic Acid. *Clin Transl Gastroenterol.* 2017;8(6):e100.
150. Christensen E, Neuberger J, Crowe J, et al. Beneficial effect of azathioprine and prediction of prognosis in primary biliary cirrhosis. Final results of an international trial. *Gastroenterology.* 1985;89(5):1084-1091.
151. Christensen E, Altman DG, Neuberger J, De Stavola BL, Tygstrup N, Williams R. Updating prognosis in primary biliary cirrhosis using a time-dependent Cox regression model. PBC1 and PBC2 trial groups. *Gastroenterology.* 1993;105(6):1865-1876.
152. Kamath PS, Wiesner RH, Malinchoc M, et al. A model to predict survival in patients with end-stage liver disease. *Hepatology.* 2001;33(2):464-470.
153. Chapman RW, Arborgh BA, Rhodes JM, et al. Primary sclerosing cholangitis: a review of its clinical features, cholangiography, and hepatic histology. *Gut.* 1980;21(10):870-877.
154. Hirschfield GM, Karlsen TH, Lindor KD, Adams DH. Primary sclerosing cholangitis. *Lancet.* 2013;382(9904):1587-1599.
155. Williamson KD, Chapman RW. Primary sclerosing cholangitis. *Dig Dis.* 2014;32(4):438-445.
156. European Association for the Study of the L. EASL Clinical Practice Guidelines: management of cholestatic liver diseases. *J Hepatol.* 2009;51(2):237-267.
157. Roberts MS, Angus DC, Bryce CL, Valenta Z, Weissfeld L. Survival after liver transplantation in the United States: a disease-specific analysis of the UNOS database. *Liver Transpl.* 2004;10(7):886-897.
158. Singal AK, Guturu P, Hmoud B, Kuo YF, Salameh H, Wiesner RH. Evolving frequency and outcomes of liver transplantation based on etiology of liver disease. *Transplantation.* 2013;95(5):755-760.

159. Tamura S, Sugawara Y, Kaneko J, Matsui Y, Togashi J, Makuuchi M. Recurrence of primary sclerosing cholangitis after living donor liver transplantation. *Liver Int.* 2007;27(1):86-94.
160. Gautam M, Cheruvattath R, Balan V. Recurrence of autoimmune liver disease after liver transplantation: a systematic review. *Liver Transpl.* 2006;12(12):1813-1824.
161. Charatcharoenwitthaya P, Lindor KD. Recurrence of primary sclerosing cholangitis: what do we learn from several transplant centers? *Liver Transpl.* 2008;14(2):130-132.
162. Campsen J, Zimmerman MA, Trotter JF, et al. Clinically recurrent primary sclerosing cholangitis following liver transplantation: a time course. *Liver Transpl.* 2008;14(2):181-185.
163. Alabraba E NP, Gunson B, et al. A re-evaluation of the risk factors for the recurrence of primary sclerosing cholangitis in liver allografts. *Liver Transplant*. 2009;15:330e340.
164. Broome U, Olsson R, Loof L, et al. Natural history and prognostic factors in 305 Swedish patients with primary sclerosing cholangitis. *Gut.* 1996;38(4):610-615.
165. Abdalian R, Heathcote EJ. Sclerosing cholangitis: a focus on secondary causes. *Hepatology.* 2006;44(5):1063-1074.
166. Deneau MR, El-Matary W, Valentino PL, et al. The natural history of primary sclerosing cholangitis in 781 children: A multicenter, international collaboration. *Hepatology.* 2017;66(2):518-527.
167. Boberg KM, Chapman RW, Hirschfield GM, et al. Overlap syndromes: the International Autoimmune Hepatitis Group (IAIHG) position statement on a controversial issue. *J Hepatol.* 2011;54(2):374-385.
168. Molodecky NA, Kareemi H, Parab R, et al. Incidence of primary sclerosing cholangitis: a systematic review and meta-analysis. *Hepatology.* 2011;53(5):1590-1599.
169. Boonstra K, Weersma RK, van Erpecum KJ, et al. Population-based epidemiology, malignancy risk, and outcome of primary sclerosing cholangitis. *Hepatology.* 2013;58(6):2045-2055.
170. Lee YM, Kaplan MM. Primary sclerosing cholangitis. *N Engl J Med.* 1995;332(14):924-933.
171. Lunder AK, Hov JR, Borthne A, et al. Prevalence of Sclerosing Cholangitis Detected by Magnetic Resonance Cholangiography in Patients With Long-term Inflammatory Bowel Disease. *Gastroenterology.* 2016;151(4):660-669 e664.
172. Fraga M, Fournier N, Safroneeva E, et al. Primary sclerosing cholangitis in the Swiss Inflammatory Bowel Disease Cohort Study: prevalence, risk factors, and long-term follow-up. *Eur J Gastroenterol Hepatol.* 2017;29(1):91-97.
173. Aadland E, Schrumpf E, Fausa O, et al. Primary sclerosing cholangitis: a long-term follow-up study. *Scand J Gastroenterol.* 1987;22(6):655-664.
174. Fausa O, Schrumpf E, Elgjo K. Relationship of inflammatory bowel disease and primary sclerosing cholangitis. *Semin Liver Dis.* 1991;11(1):31-39.
175. Loftus EV, Jr., Harewood GC, Loftus CG, et al. PSC-IBD: a unique form of inflammatory bowel disease associated with primary sclerosing cholangitis. *Gut.* 2005;54(1):91-96.
176. Bambha K, Kim WR, Talwalkar J, et al. Incidence, clinical spectrum, and outcomes of primary sclerosing cholangitis in a United States community. *Gastroenterology.* 2003;125(5):1364-1369.
177. Halliday JS, Djordjevic J, Lust M, et al. A unique clinical phenotype of primary sclerosing cholangitis associated with Crohn's disease. *J Crohns Colitis.* 2012;6(2):174-181.
178. Schaeffer DF, Win LL, Hafezi-Bakhtiari S, Cino M, Hirschfield GM, El-Zimaity H. The phenotypic expression of inflammatory bowel disease in patients with primary sclerosing cholangitis differs in the distribution of colitis. *Digestive diseases and sciences.* 2013;58(9):2608-2614.

179. Verdonk RC, Dijkstra G, Haagsma EB, et al. Inflammatory bowel disease after liver transplantation: risk factors for recurrence and de novo disease. *Am J Transplant.* 2006;6(6):1422-1429.
180. Angulo P, Lindor KD. Primary sclerosing cholangitis. *Hepatology.* 1999;30(1):325-332.
181. Schrumpf E, Fausa O, Elgjo K, Kolmannskog F. Hepatobiliary complications of inflammatory bowel disease. *Semin Liver Dis.* 1988;8(3):201-209.
182. de Vries AB JM, Blokzijl H, Weersma RK. . Distinctive inflammatory bowel disease phenotype in primary sclerosing cholangitis. *World J Gastroenterol 2015;.* 2015;21(6):1956-1971.
183. Portmann B, Zen Y. Inflammatory disease of the bile ducts-cholangiopathies: liver biopsy challenge and clinicopathological correlation. *Histopathology.* 2012;60(2):236-248.
184. Andersen IM, Tengesdal G, Lie BA, Boberg KM, Karlsen TH, Hov JR. Effects of coffee consumption, smoking, and hormones on risk for primary sclerosing cholangitis. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association.* 2014;12(6):1019-1028.
185. Hov JR, Kummen M. Intestinal microbiota in primary sclerosing cholangitis. *Curr Opin Gastroenterol.* 2017;33(2):85-92.
186. Bergquist A, Montgomery SM, Bahmanyar S, et al. Increased risk of primary sclerosing cholangitis and ulcerative colitis in first-degree relatives of patients with primary sclerosing cholangitis. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association.* 2008;6(8):939-943.
187. Karlsen TH, Franke A, Melum E, et al. Genome-wide association analysis in primary sclerosing cholangitis. *Gastroenterology.* 2010;138(3):1102-1111.
188. Folseraas T, Melum E, Rausch P, et al. Extended analysis of a genome-wide association study in primary sclerosing cholangitis detects multiple novel risk loci. *J Hepatol.* 2012;57(2):366-375.
189. 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(6):670-675.
190. Ellinghaus D, Folseraas T, Holm K, et al. Genome-wide association analysis in primary sclerosing cholangitis and ulcerative colitis identifies risk loci at GPR35 and TCF4. *Hepatology.* 2013;58(3):1074-1083.
191. van Milligen de Wit AW, van Deventer SJ, Tytgat GN. Immunogenetic aspects of primary sclerosing cholangitis: implications for therapeutic strategies. *Am J Gastroenterol.* 1995;90(6):893-900.
192. Woolf GM, Vierling JM. Disappearing intrahepatic bile ducts: the syndromes and their mechanisms. *Semin Liver Dis.* 1993;13(3):261-275.
193. Broome U, Hultcrantz R, Scheynius A. Lack of concomitant expression of ICAM-1 and HLA-DR on bile duct cells from patients with primary sclerosing cholangitis and primary biliary cirrhosis. *Scand J Gastroenterol.* 1993;28(2):126-130.
194. Bloom S, Fleming K, Chapman R. Adhesion molecule expression in primary sclerosing cholangitis and primary biliary cirrhosis. *Gut.* 1995;36(4):604-609.
195. Das KM, Vecchi M, Sakamaki S. A shared and unique epitope(s) on human colon, skin, and biliary epithelium detected by a monoclonal antibody. *Gastroenterology.* 1990;98(2):464-469.
196. Mueller T, Beutler C, Pico AH, et al. Enhanced innate immune responsiveness and intolerance to intestinal endotoxins in human biliary epithelial cells contributes to chronic cholangitis. *Liver Int.* 2011;31(10):1574-1588.

197. Fickert P, Fuchsbichler A, Wagner M, et al. Regurgitation of bile acids from leaky bile ducts causes sclerosing cholangitis in Mdr2 (Abcb4) knockout mice. *Gastroenterology.* 2004;127(1):261-274.
198. Jorgensen KK, Grzyb K, Lundin KE, et al. Inflammatory bowel disease in patients with primary sclerosing cholangitis: clinical characterization in liver transplanted and nontransplanted patients. *Inflamm Bowel Dis.* 2012;18(3):536-545.
199. Grant AJ, Lalor PF, Hubscher SG, Briskin M, Adams DH. MAdCAM-1 expressed in chronic inflammatory liver disease supports mucosal lymphocyte adhesion to hepatic endothelium (MAdCAM-1 in chronic inflammatory liver disease). *Hepatology.* 2001;33(5):1065-1072.
200. Grant AJ, Lalor PF, Salmi M, Jalkanen S, Adams DH. Homing of mucosal lymphocytes to the liver in the pathogenesis of hepatic complications of inflammatory bowel disease. *Lancet.* 2002;359(9301):150-157.
201. Van Nieuwkerk CM, Elferink RP, Groen AK, et al. Effects of Ursodeoxycholate and cholate feeding on liver disease in FVB mice with a disrupted mdr2 P-glycoprotein gene. *Gastroenterology.* 1996;111(1):165-171.
202. Perez MJ, Briz O. Bile-acid-induced cell injury and protection. *World J Gastroenterol.* 2009;15(14):1677-1689.
203. Beuers U, Bilzer M, Chittattu A, et al. Tauroursodeoxycholic acid inserts the apical conjugate export pump, Mrp2, into canalicular membranes and stimulates organic anion secretion by protein kinase C-dependent mechanisms in cholestatic rat liver. *Hepatology.* 2001;33(5):1206-1216.
204. Beuers U, Nathanson MH, Isales CM, Boyer JL. Tauroursodeoxycholic acid stimulates hepatocellular exocytosis and mobilizes extracellular Ca++ mechanisms defective in cholestasis. *J Clin Invest.* 1993;92(6):2984-2993.
205. Beuers U, Throckmorton DC, Anderson MS, et al. Tauroursodeoxycholic acid activates protein kinase C in isolated rat hepatocytes. *Gastroenterology.* 1996;110(5):1553-1563.
206. Paumgartner G, Beuers U. Ursodeoxycholic acid in cholestatic liver disease: mechanisms of action and therapeutic use revisited. *Hepatology.* 2002;36(3):525-531.
207. Martinez JD, Stratagoules ED, LaRue JM, et al. Different bile acids exhibit distinct biological effects: the tumor promoter deoxycholic acid induces apoptosis and the chemopreventive agent ursodeoxycholic acid inhibits cell proliferation. *Nutrition and cancer.* 1998;31(2):111-118.
208. Lindor KD. Ursodiol for primary sclerosing cholangitis. Mayo Primary Sclerosing Cholangitis-Ursodeoxycholic Acid Study Group. *N Engl J Med.* 1997;336(10):691-695.
209. Lindor KD, Kowdley KV, Luketic VA, et al. High-dose ursodeoxycholic acid for the treatment of primary sclerosing cholangitis. *Hepatology.* 2009;50(3):808-814.
210. Beuers U, Spengler U, Kruis W, et al. Ursodeoxycholic acid for treatment of primary sclerosing cholangitis: a placebo-controlled trial. *Hepatology.* 1992;16(3):707-714.
211. Stiehl A, Walker S, Stiehl L, Rudolph G, Hofmann WJ, Theilmann L. Effect of ursodeoxycholic acid on liver and bile duct disease in primary sclerosing cholangitis. A 3-year pilot study with a placebo-controlled study period. *J Hepatol.* 1994;20(1):57-64.
212. Chapman RW. Medical treatment of primary sclerosing cholangitis with ursodeoxycholic acid. *Dig Liver Dis.* 2003;35(5):306-308.
213. 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(5):1464-1472.

214. de Vries EM, Wang J, Leeflang MM, et al. Alkaline Phosphatase at Diagnosis of Primary Sclerosing Cholangitis and One Year Later: Evaluation of Prognostic Value. *Liver Int.* 2016.
215. Poropat G, Giljaca V, Stimac D, Gluud C. Bile acids for primary sclerosing cholangitis. *Cochrane Database Syst Rev.* 2011(1):CD003626.
216. Christensen B, Micic D, Gibson PR, et al. Vedolizumab in patients with concurrent primary sclerosing cholangitis and inflammatory bowel disease does not improve liver biochemistry but is safe and effective for the bowel disease. *Aliment Pharmacol Ther.* 2018;47(6):753-762.
217. Lim TY, Pavlidis P, Gulati S, et al. Vedolizumab in Inflammatory Bowel Disease Associated with Autoimmune Liver Disease Pre- and Postliver Transplantation: A Case Series. *Inflamm Bowel Dis.* 2016;22(10):E39-40.
218. Lynch KD, Chapman RW, Keshav S, et al. Effects of Vedolizumab in Patients With Primary Sclerosing Cholangitis and Inflammatory Bowel Diseases. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association.* 2020;18(1):179-187 e176.
219. Caron B, Peyrin-Biroulet L, Pariente B, et al. Vedolizumab Therapy is Ineffective for Primary Sclerosing Cholangitis in Patients With Inflammatory Bowel Disease: A GETAID Multicentre Cohort Study. *J Crohns Colitis.* 2019;13(10):1239-1247.
220. Fickert P, Hirschfield GM, Denk G, et al. norUrsodeoxycholic acid improves cholestasis in primary sclerosing cholangitis. *J Hepatol.* 2017;67(3):549-558.
221. Kowdley KV, Vuppalanchi R, Levy C, et al. A randomized, placebo-controlled, phase II study of obeticholic acid for primary sclerosing cholangitis. *J Hepatol.* 2020.
222. Tischendorf JJ, Hecker H, Kruger M, Manns MP, Meier PN. Characterization, outcome, and prognosis in 273 patients with primary sclerosing cholangitis: A single center study. *Am J Gastroenterol.* 2007;102(1):107-114.
223. Claessen MM, Vleggaar FP, Tytgat KM, Siersema PD, van Buuren HR. High lifetime risk of cancer in primary sclerosing cholangitis. *J Hepatol.* 2009;50(1):158-164.
224. Bergquist A, Ekbom A, Olsson R, et al. Hepatic and extrahepatic malignancies in primary sclerosing cholangitis. *J Hepatol.* 2002;36(3):321-327.
225. Miros M, Kerlin P, Walker N, Harper J, Lynch S, Strong R. Predicting cholangiocarcinoma in patients with primary sclerosing cholangitis before transplantation. *Gut.* 1991;32(11):1369-1373.
226. Nashan B, Schlitt HJ, Tusch G, et al. Biliary malignancies in primary sclerosing cholangitis: timing for liver transplantation. *Hepatology.* 1996;23(5):1105-1111.
227. Bergeron V, Vienne A, Sokol H, et al. Risk factors for neoplasia in inflammatory bowel disease patients with pancolitis. *Am J Gastroenterol.* 2010;105(11):2405-2411.
228. Broome U, Lofberg R, Veress B, Eriksson LS. Primary sclerosing cholangitis and ulcerative colitis: evidence for increased neoplastic potential. *Hepatology.* 1995;22(5):1404-1408.
229. Zheng HH, Jiang XL. Increased risk of colorectal neoplasia in patients with primary sclerosing cholangitis and inflammatory bowel disease: a meta-analysis of 16 observational studies. *Eur J Gastroenterol Hepatol.* 2016;28(4):383-390.
230. Brentnall TA, Haggitt RC, Rabinovitch PS, et al. Risk and natural history of colonic neoplasia in patients with primary sclerosing cholangitis and ulcerative colitis. *Gastroenterology.* 1996;110(2):331-338.
231. Shetty K, Rybicki L, Brzezinski A, Carey WD, Lashner BA. The risk for cancer or dysplasia in ulcerative colitis patients with primary sclerosing cholangitis. *Am J Gastroenterol.* 1999;94(6):1643-1649.

232. Marchesa P, Lashner BA, Lavery IC, et al. The risk of cancer and dysplasia among ulcerative colitis patients with primary sclerosing cholangitis. *Am J Gastroenterol.* 1997;92(8):1285-1288.
233. Lindberg BU, Broome U, Persson B. Proximal colorectal dysplasia or cancer in ulcerative colitis. The impact of primary sclerosing cholangitis and sulfasalazine: results from a 20-year surveillance study. *Diseases of the colon and rectum.* 2001;44(1):77-85.
234. Soetikno RM, Lin OS, Heidenreich PA, Young HS, Blackstone MO. Increased risk of colorectal neoplasia in patients with primary sclerosing cholangitis and ulcerative colitis: a meta-analysis. *Gastrointest Endosc.* 2002;56(1):48-54.
235. Hanouneh IA, Macaron C, Lopez R, Zein NN, Lashner BA. Risk of colonic neoplasia after liver transplantation for primary sclerosing cholangitis. *Inflamm Bowel Dis.* 2012;18(2):269-274.
236. Singh S, Edakkanambeth Varayil J, Loftus EV, Jr., Talwalkar JA. Incidence of colorectal cancer after liver transplantation for primary sclerosing cholangitis: a systematic review and meta-analysis. *Liver Transpl.* 2013;19(12):1361-1369.
237. Dvorchik I, Subotin M, Demetris AJ, et al. Effect of liver transplantation on inflammatory bowel disease in patients with primary sclerosing cholangitis. *Hepatology.* 2002;35(2):380-384.
238. de Vries EM, Wang J, Williamson KD, et al. A novel prognostic model for transplant-free survival in primary sclerosing cholangitis. *Gut.* 2017.
239. Goode EC, Clark AB, Mells GF, et al. Factors Associated With Outcomes of Patients With Primary Sclerosing Cholangitis and Development and Validation of a Risk Scoring System. *Hepatology.* 2019;69(5):2120-2135.
240. Eaton JE, Vesterhus M, McCauley BM, et al. Primary Sclerosing Cholangitis Risk Estimate Tool (PREsTo) Predicts Outcomes in PSC: A Derivation & Validation Study Using Machine Learning. *Hepatology.* 2018.
241. Ponsioen CY, Chapman RW, Chazouilleres O, et al. Surrogate endpoints for clinical trials in primary sclerosing cholangitis: Review and results from an International PSC Study Group consensus process. *Hepatology.* 2016;63(4):1357-1367.
242. Singal AK, Stanca CM, Clark V, et al. Natural history of small duct primary sclerosing cholangitis: a case series with review of the literature. *Hepatol Int.* 2011;5(3):808-813.
243. Bjornsson E, Olsson R, Bergquist A, et al. The natural history of small-duct primary sclerosing cholangitis. *Gastroenterology.* 2008;134(4):975-980.
244. Weismuller TJ, Trivedi PJ, Bergquist A, et al. Patient Age, Sex, and Inflammatory Bowel Disease Phenotype Associate With Course of Primary Sclerosing Cholangitis. *Gastroenterology.* 2017;152(8):1975-1984 e1978.
245. Floreani A, Rizzotto ER, Ferrara F, et al. Clinical course and outcome of autoimmune hepatitis/primary sclerosing cholangitis overlap syndrome. *Am J Gastroenterol.* 2005;100(7):1516-1522.
246. Bjornsson E, Chari S, Silveira M, et al. Primary sclerosing cholangitis associated with elevated immunoglobulin G4: clinical characteristics and response to therapy. *Am J Ther.* 2011;18(3):198-205.
247. Mendes FD, Jorgensen R, Keach J, et al. Elevated serum IgG4 concentration in patients with primary sclerosing cholangitis. *Am J Gastroenterol.* 2006;101(9):2070-2075.
248. Alswat K, Al-Harthy N, Mazrani W, Alshumrani G, Jhaveri K, Hirschfield GM. The spectrum of sclerosing cholangitis and the relevance of IgG4 elevations in routine practice. *Am J Gastroenterol.* 2012;107(1):56-63.

249. 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(5):1954-1963.
250. Benito de Valle M, Muller T, Bjornsson E, et al. The impact of elevated serum IgG4 levels in patients with primary sclerosing cholangitis. *Dig Liver Dis.* 2014;46(10):903-908.
251. Marelli L, Xirouchakis E, Kalambokis G, Cholongitas E, Hamilton MI, Burroughs AK. Does the severity of primary sclerosing cholangitis influence the clinical course of associated ulcerative colitis? *Gut.* 2011;60(9):1224-1228.
252. Navaneethan U, Venkatesh PG, Mukewar S, et al. Progressive primary sclerosing cholangitis requiring liver transplantation is associated with reduced need for colectomy in patients with ulcerative colitis. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association.* 2012;10(5):540-546.
253. Navaneethan U, Venkatesh PG, Lashner BA, Shen B, Kiran RP. The impact of ulcerative colitis on the long-term outcome of patients with primary sclerosing cholangitis. *Aliment Pharmacol Ther.* 2012;35(9):1045-1053.
254. Nordenvall C, Olen O, Nilsson PJ, et al. Colectomy prior to diagnosis of primary sclerosing cholangitis is associated with improved prognosis in a nationwide cohort study of 2594 PSC-IBD patients. *Aliment Pharmacol Ther.* 2018;47(2):238-245.
255. Wiesner RH, Grambsch PM, Dickson ER, et al. Primary sclerosing cholangitis: natural history, prognostic factors and survival analysis. *Hepatology.* 1989;10(4):430-436.
256. Farrant JM, Hayllar KM, Wilkinson ML, et al. Natural history and prognostic variables in primary sclerosing cholangitis. *Gastroenterology.* 1991;100(6):1710-1717.
257. Dickson ER, Murtaugh PA, Wiesner RH, et al. Primary sclerosing cholangitis: refinement and validation of survival models. *Gastroenterology.* 1992;103(6):1893-1901.
258. Kim WR, Therneau TM, Wiesner RH, et al. A revised natural history model for primary sclerosing cholangitis. *Mayo Clin Proc.* 2000;75(7):688-694.
259. 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(3):652-657.
260. de Vries EM, de Krijger M, Farkkila M, et al. Validation of the prognostic value of histologic scoring systems in primary sclerosing cholangitis: An international cohort study. *Hepatology.* 2017;65(3):907-919.
261. Ji SG, Juran BD, Mucha S, et al. Genome-wide association study of primary sclerosing cholangitis identifies new risk loci and quantifies the genetic relationship with inflammatory bowel disease. *Nat Genet.* 2017;49(2):269-273.
262. Eaton JE, Talwalkar JA, Lazaridis KN, Gores GJ, Lindor KD. Pathogenesis of primary sclerosing cholangitis and advances in diagnosis and management. *Gastroenterology.* 2013;145(3):521-536.
263. Worthington J, Chapman R. Primary sclerosing cholangitis. *Orphanet J Rare Dis.* 2006;1:41.
264. 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;26(12):1374-1379.
265. Younossi ZM, Afendy A, Stepanova M, et al. Development and validation of a primary sclerosing cholangitis-specific patient-reported outcomes instrument: The PSC PRO. *Hepatology.* 2018;68(1):155-165.
266. Brensilver HL, Kaplan MM. Significance of elevated liver alkaline phosphatase in serum. *Gastroenterology.* 1975;68(6):1556-1562.

267. Kaplan MM, Righetti A. Induction of rat liver alkaline phosphatase: the mechanism of the serum elevation in bile duct obstruction. *J Clin Invest.* 1970;49(3):508-516.
268. Chapman R, Fevery J, Kalloo A, et al. Diagnosis and management of primary sclerosing cholangitis. *Hepatology.* 2010;51(2):660-678.
269. Poupon R. Liver alkaline phosphatase: a missing link between choleresis and biliary inflammation. *Hepatology.* 2015;61(6):2080-2090.
270. Stanich PP, Bjornsson E, Gossard AA, Enders F, Jorgensen R, Lindor KD. Alkaline phosphatase normalization is associated with better prognosis in primary sclerosing cholangitis. *Dig Liver Dis.* 2011;43(4):309-313.
271. Al Mamari S, Djordjevic J, Halliday JS, Chapman RW. 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(2):329-334.
272. Lindstrom L, Hultcrantz R, Boberg KM, Friis-Liby I, Bergquist A. Association between reduced levels of alkaline phosphatase and survival times of patients with primary sclerosing cholangitis. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association.* 2013;11(7):841-846.
273. Rupp C, Rossler 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;40(11-12):1292-1301.
274. Goode E, Srivastava B, Clark A, Spiess K, Hirschfield G, WT G. Early clincial features associated with long-term risk of transplantation in primary sclerosing cholangitis: results from the UK-PSC Consortium. *Journal of Hepatology.* 2015;62:S230.
275. Pugh RN, Murray-Lyon IM, Dawson JL, Pietroni MC, Williams R. Transection of the oesophagus for bleeding oesophageal varices. *Br J Surg.* 1973;60(8):646-649.

CHAPTER 2

2

MILDER DISEASE STAGE IN PATIENTS WITH PRIMARY BILIARY CHOLANGITIS OVER A 44-YEAR PERIOD: A CHANGING NATURAL HISTORY

Murillo Perez CF, Goet JC, Lammers WJ, Gulamhusein A, van Buuren HR, Ponsioen CY, Carbone M, Mason A, Corpechot C, Invernizzi P, Mayo MJ, Battezzat PM, Floreani A, Parés A, Nevens F, Kowdley KV, Bruns T, Dalekos GN, Thorburn D, Hirschfield GM, LaRusso NF, Lindor KD, Zachou K, Poupon R, Trivedi PJ, Verhelst X, Janssen HLA, Hansen BE - *on behalf of the Global PBC Study Group*

Hepatology 2018; 67(5):1920-1930

ABSTRACT

Changes over time in the presenting features and clinical course of patients with primary biliary cholangitis are poorly described. We sought to describe temporal trends in patient and disease characteristics over a 44-year period across a large international primary biliary cholangitis cohort of 4,805 patients diagnosed between 1970 and 2014, from 17 centers across Europe and North America. Patients were divided into five cohorts according to their year of diagnosis: 1970-1979 (n = 143), 1980-1989 (n = 858), 1990-1999 (n = 1,754), 2000-2009 (n = 1,815), and ≥2010 (n = 235). Age at diagnosis, disease stage, response to ursodeoxycholic acid, and clinical outcomes were compared. Mean age at diagnosis increased incrementally by 2-3 years per decade from 46.9 ± 10.1 years in the 1970s to 57.0 ± 12.1 years from 2010 onward ($P < 0.001$). The female to male ratio (9:1) and antimitochondrial antibody positivity (90%) were not significantly variable. The proportion of patients presenting with mild biochemical disease (according to Rotterdam staging) increased from 41.3% in the 1970s to 72.2% in the 1990s ($P < 0.001$) and remained relatively stable thereafter. Patients with a mild histological stage at diagnosis increased from 60.4% (1970-1989) to 76.5% (1990-2014) ($P < 0.001$). Correspondingly, response to ursodeoxycholic acid according to Paris-I criteria increased; 51.7% in the 1970s and 70.5% in the 1990s ($P < 0.001$). Recent decades were also characterized by lower decompensation rates (18.5% in the 1970s to 5.8% in the 2000s, $P < 0.001$) and higher 10-year transplant-free survival (48.4%, 68.7%, 79.7%, and 80.1% for each respective cohort; $P < 0.001$).

Conclusion: In recent decades, a pattern of primary biliary cholangitis presentation consistent with an older age at diagnosis alongside reduced disease severity has been noted; the observed trends may be explained by an increase in routine testing of liver function and/or a changing environmental trigger.

STUDY HIGHLIGHTS

What is known:
- Little data are available regarding possible changes over time in clinical presentation and course of the disease in patients with primary biliary cholangitis

What is new here:
- Patients diagnosed in recent decades are older, have milder disease and respond more favorably to treatment with UDCA

INTRODUCTION

Primary biliary cholangitis (PBC) is a chronic autoimmune liver disease characterized by inflammation and destruction of the small intralobular bile ducts[1-3]. The disease mainly affects middle-aged women and has a slow, progressive course that may lead to fibrosis, cirrhosis, and liver failure requiring liver transplantation. The standard treatment for PBC is ursodeoxycholic acid (UDCA) as its long-term use improves liver biochemistry, delays histological progression, and may improve transplant-free survival [4-6]. However, up to 40% of patients can have an inadequate response to UDCA that is associated with reduced transplant-free survival[4,7-9].

PBC is a rare disease with multiple studies reporting an increase in its incidence and prevalence in recent years[10-18]. In a systematic review conducted by Boonstra *et al.* (2012) the incidence of PBC varied from 0.33 to 5.8 per 100,000/year, yet its temporal trends are conflicting as some studies suggest an increase[11,12] while others do not substantiate this finding[19,20]. The prevalence ranged from 1.91 to 40.2 per 100,000 and all investigated studies reported an increase[10]. An increase in prevalence impacts how PBC contributes to the health care system and may be a result of multiple societal and disease factors. It is important to note that initial reports of an increasing prevalence began during the off-label use of UDCA period, which suggests that the increased prevalence in the UDCA-era may be due to prolonged survival[11,14,16]. Correspondingly, the absolute number of liver transplantations for PBC has decreased in Europe and the United States since the introduction of UDCA in the early 1990s[3,9,21-23].

In addition to epidemiological changes, the clinical presentation of PBC has also changed over the years. Whereas most patients presented with an advanced histological stage in earlier decades, nowadays most patients present during an asymptomatic stage[24,25]. Therefore, the underlying assumption that PBC, as a disease, is a static entity may not be accurate. We used a representative large cohort of patients with PBC to assess how disease presentation and prognosis have changed over the last nearly 50 years. In doing so, we provide long-term insight into the changing nature of PBC in clinical practice.

PATIENTS AND METHODS

Population and study design

This was a retrospective study based on patient data retrieved from the Global PBC Study Group database, of which characteristics have been described in previous publications[26,27]. The database comprises long-term follow-up cohorts from 17 centers across North America and Europe. UDCA-treated and non-treated patients aged ≥18 with an established PBC diagnosis from 1970 to 2014, according to internationally accepted guidelines, were included in the study[3,28,29]. (3,28,29). Patients with either a short follow-up (<6 months), an unknown date of important clinical events, an overlap syndrome, or another concomitant liver disease were excluded. Completeness and accuracy of the database was established through visits to individual centers. This study was conducted in accordance with the 1975 Declaration of Helsinki. The protocol was approved by the

institutional research board of the corresponding center and at all participating centers as per local regulations.

Data collection

In the established database, study entry (baseline) was the date of UDCA therapy initiation or the date of first visit for non-treated patients. The following demographic and clinical data were available at study entry: sex, date of birth, date of diagnosis, anti-mitochondrial antibody (AMA) serological status, liver histology, biochemical disease stage, and UDCA therapy (if received and dosage). In addition, the following laboratory values were available at study entry and every 6-12 months until the end of follow-up: alkaline phosphatase, aspartate aminotransferase, alanine aminotransferase, total bilirubin, albumin, and platelet count. Histology was considered if the liver biopsy was completed within 24 months of diagnosis date and dichotomised according to Ludwig et al.[30] and Scheuer's[31] classification; specifically, as mild (stage I and II) and advanced (stage III and IV). The Rotterdam criteria were used to determine patients' biochemical stage. According to these criteria, mild stage is defined as normal bilirubin and albumin, moderate stage is defined as abnormal bilirubin or albumin, and advanced stage is defined as abnormal bilirubin and albumin[9,32]. Baseline aspartate aminotransferase/platelet ratio index, an independent predictor of transplant-free survival, was calculated to stratify patients at risk of liver transplantation and death based on a threshold of 0.54[33]. The first occurrence of hepatic decompensation (ascites, variceal bleeding, or hepatic encephalopathy), hepatocellular carcinoma (HCC), liver transplantation, or all-cause mortality were also retrieved.

In patients that received therapy, biochemical response to UDCA was determined according to Barcelona, Paris-I, Rotterdam, Toronto, and Paris-II criteria[7-9 34,35]. In addition, the GLOBE score was compared to age-specific thresholds to determine UDCA-response[26]. Patients were considered responders if their GLOBE score did not surpass their age-specific threshold.

Statistical Analysis

Patients diagnosed between 1970 and 2014 were divided into five cohorts according to their year of diagnosis: 1970-1979, 1980-1989, 1990-1999, 2000-2009, and ≥2010. To compare patient and disease characteristics across the five cohorts, we conducted Chi-square tests for categorical variables and analyses of variance for continuous data. A p-value less than 0.05 was considered significant for all statistical analyses. Significant results were further analysed to correct for any possible confounding variables and to assess the influence of other explanatory variables on the outcome measure. A multivariable logistic regression was applied to binary outcomes, such as biochemical response to UDCA, biochemical disease stage (moderate and advanced disease stage grouped as advanced), and histological stage (odds ratio with 95% confidence interval [CI]).

For time-to-event analyses, patients diagnosed from 2010 onward were excluded due to a shorter follow-up period than the other cohorts. Patients without an event and those who were lost to follow-up were censored at their last visit. The rates of hepatic decompensation, HCC, and liver transplant-free survival were assessed by Kaplan-Meier

estimates and compared across decades using the log-rank test. If decompensation occurred within the first year of study entry, the patient was excluded from the time-to-event analysis for decompensation. Transplant-free survival was compared across decades in the PBC population and within each decade to an age- and gender-matched Dutch population. These outcomes were also estimated by Cox proportional hazards' modelling (hazards ratio [HR] with 95% CI).

Demographic and clinical characteristics are presented as count (percentage) for categorical data and mean ± standard deviation (SD) for continuous variables. Laboratory values are presented as median (interquartile range [IQR]). Data that were not normally distributed were log transformed for the analyses. All analyses were two-sided and were performed using IBM SPSS Statistics for Windows, version 24.0 (IBM Corp., Armonk, NY).

RESULTS

Study population characteristics

A total of 4805 PBC patients, diagnosed between 1970 and 2014, were included and divided into five cohorts according to their year of diagnosis (**Table 1, Supplementary Table 1**). 143 patients were diagnosed from 1970 to 1979, 858 patients from 1980 to 1989, 1754 patients from 1990 to 1999, 1815 patients from 2000 to 2009, and 235 patients from 2010 onward. The characteristics of each cohort are presented in **Table 1**. The median follow-up for the five respective cohorts were: 6.7 years (IQR 3.0-14.3), 8.9 years (IQR 4.0-14.7), 10.0 years (IQR 6.0-13.9), 5.6 years (IQR 3.4-8.3), and 1.6 years (IQR 1.0-2.1). The mean time from diagnosis to study entry was variable for each cohort: 11.1 years (SD 7.0) for the 1970s, 5.1 years (SD 4.5) for the 1980s, 1.4 years (SD 2.3) for the 1990s, 0.4 years (SD 1.1) for the 2000s, and 0.1 years (SD 0.2) from 2010 onward. To consider this variation, all analyses were repeated in a sub-group of patients (n=3518) with a maximum two-year lag between diagnosis and study entry, which included 14%, 29%, 76%, 93%, and 100% of patients from the main analysis in each respective cohort (**Supplementary Table 2**).

Age and sex trends

The mean age at diagnosis increased incrementally from 46.9±10.1 years in the 1970s to 57.0±12.1 years from 2010 onward ($p<0.001$, **Figure 1A**). This trend was consistent across center, sex, and biochemical disease stage (**Supplementary Figure 1 A-C**). The effect of calendar time on the increase in age at diagnosis remained significant ($p<0.001$) after correcting for sex (**Supplementary Table 3**). Furthermore, the age distribution of patients notably changed over the investigated decades ($p<0.001$, **Figure 1B**). The proportion of patients aged 50-59 years at diagnosis remained relatively stable across the years, whereas the proportion of patients <50 years of age decreased and patients ≥60 years of age increased. There was no significant temporal trend in the female to male ratio, which remained approximately 9:1 (**Table 1**).

Liver biochemistry and serological status

The proportion of patients that were AMA-positive did not significantly differ across the investigated decades (**Table 1**). Median alkaline phosphatase and bilirubin values (×

upper limit of normal) at study entry decreased, while circulating platelet counts were noted to increase ($p<0.001$), which collectively suggests a less advanced disease stage. The proportion of patients with alkaline phosphatase values below 2 × the upper limit of normal increased gradually from 30.0% in the 1970s to 63.1% from 2010 onward ($p<0.001$) (**Figure 2A**). The proportion of patients with normal serum bilirubin concentrations also increased from 51.1% in the 1970s to 77.6% in the 1990s, after which it remained relatively stable ($p<0.001$) (**Figure 2B**). Furthermore, a reduced percentage of patients with aspartate aminotransferase/platelet ratio index >0.54 at study entry was observed (**Table 1**).

Table 1. Demographic and clinical characteristics of PBC patients at study entry over calendar time

Baseline characteristics	1970-1979 (n=143)	1980-1989 (n=858)	1990-1999 (n=1754)	2000-2009 (n=1815)	≥2010 (n=235)	p-value
Age at diagnosis, y[a]	46.9 (10.1)	50.1 (10.7)	52.8 (11.5)	55.0 (12.5)	57.0 (12.1)	<0.001
Female	131 (91.6)	775 (90.3)	1593 (90.8)	1619 (89.2)	207 (88.1)	0.396
AMA-positive[b]	123/140 (87.9)	763/842 (90.6)	1565/1704 (91.8)	1599/1765 (90.6)	213/235 (90.6)	0.449
Laboratory values[c]						
Serum ALP (×ULN)	2.99 (1.85-4.77)	3.20 (1.95-5.23)	2.03 (1.30-3.56)	1.79 (1.19-3.05)	1.55 (1.08-2.93)	<0.001
Serum bilirubin (×ULN)	0.93 (0.60-2.1)	0.81 (0.52-1.30)	0.64 (0.47-1.00)	0.60 (0.41-0.95)	0.59 (0.41-1.0)	<0.001
Serum AST (×ULN)	1.59 (1.06-2.32)	1.95 (1.20-2.77)	1.35 (0.87-2.20)	1.30 (0.90-2.00)	1.29 (0.85-2.07)	<0.001
Serum ALT (×ULN)	1.30 (0.85-2.47)	2.00 (1.3-3.1)	1.66 (1.03-2.68)	1.42(0.90-2.27)	1.32 (0.75-2.38)	<0.001
Serum albumin (×LLN)	1.11 (0.99-1.21)	1.16 (1.06-1.26)	1.14 (1.06-1.23)	1.14 (1.06-1.23)	1.14 (1.03-1.23)	0.005
Platelet count ($\times 10^9$/L)	194 (127-242.5)	224 (165-275)	238 (185-289)	258 (204-311)	237 (174.5-291)	<0.001
APRI (>0.54)[d]	61 (76.3)	260 (69.0)	456 (52.3)	476 (47.4)	85 (54.1)	<0.001
Biochemical disease stage[e]	121/143	627/859	985/1755	1073/1816	152/235	<0.001
Mild	50/121 (41.3)	370/627 (59.0)	711/985 (72.2)	757/1073 (70.5)	106/152 (69.7)	
Moderately advanced	51/121 (42.1)	196/627 (31.3)	205/985 (20.8)	238/1073 (22.2)	27/152 (17.8)	
Advanced	20/121 (16.5)	61/627 (9.7)	69/985 (7.0)	78/1073 (7.3)	19/152 (12.5)	
Histological disease stage[f]	326/1001		948/1754	943/2050		<0.001
Mild (I or II)	197 (60.4)		634 (66.9)	721 (76.5)		
Advanced (III or IV)	129 (39.6)		314 (33.1)	222 (23.5)		
UDCA-treated[g]	78/139 (56.1)	735/832 (88.3)	1605/1737 (92.4)	1563/1789 (87.4)	195/230 (84.8)	<0.001

Data represented as mean (standard deviation), n (%), or median (interquartile range). AMA, antimitochondrial antibody; ALP, alkaline phosphatase; AST, aspartate aminotransferase; ULN, upper limit of normal; ALT, alanine aminotransferase; LLN, lower limit of normal; APRI, AST to platelet ratio index; UDCA, ursodeoxycholic acid.
[a]Age at diagnosis not available for one patient in 2000-2009 cohort. [b]AMA status was available for 4686 (97.5%) patients.
[c]ALP, bilirubin, AST, and ALT were log transformed prior to analyses and availability for laboratory values is as follows: ALP: 3560 (74.1%); Bilirubin: 3595 (74.8%); AST: 3460 (72.0%); ALT: 3007 (62.6%); Albumin: 3039 (63.2%); Platelet count: 2769 (57.6%)
[d]The cut-point APRI >0.54 at baseline is predictive of liver transplantation or death (33) [e]Biochemical disease stage classification according to Rotterdam criteria (9) was available in 2958 (61.6%) patients. [f]Histological disease stage at diagnosis according to Ludwig et al. and Scheuer (30,31) classification was available in 2217 (46.1%) patients.
[g]UDCA therapy status was available for 4727 patients (98.4%).

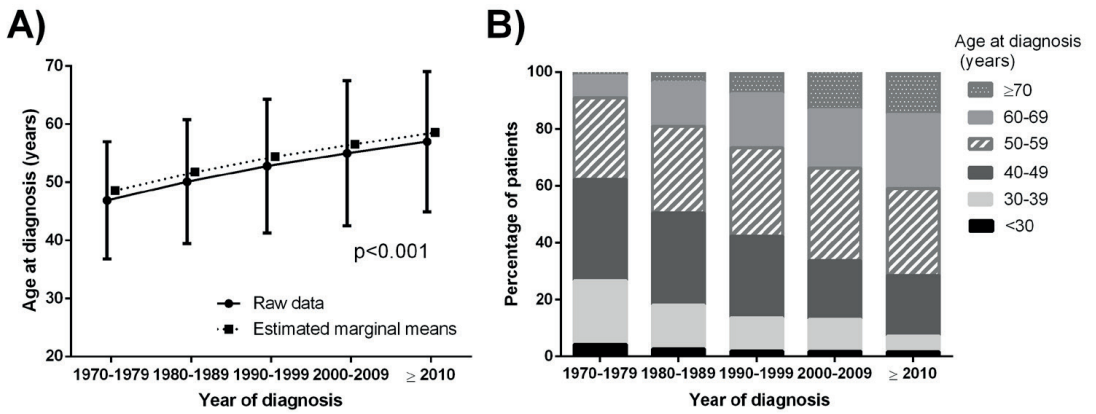


Figure 1. Age at diagnosis of PBC patients across different decades. A) Mean age (± standard deviation) at diagnosis (*dots*) and estimated marginal means (*squares*) obtained after adjusting for sex. **B)** The distribution of age groups over calendar time.

Trends in biochemical and histological disease stage

There was a gradual increase in the proportion of patients presenting with a mild biochemical disease stage from the 1970s to 1990s, and remained stable thereafter (p<0.001, **Figure 2C**). In a multivariable logistic regression, calendar time was a significant predictor for biochemical disease stage (p<0.001) after adjusting for sex and age at diagnosis. Earlier decades were associated with an advanced biochemical disease stage.

Out of 2831 patients who underwent liver biopsy at diagnosis, 2217 patients had histological disease stage available and were included in a subgroup analysis that combined cohorts due to the limited number of biopsies in the first and last cohorts. There were 326 biopsies from 1970-1989, 948 biopsies from 1990-1999, and 943 from 2000-2014. The proportion of patients with a mild histological stage (I or II) at diagnosis increased with time **(Table 1, Figure 2D)**. In a multivariable logistic regression, calendar time was a significant predictor for histological stage after adjusting for sex and age at diagnosis (p<0.001).

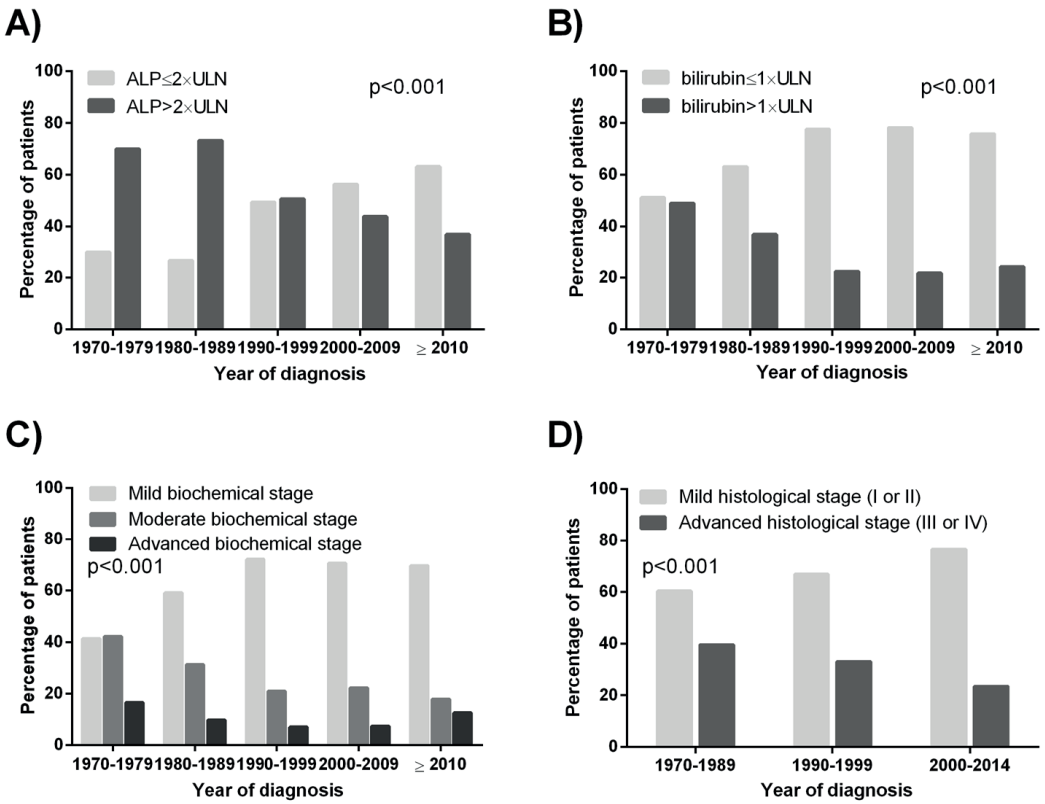


Figure 2. Study entry characteristics associated with disease severity of patients diagnosed in different decades. A) Percentage of patients with alkaline phosphatase (ALP) above or below 2 times the upper limit of normal (×ULN). **B)** Percentage of patients with bilirubin above or below 1×ULN. C) Percentage of patients corresponding to each biochemical stage according to Rotterdam criteria (9); mild (normal albumin and bilirubin), moderate (abnormal albumin or bilirubin), advanced (abnormal albumin and bilirubin). **D)** Percentage of patients corresponding to each histological stage at diagnosis according to Ludwig et al.'s[30] and Scheuer's[31] classification: mild (stage I and II) or advanced (stage III and IV).

Trends in UDCA-response rates

The proportion of patients that ever received UDCA increased across the investigated decades (p<0.001, **Table 1**). In patients that received UDCA, the median number of years between diagnosis and the start of UDCA therapy decreased across the respective cohorts (1970s to ≥2010): 12.6 years (IQR 10.6-16.1), 4.4 years (IQR 2.1-8.1), 0.23 years (IQR 0.0-2.0), 0.05 years (IQR 0.0-0.41), and 0.0 years (IQR 0.0-0.04). Additionally, the median initial dosage of UDCA received by patients across the five respective cohorts increased: 9.4 mg/kg/day (IQR 8.5-11.0), 10.0 mg/kg/day (IQR 8.7-13.7), 12.2 mg/kg/day (IQR 9.2-14.7), 13.5 mg/kg/day (IQR 11.1-15.3), 13.3 mg/kg/day (IQR 11.1-15.1).

The proportion of UDCA-responders according to Paris-I, Toronto, Paris-II, Rotterdam, and GLOBE score criteria increased over the investigated decades (p<0.001), but not according to Barcelona criteria (**Figure 3, Supplementary Table 4**). Importantly, this trend remained true in patients who did not meet the individual criteria at baseline (**Supplementary Table**

5). In a multivariable logistic regression, calendar time was not a significant predictor for UDCA-response according to Paris-I criteria (**Table 2**).

Table 2. Multivariable logistic regression for the attainment of biochemical response according to Paris-I[a] (n=2283)

Variable	OR	95% CI	p-value
Male sex	0.90	0.63-1.29	0.58
Year of diagnosis			0.67
1970-1979	1.00		
1980-1989	0.80	0.37-1.71	0.66
1990-1999	1.01	0.44-2.37	0.96
2000-2009	0.97	0.40-2.32	0.94
≥2010	0.92	0.33-2.57	0.88
Age at diagnosis			0.04
<30	1.00		
30-39	1.29	0.53-3.15	0.57
40-49	1.41	0.60-3.33	0.44
50-59	1.95	0.82-4.59	0.13
60-69	2.06	0.86-4.96	0.11
≥70	2.06	0.82-5.21	0.13
Log bilirubin (×ULN)	0.01	0.01-0.02	<0.001
Log ALP (×ULN)	0.12	0.08-0.18	<0.001
Difference between diagnosis and study entry (years)	0.98	0.94-1.03	0.44

OR, odds ratio; CI, confidence interval; ULN, upper limit of normal; ALP, alkaline phosphatase. [a]Response rate according to Paris-I is defined as: ALP ≤3 ×ULN, AST ≤2 ×ULN, and normal bilirubin after 1 year of UDCA therapy.

Response was associated with an increased age at diagnosis, and lower alkaline phosphatase and bilirubin levels (p<0.001). Additionally, calendar time was also not a significant predictor for UDCA-response according to Toronto, Paris-II, Rotterdam, and GLOBE score criteria (results not shown).

Decompensation, HCC, and transplant-free survival

The 10-year incidence rate of hepatic decompensation (ascites, variceal bleeding, or hepatic encephalopathy, whichever came first) decreased over time: 18.5% in the 1970s, 13.7% in the 1980s, 8.5% in the 1990s, and 5.8% in the 2000s (**Figure 4Ai**). All pairwise comparisons were significantly different, except the difference between the 1970s and 1980s cohorts (p=0.45). In a multivariable Cox regression, a temporal trend of lower decompensation risk was observed after adjusting for sex and age at diagnosis (**Figure 4Bi**, p=0.07). Calendar time as a continuous variable was a significant predictor for hepatic decompensation (HR per 10-year increase: 0.57, 95% CI 0.44-0.75, p<0.001).

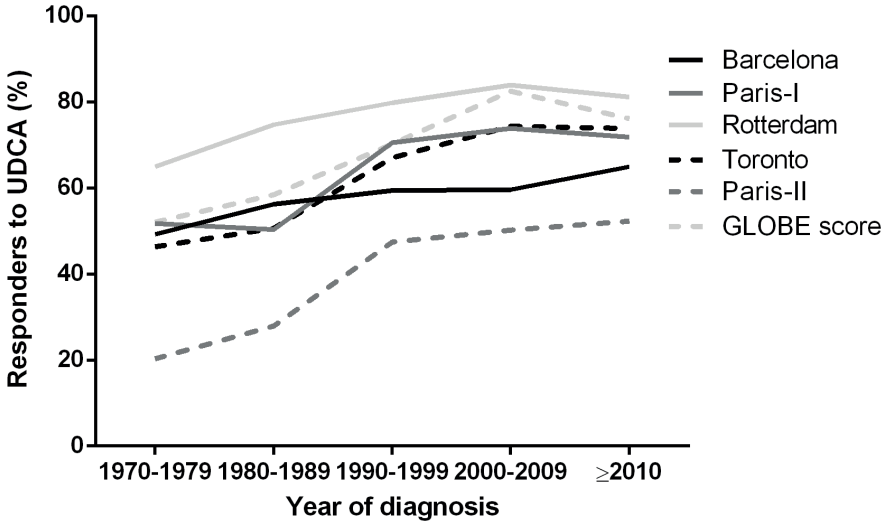


Figure 3. Response rates to ursodeoxycholic acid (UDCA) therapy over calendar time. Response was determined according to various published criteria: Barcelona, Paris-I, Rotterdam, Toronto, Paris-II, and the GLOBE score (7-9,26,34,35). Response rates according to all criteria were significantly different over calendar time (p<0.001), except Barcelona criteria (p=0.19).

The 10-year HCC incidence rates across the investigated decades were: 10.3%, 4.0%, 2.1%, and 2.3%, respectively (**Figure 4Aii**). The Kaplan-Meier estimate of cumulative HCC incidence was significantly higher in the 1970s compared to the 1980s (p=0.01), 1990s (p<0.001), and 2000s (p<0.001). In a multivariable Cox regression, calendar time was not a significant predictor for HCC risk (p=0.68) after adjusting for sex, age at diagnosis, and UDCA treatment (**Figure 4Bii**).

The 10-year liver-related death rate decreased from 1970-2009: 34.6%, 13.2%, 5.6%, and 6.4% (p<0.001). Furthermore, the 10-year transplant-free survival rate improved over the four respective investigated decades: 48.4%, 68.7%, 79.7%, and 80.1% (**Figure 4Aiii**). There was a significant difference in transplant-free survival between the 1970s and 1980s (p<0.001), and between the 1980s and 1990s (p<0.001). However, the transplant-free survival rates between the 1990s and 2000s were equivalent (p=0.80). In a multivariable Cox regression, calendar time remained an independent predictor of transplant-free survival, and earlier decades were associated with an increased risk for liver transplantation and all-cause mortality (**Figure 4Biii, Supplementary Table 6**). Furthermore, the 10-year transplant-free survival of PBC patients has improved even when compared to an age- and gender-matched general population (1970s: HR 4.38, 95% CI 3.54-5.43, p<0.001; 1980s: HR 2.90, 95% CI 2.60-3.24, p<0.001; 1990s: HR 2.14, 95% CI 1.94-2.36, p<0.001; 2000s: HR 1.93, 95% CI 1.69-2.21, p<0.001).

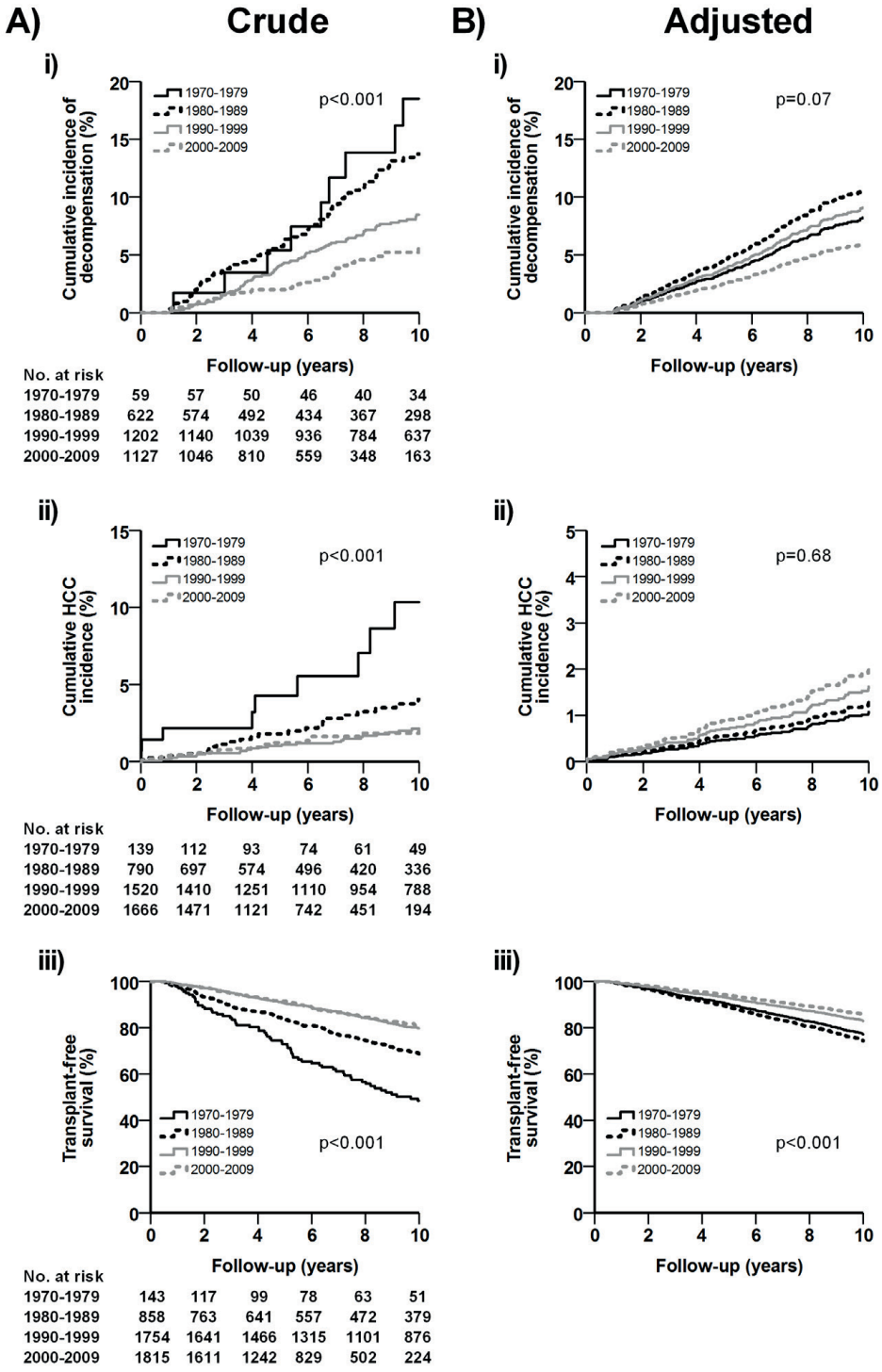


Figure 4. Time-to-event analyses of decompensation, hepatocellular carcinoma (HCC), and liver transplantation or death over calendar time. A) Kaplan-Meier (crude) and **B)** Multivariable Cox regression (adjusted) estimates of **i)** cumulative incidence of decompensation, **ii)** cumulative incidence of hepatocellular carcinoma (HCC), and **iii)** transplant-free survival.

DISCUSSION

In this study of a large, internationally representative cohort of PBC patients, we demonstrate that patients diagnosed in recent decades are older and have a milder disease stage compared to patients diagnosed in earlier decades. In addition, more patients respond favourably to UDCA therapy and have improved transplant-free survival. To the best of our knowledge no previous study has reported on these PBC trends. These results provide unique insight into the possible changing natural history of PBC over the last five decades. It is noteworthy to mention that similar results have been observed in a study from Sweden that included 246 patients diagnosed with primary sclerosing cholangitis between 1984 and 2004. Bergquist *et al.* reported an increase in age at diagnosis and lower frequency of symptoms in patients diagnosed after 1998[36].

Although some of the observed trends could be potentially attributed to more sensitive AMA tests that detect the disease at an earlier stage, we speculate that any changes in AMA testing have not had a major impact in the observed temporal trends. The conventional method of AMA detection is indirect immunofluorescence, yet there has been an increase in ELISA-based assays and immunoblotting that have led to greater sensitivity and specificity[37]. These improvements would translate to an increase in the proportion of AMA-positive patients, however this has remained unchanged.

We demonstrate a 10-year increase in the mean age at diagnosis from 1970 to 2014. A similar increase has been reported previously in the Canadian PBC population, in which prevalent cases in 1996 had a median age of 53, whereas prevalent cases in 2002 had a median age of 57[18]. These numbers coincide with the findings from our study, in which the mean age at diagnosis in the 1990s and 2000s is 52.8 and 55.0 years, respectively. Furthermore, an increased proportion of patients diagnosed in recent years are over 50 years of age and account for 71.5% of patients diagnosed on 2010 and beyond. Comparable results were found within the UK-PBC cohort, in which 75% of patients prevalent between 2008 and 2010 were over 50 years of age[38].

The increase in age may be attributed to the general aging of the population, as the median age in Northern America and Europe has reportedly increased from 30 in 1970 to 40 in 2015[39]. This represents a 10-year increase over a 45-year period, which is similar to the 10-year increase in age at diagnosis we observe over a 44-year interval. Furthermore, the 34% absolute increase of PBC patients 50 years old and above from 1970 to 2014 was greater than that of the general population, which was only 11% (25% in 1970 to 36% in 2015). The increase in age may also be attributed to differences in the trigger for a PBC diagnosis over the years. Although we are not able to assess the symptoms in our cohort, we speculate that patients in recent decades are predominantly asymptomatic and are therefore diagnosed when they see their physician to undergo routine testing of liver function, which occurs more frequently in older individuals. Conversely, younger patients in earlier decades were more likely to develop symptoms, which led to their diagnoses[40,41]. Lastly, the increase in age may be disease-specific and represent a shift in the natural history of PBC towards a new older at-risk population, considering the increase in age was observed irrespective of biochemical disease stage. It can also be speculated that the

later onset is a result of a prolonged subclinical disease period and potentially a delayed exposure to an unknown environmental trigger due to temporal changes in lifestyle.

An older age at diagnosis is clinically important because it has been associated with an increased likelihood of meeting Paris-I criteria for response to UDCA (38). Similarly, we found an older age at diagnosis to be an independent predictor of Paris-I response, yet calendar time was not a significant predictor. These results indicate the increase in age at diagnosis may be an important factor contributing to the increase in UDCA-response rather than calendar time itself. Furthermore, the low response rates observed in earlier decades can be a result of inadequate UDCA dosages and the delay in treatment. The importance of an adequate UDCA dosage of 13-15mg/kg per day has been emphasized in a study that found 40% of UDCA-non-responders in whom the dosage was increased became responders[42,43].

In recent decades, patients present at an older age, yet they have milder biochemical and histological disease stage. Improved disease severity might be explained by an earlier detection of PBC due to improved disease awareness leading to liver function tests and AMA assays[44,45]. The histological disease stage at diagnosis has important prognostic implications for UDCA-response and survival. Advanced histological stages are associated with an increased risk of treatment failure[8]. In addition, the survival of UDCA-treated patients in stage I/II is similar to that of an age- and sex-matched control population, while the probability of liver transplantation or death is significantly increased in patients with advanced histological stages[46].

Although a decrease in the number of liver transplantations for PBC has been reported over the years[22], an improvement in transplant-free survival has not been previously documented. In a Canadian population-based study of patients diagnosed between 1996 and 2002, Myers *et al.* (2009) did not observe a significant difference in survival according to year of diagnosis[18]. The lack of difference in survival may be attributed to the small interval of study, which only spanned six years. The reported increase in median age of the general population well reflects an increase in life expectancy over time[39]; therefore transplant-free survival was compared to the general population. Our study showed that transplant-free survival improved over a 44-year period, even when compared to the general population, and supports its potential role in the increased prevalence of PBC.

The inclusion of a large population of PBC patients from different geographical regions, long-term follow-up, and broad study period are some of the strengths of our study. However, some limitations need to be considered. First, the 1970s and 1980s cohorts were susceptible to a delay in documentation since study entry can be many years after the date of diagnosis in these cohorts. As such, the difference in years between these two dates was included in all multivariable analyses and we assessed a sub-group of patients with a maximum two-year difference. The same trends emerged in the sub-group analyses, thus excluding the possibility that the delay in documentation is the reason for an advanced disease in the early cohorts. Second, due to the retrospective nature of the study, biochemical data was not available for all patients and thus response to UDCA could not be determined for all patients. To account for missing laboratory values, all

analyses were repeated in an imputed dataset and revealed similar results. Lastly, the trends observed in our study cohort could not be assessed for correlations with symptom profiles or various environmental factors previously associated with PBC, such as smoking, age at first pregnancy, or the use of hormonal replacement therapy[47]. Even though the trends observed may be due to a selection of patients whose diagnosis is triggered by symptoms or complications in earlier decades rather than routine liver function tests as in recent decades, we describe the presenting characteristics of a typical PBC patient seen by physicians and how they have changed over time. The observed temporal trends warrant further investigation in other PBC populations to determine whether they are universally applicable and to explore the potential influence of a changing environmental trigger.

In conclusion, we demonstrate a 10-year increase in age at diagnosis accompanied by milder disease severity at presentation of PBC patients. These findings provide the most comprehensive evidence of a changing natural history of PBC to date.

REFERENCES

1. Kaplan MM, Gershwin ME. Primary biliary cirrhosis. *N Engl J Med.* 2005;353(12):1261-1273.
2. Poupon R. Primary biliary cirrhosis: a 2010 update. *J Hepatol.* 2010;52(5):745-758.
3. Lindor KD, Gershwin ME, Poupon R, et al. Primary biliary cirrhosis. *Hepatology.* 2009;50(1):291-308.
4. Pratt DS. Primary Biliary Cholangitis--A New Name and a New Treatment. *N Engl J Med.* 2016;375(7):685-687.
5. Poupon RE, Lindor KD, Cauch-Dudek K, Dickson ER, Poupon R, Heathcote EJ. Combined analysis of randomized controlled trials of ursodeoxycholic acid in primary biliary cirrhosis. *Gastroenterology.* 1997;113(3):884-890.
6. Corpechot C, Carrat F, Poupon R, Poupon RE. Primary biliary cirrhosis: incidence and predictive factors of cirrhosis development in ursodiol-treated patients. *Gastroenterology.* 2002;122(3):652-658.
7. Pares A, Caballeria L, Rodes J. Excellent long-term survival in patients with primary biliary cirrhosis and biochemical response to ursodeoxycholic Acid. *Gastroenterology.* 2006;130(3):715-720.
8. Corpechot C, Abenavoli L, Rabahi N, et al. Biochemical response to ursodeoxycholic acid and long-term prognosis in primary biliary cirrhosis. *Hepatology.* 2008;48(3):871-877.
9. Kuiper EM, Hansen BE, de Vries RA, et al. Improved prognosis of patients with primary biliary cirrhosis that have a biochemical response to ursodeoxycholic acid. *Gastroenterology.* 2009;136(4):1281-1287.
10. Boonstra K, Beuers U, Ponsioen CY. Epidemiology of primary sclerosing cholangitis and primary biliary cirrhosis: a systematic review. *J Hepatol.* 2012;56(5):1181-1188.
11. Metcalf JV, Bhopal RS, Gray J, Howel D, James OF. Incidence and prevalence of primary biliary cirrhosis in the city of Newcastle upon Tyne, England. *Int J Epidemiol.* 1997;26(4):830-836.
12. Boonstra K, Kunst AE, Stadhouders PH, et al. Rising incidence and prevalence of primary biliary cirrhosis: a large population-based study. *Liver Int.* 2014;34(6):e31-38.
13. Eriksson S, Lindgren S. The prevalence and clinical spectrum of primary biliary cirrhosis in a defined population. *Scand J Gastroenterol.* 1984;19(7):971-976.
14. Myszor M, James OF. The epidemiology of primary biliary cirrhosis in north-east England: an increasingly common disease? *Q J Med.* 1990;75(276):377-385.
15. Remmel T, Remmel H, Uibo R, Salupere V. Primary biliary cirrhosis in Estonia. With special reference to incidence, prevalence, clinical features, and outcome. *Scand J Gastroenterol.* 1995;30(4):367-371.
16. James OF, Bhopal R, Howel D, Gray J, Burt AD, Metcalf JV. Primary biliary cirrhosis once rare, now common in the United Kingdom? *Hepatology.* 1999;30(2):390-394.
17. Pla X, Vergara M, Gil M, et al. Incidence, prevalence and clinical course of primary biliary cirrhosis in a Spanish community. *Eur J Gastroenterol Hepatol.* 2007;19(10):859-864.
18. Myers RP, Shaheen AA, Fong A, et al. Epidemiology and natural history of primary biliary cirrhosis in a Canadian health region: a population-based study. *Hepatology.* 2009;50(6):1884-1892.
19. McNally RJ, James PW, Ducker S, Norman PD, James OF. No rise in incidence but geographical heterogeneity in the occurrence of primary biliary cirrhosis in North East England. *Am J Epidemiol.* 2014;179(4):492-498.

20. Baldursdottir TR, Bergmann OM, Jonasson JG, Ludviksson BR, Axelsson TA, Bjornsson ES. The epidemiology and natural history of primary biliary cirrhosis: a nationwide population-based study. *Eur J Gastroenterol Hepatol.* 2012;24(7):824-830.
21. Gross RG, Odin JA. Recent advances in the epidemiology of primary biliary cirrhosis. *Clin Liver Dis.* 2008;12(2):289-303; viii.
22. Lee J, Belanger A, Doucette JT, Stanca C, Friedman S, Bach N. Transplantation trends in primary biliary cirrhosis. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association.* 2007;5(11):1313-1315.
23. Hirschfield GM, Mason A, Luketic V, et al. Efficacy of obeticholic acid in patients with primary biliary cirrhosis and inadequate response to ursodeoxycholic acid. *Gastroenterology.* 2015;148(4):751-761 e758.
24. Christensen E, Crowe J, Doniach D, et al. Clinical pattern and course of disease in primary biliary cirrhosis based on an analysis of 236 patients. *Gastroenterology.* 1980;78(2):236-246.
25. Locke GR, 3rd, Therneau TM, Ludwig J, Dickson ER, Lindor KD. Time course of histological progression in primary biliary cirrhosis. *Hepatology.* 1996;23(1):52-56.
26. Lammers WJ, Hirschfield GM, Corpechot C, et al. Development and Validation of a Scoring System to Predict Outcomes of Patients With Primary Biliary Cirrhosis Receiving Ursodeoxycholic Acid Therapy. *Gastroenterology.* 2015;149(7):1804-1812 e1804.
27. Lammers WJ, van Buuren HR, Hirschfield GM, 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.* 2014;147(6):1338-1349.
28. Nguyen DL, Juran BD, Lazaridis KN. Primary biliary cirrhosis. *Best Pract Res Clin Gastroenterol.* 2010;24(5):647-654.
29. EASL Clinical Practice Guidelines: The diagnosis and management of patients with primary biliary cholangitis. *J Hepatol.* 2017;67(1):145-172.
30. Ludwig J, Dickson ER, McDonald GS. Staging of chronic nonsuppurative destructive cholangitis (syndrome of primary biliary cirrhosis). *Virchows Arch A Pathol Anat Histol.* 1978;379(2):103-112.
31. Scheuer P. Primary biliary cirrhosis. *Proc R Soc Med.* 1967;60(12):1257-1260.
32. ter Borg PC, Schalm SW, Hansen BE, van Buuren HR, Dutch PBCSG. Prognosis of ursodeoxycholic Acid-treated patients with primary biliary cirrhosis. Results of a 10-yr cohort study involving 297 patients. *Am J Gastroenterol.* 2006;101(9):2044-2050.
33. Trivedi PJ, Bruns T, Cheung A, et al. Optimising risk stratification in primary biliary cirrhosis: AST/platelet ratio index predicts outcome independent of ursodeoxycholic acid response. *Journal of Hepatology.* 2014;60(6):1249-1258.
34. Kumagi T, Guindi M, Fischer SE, et al. Baseline ductopenia and treatment response predict long-term histological progression in primary biliary cirrhosis. *Am J Gastroenterol.* 2010;105(10):2186-2194.
35. Corpechot C, Chazouilleres O, Poupon R. Early primary biliary cirrhosis: biochemical response to treatment and prediction of long-term outcome. *J Hepatol.* 2011;55(6):1361-1367.
36. Bergquist A, Said K, Broome U. Changes over a 20-year period in the clinical presentation of primary sclerosing cholangitis in Sweden. *Scand J Gastroenterol.* 2007;42(1):88-93.
37. Oertelt S, Rieger R, Selmi C, et al. A sensitive bead assay for antimitochondrial antibodies: Chipping away at AMA-negative primary biliary cirrhosis. *Hepatology.* 2007;45(3):659-665.

38. Carbone M, Mells GF, Pells G, et al. Sex and age are determinants of the clinical phenotype of primary biliary cirrhosis and response to ursodeoxycholic acid. *Gastroenterology.* 2013;144(3):560-569 e567; quiz e513-564.
39. **United Nations.** World population prospects: the 2017 revision. https://esa.un.org/unpd/wpp/Publications/Files/WPP2017_Key-Findings.pdf.
40. Mahl TC, Shockcor W, Boyer JL. Primary biliary cirrhosis: survival of a large cohort of symptomatic and asymptomatic patients followed for 24 years. *J Hepatol.* 1994;20(6):707-713.
41. Prince M, Chetwynd A, Newman W, Metcalf JV, James OF. Survival and symptom progression in a geographically based cohort of patients with primary biliary cirrhosis: follow-up for up to 28 years. *Gastroenterology.* 2002;123(4):1044-1051.
42. Lammers WJ, Leeman M, Ponsioen CI, et al. How the concept of biochemical response influenced the management of primary biliary cholangitis over time. *Neth J Med.* 2016;74(6):240-246.
43. Angulo P, Dickson ER, Therneau TM, et al. Comparison of three doses of ursodeoxycholic acid in the treatment of primary biliary cirrhosis: a randomized trial. *J Hepatol.* 1999;30(5):830-835.
44. Prince MI, Chetwynd A, Craig WL, Metcalf JV, James OF. Asymptomatic primary biliary cirrhosis: clinical features, prognosis, and symptom progression in a large population based cohort. *Gut.* 2004;53(6):865-870.
45. European Association for the Study of the L. EASL Clinical Practice Guidelines: management of cholestatic liver diseases. *J Hepatol.* 2009;51(2):237-267.
46. Corpechot C, Carrat F, Bahr A, Chretien Y, Poupon RE, Poupon R. The effect of ursodeoxycholic acid therapy on the natural course of primary biliary cirrhosis. *Gastroenterology.* 2005;128(2):297-303.
47. Gershwin ME, Selmi C, Worman HJ, et al. Risk factors and comorbidities in primary biliary cirrhosis: a controlled interview-based study of 1032 patients. *Hepatology.* 2005;42(5):1194-1202.

SUPPLEMENTARY DATA

Supplementary Table 1. Distribution of PBC patients across calendar time and center

Center	1970-1979 (n=143)	1980-1989 (n=858)	1990-1999 (n=1754)	2000-2009 (n=1815)	≥2010 (n=235)	Total N=4805
North Europe						
Rotterdam, The Netherlands (1973-2012)[a]	25 (17.5)	122 (14.2)	274 (15.6)	361 (19.9)	37 (15.7)	819
Leuven, Belgium (1974-2011)[b]	5 (3.5)	20 (2.3)	44 (2.5)	64 (3.5)	13 (5.5)	146
Ghent, Belgium (1991-2014)[c]	0	0	4 (0.2)	14 (0.8)	6 (2.6)	24
Paris, France (1974-2001)[b]	11 (7.7)	209 (24.4)	113 (6.4)	14 (0.8)	0	347
London, UK (1972-2007)[b]	11 (7.7)	31 (3.6)	68 (3.9)	26 (1.4)	0	136
Birmingham, UK (1972-2011)[b]	1 (0.7)	4 (0.5)	79 (4.5)	264 (14.5)	14 (6.0)	362
Jena, Germany (1979-2013)[c]	1 (0.7)	5 (0.6)	38 (2.2)	53 (2.9)	24 (10.2)	121
South Europe						
Milan, Italy (1970-2005)[b,d]	71 (49.7)	217 (25.3)	183 (10.4)	62 (3.4)	0	533
Padua, Italy (1972-2012)[b]	3 (2.1)	38 (4.4)	102 (5.8)	99 (5.5)	28 (11.9)	270
Barcelona, Spain (1971-2005)[b]	3 (2.1)	51 (5.9)	147 (8.4)	68 (3.7)	0	269
Larissa, Greece (1991-2014)[c]	0	0	1 (0.1)	76 (4.2)	23 (9.8)	100
North America						
Rochester, USA (1970-2012)[b]	2 (1.4)	11 (1.3)	245 (14)	352 (19.4)	69 (29.4)	679
Toronto, Canada (1974-2010)[b]	9 (6.3)	87 (10.1)	229 (13.1)	257 (14.2)	1 (0.4)	583
Texas, USA (1977-2011)[b]	1 (0.7)	62 (7.2)	209 (11.9)	44 (2.4)	10 (4.3)	326
Edmonton, Canada (1989-2007)[b]	0	1 (0.1)	13 (0.7)	42 (2.3)	0	56
Seattle, USA (1995-2012)[b]	0	0	5 (0.3)	19 (1)	10 (4.3)	34

Data represented as n (% within corresponding decade).
[a]Comprised of centers across the Netherlands (mainly secondary centers).
[b]Tertiary center. [c]Secondary center. [d]Comprised of two centers.

Supplementary Table 2. Calendar time trends in patients with a maximum lag of 2 years between diagnosis and study entry

Characteristics	1970-1979 (n=20)	1980-1989 (n=245)	1990-1999 (n=1331)	2000-2009 (n=1687)	≥2010 (n=235)	p-value
Age at diagnosis, y[a]	49.3 (12.9)	52.3 (11.7)	52.9 (11.6)	55.0 (12.6)	57.0 (12.1)	<0.001
Female	18 (90)	220 (89.8)	1210 (90.9)	1509 (89.4)	207 (88.1)	0.60
AMA-positive[b]	16 (84.2)	217 (90.0)	1190 (91.8)	1487 (90.7)	213 (90.6)	0.63
Laboratory values[c]						
Serum ALP (×ULN)	3.05 (1.15-7.32)	3.76 (2.04-6.50)	2.14 (1.33-3.69)	1.83 (1.21-3.08)	1.55 (1.08-2.93)	<0.001
Serum bilirubin (×ULN)	1.3 (0.59-4.56)	0.74 (0.47-1.27)	0.62 (0.47-1.00)	0.60 (0.41-0.97)	0.59 (0.41-1.00)	0.001
Serum AST (×ULN)	1.47 (0.91-1.80)	1.8 (1.13-2.6)	1.43 (0.94-2.27)	1.32 (0.92-2.03)	1.29 (0.85-2.07)	<0.001
Serum ALT (×ULN)	0.98 (0.53-1.64)	1.95 (1.19-3.00)	1.71 (1.06-2.75)	1.46 (0.91-2.33)	1.32 (0.75-2.38)	<0.001
Serum albumin (×LLN)	1.04 (0.92-1.15)	1.11 (1.03-1.25)	1.14 (1.06-1.23)	1.14 (1.06-1.23)	1.14 (1.03-1.23)	0.038
Platelet count (×109/L)	203 (187-244)	256 (194-305)	242 (190-295)	257 (204-310)	237 (175-291)	0.001
APRI (>0.54)[d]	10 (71.4)	56 (60.2)	359 (54.9)	454 (47.8)	85 (54.1)	0.009
Biochemical disease stage[e]						<0.001
Mild	5 (27.8)	100 (59.2)	515 (72.4)	707 (70.2)	106 (69.7)	
Moderately advanced	7 (38.9)	52 (30.8)	145 (20.4)	228 (22.6)	27 (17.8)	
Advanced	6 (33.3)	17 (10.1)	51 (7.2)	72 (7.1)	19 (12.5)	
UDCA-treated[g]	0	172 (76.1)	1208 (91.8)	1447 (87.1)	195 (84.8)	<0.001
UDCA dosage (mg/kg per day)[h]	-	11.7 (3.9)	11.9 (3.5)	13.3 (3.3)	13.1 (3.1)	<0.001
Response to UDCA[i]						
Toronto	-	43/83 (51.8)	427/616 (69.3)	523/710 (73.7)	48/65 (73.8)	<0.001
Paris-I	-	61/113 (54.0)	610/837 (72.9)	725/977 (74.2)	107/149 (71.8)	<0.001
Barcelona	-	73/113 (64.6)	518/795 (65.2)	633/1058 (59.8)	100/155 (64.5)	0.12
Paris-II	-	36/116 (31.0)	432/861 (50.2)	522/1038 (50.3)	81/155 (52.3)	<0.001
Rotterdam	-	87/106 (82.1)	424/516 (82.2)	561/668 (84.0)	90/111 (81.1)	0.79
GLOBE score	-	27/40 (67.5)	162/211 (76.8)	360/435 (82.8)	67/88 (76.1)	0.047

Characteristics	1970-1979 (n=20)	1980-1989 (n=245)	1990-1999 (n=1331)	2000-2009 (n=1687)	≥2010 (n=235)	p-value
Kaplan-Meier estimates						
10-year decompensation rate (%)	-	7.9	7.1	5.6	-	0.49
10-year HCC incidence rate (%)	-	3.0	1.2	2.4	-	0.16
10-year transplant-free survival (%)	40.1	72.0	87.6	87.1	-	<0.001
10-year liver-related death (%)	53.2	14.0	4.9	6.5	-	<0.001

Data represented as mean (standard deviation), n (%), or median (interquartile range).
AMA, antimitochondrial antibody; ALP, alkaline phosphatase; ULN, upper limit of normal; AST, aspartate aminotransferase; ALT, alanine aminotransferase; LLN, lower limit of normal; APRI, AST to platelet ratio index; UDCA, ursodeoxycholic acid; HCC, hepatocellular carcinoma.
[a]Age at diagnosis not available for one patient in 2000-2009 cohort.
[b]AMA status was available for 3430 (97.5%) patients.
[c]ALP, bilirubin, AST, and ALT were log transformed prior to analyses and availability for laboratory values is as follows:
ALP: 2662 (75.7%); Bilirubin: 2586 (73.5%); AST: 2593 (73.7%); ALT: 2271 (64.6%); Albumin: 2123 (60.3%); Platelet count: 1998 (56.8%)
[d]The cut-point APRI >0.54 at baseline is predictive of liver transplantation or death (33)
[e]Biochemical disease stage classification according to Rotterdam criteria (9) was available in 2057 (58.5%) patients.
[f]UDCA therapy status was available for 3452 patients (98.1%).
[g]UDCA dosage was available for 1319 (43.6%) of UDCA-treated patients.
[h]Response was determined based on the availability of laboratory values at 1 year of UDCA therapy. Response according to Toronto criteria was calculated after 2 years of UDCA therapy.

Supplementary Table 3. Factorial ANOVA analysis of age at diagnosis over calendar time adjusting for sex

N=4804	Beta coefficient	Lower 95% CI	Upper 95% CI	p-value
Male	4.03	2.93	5.14	<0.001
Female	0.00			
Year of diagnosis 1970-1980	-10.00	-12.43	-7.57	<0.001
Year of diagnosis 1980-1990	-6.83	-8.51	-5.14	<0.001
Year of diagnosis 1990-2000	-4.17	-5.76	-2.58	<0.001
Year of diagnosis 2000-2010	-2.00	-3.58	-0.41	0.014
Year of diagnosis ≥2010	0.00			

ANOVA, analysis of variance; CI, Confidence interval.

Supplementary Table 4. Response rate in UDCA-treated patients according to various published criteria over calendar time

Response criterion[a]	1970-1979 (n=78)	1980-1989 (n=735)	1990-1999 (n=1605)	2000-2009 (n=1563)	≥2010 (n=195)	p-value
Barcelona	30/61 (49.2)	277/493 (56.2)	630/1062 (59.3)	674/1131 (59.6)	100/155 (64.9)	0.185
Paris-I	31/60 (51.7)	268/533 (50.3)	790/1121 (70.5)	773/1047 (73.8)	107/149 (71.8)	<0.001
Rotterdam	37/57 (64.9)	358/479 (74.7)	595/746 (79.8)	601/716 (83.9)	90/111 (81.1)	<0.001
Toronto	19/41 (46.3)	200/395 (50.6)	570/851 (67.0)	564/759 (74.3)	48/65 (73.8)	<0.001
Paris-II	13/64 (20.3)	151/542 (27.9)	548/1157 (47.4)	563/1121 (50.2)	81/155 (52.3)	<0.001
GLOBE score[b]	13/25 (52.0)	111/190 (58.4)	200/285 (70.2)	382/463 (82.5)	67/88 (76.1)	<0.001

Data represented as n (%).
UDCA, ursodeoxycholic acid.
[a]Response was determined based on the availability of laboratory values at 1 year of UDCA therapy. Response according to Toronto criteria was calculated after 2 years of UDCA therapy.
[b]Response according to the GLOBE score was established when the calculated value did not surpass the age-specific threshold (26).

Supplementary Table 5. Response rate over calendar time in UDCA-treated patients who did not meet criteria at baseline

Response criterion[a]	1970-1979 (n=78)	1980-1989 (n=735)	1990-1999 (n=1605)	2000-2009 (n=1563)	≥2010 (n=195)	p-value
Paris-I	12/40 (30.0)	122/344 (35.5)	202/436 (46.3)	215/410 (52.4)	28/58 (48.3)	<0.001
Rotterdam	9/29 (31.0)	121/242 (50.0)	176/327 (53.8)	264/379 (69.7)	43/64 (67.2)	<0.001
Toronto	12/34 (35.3)	115/284 (40.5)	209/395 (52.9)	209/352 (59.4)	12/24 (50.0)	<0.001
Paris-II	10/57 (17.5)	106/448 (23.7)	235/695 (33.8)	274/705 (38.9)	36/93 (38.7)	<0.001

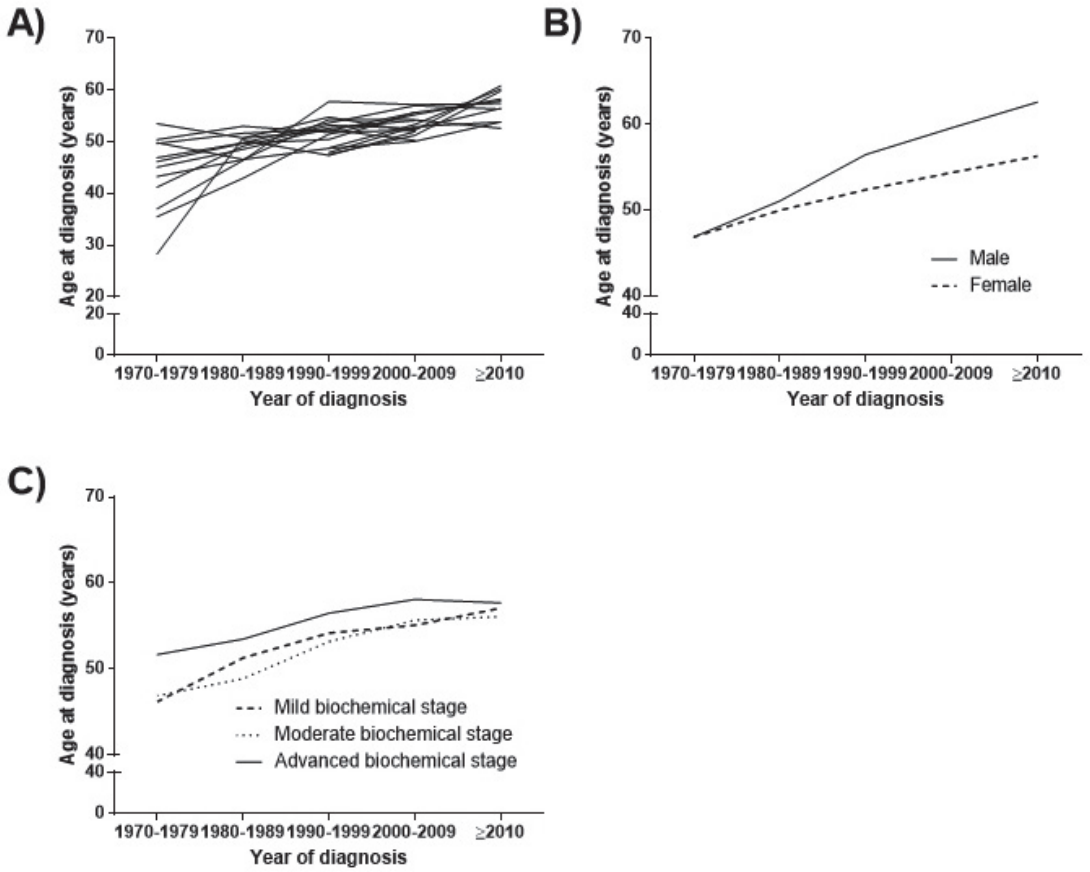
Data represented as n (%).
UDCA, ursodeoxycholic acid.
[a]Response was determined based on the availability of laboratory values at 1 year of UDCA therapy. Response according to Toronto criteria was calculated after 2 years of UDCA therapy.

Supplementary Table 6. Multivariable Cox regression analysis of 10-year transplant-free survival (n=3354)

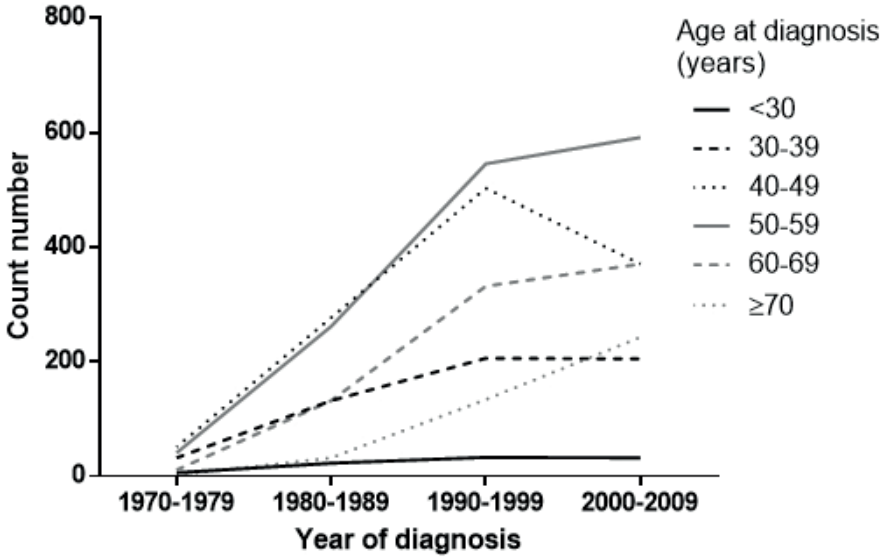
Variable	HR	95% CI	p-value
Male sex	1.11	0.89-1.40	0.350
UDCA	0.55	0.45-0.68	<0.001
Year of diagnosis			<0.001
1970-1979	1.00		
1980-1989	1.14	0.81-1.60	0.454
1990-1999	0.72	0.49-1.06	0.095
≥2010	0.60	0.40-0.89	0.011
Age at diagnosis			<0.001
<30	1.00		
30-39	1.45	0.58-3.63	0.423
40-49	2.31	0.95-5.63	0.066
50-59	2.34	0.96-5.71	0.061
60-69	4.46	1.82-10.89	0.001
>70	8.52	3.45-21.07	<0.001
Log bilirubin (×ULN)	12.8	10.6-15.4	<0.001
Difference between diagnosis and study entry (years)	1.06	1.03-1.08	<0.001

HR, hazard ratio; CI, confidence interval; UDCA, ursodeoxycholic acid; ULN, upper limit of normal.

Supplementary Figure 1. Mean age at diagnosis over calendar time stratified by A) Center (each line corresponds to an individual center); B) Sex; and C) Biochemical disease stage.



Supplementary Figure 2. Absolute number of patients according to age at diagnosis and over calendar time.



CHAPTER 3

3

FACTORS ASSOCIATED WITH PROGRESSION AND OUTCOMES OF EARLY STAGE PRIMARY BILIARY CHOLANGITIS

Goet JC*, Gatselis NK*, Zachou K, Lammers WJ, Janssen HLA, Hirschfield GM, Corpechot C, Lindor KD, Invernizzi P, Mayo MJ, Battezzati PM, Floreani A, Parés A, Lygoura V, Nevens F, Mason AL, Kowdley KV, Ponsioen CY, Bruns T, Thorburn D, Verhelst X, Harms MH, van Buuren HR, Hansen BE, Dalekos GN - *on behalf of the Global PBC Study Group*

*shared first authorship

Clinal Gastroenterology and Hepatology 2020; 18(3):684-692

ABSTRACT

Background & Aims: Patients usually receive a diagnosis of primary biliary cholangitis (PBC) at an early stage, based on biochemical analyses. We investigated the proportions of these patients who progress to moderate or advanced PBC and factors associated with progression and patient survival.

Methods: We obtained data from 1615 patients (mean age, 55.4 years) with early-stage PBC (based on their normal levels of albumin and bilirubin), collected at the time of initial evaluation or treatment, from the Global PBC Study Group database (comprising patients at 19 liver centers in North American and European countries). We collected data from healthcare evaluations on progression to moderate PBC (abnormal level of bilirubin or albumin) or advanced-stage PBC (abnormal level of both). The median follow-up time was 7.9 years. The composite endpoint was decompensation, hepatocellular carcinoma, liver transplantation or death.

Results: Of the 1615 patients identified with early-stage PBC, 904 developed moderate PBC and 201 developed advanced disease over the study period. Proportions of patients who transitioned to moderate PBC at 1, 3, and 5 years were 12.9%, 30.2%, and 45.8%. The proportions of these patients who then transitioned to advanced PBC 1, 3, and 5 years later were 3.4%, 12.5%, and 16.0%. During the follow-up period, 236 patients had a clinical event. Proportions of patients with moderate PBC and event-free survival were 97.9%, 95.1%, and 91.5% at 1, 3, and 5 years, and proportions of patients with advanced PBC and event-free survival were 90.6%, 71.2%, and 58.3% at 1, 3, and 5 years later, respectively. Variables associated with transition from early to moderate PBC included baseline levels of bilirubin, albumin, and alkaline phosphatase; aspartate to alanine aminotransferase ratio; platelet count; and treatment with ursodeoxycholic acid. Transitions from early to moderate PBC and from moderate to advanced PBC were associated with higher probabilities of a clinical event (time-dependent hazard ratios, 3.0; 95% CI, 2.0–4.5 and 4.6; 95% CI, 3.5–6.2).

Conclusions: Approximately half of patients with early-stage PBC progress to a more severe stage within 5 years. Progression is associated with increased risk of a clinical event, underlining the importance of surveillance for patients with early-stage PBC.

STUDY HIGHLIGHTS

What is known:
- Few studies have specifically explored predictive factors for progression of disease in patients with early stage PBC

What is new here:
- Pre-treatment levels of serum bilirubin, alkaline phosphatase, AST to ALT ratio and platelet count predict the risk of progression to more advanced disease stages. Treatment with UDCA decreases this risk

INTRODUCTION

Primary biliary cholangitis (PBC) is a slowly progressive cholestatic liver disease that may lead to cirrhosis and liver failure, requiring liver transplantation (LT). Currently, ursodeoxycholic acid (UDCA) is the first line therapy, but new therapies are becoming available[1]. The clinical course of PBC varies greatly and its clinical presentation has changed substantially over recent decades. In the past, most patients were diagnosed at symptomatic and advanced disease stages (extensive fibrosis or cirrhosis)[2], but the majority of PBC patients seen in recent clinical practice are diagnosed at asymptomatic and earlier stages[3-53-5]. This shift in clinical presentation may be due to higher disease awareness[1], improved diagnostic assays for antimitochondrial antibodies (AMA) detection[6], and/or more routine testing of liver function tests. Regardless, given the shift, relying solely on hard clinical endpoints, such as death or LT, may not be feasible in clinical studies of patients with PBC[7].

The long-term survival of PBC patients with an early biochemical stage (defined as normal albumin and normal bilirubin based on the Rotterdam criteria) is generally comparable to the survival of general population[8]. These patients are therefore perceived as having a low-risk of developing progressive disease. However, patients identified as low-risk at the beginning of the disease may still progress to a moderate (abnormal albumin or bilirubin) or advanced stage (both abnormal albumin and bilirubin) during follow-up, which may be associated with worse long-term survival. Knowledge regarding predictive factors for biochemical transitions, as well as their impact on prognosis, may aid in identifying patients that are likely to progress over time. Therefore, in a cohort of PBC patients with early biochemical stage, we examined the proportions who progressed to moderate or advanced PBC and factors associated with progression and patient survival.

METHODS

Patients

This study was a subgroup analysis of patients included in the Global PBC Study Group (GPBCSG) database; a multicenter collaboration between 19 liver centers from 12 North American and European countries. All patients had an established diagnosis of PBC[1,9,10]. Follow up data were prospectively collected.

For the current study, only those patients with biochemically early disease at baseline according to the Rotterdam criteria[11] were included and evaluated for transition to biochemically moderate and/or advanced disease during follow-up. Both UDCA-treated and untreated patients were included. Early stage was defined by normal bilirubin and albumin levels, moderately advanced disease was defined by an abnormal bilirubin or albumin level, and advanced disease was defined by abnormal bilirubin and albumin levels. Exclusion criteria were missing laboratory data at baseline, unknown start date of treatment with UDCA and/or last follow-up date, short follow-up (<6 months or only baseline visit available), or a concomitant liver disease including autoimmune hepatitis/PBC variant and alcoholic liver disease. Collected clinical and laboratory data included sex, age, diagnosis of PBC, liver histology, treatment (type of medication, dosage and

duration), duration and last date of follow-up, baseline AMA status, laboratory values (serum ALP, total bilirubin, albumin, AST, ALT, and platelets) and outcomes (death and cause of death, liver transplantation, hepatocellular carcinoma (HCC), ascites, variceal bleeding and hepatic encephalopathy).

Liver histology performed within 1 year of study entry or documented cirrhosis before study entry was classified as a baseline biopsy. Histological data was assessed for severity according to Ludwig[12] and Scheuer's classification[13].

This study was conducted in accordance with the protocol and principles of the 1975 Declaration of Helsinki. The protocol was approved by the Institutional Research Board of the corresponding center, and at each participating center, in accordance with local regulations.

Endpoints

The endpoints included progression to moderate PBC (abnormal level of bilirubin or albumin) or advanced-stage PBC (abnormal level of both) and progression to composite clinical endpoint, defined by either LT, ascites, variceal bleeding, hepatic encephalopathy, development of hepatocellular carcinoma (HCC) or death, whichever occurred first.

Statistical analysis

Baseline visit was considered as the start date of UDCA or the date of initial evaluation in case of untreated patients. Patients that remained stable at biochemical early disease during follow-up or patients who did not reach any of the components of the composite clinical event were censored at their last follow-up visit. To determine the overall impact of transitions between biochemical stages on event-free survival, the time until patients transitioned from early- to moderate, and early- to advanced biochemical stage, was modeled as time-dependent covariate.

Biochemical data were not available for all patients. Missing data varied between 8% for ALP and 26.6% for AST values; it is probable that stable patients had less laboratory measurements, contrary to the closer follow-up of patients with an accelerated progression, thereby creating a bias in the availability of data. Due to the extensive data collection effort for the GPBCSG, laboratory data was available in total 75,000 visits between UDCA initiation/date of initial evaluation until the end of follow-up or the occurrence of a clinical event. As such, the trajectory of the lab values over time was used as prior information for the imputation of missing values as well as the strong correlation between multiple different lab values and the endpoint. Therefore, we used an imputed dataset for our primary analysis to ensure full detection of biochemical transitions. In detail, SAS (SAS Proc. MI, MCMC method; SAS software, version 9.3, SAS Institute, Cary, NC) was used to generate 10 imputed datasets of laboratory results at yearly time points between start of UDCA treatment or initial evaluation up to 15 years of follow-up. Missing data were considered as missed at random. Rubin's rules were used for estimation of the parameters and the standard error[14,1510,11]. The imputation model included baseline variables that were potentially predictive for outcomes in PBC (e.g. year of diagnosis, age) as well as the outcomes themselves. Only continues biochemical variables were imputed.

Univariate and multivariate Cox regression analyses was performed to assess the impact of various factors on the rate of biochemical transition from biochemically early to moderate, and from moderate to advanced disease. In the analyses of factors associated with transition from moderate to advanced disease, laboratory parameters correspond to the time-point of moderate disease development. In multivariable analyses, the model with the lowest Akaike Information Criteria was chosen. The effect of albumin and bilirubin within the normal range are presented on a continuous scale and by a binary split at the median for reason of clinical interpretation and the aim of comparing groups of equal sample size. In addition, transition rates were assessed and compared between patients with baseline GLOBE scores above the age-specific GLOBE score threshold and those remaining below this threshold using Kaplan-Meier estimates. The calculation and use of GLOBE score thresholds have been described previously with patients above the threshold presenting a significantly worse survival compared to matched individuals from the general population[16].

Normally distributed data are presented as mean ± standard deviation (SD) and skewed distributed data as median and interquartile range (IQR). Where indicated, continuous variables underwent natural logarithmic transformation to correct for nonlinearity. All analyses were 2-sided. $P<0.05$ was considered statistically significant. Statistics analyses were performed with IBM SPSS Statistics 22.0 (SPSS Inc, Chicago, IL).

RESULTS

Study populations characteristics

The patient population consisted of 1615 patients with biochemically early PBC (**Supplementary Figure 1**). Mean (SD) age at study entry was 55.4 (11.9) years. The median (IQR, range) total follow-up period was 7.9 (4.3-12.5, 0.5 - 25) years. Baseline patient characteristics are shown in **Table 1**. The majority of patients (1415/1615; 87.6%) were treated with UDCA. Histological disease stage was available for 798 (49%) patients of whom most had early disease (stage I or II, 623/798; 78%).

Of the 1415 UDCA-treated patients, a total of 1383 had a follow-up longer than one year, with biochemical non-response rates, according to various criteria, ranging from 14.0% with Paris I criteria to 59.9% with Barcelona criteria. According to GLOBE score, only 11.0% of patients presenting with early-stage PBC had an estimated survival that was worse than that of general population at one year of UDCA therapy.

Table 1. Baseline study population characteristics

	Total cohort N=1615
Age at diagnosis (years)	53.4 (12.0)
Age at study entry (years)	55.4 (11.9)
Year of diagnosis (median, IQR)	1997 (1990-2004)
Year of diagnosis, range	1961-2014
Female, n (%)	1480 (91.6)
AMA+, n (%)	1459 (90.3)
UDCA treated, n (%)	1415 (87.6)
Biopsy stage[a], n (%)	
I	381 (47.7)
II	242 (30.3)
III	119 (14.9)
IV	56 (7.0)
Serum total bilirubin ×ULN	0.52 (0.40-0.70)
Serum albumin ×LLN	1.19 (1.11-1.26)
Serum ALP ×ULN	1.93 (1.25-3.11)
Serum AST ×ULN	1.28 (0.90-1.87)
Serum ALT ×ULN	1.50 (0.95-2.33)
Serum platelets $\times 10^3/mm^3$	255 (203-366)

AMA, anti-mitochondrial antibodies; ULN, upper limit of normal; ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; LLN, lower limit of normal.
Data represented as mean (standard deviation) and median (interquartile range).
[a]Baseline biopsies (obtained within 1 year of start of UDCA) were available in 798 of 1615 (49%) patients. Histological stage according to Ludwig[13] and Scheuer[14] classification.

Transitions in biochemical stage

During the median 7.9 (IQR 4.3-12.5) years of follow-up, 904 patients transitioned from biochemically early to moderately advanced stage. Sequentially, 201 out of 904 patients transitioned from moderate to advanced disease (**Figure 1**). Thirty-four patients transitioned directly from early- to advanced stage. For these patients, we assumed that the transition took place via moderate stage. Overall, the transition rates from early- to moderate stage were 12.9% at 1 year, 30.2% at 3 years, and 45.8% at 5 years of follow-up (**Figure 2A**). The cumulative transition rates from moderate to advanced (201/904) stage were 3.4% at 1 year, 12.5% at 3 years, and 16.0% at 5 years of follow-up (**Figure 2B**). The median time to transition to moderate stage was 2.5 (IQR 1-5, range 0.5-15) years. Median time from moderate to advanced stage was 1.5 (IQR 0.5-3.5, range 0.5-11.5) years.

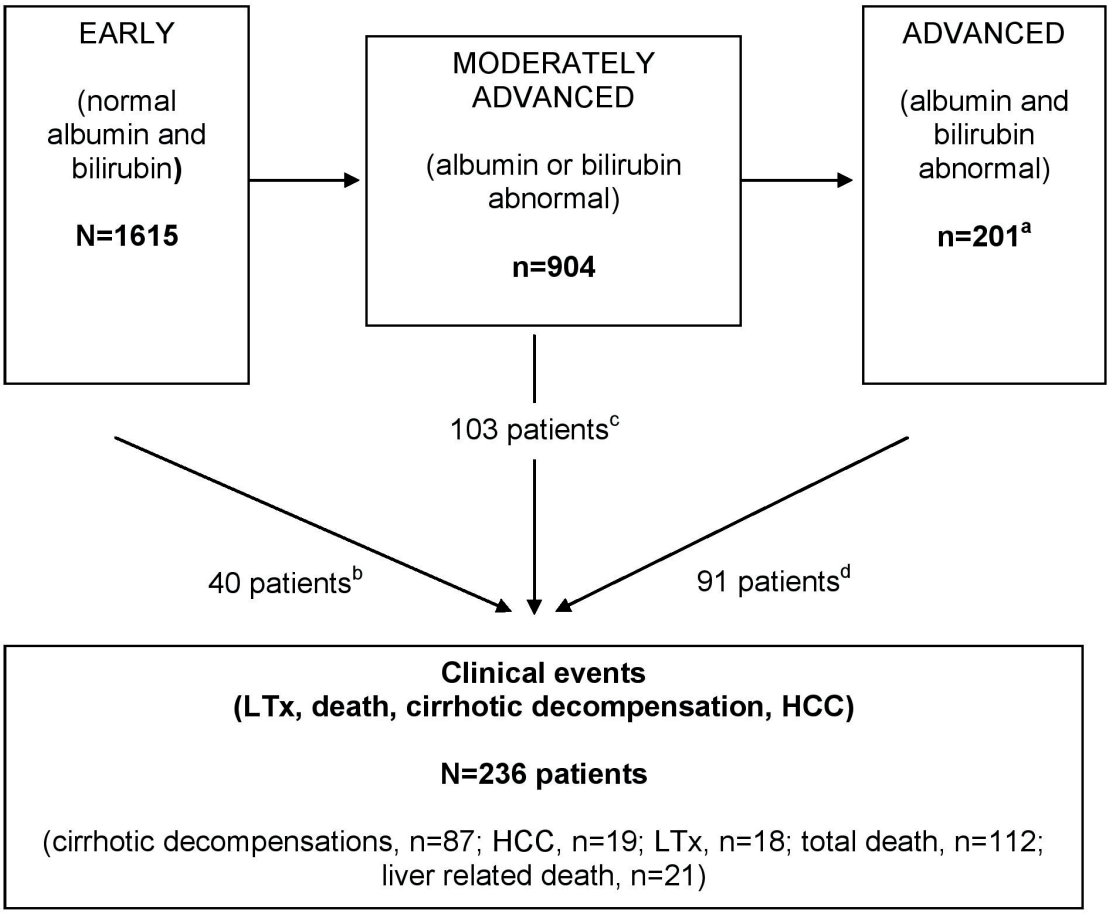


Figure 1. Transitions of patients according to their biochemically stage during follow-up and clinical events. HCC; hepatocellular carcinoma, LTx; liver transplantation. [a]34 patients transited directly to advanced stage. For these patients, we assumed they transited gradually through moderate- to advanced stage.
[b]8 cirrhotic decompensations; 0 HCC; 0 LTx; 31 Death (5 liver-related). [c]40 cirrhotic decompensations; 11 HCC; 7 LTx; 47 Death (7 liver-related). [d]39 decompensations; 8 HCC; 11 LTx; 34 Death (9 liver-related).

Factors associated with transition from biochemical early to moderately advanced disease stage

In univariate analyses, baseline factors associated with progression from biochemically early to moderate disease were male sex, higher age at study entry, bilirubin, albumin, alkaline phosphatase (ALP), and aspartate/alanine transaminase (AST/ALT) ratio. Patients with a more recent diagnosis and higher platelets at baseline were less likely to transition to moderate stage (**Table 2**). Biopsy (available for 798 patients) stages III and IV were associated with a higher probability of disease progression compared to stage I (HR 1.7 [95% CI: 1.3-2.2; *P*<.001] and HR 1.8 [95% CI: 1.3-2.5; *P*<.001], respectively), while stage II was not associated with a higher probability of disease progression (*P* = .14). For patients with histological stage I or II, the transition rates from early- to moderate stage at 1, 3, and 5 year(s) were 10.9%, 31.5% and 43.6%, respectively. The transition rates were higher for those with stage III or IV: 18.4%, 41.5% and 60.2% at 1, 3, and 5 year(s), respectively.

In multivariable analyses, all variables except age at study entry and year of diagnosis remained significantly associated with transition to moderate stage (**Table 2**). Bilirubin and albumin at a binary split of >0.5xULN and ≤1.2xULN, respectively, were significantly associated with biochemical transition in multivariable analyses (HR 1.5 [95% CI: 1.3-1.8; *P*<.001] and HR 1.6 [95% CI: 1.4-1.9; *P* < .001], respectively). Baseline AST/ALT ratio (HR 1.3; 95% CI: 1.1-1.5) and ALP levels (HR 1.3; 95% CI: 1.2-1.5) were positive predictors of progression, while patients treated with UDCA had a lower transition rate (HR 0.70; 95% CI: 0.57-0.86). In addition, ALP >1.67xULN was associated with transition to moderate stage (HR 1.4; 95% CI: 1.2-1.6) (**Supplementary Figure 2**). Bilirubin, albumin, ALP, AST/ALT ratio and platelets retained their prognostic value for biochemical transition in the subgroup of UDCA-treated patients (**Supplementary Table 1**).

A total of 242 out of 1615 with normal albumin and bilirubin at baseline had GLOBE score values beyond the age-specific GLOBE score threshold. These patients were more likely to progress to moderate stage (HR 1.8; 95% CI 1.6-2.1, **Supplementary Figure 3**).

In subgroup-analyses including only histologically proven early stage (I-II) patients, all factors except platelets and AST/ALT ratio remained as independent predictors of transition from early to moderate biochemical stage (**Supplementary Table 2**).

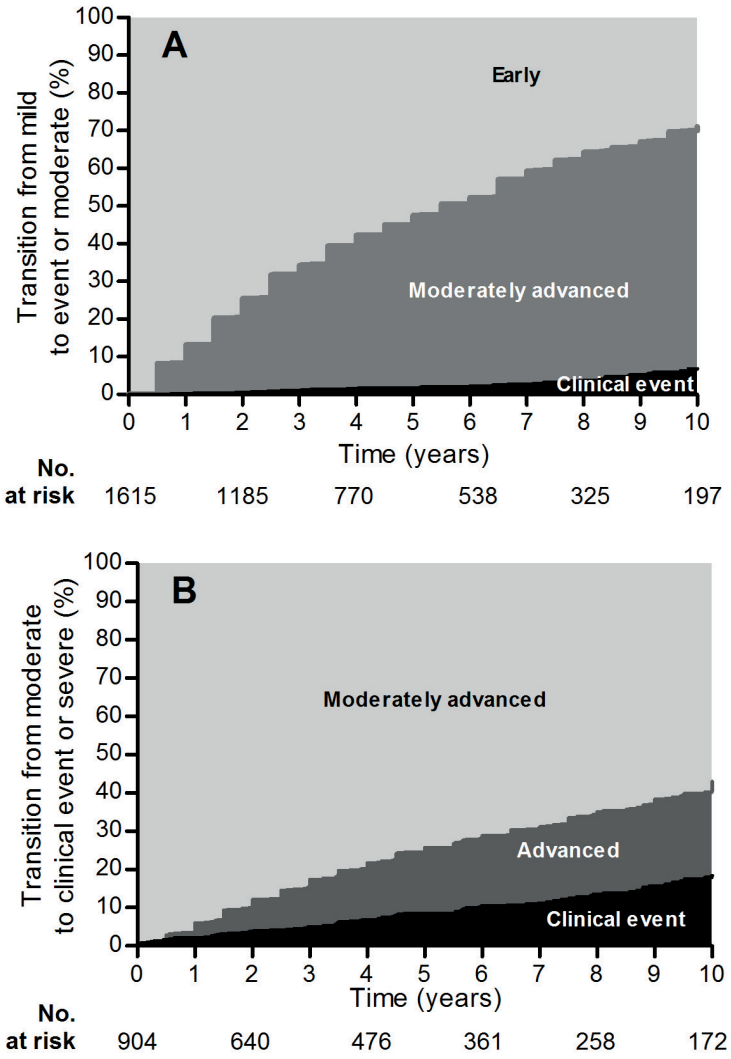


Figure 2. Cumulative incidence of biochemical transition and events. Kaplan- Meier estimates of transition from mild to a moderate stage or event (A), and transition from moderate to advanced (severe) disease stage or an event (B).

Factors associated with transition from biochemical moderately advanced to advanced disease stage

In 904 patients that transited to moderate stage, univariate analyses revealed older age, total bilirubin, albumin, ALP, and transaminases at the time point of transition to moderate stage as predictive factors for subsequent transition to advanced disease (**Table 3**). In contrast, UDCA-treated patients and those with higher platelets were less likely to transition to advanced disease. While multivariate analyses rendered age and platelets non-significant, all other variables remained associated with biochemical transition to advanced stage, with UDCA being associated with a lower probability of transition (HR 0.57; 95% CI: 0.40-0.82). ALP levels >1.67xULN were significantly associated with transition to advanced stage (HR 2.5; 95% CI: 1.9-3.3) (**Supplementary Figure 2**). Albumin levels (binary split of <1xLLN) were not associated with biochemical transition whereas bilirubin levels (at a binary split of >1xULN, HR 2.0; 95% CI: 1.4-1.7) were associated with transition to advanced stage. Patients with GLOBE score values above the threshold (568/904) had higher rates of progression to advanced disease than those below the GLOBE score threshold (HR 3.0; 95% CI 2.0-4.3, **Supplementary Figure 3**). Bilirubin, albumin, ALP, and AST/ALT ratio were also associated with biochemical transition to advanced stage in the subgroup of UDCA-treated patients (**Supplementary Table 3**).

Table 2. Baseline factors associated with the transition from biochemically early to moderately advanced disease

	Univariate analyses				Multivariate analysesa			
	HR	**95% CI**		***P***	**HR**	**95% CI**		***P***
Male sex	1.27	1.01	1.60	.043	1.27	1.01	1.60	.045
Age at entry, *per 10 years*	1.06	1.00	1.12	.044	-	-	-	-
Year of diagnosis, *per decade*	0.86	0.78	0.94	<.001	-	-	-	-
UDCA usage	0.90	0.73	1.09	.279	0.70	0.57	0.86	.001
Bilirubin x ULN[a]	2.14	1.78	2.58	<.001	1.86	1.53	2.26	<.001
Albumin x LLN	0.07	0.04	0.12	<.001	0.08	0.04	0.16	<.001
ALP x ULN[a]	1.40	1.28	1.55	<.001	1.33	1.20	1.48	<.001
AST/ALT ratio[a]	1.33	1.11	1.58	.002	1.27	1.06	1.53	<.001
Platelets, *per 10 units (x10^3/mm3) increase*	0.98	0.97	0.99	<.001	0.99	0.98	0.99	.007

ALP, alkaline phosphatase, AST, aspartate aminotransferase; ALT, alanine aminotransferase; ULN, upper limit of normal; LLN, lower limit of normal.
[a]These biochemical variables were transformed with natural logarithm.

Table 3. Factors associated with the transition from biochemically moderately advanced to advanced disease (n=904)

	Univariate analyses				Multivariable analyses			
	HR	**95% CI**		***P***	**HR**	**95% CI**		***P***
Male sex	1.22	0.78	1.92	.390	-	-	-	-
Age, *per 10 years*[a]	1.23	1.11	1.35	.001	-	-	-	-
Year of diagnosis, *per decade*[b]	0.86	0.69	1.05	.119	-	-	-	-
UDCA usage	0.45	0.31	0.63	<.001	0.57	0.40	0.82	.002
Bilirubin x ULN[b,c]	2.69	2.09	3.47	<.001	3.87	3.04	4.94	<.001
Albumin x LLN	0.36	0.15	0.94	.036	0.02	0.01	0.05	< .001
ALP x ULN[b,c]	2.21	1.80	2.71	<.001	2.05	1.65	2.54	< .001
AST/ALT ratio[b,c]	1.51	1.08	2.12	.016	1.66	1.17	2.35	.004
Platelets, *per 10 units ($x10^3$/mm3) increase*	0.98	0.96	0.99	<.001	-	-	-	-

UDCA, ursodeoxycholic acid; ALP, alkaline phosphatase; AST, aspartate aminotransferase; ALT, alanine aminotransferase

[a]Age at which biochemical moderate stage was reached.

[b]These biochemical variables were transformed with natural logarithm.

[c]Laboratory parameters correspond to the time-point of moderate disease development.

The impact of biochemical transition on events

During follow-up, 236 patients developed at least one clinical event: 87decompensations, 19 HCC, 18 LT and 112 deaths (21 liver-related deaths) (**Figure 1**). Overall, the event-free survival for patients remaining in early biochemical disease stage at 1, 3, 5 and 10 years was 99.8%, 98.9%, 98.3% and 93.2%, respectively (**Figure 3**). Once patients reached a moderate stage (n=904), the event-free survival rates at 1, 3 and 5 year(s) of follow-up were 97.9%, 95.1% and 91.5%, respectively. For patients that consequently transited to advanced biochemical stage (n=201), these rates were 90.6%, 71.2% and 58.3%, respectively (**Figure 3**). In univariate and multivariate cox-regression, in which biochemical transition was modeled as time-dependent covariate, a higher probability of a clinical event during follow-up was found for patients that transited from biochemical early to moderate stage (HR 3.0; 95% CI: 2.0-4.5), from biochemical moderate to advanced stage (HR 4.6; 95% CI: 3.5-6.2), and overall from early to advanced disease stage (HR 14.1; 95% CI: 9.3-21.4, *P*<.001) (**Supplementary Table 4**).

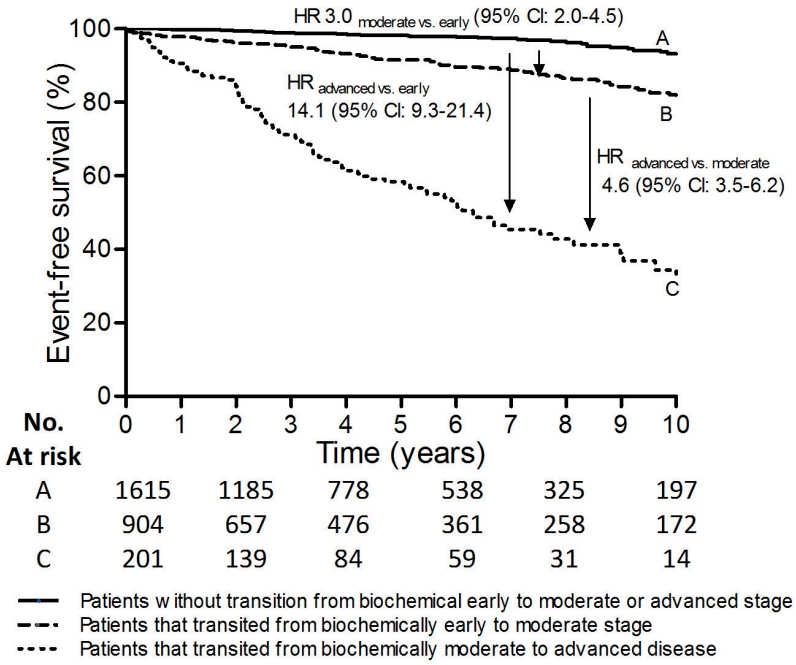


Figure 3. Event-free survival in a clock-reset approach in patients that transit from early - to moderate - to advanced biochemical disease stage. Patients that did not transit in biochemical stage remain in line A. Patients who transited from early to moderate stage were switched to a new survival curve (B), which was then reset as time 0 for their further follow-up. Patients that then transited from moderate to advanced stage were switched to survival curve (C). All hazard ratios were obtained by considering biochemical transition as a time-dependent covariate in cox-regression analyses.

DISCUSSION

In this study of a large cohort of more than 1500 PBC patients from European and North American centres, we found that approximately one out of two patients presenting with an early biochemical disease stage transitioned towards a moderate biochemical stage within 5 years of follow-up. Almost one in six of these patients that reach a moderate stage eventually transitioned to an advanced stage within the next 5 years. These transitions were associated with an increased risk of clinical events, underlining the importance of clinical surveillance even in early stage PBC patients.

We were also able to identify the patients who are more likely to transition into moderately advanced and advanced stage. UDCA treatment was associated with lower rates of biochemical transition (as well as the composite endpoint) during follow-up. Extended data from several studies indicate that UDCA improves the natural history of PBC, even when administered in the early stages[3,4,8]. The GPBCSG recently showed that UDCA treated patients have a lower risk of LT or death than untreated patients, and more importantly, the benefit was observed in both responders and non-responders[17]. Our data point in the same direction, by showing UDCA treatment in patients with early biochemical stage may reduce progression to more advanced stages. Taking into account that UDCA is recommended for all PBC patients including those at early stages[1], we conducted a sub-

analysis including only UDCA-treated patients with similar results in regards to key risk factors (**Supplementary Tables 1 and 3**).

Similar to findings in previous studies, ALP levels were indicative of disease progression. In PBC, ALP is considered as one of the most robustly validated markers of disease activity. In meta-analyses of 4845 patients, we previously showed that ALP (>2.0xULN) was strongly associated with LT and death across various subgroups[18]. Moreover, ALP is an important component of most biochemical criteria that assess treatment response to UDCA after one year of therapy. Our results indicate that higher ALP levels are not only associated with hard clinical endpoints, but also with biochemical transition.

In accordance with previous studies, we found platelet count and AST/ALT ratio to be associated with biochemical disease progression, particularly in patients with advanced histological stage (III-IV)[19-25]. Platelet count is generally considered a marker of portal hypertension and is of particular importance in discriminating non-cirrhotic from cirrhotic patients with normal bilirubin and albumin levels. Our results emphasize the importance of platelet count and AST/ALT ratio in identifying PBC patients that are likely to progress.

In contrast to previous studies, male sex was not consistently associated with disease progression[26-28]. One possible explanation for this is that only patients with mild PBC were included in our study, while in a previous study, the negative impact of male gender appeared to be limited only to patients with advanced disease[26]. Alternatively, the use of AST/ALT ratio in our multivariable analyses could have rendered male sex a non-significant factor. Prior studies have documented relatively higher rates of alcohol consumption in men, which is characterized by an increased AST/ALT ratio[4,28]. This finding coincides with a recent Greek clinical study, where, after adjusting for other confounding factors including alcohol consumption, male sex did not independently pose a greater risk for disease progression during follow-up[29]. Of note, in accordance with our findings, a recent study from GPBCSG did not associate sex with response and transplant-free survival[30].

Strengths of our study are the inclusion of a large cohort of PBC patients from different geographical areas as well as long-term follow-up with many clinical events. This increases the reliability and generalizability of our results. The GPBCSG database captured all patient-visits and represents the current clinical practices in the participating centers and the results are therefore highly relevant to clinicians working with PBC patients. Although few laboratory data were missing, the results presented are based on the imputed databases. To support our findings, a sensitivity analysis that excluded patients with missing data yielded similar results (**Supplementary Table 5 and 6**). No data on co-factors (e.g. alcohol consumption or obesity) were available in the GPBCSG database[8,29]. However, patients with alcoholic liver disease were excluded. Nevertheless, future studies taking into account the afore-mentioned co-factors are warranted to identify patient groups with high risk of disease progression.

A question that remains is how often we should perform laboratory exams in early PBC patients. Personalized risk stratification, using biochemical response markers or prognostic models following one year of UDCA therapy can identify patients at risk of progressive

disease[1,16,31]. EASL clinical practice guidelines for PBC recommend that all patients should have life-long follow-up, recognizing that patients have different disease courses and may require varied levels of attention[1]. Patients with higher baseline GLOBE score had higher biochemical transition rates. Baseline GLOBE score calculation can therefore aid in the identification of patients that need closer surveillance, even those with early biochemical stage. Our results also indicate that transitions in biochemical stage are important markers of disease progression that could be incorporated in disease staging. Transitions to moderately advanced and advanced disease stage may justify a change in follow-up regimen with a closer surveillance.

In conclusion, our internationally representative study provides a comprehensive overview of the natural history of PBC patients with early disease stage, showing that almost one out of two patients with early biochemical disease will transit to moderately advanced disease and, approximately one sixth of them can progress to advanced stage. These transitions are associated with an increased probability of clinical events. The findings underline the importance of clinical surveillance in PBC patients with early biochemical disease stage.

REFERENCES

1. EASL Clinical Practice Guidelines: The diagnosis and management of patients with primary biliary cholangitis. *J Hepatol.* 2017;67(1):145-172.
2. Prince M, Chetwynd A, Newman W, Metcalf JV, James OF. Survival and symptom progression in a geographically based cohort of patients with primary biliary cirrhosis: follow-up for up to 28 years. *Gastroenterology.* 2002;123(4):1044-1051.
3. Floreani A, Caroli D, Variola A, et al. A 35-year follow-up of a large cohort of patients with primary biliary cirrhosis seen at a single centre. *Liver Int.* 2011;31(3):361-368.
4. Gatselis NK, Zachou K, Lygoura V, et al. Geoepidemiology, clinical manifestations and outcome of primary biliary cholangitis in Greece. *Eur J Intern Med.* 2017;42:81-88.
5. Murillo Perez F, Goet JC, Lammers WJ, et al. Milder disease stage in patients with primary biliary cholangitis over a 44-year period: A changing natural history. *Hepatology.* 2017; doi: 10.1002/hep.29717. [Epub ahead of print].
6. Gatselis NK, Dalekos GN. Molecular diagnostic testing for primary biliary cholangitis. *Expert Rev Mol Diagn.* 2016;16(9):1001-1010.
7. Prince MI, James OF. The epidemiology of primary biliary cirrhosis. *Clin Liver Dis.* 2003;7(4):795-819.
8. ter Borg PC, Schalm SW, Hansen BE, van Buuren HR, Dutch PBCSG. Prognosis of ursodeoxycholic Acid-treated patients with primary biliary cirrhosis. Results of a 10-yr cohort study involving 297 patients. *Am J Gastroenterol.* 2006;101(9):2044-2050.
9. European Association for the Study of the L. EASL Clinical Practice Guidelines: management of cholestatic liver diseases. *J Hepatol.* 2009;51(2):237-267.
10. Lindor KD, Gershwin ME, Poupon R, et al. Primary biliary cirrhosis. *Hepatology.* 2009;50(1):291-308.
11. Kuiper EM, Hansen BE, de Vries RA, et al. Improved prognosis of patients with primary biliary cirrhosis that have a biochemical response to ursodeoxycholic acid. *Gastroenterology.* 2009;136(4):1281-1287.
12. Ludwig J, Dickson ER, McDonald GS. Staging of chronic nonsuppurative destructive cholangitis (syndrome of primary biliary cirrhosis). *Virchows Arch A Pathol Anat Histol.* 1978;379(2):103-112.
13. Scheuer P. Primary biliary cirrhosis. *Proc R Soc Med.* 1967;60(12):1257-1260.
14. Little RJA, Rubin DB. *Statistical analysis with missing data.* New York: John Wiley & Sons; 1987.
15. Rubin D. Multiple imputation aft er 18+ years. *J Am Stat Assoc.* 1996;91: 473 – 489.
16. Lammers WJ, Hirschfield GM, Corpechot C, et al. Development and Validation of a Scoring System to Predict Outcomes of Patients With Primary Biliary Cirrhosis Receiving Ursodeoxycholic Acid Therapy. *Gastroenterology.* 2015;149(7):1804-1812 e1804.
17. van der Meer AJ, Harms MH, Corpechot C, et al. Ursodeoxycholic Acid is Associated with a Prolonged Transplant-free Survival in All Patients with Primary Biliary Cholangitis - there is no such thing as non-response. *Hepatology.* 2017;66(1):suppl 155A.
18. Lammers WJ, van Buuren HR, Hirschfield GM, 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.* 2014;147(6):1338-1349.
19. Trivedi PJ, Bruns T, Cheung A, et al. Optimising risk stratification in primary biliary cirrhosis: AST/platelet ratio index predicts outcome independent of ursodeoxycholic acid response. *Journal of Hepatology.* 2014;60(6):1249-1258.

20. Joshita S, Umemura T, Ota M, Tanaka E. AST/platelet ratio index associates with progression to hepatic failure and correlates with histological fibrosis stage in Japanese patients with primary biliary cirrhosis. *J Hepatol.* 2014;61(6):1443-1445.
21. Patanwala I, McMeekin P, Walters R, et al. A validated clinical tool for the prediction of varices in PBC: the Newcastle Varices in PBC Score. *J Hepatol.* 2013;59(2):327-335.
22. Trivedi PJ, Lammers WJ, van Buuren HR, et al. Stratification of hepatocellular carcinoma risk in primary biliary cirrhosis: a multicentre international study. *Gut.* 2015.
23. Nyblom H, Bjornsson E, Simren M, Aldenborg F, Almer S, Olsson R. The AST/ALT ratio as an indicator of cirrhosis in patients with PBC. *Liver Int.* 2006;26(7):840-845.
24. Reig A, Pares A. Low platelet count influences the performance of the different biochemical criteria of rsodeoxycholic acid therapy response in primary biliary cholangitis. *J Hepatol* 2017;66:S355.
25. Trivedi PJ, Corpechot C, Pares A, Hirschfield GM. Risk stratification in autoimmune cholestatic liver diseases: Opportunities for clinicians and trialists. *Hepatology.* 2016;63(2):644-659.
26. Carbone M, Mells GF, Pells G, et al. Sex and age are determinants of the clinical phenotype of primary biliary cirrhosis and response to ursodeoxycholic acid. *Gastroenterology.* 2013;144(3):560-569 e567; quiz e513-564.
27. Cheung A LW, Hirschfield GM, Invernizzi P, Mason A, Ponsioen CY, Floreanim, A, Corpechot C, Mayo MJ, Pares A, Battezzati PM, Nevens F, Thorburn D, Kowdley KV, Trivedi PJ, Kumagi T, Lleo A, LaRusso NF, Boonstra K, Cazzagon N, Franceschet I, Poupon R, Caballeria J, Lindor KD, Hansen BE, Janssen HLA, Van Buuren HR. Age, bilirubin and albumin, regardless of sex, are the strongest independent predictors of biochemical response and transplantation-free survival in patients with primary biliary cirrhosis. *J Hepatol.* 2015;62:S798.
28. Harada K, Hirohara J, Ueno Y, et al. Incidence of and risk factors for hepatocellular carcinoma in primary biliary cirrhosis: national data from Japan. *Hepatology.* 2013;57(5):1942-1949.
29. Azariadis K, Lygoura V, Zachou K, et al. Characteristics of Greek male patients with primary biliary cholangitis. *J Hepatol* 2017;66:S552.
30. Cheung AC, Lammers WJ, Murillo Perez CF, et al. Effects of Age and Sex of Response to Ursodeoxycholic Acid and Transplant-free Survival in Patients With Primary Biliary Cholangitis. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association.* 2019;17(10):2076-2084 e2072.
31. Carbone M, Sharp SJ, Flack S, et al. The UK-PBC risk scores: Derivation and validation of a scoring system for long-term prediction of end-stage liver disease in primary biliary cirrhosis. *Hepatology.* 2015:doi: 10.1002/hep.28017.

SUPPLEMENTARY DATA

Supplementary Table 1. Baseline factors associated with the transition from biochemically early to moderately advanced disease in UDCA-treated patients only (n=1415)

	Univariate analyses				Multivariable analyses[a]			
	HR	**95% CI**		***P***	**HR**	**95% CI**		***P***
Male sex	1.33	1.05	1.70	0.02	1.28	0.99	1.64	0.057
Age at entry, *per 10 years*	1.04	0.98	1.11	0.176	-	-	-	-
Year of diagnosis, *per decade*	0.88	0.80	0.97	<0.01	-	-	-	-
Bilirubin x ULN[a]	2.05	1.69	2.50	<0.001	1.82	1.48	2.25	<0.001
Albumin x LLN	0.07	0.04	0.14	<0.001	0.10	0.05	0.20	<0.001
ALP x ULN[a]	1.33	1.20	1.48	<0.001	1.30	1.16	1.46	<0.001
AST/ALT ratio[a]	1.33	1.11	1.61	<0.01	1.31	1.08	1.60	0.006
Platelets, *per 10 units (x10^3/mm3) increase*	0.99	0.99	0.99	<0.01	0.99	0.98	0.99	0.031

ALP, alkaline phosphatase; AST, aspartate aminotransferase; ALT, alanine aminotransferase; ULN, upper normal limit of normal; LLN, lower limit of normal.
[a]These biochemical variables were transformed with natural logarithm.

Supplementary Table 2. Baseline factors associated with the transition from biochemically early to moderately advanced disease in patients with biopsy stage I or II (n=623).

	Univariate analyses				Multivariable analyses[a]			
	HR	**95% CI**		***P***	**HR**	**95% CI**		***P***
Male sex	1.30	0.91	1.85	0.153	-	-	-	-
Age at entry, *per 10 years*	1.06	0.96	1.16	0.232	-	-	-	-
Year of diagnosis, *per decade*	0.81	0.68	0.94	0.002	0.89	0.75	1.03	0.098
UDCA usage[a]	0.93	0.59	1.48	0.772	-	-	-	-
Bilirubin x ULN[b]	1.94	1.43	2.62	<0.001	1.75	1.28	1.38	<0.001
Albumin x LLN	0.12	0.04	0.38	<0.001	0.14	0.05	0.41	<0.001
ALP x ULN[b]	1.35	1.15	1.58	<0.001	1.23	1.05	1.45	0.012
AST/ALT ratio[b]	1.21	0.90	1.61	0.208	-	-	-	-
Platelets, *per 10 units (x10^3/mm3) increase*	0.99	0.98	1.01	0.366	-	-	-	-

UDCA, ursodeoxycholic acid; ALP, alkaline phosphatase; AST, aspartate aminotransferase; ALT, alanine aminotransferase; ULN, upper normal limit of normal; LLN, lower limit of normal.
[a]UDCA was not entered into multivariable analysis, because only 34 patients were not treated with UDCA.
[b]These biochemical variables were transformed with natural logarithm.

Supplementary Table 3. Factors associated with the transition from biochemically moderately advanced to advanced disease in UDCA-treated patients only (n=794)

	Univariate analyses				Multivariable analysesa			
	HR	**95% CI**		***P***	**HR**	**95% CI**		***P***
Male sex	1.38	0.85	2.22	0.192	-	-	-	-
Age at entry, *per 10 years*[a]	1.19	1.06	1.33	0.011	-	-	-	-
Year of diagnosis, *per decade*	0.89	0.68	1.10	0.276	-	-	-	-
Bilirubin x ULN[b,c]	2.57	1.95	3.40	<0.001	3.74	2.88	4.85	<0.001
Albumin x LLN[c]	0.39	0.14	1.10	0.076	0.02	0.01	0.06	<0.001
ALP x ULN[b,c]	1.94	1.54	2.44	<0.001	1.98	1.54	2.53	<0.001
AST/ALT ratio[b,c]	1.43	0.98	2.09	0.067	1.72	1.14	2.59	0.004
Platelets, *per 10 units (x10³/mm3) increase*	0.99	0.99	1.00	0.239	-	-	-	-

ALP, alkaline phosphatase; AST, aspartate aminotransferase; ALT, alanine aminotransferase
[a]Age at which biochemical moderate stage was reached.
[b]These biochemical variables were transformed with natural logarithm.
[c]Laboratory parameters correspond to the time-point of moderated disease development.

Supplementary Table 4. The association of time-dependent transition with clinical events

	Univariate analyses				Multivariable analyses			
	HR	**95% CI**		***P***	**HR**	**95% CI**		***P***
Moderate vs. early biochemical stage[a]	3.37	2.28	4.97	< .001	3.05	2.06	4.51	< .001
Advanced vs. moderate biochemical stage[a]	5.38	4.05	7.13	< .001	4.63	3.47	6.17	< .001
Advanced vs. early biochemical stage[a]	18.12	12.10	27.12	< .001	14.11	9.32	21.36	< .001
Male sex	1.30	0.85	1.98	0.224	-	-	-	-
UDCA treatment	0.49	0.36	0.68	< .001	-	-	-	-
Age at entry, *per 10 years*	1.57	1.43	1.68	< .001	1.42	1.29	1.55	< .001
Year of diagnosis, *per decade*	0.79	0.62	0.96	.006	-	-	-	-
Bilirubin x ULN[b]	1.94	1.35	2.79	< .001	-	-	-	-
Albumin x LLN	0.18	0.06	0.57	< .001	-	-	-	-
ALP x ULN[b]	1.56	1.28	1.89	< .001	1.85	1.17	2.95	.009
AST/ALT ratio[b]	1.79	1.28	2.52	< .001	-	-	-	-
Platelets, *per 10 units (x10³/mm³) increase*	0.96	0.95	0.98	< .001	0.98	0.97	0.99	.045

ALP, alkaline phosphatase, AST, aspartate aminotransferase; ALT, alanine aminotransferase; ULN, upper limit of normal; LLN, lower limit of normal.
[a] HR obtained by considering biochemical transition as a time-dependent covariate.
[b] These biochemical variables were transformed with natural logarithm.

Supplementary Table 5. Baseline factors associated with the transition from biochemically early to moderately advanced disease in the original dataset

		Univariate analyses			Multivariable analysesa			
	N	**HR**	**95% CI**	***P***	**HR**	**95% CI**	***P***	
Male sex	1615	1.01	1.60	0.043	1.01	-	-	-
Age at entry, *per 10 years*	1615	1.00	1.12	0.044	1.00	-	-	-
Year of diagnosis, *per decade*	1615	0.86	0.78	0.78	<0.001	-	-	-
UDCA usage	1615	0.90	0.73	1.09	0.279	0.67	0.45	1.01	0.057
Bilirubin [b]	1615	2.14	1.78	2.58	<0.001	1.84	1.46	2.31	<0.001
Albumin	1615	0.07	0.04	0.12	<0.001	0.04	0.02	0.09	<0.001
ALP x ULN[b]	1486	1.44	1.30	1.60	<0.001	1.20	1.05	1.36	0.002
AST x ULN[b]	1406	1.44	1.26	1.65	<0.001				
ALT x ULN	1338	1.15	1.03	1.29	0.015				
AST/ALT ratio	1322	1.37	1.12	1.67	0.002	-	-	-	-
Platelets, *per 10 units increase*	1202	0.98	0.97	0.99	<0.001	0.99	0.99	0.99	0.011

ALP, alkaline phosphatase, AST, aspartate aminotransferase; ALT, alanine aminotransferase; ULN, upper limit of normal; LLN, lower limit of normal.

[a] This model was constructed for 1042 patients with complete data on all variables.

[b] These biochemical variables were transformed with natural logarithm

ALP, alkaline phosphatase, AST, aspartate aminotransferase; ALT, alanine aminotransferase; ULN, upper limit of normal; LLN, lower limit of normal.

Supplementary Table 6. The association of time-dependent transition with clinical events in original dataset

		Univariate analyses				Multivariable analyses [a]			
	N	**HR**	**95% CI**		***P***	**HR**	**95% CI**		***P***
Moderate vs. early biochemical stage[a]	1615	3.37	2.28	4.97	< .001	2.54	1.58	4.09	< .001
Advanced vs. moderate biochemical stage[a]	1615	5.38	4.05	7.13	< .001	4.09	2.83	5.93	< .001
Advanced vs. early biochemical stage[a]	1615	18.12	12.10	27.12	< .001	10.42	6.23	17.42	< .001
Male sex	1615	1.30	0.85	1.98	.224	-	-	-	-
UDCA treatment	1615	0.49	0.36	0.68	< .001	-	-	-	-
Age at entry, *per 10 years*	1615	1.57	1.43	1.68	< .001	1.56	1.40	1.72	< .001
Year of diagnosis, *per decade*	1615	0.79	0.62	0.96	.006	-	-	-	-
Bilirubin x ULN[b]	1615	1.94	1.35	2.79	< .001	-	-	-	-
Albumin x LLN	1615	0.18	0.06	0.57	< .001	-	-	-	-
ALP x ULN[c]	1486	1.59	1.31	1.95	< .001	1.56	1.19	2.04	.001
AST x ULN[c]	1406	1.49	1.14	1.93	.003				
ALT x ULN[c]	1338	1.06	0.84	1.33	.637				
AST/ALT ratio[c]	1322	1.92	1.30	2.82	.001	-	-	-	-
Platelets, *per 10 units increase*	1202	0.96	0.94	0.98	< .001	0.97	0.95	0.99	.007

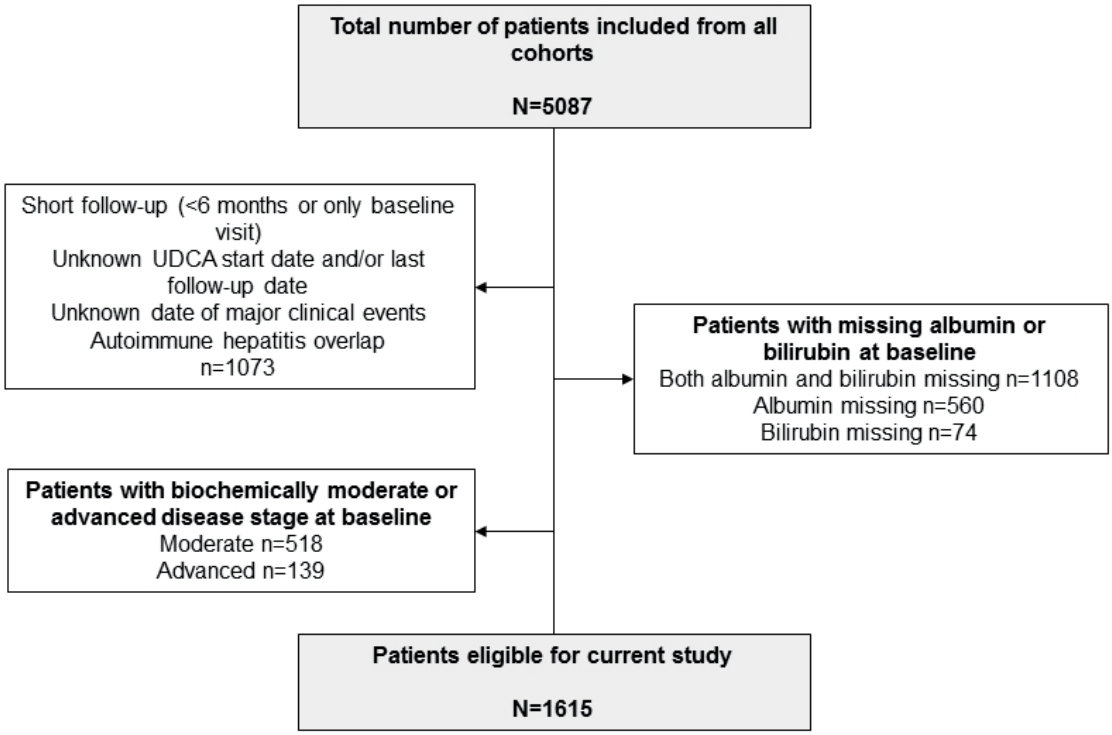
ALP, alkaline phosphatase, AST, aspartate aminotransferase; ALT, alanine aminotransferase; ULN, upper limit of normal; LLN, lower limit of normal.

[a] This model was constructed for 1042 patients with complete data on all variables.

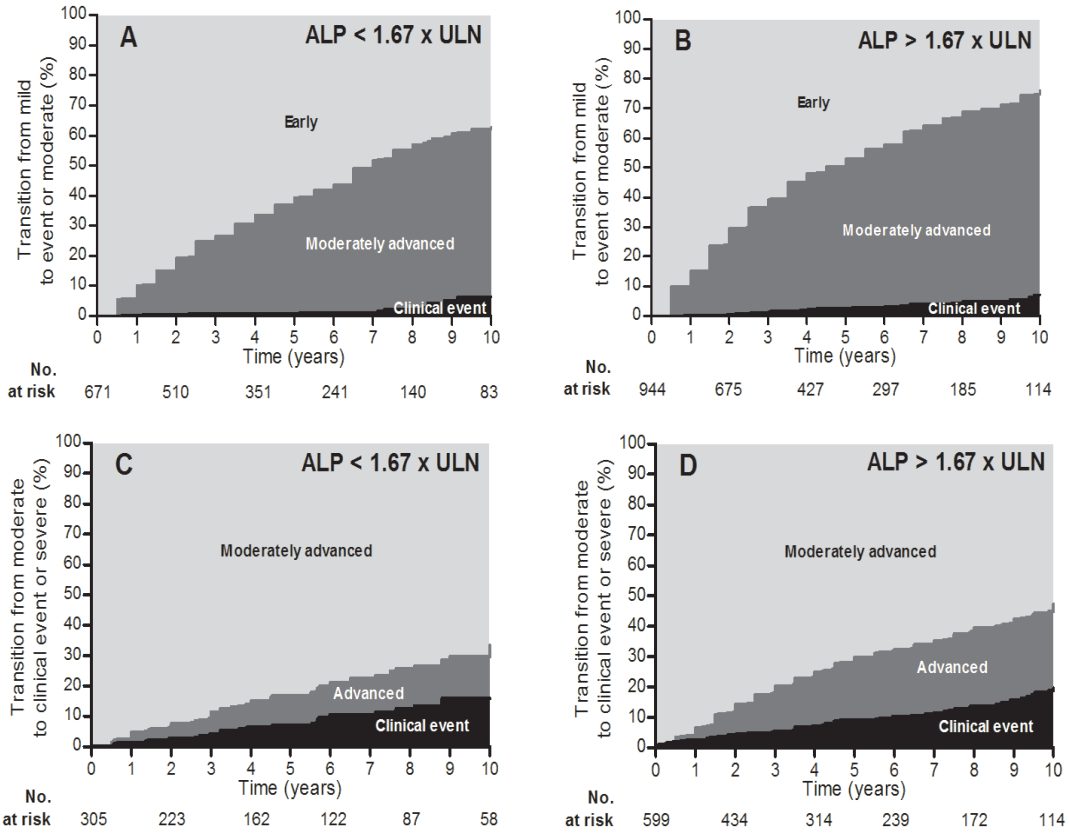
[b] HR obtained by considering biochemical transition as a time-dependent covariate.

[c] These biochemical variables were transformed with natural logarithm.

Supplementary Figure 1. Flowchart of patients included in the study

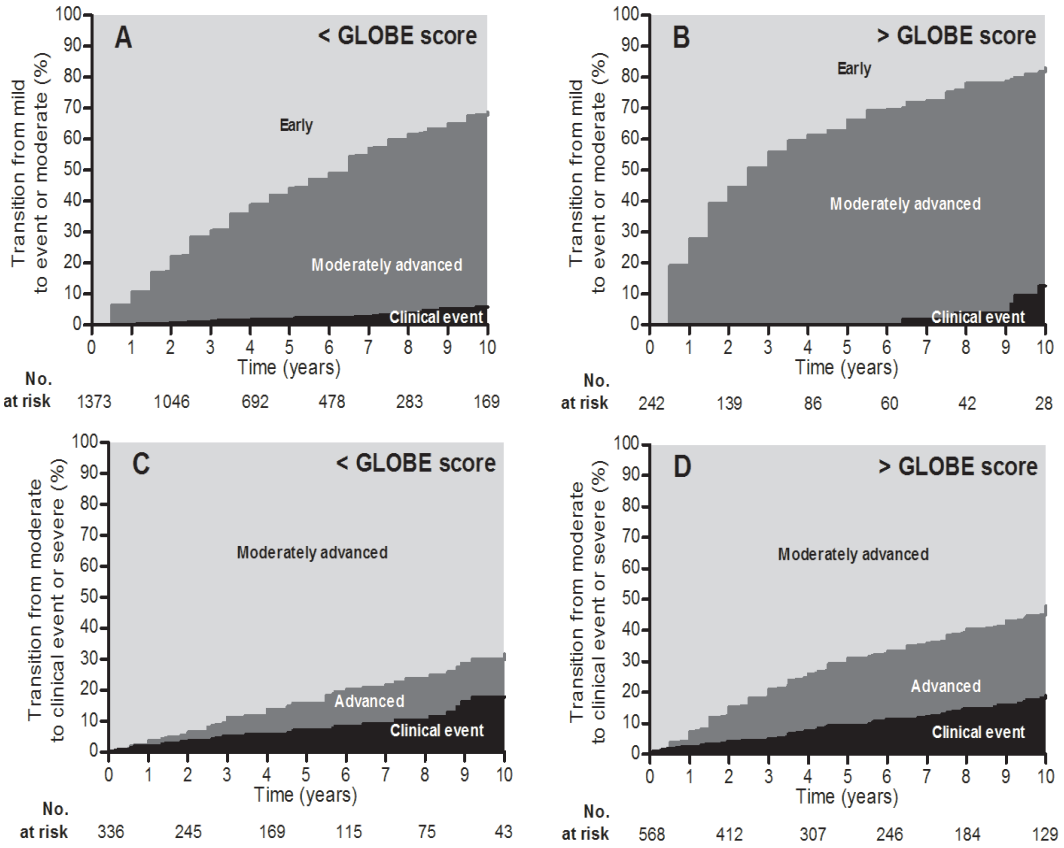


Supplementary Figure 2. Cumulative incidence of biochemical transition and events stratified by alkaline phosphatase



Kaplan Meier estimates of transition from early to a moderately advanced disease stage or event (A and B), and transition from moderately advanced to advanced disease stage or an event (C and D), stratified according to ALP.

Supplementary Figure 3. Cumulative incidence of biochemical transition and events stratified by GLOBE score



Kaplan Meier estimates of transition from early to a moderately advanced disease stage or event (A and B), and transition from moderately advanced to advanced disease stage or an event (C and D), stratified according to GLOBE score.

CHAPTER 5

5

A COMPARISON OF PROGNOSTIC SCORES (MAYO, UK-PBC and GLOBE) IN PRIMARY BILIARY CHOLANGITIS

Goet JC, Murrilo Perez CF, Harms MH, Floreani A, Cazzagon N, Bruns T, Prechter F, Dalekos GN, Verhelst X, Gatselis NK, Linder KD, Lammers WJ, Gulamhusein A, Reig A, Carbone M, Nevens F, Hirschfield GM, van der Meer AJ, van Buuren HR, Hansen BE*, Parés A* - *on behalf of the Global PBC Study Group*

*shared last authorship

American Journal of Gastroenterology, 2021 - accepted

ABSTRACT

Background: Comparative data on scores that predict outcome in primary biliary cholangitis (PBC) are scarce. We aimed to assess and compare the prognostic value of the Mayo risk score (MRS, 1989 and 1994), UK-PBC and GLOBE scores in a large international cohort of patients with PBC.

Methods: Ursodeoxycholic acid (UDCA)-treated patients from 7 centers participating in the GLOBAL PBC Study Group were included. The discriminatory performance of the scores was assessed with C-statistics at yearly intervals up to 5 years. MELD was included for comparison. Prediction accuracy was assessed by comparing predicted survival and actual survival in Kaplan-Meier analyses.

Results: 1100 UDCA-treated PBC patients were included, with a mean (SD) age of 53.6 (12.0) years, of whom 1003 (91%) were female. During a median follow-up of 7.6 (IQR 4.1-11.7) years, 42 patients underwent LT and 127 patients died. At 1 year, the C-statistic for MELD was 0.68 (95% confidence interval [CI] 0.64-0.72), 0.74 (95% CI 0.67-0.80) for UK-PBC, 0.76 (95% CI 0.72-0.81) for MRS (1989 and 1994), and 0.80 (95% CI 0.76-0.84) for GLOBE score. The GLOBE score showed superior discriminatory performance but differences were not statistically different. For all scores, discriminatory performance increased in those with bilirubin >0.6×ULN and advanced fibrosis estimated with FIB-4. The predicted (median) minus observed 5-year transplant-free survival was +0.4% and +2.5% for the MRS (1989) and GLOBE, respectively.

Conclusion: All prognostic scores developed for PBC (GLOBE score, UK-PBC and MRS) demonstrated comparable discriminating performance for LT or death, as well as good prediction accuracy.

STUDY HIGHLIGHTS

What is known:
- The Mayo Risk score, UK-PBC, and GLOBE score predict clinical outcomes in patients with primary biliary cholangitis
- These scores were developed in varying patient populations and with varying treatment status

What is new here:
- Prediction of clinical outcomes by Mayo Risk score, UK-PBC, and GLOBE is equivalent in UDCA-treated patients. Implementation of risk scores in PBC should be based on clinical context

INTRODUCTION

Primary biliary cholangitis (PBC) is a chronic cholestatic liver disease that predominantly affects middle-aged women[1,2]. PBC is a usually slowly progressive disorder, potentially leading to cirrhosis, liver failure requiring liver transplantation (LT), or death[1,2]. On an individual level, patients are nowadays often asymptomatic at diagnosis while the clinical course and response to therapy vary greatly[3,4].

Over the past decades, several risk scores have been proposed in PBC that can estimate a patient's risk of adverse outcomes and that can aid in the process of patient counselling and medical management, in particular with respect to treatment decisions and timing of liver transplantation. The Mayo Risk Score (MRS) is a frequently used model to predict survival probability with an initial intended application in the selection and timing of LT. This score was originally developed in untreated patients with PBC to predict survival up to 7 years[5], later adapted to predict short-term survival at 2 years and for use at any point during follow-up[6], and eventually abbreviated to quickly estimate the risk score[7]. Data regarding the prognostic performance of the MRS in ursodeoxycholic (UDCA)-treated patients is conflicting[8–12].

A more general model currently used to allocate patients for liver transplantation is the Model for End-stage Liver Disease (MELD). The MELD score was originally developed to predict survival in cirrhotic patients who underwent placement of a transjugular intrahepatic portosystemic shunt[13] and later modified and validated for the prediction of short-term survival in patients with cirrhosis with varying disease severity and etiology, including PBC[14]. To date, data on the appropriateness of the MELD score for risk stratification in the context of medical treatment in PBC patients are lacking.

More recently, two new models were introduced. The UK-PBC group developed a new scoring system for long-term prediction of liver transplantation and liver-related death with the best fitting model comprising baseline albumin and platelet count, as well as bilirubin, transaminases, and alkaline phosphatase, after 12 months of UDCA[15] they do not take other prognostic variables into account, such as the stage of the liver disease. We sought to improve existing long-term prognostic models of PBC using data from the UK-PBC Research Cohort. We performed Cox's proportional hazards regression analysis of diverse explanatory variables in a derivation cohort of 1,916 UDCA-treated participants. We used nonautomatic backward selection to derive the best-fitting Cox model, from which we derived a multivariable fractional polynomial model. We combined linear predictors and baseline survivor functions in equations to score the risk of a liver transplant or liver-related death occurring within 5, 10, or 15 years. We validated these risk scores in an independent cohort of 1,249 UDCA-treated participants. The best-fitting model consisted of the baseline albumin and platelet count, as well as the bilirubin, transaminases, and alkaline phosphatase, after 12 months of UDCA. In the validation cohort, the 5-, 10-, and 15-year risk scores were highly accurate (areas under the curve: >0.90. The GLOBE score comprises age, bilirubin, albumin, alkaline phosphatase, and platelet count as independent predictors of LT or death in UDCA-treated PBC patients[16]. The performance

of the UK-PBC risk score and GLOBE score as compared to the MRS in UDCA-treated PBC patients is not known.

In the current study, we aimed to assess and compare the performance of these prognostic scores developed for PBC in an international cohort of UDCA-treated PBC patients, while also taking into consideration the MELD score.

PATIENTS AND METHODS

Population and study design

Patients' data was derived from the GLOBAL PBC Study Group database (GPBCsg) Characteristics of the GPBCsg's cohort, comprising long-term follow-up data of 18 liver units across Europe and North America, have been described elsewhere[16]. For the current study, patients' data was derived from 7 centers from the Global PBC Study Group database: Toronto Centre for Liver Disease, University of Toronto, Canada; University of Padua, Padua, Italy; University of Thessaly, Larissa, Greece; University of Jena, Jena, Germany; University of Barcelona, Barcelona, Spain; Ghent University Hospital, Ghent, Belgium, Erasmus University Medical Center, Rotterdam, The Netherlands. UDCA-treated patients with an established diagnosis of PBC in accordance with internationally accepted guidelines were included[17,18]. Patients were excluded if the follow-up was less than 6 months and/or less than 2 recorded visits, the date of start of treatment or date of major clinical events was unknown, or in the case of concomitant liver disease.

Data collection

The following clinical data were collected for the original cohort: sex, age, date of PBC diagnosis, liver histology, treatment (type of medication, dosage and duration), last follow-up date, and clinical outcomes (death, cause of death, liver transplantation). Previously collected laboratory data collected included: baseline antimitochondrial antibody status, and baseline and yearly laboratory values (serum alkaline phosphatase [ALP], total bilirubin, albumin, aspartate aminotransferase [AST], alanine aminotransferase [ALT], and platelet count). Stage of disease was defined biochemically. Biochemical disease stage was classified according to Rotterdam criteria[11], namely mild (normal bilirubin and albumin), moderately advanced (abnormal bilirubin or albumin) and advanced disease (both abnormal bilirubin and albumin). Data in the original cohort were collected up to December 31st, 2012[19]. For three centres (University of Jena, Jena, Germany; University of Thessaly, Larissa, Greece; Ghent University Hospital, Ghent, Belgium), data were collected up to December 31st, 2015. To enable calculation of all risk scores, additional information was collected on dialysis treatment, use of diuretics, presence of peripheral edema, serum creatinine, prothrombin time (PT), and international normalized ratio (INR). In case a physical examination was documented and there was an absence of documented edema we presumed "no edema".

Extensive efforts were made to ensure completeness and reliability of the data, including center visits for paper and electronic chart review. This study was conducted in accordance with the protocol and the principles of the Declaration of Helsinki. The protocol was approved by the Institutional Research Board of the corresponding centre, and at each participating centre in accordance with local regulations.

Statistical analyses

Baseline was set at start of UDCA therapy. The primary endpoint was defined as a composite of either liver transplantation or death. Patients without documented events during follow-up were censored at their last follow-up visit. The 1989 MRS was calculated using the formula: 0.0394 x age + 0.8707 x ln(bilirubin [mg/dl]) + 2.380 x ln(PT)+ 0.8592 x edema - 2.533 x ln(albumin [g/dl]). The 1994 MRS was calculated with the formula: 0.051 x age + 1.209 x ln(bilirubin [mg/dl]) + 2.754 x ln(PT)+ 0.675*edema - 3.304 x ln(albumin [g/dl]). In cases when PT was missing 6.843 x ln(INR) was used instead of PT. Edema was coded as 0 for no edema and no diuretic therapy; 0.5 for edema present without diuretic therapy or edema resolved with diuretic therapy; and 1 for edema despite diuretic therapy. For comparative purposes, we included MELD and calculated lab MELD score using the formula: 10 x 0.957 x Log_e (creatinine [mg/dL]) + 0.378 x Log_e (bilirubin [mg/dL]) + 1.120 x Log_e (INR) + 0.643. Laboratory values less than 1.0 were set to 1.0 in the calculation; maximum serum creatinine in the equation was 4.0 mg/dL; lab MELD scores exceeding 40 were adjusted to 40.[20] The GLOBE score was calculated using the formula: 0.044378 x $age^{at\ start\ of\ UDCA\ therapy}$ + 0.335648 x ln($ALP^{1\ year\ UDCA}$/upper limit of normal [ULN]) + 0.93982 x ln($bilirubin^{1\ year\ UDCA}$/ULN) - 0.002581 x Platelet $count^{1\ year\ UDCA}$ per 10^9/L - 2.266708 x $albumin^{1\ year\ UDCA}$/lower limit of normal (LLN) + 1.216865. The UK-PBC score was calculated as follows: r= 0.0287854*(alp12*ULN-1.722136304)-0.0422873*(((alt ast12*ULN/10)^-1)-8.675729006)+ 1.4199*(ln(bili12*ULN/10)+2.709607778) -1.960303*(alb0*LLN-1.17673001)–0.4161954*(plt0*LLN- 1.873564875)).

These scores were calculated at yearly intervals up to 5 years after initiation of UDCA therapy. We used descriptive statistics, including boxplots, to visualize the various risk score indices during follow-up in patients that would eventually have a composite endpoint of liver transplantation or death in comparison to patients alive at the end of follow-up.

Validity of the prediction models was assessed based on discrimination and calibration of the models. Discrimination is the ability to categorize those with and without the outcome of interest based on predictive values[21]. Calibration is the measure of how accurately the predicted outcome matches the observed outcome[21]. At yearly time points, Cox proportional hazards regressions were conducted and the overall discriminative performance for the different scores was calculated with concordance statistic (C-statistic). Cox regression analyses were performed to assess the additional value of combining risk prediction models in estimating the risk of liver transplantation or death with application at 1 year of UDCA. In addition, the C-statistic for various combinations of risk prediction models was assessed.

Sub-analyses of discriminative ability for the various risk prediction models was performed in patients with bilirubin ≤ 0.6 x ULN compared to those with bilirubin values > 0.6 x ULN at baseline and 1 year of UDCA, as this threshold was associated with increased risk for liver transplantation and death[22]. In addition, to assess the performance of the various risk prediction models in those with no or low fibrosis stage (stage 1 and 2) versus those with advanced fibrosis (stage 3 and 4), patients were stratified according to Fibrosis-4 (FIB-4) Index for Liver Fibrosis[23]. Patients with a FIB-4≥ 1.8 were considered to have advanced (stage 3 and 4) fibrosis[24].

Model calibration for the MRS, MELD, UK-PBC and GLOBE score was assessed graphically by comparing observed transplant-free survival from Kaplan-Meier estimates with transplant-free survival predicted by the risk prediction models at 1 year of UDCA. The calibration for the UK-PBC survival estimates were not included in this analysis, as it relates to liver-related death survival rather than transplant-free survival.

To account for missing values SAS version 9.4 (SAS Institute Inc., Cary, NC, SAS Proc MI, MCMC method) was used to generate 10 imputed datasets of laboratory results at yearly time points between initiation of UDCA therapy and 5 years of follow-up, as described in a previous study, as described in a previous study[25–28]. This method uses chained equation to simultaneous impute all missing values drawing from the distribution of known values. Missing data were considered to be missing at random. Rubin's rules were used for estimation of the parameters and the standard error[28]. The imputation model included baseline variables that were potentially predictive for outcomes in PBC (e.g. year of diagnosis, age) as well as the outcomes themselves. In cases where PT was missing, we assumed normal PT and INR values when albumin and bilirubin were within the normal range. Subsequently, the missing PT and INR values were imputed by multiple imputation as previously described. Data are presented as median and interquartile range (IQR) for continuous variables.

RESULTS

Study population characteristics

A total of 1100 UDCA-treated PBC patients were included, with a mean age at start of follow-up of 53.6 (SD 12.0) years, of whom 1003 (91%) were females. Clinical and biochemical patient characteristics at initiation of UDCA therapy are shown in **Table 1**. Median follow-up was 7.6 (IQR 4.1-11.7) years. During follow-up, a total of 169 patients experienced a clinical endpoint, 42 underwent liver transplantation and 127 patients died. In 86/127 (67.7%) patients the cause of death was considered liver related. For the current study population, the 5-, 10-, and 15-year transplant-free survival rates were 93.4%, 83.8%, and 75.6% respectively, as shown in **Figure 1**.

At initiation of UDCA therapy, 215 (19.6%) patients had serum bilirubin values above the ULN and 107 (9.7%) had albumin values below the LLN. The patient population consisted of 816 (74.2%) patients with biochemically early disease stage according to Rotterdam criteria (normal albumin and bilirubin), 241 (21.9%) had moderately advanced disease stage (abnormal albumin or bilirubin), and 43 (3.9%) had advanced disease stage (abnormal albumin and bilirubin).

At the start of UDCA therapy, the median (IQR) score for MRS (1989 model), MRS (1994 model), MELD, and GLOBE was 3.94 (3.38-4.58), 4.24 (3.50-5.05), 7.00 (6.00-9.00), and 0.02 (-0.64-0.75), respectively (**Table 1**). Median scores of the various risk score indices at initiation of UDCA therapy and 5 years thereafter in patients that developed a clinical endpoint versus those that were still alive at the end of follow-up are shown in **Figure 2**.

Discriminatory performance of the Mayo risk score, MELD, UK-PBC and GLOBE scores

At baseline, the overall discriminatory performance of the GLOBE score, expressed as the C-statistic, for predicting the risk of death or liver transplantation was 0.78 (95% confidence interval [CI] 0.74-0.82) versus 0.77 (95% CI 0.73-0.81) for the MRS (1989 and 1994) and 0.68 (95% CI 0.65-0.71) for the MELD score (**Supplementary Table 1**). At 1 year of UDCA therapy the C-statistic for the GLOBE score was 0.80 (95% CI 0.76-0.84), 0.76 (95% CI 0.72-0.81) for MRS (1989 and 1994), 0.68 (95% CI 0.64-0.72) for the MELD score, and 0.74 (95% CI 0.67-0.80) for UK-PBC. The performance of MELD, as assessed with c-statistics, was statistically significantly lower compared to the remaining scores. In the 5 years after initiation of UDCA therapy, the difference in discriminatory performance for the various risk prediction models remained comparable (**Table 2** and **Supplementary Table 1**). While the performance of the GLOBE score was statistically different from that of UK-PBC for the prediction of liver transplantation and death at 1 year (P=0.02), there were no statistically significant differences between these scores for the prediction of liver-related death or liver transplantation at 1 year of UDCA therapy, which was 0.81 (95% CI 0.77-0.86) for the GLOBE score and 0.81 (95% CI 0.76-0.85) for UK-PBC (P=0.45) (**Supplementary Table 1**).

Table 1. Baseline cohort characteristics

	Total cohort N=1100
Center, n, (%)	
Rotterdam, the Netherlands	88 (8.0)
Barcelona, Spain	27 (2.5)
Padua, Italy	240 (21.8)
Toronto, Canada	487 (44.3)
Larissa, Greece	210 (19.1)
Jena, Germany	39 (3.5)
Ghent, Belgium	9 (0.8)
Age at diagnosis, years	52.4 (12.1)
Age at start of follow-up	53.6 (12.0)
Female, n (%)	1003 (91.2)
AMA+, n (%)	995 (90.5)
Year of diagnosis	2001 (1995-2006)
Year of diagnosis, range	1971-2015
Serum total bilirubin ×ULN	0.60 (0.45-0.90)
Serum ALP ×ULN	1.93 (1.19-3.54)
Serum AST ×ULN	1.49 (1.00-2.35)
Serum ALT ×ULN	1.62 (1.00-2.64)
Serum albumin ×LLN	1.17 (1.08-1.27)
Serum platelets $\times 10^9$/L	253 (94.2)
Serum creatinine ×ULN	0.74 (0.58-0.86)
PT (sec)	12.0 (11.00-13.00)
INR	1.00 (0.93-1.08)
MRS 1989	3.94 (3.38-4.58)
MRS 1994	4.24 (3.50-5.05)
MELD	7.00 (6.00-9.00)
GLOBE Score	0.02 (-0.64-0.75)
Death	127
Liver-related death	86
Liver transplantation	42

Abbreviations: AMA, anti-mitochondrial antibodies; ULN, upper limit of normal; ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; LLN, lower limit of normal; PT, prothrombin time; INR, international normalized ratio.
Data represented as mean (standard deviation) and median (interquartile range) unless specified otherwise.

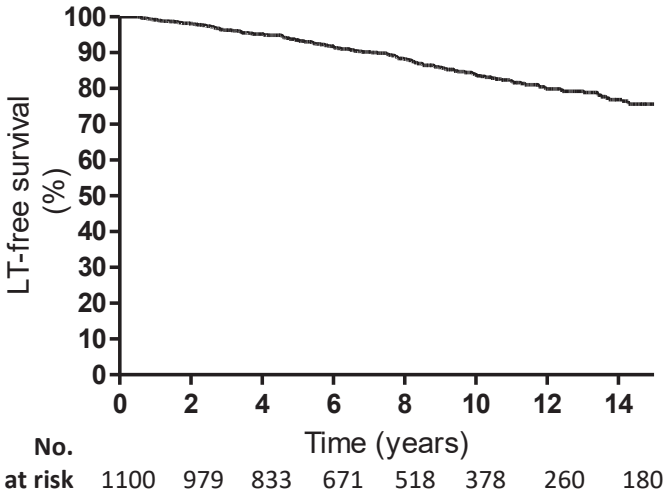


Figure 1. Kaplan-Meier estimate of survival in this cohort

Table 2. Discriminative performance of the various risk prediction scores calculated after 1, 3 and 5 years of UDCA therapy

Risk prediction model	C-statistic at various follow-up time points (95% CI)		
	1 year of UDCA	3 years of UDCA	5 years of UDCA
MRS 1989	0.76 (0.72 - 0.81)	0.82 (0.77 - 0.87)	0.80 (0.74 - 0.86)
MRS 1994	0.76 (0.72 - 0.81)	0.82 (0.78 - 0.87)	0.81 (0.75 - 0.86)
MELD	0.68 (0.64 - 0.72)	0.76 (0.71 - 0.80)	0.70 (0.66 - 0.75)
UK-PBC	0.74 (0.67 - 0.80)	0.78 (0.72 - 0.84)	0.80 (0.75 - 0.86)
GLOBE score	0.80 (0.76 - 0.84)	0.83 (0.78 - 0.88)	0.84 (0.79 - 0.90)

Abbreviations: CI, confidence interval; UDCA, ursodeoxycholic acid.

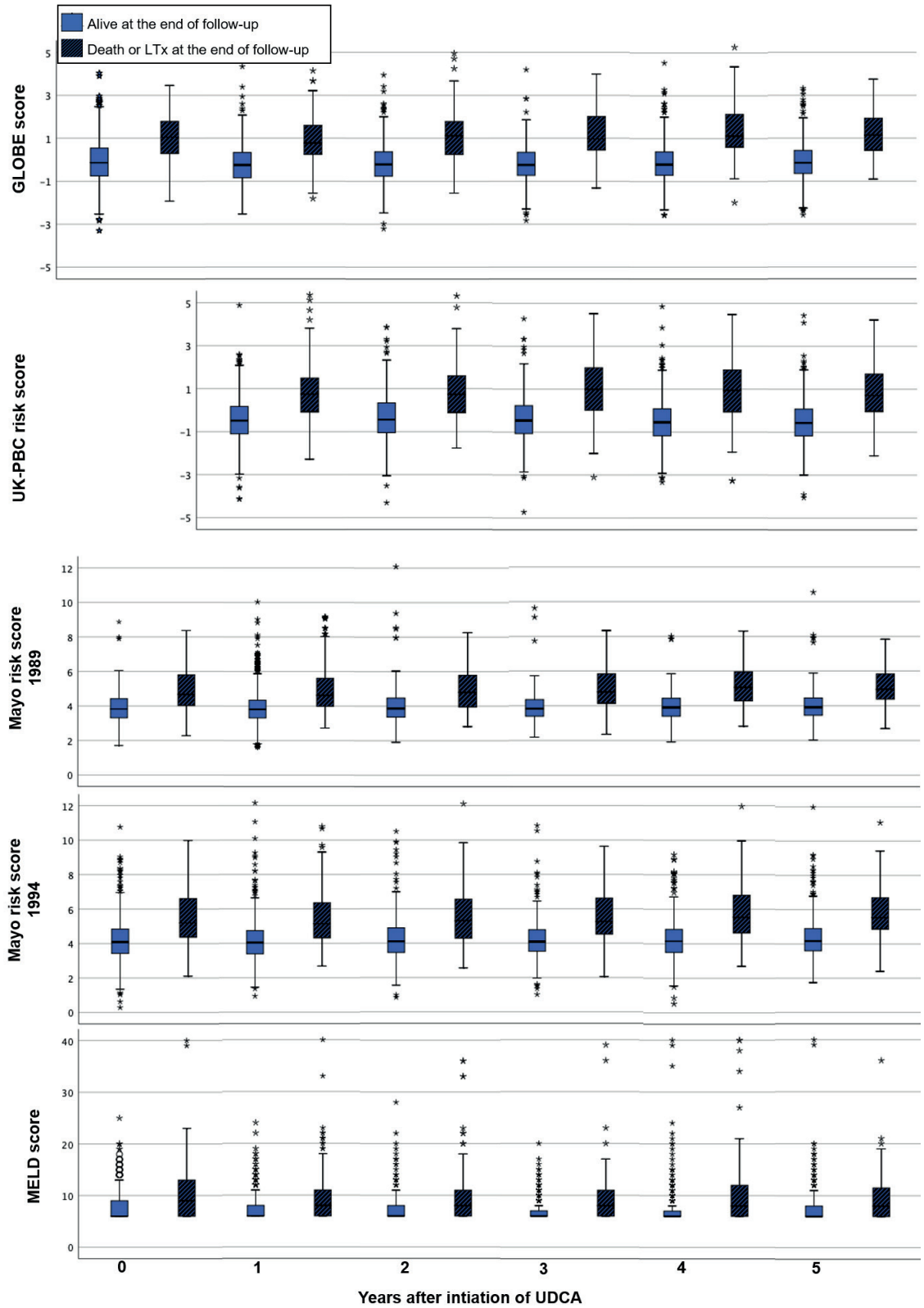


Figure 2. Boxplots of the various risk prediction scores from initiation of UDCA therapy to 5 years according to whether they experienced a clinical outcome at the end of follow-up

Sub-analyses of the discriminatory ability in patients with bilirubin values ≤ 0.6 x ULN and those with bilirubin values > 0.6 x ULN at baseline and 1 year of UDCA showed that, in general, all scores had better discriminative performance in patients with bilirubin values > 0.6 x ULN (**Table 3**). Sub-analyses according to FIB-4 were also performed, in which a total of 387 (35.2%) patients had FIB-4 scores > 1.8 at initiation of UDCA therapy indicating advanced fibrosis. At 1 year of UDCA therapy, 253/905 (28.0%) patients met the threshold for advanced fibrosis. Discriminatory ability of the risk scores stratified according to FIB-4 demonstrated that the performance is higher in those with FIB-4 ≥ 1.8.

Combined performance of the Mayo risk score, MELD, UK-PBC and GLOBE scores

In univariable Cox regression analyses, the prognostic indexes of all individual scores were significantly associated with death or liver transplantation (**Table 4**). In a multivariable analysis that included all respective scores with the exclusion of MRS 1989, only the GLOBE score (hazard ratio (HR) 2.36 [95% confidence interval (CI): 1.71-3.27) *P*<.001]) and MRS 1994 (HR 1.28 [95% CI: 1.06-1.55; *P*=.01]) remained significantly associated with death or liver transplantation.

Addition of the MRS, MELD, or UK-PBC to the GLOBE score did not result in an increase in discriminatory performance, which remained at 0.80 (**Table 5**). Combining the UK-PBC score with the MRS, MELD or GLOBE resulted in an increase in C-statistic ranging from 0.01 to 0.06, with the highest increase observed from the addition of the GLOBE score and lowest from MELD. For various combinations of the MRS with other scores, relatively smaller changes in C-statistic were observed with the highest being from the addition of the GLOBE score (+0.04) (**Table 5**). In contrast, the addition of all scores to the MELD score yielded an increase in C-statistic, ranging from 0.07 to 0.12.

Prediction accuracy (calibration) of the Mayo risk score, MELD, UK-PBC and GLOBE scores

In **Figure 3** the observed and median predicted survival for the various risk prediction models are shown. For all models, good calibration for short-term and long-term survival was observed. In the estimates of survival, both the GLOBE and MRS 1994 tended to overestimate transplant-free survival, with the greatest deviation from observed survival at 10 years for GLOBE (3.5%) and 2 years for MRS 1994 (2.9%). MRS 1989 demonstrated the best calibration, as the difference in predicted versus observed survival was generally less than 1% at yearly intervals up to 7 years (**Supplementary Table 2**).

Table 3. Discriminative performance of the various risk prediction scores calculated at baseline and after 1 year of UDCA therapy stratified by bilirubin values and FIB-4.

	Bilirubin ≤ 0.6×ULN n=556		Bilirubin > 0.6×ULN n=544	
Model	**C-statistic**	**95% CI**	**C-statistic**	**95% CI**
Baseline (n=1100)				
MRS 1989	0.62	0.55-0.69	0.75	0.71-0.80
MRS 1994	0.62	0.56-0.69	0.75	0.70-0.79
MELD score	0.54	0.53-0.56	0.68	0.64-0.71
GLOBE score	0.72	0.63-0.81	0.75	0.70-0.79
1 year (n=905)	***n=521***		***n=384***	
MRS 1989	0.72	0.64-0.81	0.74	0.68-0.79
MRS 1994	0.72	0.63-0.81	0.73	0.67-0.79
MELD score	0.60	0.56-0.64	0.65	0.60-0.70
UK-PBC	0.62	0.51-0.72	0.70	0.64-0.76
GLOBE score	0.74	0.64-0.83	0.77	0.72-0.82
	FIB-4 < 1.8 n=713		**FIB-4 ≥ 1.8 n=387**	
Baseline (n=1100)				
MRS 1989	0.72	0.64-0.81	0.74	0.68-0.79
MRS 1994	0.72	0.63-0.81	0.73	0.67-0.79
MELD score	0.60	0.56-0.64	0.65	0.60-0.70
GLOBE score	0.72	0.63-0.81	0.75	0.70-0.79
1 year (n=905)	***n=652***		***n=253***	
MRS 1989	0.72	0.64-0.81	0.74	0.68-0.79
MRS 1994	0.72	0.63-0.81	0.73	0.67-0.79
MELD score	0.60	0.56-0.64	0.65	0.60-0.70
UK-PBC	0.62	0.51-0.72	0.70	0.64-0.76
GLOBE score	0.74	0.64-0.83	0.77	0.72-0.82

Abbreviations: UDCA, ursodeoxycholic acid; xULN, times upper limit of normal; CI, confidence interval.

Table 4. Multivariable analyses of risk prediction scores at 1 year of UDCA therapy (N=905)

	Univariate analyses			Multivariable analyses		
Prognostic score	**Hazard ratio**	**95% CI**	**p-value**	**Hazard ratio**	**95% CI**	**p-value**
MRS 1989	2.40	3.14-2.70	<0.001			
MRS 1994	1.98	1.81-2.17	<0.001	1.28	1.06-1.55	0.01
MELD	1.15	1.12-1.18	<0.001	1.02	0.97-1.07	0.37
UK-PBC	2.10	1.86-2.37	<0.001	0.99	0.82-1.21	0.99
GLOBE	3.34	2.83-3.95	<0.001	2.36	1.71-3.27	<0.001

Abbreviations: UDCA, ursodeoxycholic acid; CI, confidence interval.

Table 5. Cox regression analyses and combined discriminatory performance of prognostic scores at 1 year of UDCA therapy (N=905)

Multivariable analyses					
Prognostic score	**Hazard ratio**	**95% CI**	**p-value**	**C-statistic**	**95% CI**
GLOBE	3.09	2.35-4.05	<0.001	0.80	0.76-0.84
UK-PBC	1.07	0.87-1.33	0.52		
GLOBE	2.41	1.81-3.22	<0.001	0.80	0.75-0.84
MRS 1994	1.33	1.09-1.62	0.005		
GLOBE	3.02	2.49-3.66	<0.001	0.80	0.76-0.84
MELD	1.05	1.00-1.10	0.042		
UK-PBC	1.43	1.22-1.68	<0.001	0.78	0.73-0.82
MRS 1994	1.71	1.59-1.96	<0.001		
UK-PBC	1.92	1.64-2.25	<0.001	0.75	0.70-0.80
MELD	1.08	1.03-1.13	0.001		
MRS 1994	1.87	1.66-2.12	<0.001	0.77	0.72-0.82
MELD	1.04	0.99-1.10	0.12		

Abbreviation: UDCA, ursodeoxycholic acid;CI, confidence interval.

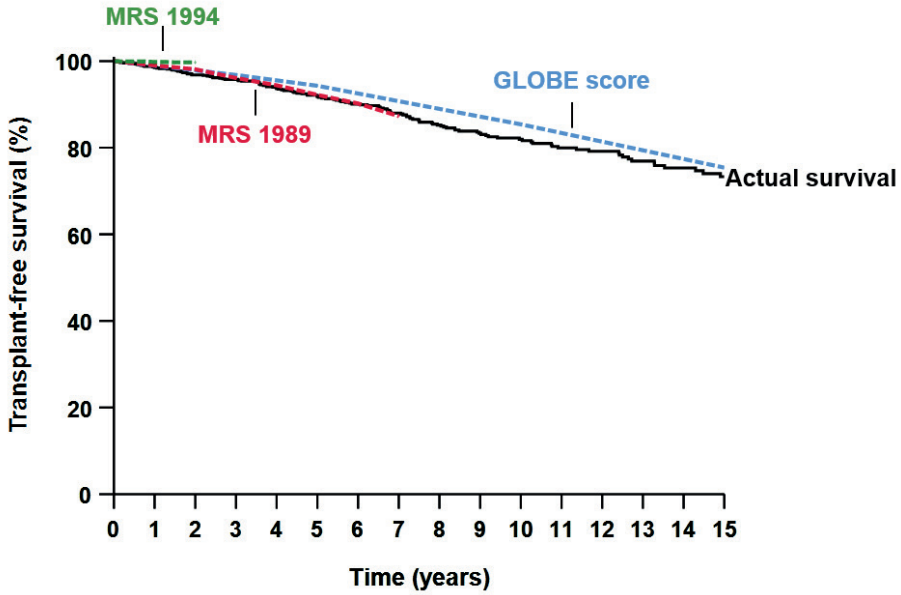


Figure 3. Predicted versus observed liver transplant-free survival for the GLOBE score and Mayo Risk Score (MRS 1989 and 1994). Figure shows prediction accuracy (calibration) of the GLOBE score and MRS up to 15 years of follow-up after 1 year of ursodeoxycholic acid (UDCA) therapy (N=905). Solid line = actual observed transplant-free survival probabilities estimated by Kaplan-Meier analyses. Dashed lines = the predicted median transplant-free survival probabilities as predicted by the GLOBE score and MRS.

DISCUSSION

In this large cohort of PBC patients, we assessed the performance of various published risk prediction models. We demonstrate that in a cohort of mainly early biochemical disease stage PBC patients, all prognostic scores evaluated (GLOBE, UK-PBC, MRS) have adequate discriminatory performance and good prediction accuracy. The discriminatory performance of these PBC-specific scores increased in those with bilirubin > 0.6 ×ULN and advanced fibrosis. Not surprisingly, our data also show that the performance of the MELD score, which was not developed for or has previously shown promise as a prognostic tool in early or non-cirrhotic liver disease, was clearly inferior to that of the PBC-specific scores.

The consistently high discriminative performance of the GLOBE score in our cohort suggests that more patients who experienced an event had a higher risk score and more patients without an event had a lower risk score than with the use of other scores. However, there were no significant differences in comparison to UK-PBC and MRS. In general, models with a C-statistic greater than 0.8 are considered good prognostic models, of which the GLOBE score was the only score to consistently reach this threshold in the prediction of transplant-free survival at various time points[29]. Secondary to the GLOBE score in discriminatory performance was the MRS (1989 and 1994). Although the MRS did not have a C-statistic above 0.8 at 1 year of UDCA therapy, the discriminatory performance increased when applied at other time points during prolonged UDCA treatment. While the MRS is the traditional risk prediction model in patients with PBC, its clinical utility may be hampered by the use of peripheral edema as a subjective parameter. It should be noted that the MELD score and MRS were derived in patients with end-stage liver disease and our cohort mainly comprised patients with biochemically early disease stage. In addition, while the MRS was developed in untreated patients with PBC, the current study included UDCA-treated patients. The prognostic value of MRS has been demonstrated in UDCA-treated patients to be associated with transplant-free survival as it stratifies patients into high-risk and low-risk groups using the original thresholds[9,10]. Given the adequate discriminatory performance and good prediction accuracy of these scores, the GLOBE and MRS can be implemented to predict overall transplant-free survival, while the clinical utility of the UK-PBC score can be aimed at predicting liver transplantation and liver-related death.

Not surprisingly, sub-group analyses showed that all risk prediction scores tended to have improved discriminatory performance in patients with bilirubin values > 0.6 x ULN compared to those with bilirubin values ≤ 0.6 x ULN. Bilirubin is one of the most robustly validated markers of disease progression in PBC and is included in all risk prediction models for PBC[19,30,31]. Bilirubin is mostly considered a "late" biomarker, i.e. elevations are seen only in late stages of the disease and increase shortly before a clinical event, and therefore may be considered less discriminatory for early detection of progression of disease and clinical outcome[30,32]. However, a recent study by the Global PBC study group showed that bilirubin values within the normal range, both at baseline and after one year of UDCA therapy, were predictive of transplant-free survival, suggesting that even increases in bilirubin values within the normal range should prompt reconsideration for second-line therapies and optimal management[22]. The threshold of 0.6 used in the current paper has been shown

to be associated with the lowest risk for liver transplantation or death, after which the risk increases[22]. Akin to the results observed for patients with bilirubin > 0.6×ULN, the various risk prediction models had better performance in those with FIB-4 levels above 1.8, which was the threshold best associated with advanced fibrosis[24]. These sub-group analyses suggest that current risk stratification tools are less accurate when used to risk stratify patients in earlier stages of disease.

Interestingly, combination of the indices of various risk prediction models in the estimation of death or liver transplantation, although statistically significant, did not result in a numerical increase in C-statistic, particularly for the GLOBE score. This suggests that, although it is not feasible to calculate multiple risk scores in clinical practice, there may some additional value of considering scores such as the MRS in addition to GLOBE. Various studies in UDCA-treated patients have reported that the MRS may underestimate survival[8,11,12]. In our study, we demonstrate that the MRS has good prediction accuracy and adequate performance and may therefore be of value in UDCA-treated patients. Theoretically, the added value of the MRS in discriminatory performance may be driven by prothrombin time and edema. However, because our cohort mainly comprises early-stage PBC patients in whom prothrombin time will be within the normal range and edema will be absent, this seems unlikely.

A strength of our study is the inclusion of a well-characterized large study population from multiple centers. Some limitations need to be considered. First, due to the retrospective nature of the current study a proportion of data was missing (**Supplementary Table 3**). To overcome this problem multiple imputation techniques were used[26]. Second, although some of the patients in this study were included in the derivation cohort of the GLOBE score, a substantial proportion (~25%) of patients not originally used in the derivation of the GLOBE score. However, sensitivity analyses of the discriminative performance of the various scores in the 25% of patients not included in the derivation of the GLOBE score yielded similar results (**Supplementary Table 4**). Third, our cohort mainly comprised early-stage disease patients. Even though our study population is representative of the majority of current PBC patients, as most patients nowadays present at early stages of disease[33], comparison of the various risk prediction models in more advanced stages of disease would be of additional value. Lastly, while the UK-PBC risk score was developed to predict a different endpoint, composed of liver-related death and liver transplantation, the discriminatory performance was also assessed for this endpoint and yielded similar results.

In conclusion, in this large cohort of mainly early disease stage PBC patients, we show that all prognostic scores developed for PBC (GLOBE, UK-PBC, MRS) have comparable performance in the prediction of clinical outcomes. Although the discriminating performance for LT or death of the GLOBE score was superior, this difference was not statistically significant compared to the other scores (MRS and UK-PBC). This is true for various time points during UDCA treatment as well as in sub-groups stratified according to biochemical and fibrosis disease stage. This suggests that implementation ought to be based on clinical context.

REFERENCES

1. Kaplan MM, Gershwin ME. Primary biliary cirrhosis. N Engl J Med 2005;353:1261-1273.
2. Poupon R. Primary biliary cirrhosis: a 2010 update. J Hepatol 2010;52:745-758.
3. Prince MI, Chetwynd A, Craig WL, et al. Asymptomatic primary biliary cirrhosis: clinical features, prognosis, and symptom progression in a large population based cohort. Gut 2004;53:865-870.
4. Floreani A, Caroli D, Variola A, et al. A 35-year follow-up of a large cohort of patients with primary biliary cirrhosis seen at a single centre. Liver Int 2011;31:361-368.
5. Dickson ER, Grambsch PM, Fleming TR, et al. Prognosis in primary biliary cirrhosis: model for decision making. Hepatology 1989;10:1-7.
6. Murtaugh PA, Dickson ER, Van Dam GM, et al. Primary biliary cirrhosis: prediction of short-term survival based on repeated patient visits. Hepatology 1994;20:126-134.
7. Kim WR, Wiesner RH, Poterucha JJ, et al. Adaptation of the Mayo primary biliary cirrhosis natural history model for application in liver transplant candidates. Liver Transplant 2000;6:489-494.
8. Parés A, Caballería L, Rodés J. Excellent long-term survival in patients with primary biliary cirrhosis and biochemical response to ursodeoxycholic acid. Gastroenterology 2006;130:715-720.
9. Kilmurry MR, Heathcote EJ, Cauch-Dudek K, et al. Is the Mayo model for predicting survival useful after the introduction of ursodeoxycholic acid treatment for primary biliary cirrhosis? Hepatology 1996;23:1148-1153.
10. Angulo P, Lindor KD, Therneau TM, et al. Utilization of the Mayo risk score in patients with primary biliary cirrhosis receiving ursodeoxycholic acid. Liver 1999;19:115-121.
11. ter Borg PCJ, Schalm SW, Hansen BE, et al. Prognosis of ursodeoxycholic acid-treated patients with primary biliary cirrhosis. Results of a 10-yr cohort study involving 297 patients. Am J Gastroenterol 2006;101:2044-2050.
12. Poupon RE, Bonnand AM, Chrétien Y, et al. Ten-year survival in ursodeoxycholic acid-treated patients with primary biliary cirrhosis. The UDCA-PBC Study Group. Hepatology 1999;29:1668-1671.
13. Malinchoc M, Kamath PS, Gordon FD, et al. A model to predict poor survival in patients undergoing transjugular intrahepatic portosystemic shunts. Hepatology 2000;31:864-871.
14. Kamath PS, Wiesner RH, Malinchoc M, et al. A model to predict survival in patients with end-stage liver disease. Hepatology 2001;33:464-470.
15. Carbone M, Sharp SJ, Flack S, et al. The UK-PBC risk scores: Derivation and validation of a scoring system for long-term prediction of end-stage liver disease in primary biliary cholangitis. Hepatology 2016;63:930-950.
16. Lammers WJ, Hirschfield GM, Corpechot C, et al. Development and validation of a scoring system to predict outcomes of patients with primary biliary cirrhosis receiving ursodeoxycholic acid therapy. Gastroenterology 2015;149:1804-1812.e4.
17. Hirschfield GM, Beuers U, Corpechot C, et al. EASL Clinical Practice Guidelines : The diagnosis and management of patients with primary biliary cholangitis. J Hepatol 2017;67:145-172.
18. Lindor KD, Gershwin ME, Poupon R, et al. Primary biliary cirrhosis. Hepatology 2009;50:291-308.

19. Lammers WJ, van Buuren HR, Hirschfield GM, 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 2014;147:1338-1349.e5.
20. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010;21:128-138.
21. Freeman RBJ, Wiesner RH, Harper A, et al. The new liver allocation system: moving toward evidence-based transplantation policy. Liver Transplant 2002;8:851-858.
22. Murillo Perez CF, Harms MH, Lindor KD, et al. Goals of Treatment for Improved Survival in Primary Biliary Cholangitis: Treatment Target Should Be Bilirubin Within the Normal Range and Normalization of Alkaline Phosphatase. Am J Gastroenterol 2020.
23. Sterling RK, Lissen E, Clumeck N, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology 2006;43:1317-1325.
24. Murillo Perez CF, Hirschfield GM, Corpechot C, et al. Fibrosis stage is an independent predictor of outcome in primary biliary cholangitis despite biochemical treatment response. Aliment Pharmacol Ther 2019;50:1127-1136.
25. Harms MH, Lammers WJ, Thorburn D, et al. Major Hepatic Complications in Ursodeoxycholic Acid-Treated Patients With Primary Biliary Cholangitis: Risk Factors and Time Trends in Incidence and Outcome. Am J Gastroenterol 2018;113:254-264.
26. Sterne JR, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009;338:b2393.
27. Rubin DB. Multiple imputation after 18+ years. J Am Statitical Assoc 1996;91:473-489.
28. Little RJA, Rubin DB. Statistical Analysis with Missing Data. New York: John Wiley & Sons, 1987.
29. Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. BMC Med Res Methodol 2013;13:33.
30. Shapiro JM, Smith H, Schaffner F. Serum bilirubin: a prognostic factor in primary biliary cirrhosis. Gut 1979;20:137-140.
31. Bonnand AM, Heathcote EJ, Lindor KD, et al. Clinical significance of serum bilirubin levels under ursodeoxycholic acid therapy in patients with primary biliary cirrhosis. Hepatology 1999;29:39-43.
32. Harms MH, Pares A, Mason AL, et al. Behavioral Patterns of Total Serum Bilirubin Prior to Major Clinical Endpoints in 3529 Patients with Primary Biliary Cholangitis. J Hepatol 2016;64:S633-S634.
33. Murillo Perez CF, Goet JC, Lammers WJ, et al. Milder disease stage in patients with primary biliary cholangitis over a 44-year period: A changing natural history. Hepatology 2018;67.

SUPPLEMENTARY DATA

Supplementary Table 1. The discriminative ability for all risk scores for liver transplantation or death between initiation of UDCA therapy up to 5 years thereafter

Model	N	C-statistic	95% CI
First calculable score in first 5 years following initiation of UDCA			
MRS 1989	1171	0.789	0.752 - 0.825
MRS 1994	1171	0.788	0.752 - 0.825
MELD score	1171	0.699	0.670 - 0.728
UK-PBC score[a]	1171	0.756	0.713 - 0.800
GLOBE score[a]	1171	0.794	0.759 - 0.830
At start of UDCA therapy			
MRS 1989	1100	0.768	0.728 - 0.809
MRS 1994	1100	0.768	0.728 - 0.808
MELD score	1100	0.681	0.652 - 0.710
GLOBE score	1100	0.779	0.739 - 0.818
At 12 months UDCA			
MRS 1989	905	0.764	0.716 - 0.813
MRS 1994	905	0.765	0.717 - 0.812
MELD score	905	0.679	0.638 - 0.720
UK-PBC score[b]	905	0.738	0.671 - 0.805
GLOBE score[b]	905	0.799	0.756 - 0.842
At 24 months UDCA			
MRS 1989	548	0.784	0.725 - 0.843
MRS 1994	548	0.786	0.727 - 0.844
MELD score	548	0.691	0.641 - 0.741
UK-PBC score[c]	548	0.741	0.691 - 0.791
GLOBE score[c]	548	0.817	0.765 - 0.868
At 36 months UDCA			
MRS 1989	648	0.819	0.769 - 0.868
MRS 1994	648	0.824	0.776 - 0.872
MELD score	648	0.757	0.711 - 0.802
UK-PBC score[d]	648	0.780	0.718 - 0.842
GLOBE score[d]	648	0.831	0.779 - 0.882
At 48 months UDCA			
MRS 1989	610	0.816	0.763 - 0.869
MRS 1994	610	0.821	0.769 - 0.873
MELD score	610	0.746	0.698 - 0.793
UK-PBC score[e]	610	0.804	0.749 - 0.859
GLOBE score[e]	610	0.844	0.797 - 0.891
At 60 months UDCA			
MRS 1989	549	0.801	0.744 - 0.8590
MRS 1994	549	0.806	0.749 - 0.8633
MELD score	549	0.704	0.657 - 0.7522
UK-PBC score[f]	549	0.778	0.716 - 0.8400
GLOBE score[f]	549	0.842	0.791 - 0.8931

Abbreviation: UDCA, ursodeoxycholic acid; CI, confidence interval.

[a]c-statistic for outcomes liver-related death or LT: 0.822 (0.785 - 0.859) for UK-PBC and 0.820 (0.784 - 0.857) for GLOBE score.
[b]c-statistic for outcomes liver-related death or LT: 0.809 (0.764 - 0.854) for UK-PBC and 0.814 (0.765 - 0.862) for GLOBE score.
[c] c-statistic for outcomes liver-related death or LT: 0.804 (0.737 - 0.870) for UK-PBC and 0.834 (0.775 - 0.893) for GLOBE score.
[d] c-statistic for outcomes liver-related death or LT: 0. 850 (0.798 - 0.903) for UK-PBC and 0.845 (0.790 - 0.901) for GLOBE score.
[e] c-statistic for outcomes liver-related death or LT: 0.870 (0.824 - 0.917) for UK-PBC and 0.860 (0.806 - 0.914) for GLOBE score.
[f] c-statistic for outcomes liver-related death or LT: 0.823 (0.764 - 0.882) for UK-PBC and 0.858 (0.805 - 0.912) for GLOBE score.

Supplementary Table 2. Difference between observed and median predicted transplant-free survival after 1 year of UDCA across different risk scores and time points

	Predicted - observed survival (%)		
Time after 1 year of UDCA	**Delta GLOBE**	**Delta MRS 1989**	**Delta MRS 1994**
1		+0.5	+1.3
2		+1.3	+2.9
3	+1.0	+0.3	
4		+0.7	
5	+2.5	+0.4	
6		+0.1	
7		-0.7	
8			
9			
10	+3.5		
11			
12			
13			
14			
15	+2.1		

Abbreviation: UDCA, ursodeoxycholic acid; MRS, Mayo risk score.

Supplementary Table 3. Proportion of patients with any missing data in the first 5 years after start of UDCA therapy

	Missing data N, %
Serum total bilirubin	36 (3.3)
Serum ALP	20 (1.8)
Serum AST	18 (1.6)
Serum ALT	20 (1.8)
Serum albumin	103 (9.4)
Serum platelets	34 (3.1)
Serum creatinine	349 (31.7)
PT[a]	203 (18.5)
INR[a]	101 (9.2)

Abbreviations: ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; PT, prothrombin time; INR, international normalized ratio.
[a] In the cases were PT was missing, we assumed normal PT and INR values when albumin and bilirubin were within the normal range. Subsequently, the missing PT and INR values were imputed by multiple imputation.

Supplementary Table 4. Sensitivity analysis of the discriminative performance of the various risk prediction scores calculated after 1 year of UDCA therapy in the 25% of cases not used in the derivation of the GLOBE score.

		C statistic (95% CI)
At 1 year of UDCA therapy - 17 events		
MRS 1989	224	0.92 (0.88-0.96)
MRS 1994	224	0.92 (0.87-0.97)
MELD score	224	0.81 (0.73-0.88)
UK-PBC score	224	0.87 (0.80-0.93)
GLOBE score	224	0.92 (0.88-0.97)

CHAPTER 6

6

VALIDATION, CLINICAL UTILITY AND LIMITATIONS OF THE AMSTERDAM-OXFORD MODEL FOR PRIMARY SCLEROSING CHOLANGITIS

Goet JC, Floreani A, Verhelst X, Cazzagon N, Perini L, Lammers WJ, de Vries AC, van der Meer AJP, van Buuren HR, Hansen BE

Journal of Hepatology 2019;71(5):992-999

ABSTRACT

Background & Aims: Recently the Amsterdam-Oxford model (AOM) was introduced; a prognostic model to assess the risk of death and/or liver transplantation (LT) in primary sclerosing cholangitis (PSC). We aimed to validate and assess the utility of the AOM.

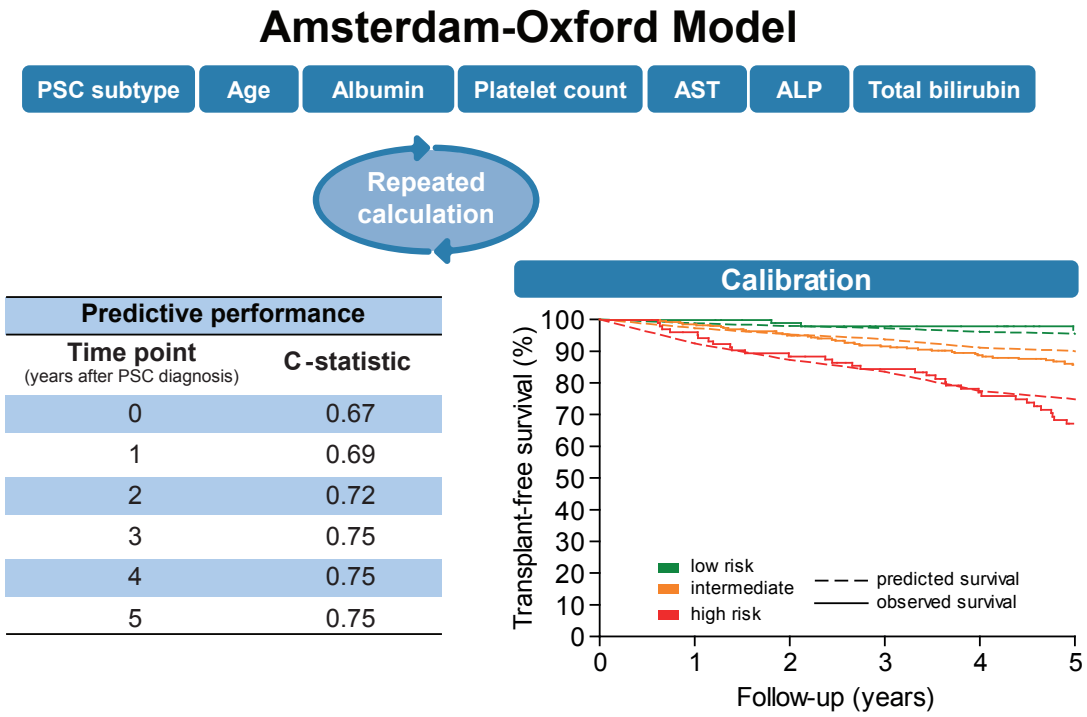
Methods: Clinical and laboratory data were collected from PSC diagnosis until the last visit or time of LT or death. The AOM was calculated at yearly intervals following PSC diagnosis. Discriminatory performance was assessed by calculation of the C-statistic and prediction accuracy by comparing the predicted survival with the observed survival in Kaplan-Meier estimates. A grid search was performed to identify the most discriminatory AOM threshold.

Results: A total of 534 PSC patients with a mean (SD) age of 39.2 (13.1) years were included. The diagnosis was large duct PSC in 466 (87%), PSC with features of autoimmune hepatitis in 52 (10%) and 16 (3%) had small-duct PSC. During the median (IQR) follow-up of 7.8 (4.0-12.6) years, 167 patients underwent LT and 65 died. The median LT-free survival was 13.2 (11.8-14.7) years. The C-statistic of the AOM ranged from 0.67 at baseline to 0.75 at 5 years of follow-up. The difference between the predicted and observed survival ranged from -1.6% at 1 year to +3.9% at 5 years of follow-up. Patients that developed AOM scores >2.0 were at significant risk of LT or death (time-dependent HR 4.09 95% CI 2.99-5.61).

Conclusions: In this large cohort of PSC patients, the AOM showed an adequate discriminative performance and good prediction accuracy at PSC diagnosis as well as during follow-up. This study further validates the AOM as a valuable risk stratification tool in PSC and extends its utility.

Lay summary
In our study we assessed whether the Amsterdam-Oxford model (AOM) is able to correctly estimate the risk of liver transplantation or death in patients with primary sclerosing cholangitis (PSC). This model uses seven objective and readily available variables to estimate prognosis for individual patients at time of PSC diagnosis. The AOM may aid in patient counselling and timing of diagnostic procedures or therapeutic interventions for complications of liver disease. In our study, we confirm the model works well at PSC diagnosis, but also when the AOM is recalculated at different time points during follow-up. This greatly improves the applicability of the model in clinical practice and for individual patients.

GRAPHICAL ABSTRACT



6

STUDY HIGHLIGHTS

What is known:
- Reliable estimates of survival are pivotal to optimize clinical management of PSC patients
- The Amsterdam-Oxford model (AOM) adequately estimates survival probabilities for PSC patients at diagnosis.

What is new here:
- The AOM also performs adequately when recalculated at other time points during the follow-up, thereby extending its utility in clinical practice

INTRODUCTION

Primary sclerosing cholangitis (PSC) is a chronic, variably progressive cholestatic liver disease characterized by inflammation of the intrahepatic and extrahepatic bile ducts, sclerosis and destruction of the biliary tract[1-4]. This leads to chronic cholestasis, biliary fibrosis and (decompensated) cirrhosis, which may eventually culminate into liver failure requiring liver transplantation; the only potential curative treatment for PSC[2,3]. Following a PSC diagnosis a median transplant-free survival of 13 years has been reported in studies from tertiary referral centers, although this may be longer in a population-based setting[5].

One of the major challenges in the management of PSC is the lack of therapies that halt disease progression. Despite the biochemical improvement reported with ursodeoxycholic acid (UDCA) treatment in PSC, a survival benefit has never been reported[6-11]. Another challenge concerns reliable estimation of prognosis in PSC, largely because of the heterogeneity in clinical course progression and the variety of outcomes ranging from end-stage liver disease to development of hepatobiliary and colorectal malignancies[12-14]. In this setting, risk prediction models that quantify the risk of future events for individual PSC patients are of critical importance for patient counselling, timely diagnostic procedures and subsequent therapeutic interventions for disease-related complications. Also, reliable risk stratification is important for the selection of patients in future drug development trials.

The Mayo risk score (MRS) is the most frequently used score to assess the short-term (4 years) mortality risk of PSC patients. However, this score was mainly derived from a cohort of patients with end-stage disease in a liver transplant centre. This may limit its applicability in early stages of disease[15]. Recently the Amsterdam-Oxford model (AOM) was introduced; a prognostic model developed in a population-based cohort to predict the long-term risk of PSC-related death and/or liver transplantation[16]. The AOM incorporates PSC subtype, age at PSC diagnosis, aspartate aminotransferase (AST), alkaline phosphatase (ALP), total bilirubin, albumin and platelet count. This score, based on these seven readily available variables, showed an adequate discriminative power and satisfactory calibration in a derivation and validation cohort. However, further validation of this population-based model is necessary to justify its application in other cohorts and centres and to extend its use at other time points during follow-up. In addition, the performance of this newly developed score has not been compared to the MRS. Therefore, we aimed to further validate the AOM and assess its utility in a large cohort of PSC patients from three tertiary centres in Europe. A secondary aim was to compare the performance of the AOM with that of the MRS.

PATIENTS AND METHODS

Population and study design

This retrospective cohort study included PSC patients from three tertiary centres in Europe: University of Padua, Italy, Ghent University Hospital, Belgium, Erasmus University Medical Center, Rotterdam, The Netherlands. Data were collected from 1984 up to June 2016 for University of Padua, from 1977 up to June 2016 for the Rotterdam University Medical Center, and from 1993 up to June 2018 for Ghent University Hospital. Complete follow-up was defined as liver transplantation or death or clinical follow-up beyond 1 January 2016 for Padua and Rotterdam, and beyond 1 January 2018 for Ghent. Patients who were diagnosed with PSC at age ≥ 18 years and in accordance with the European Association for the Study of the Liver guidelines were included[17]. Patients with follow-up less than six months with or without an event were excluded to ensure exclusion of patients that were referred because of liver failure and that were consequently diagnosed with PSC in the process of being waitlisted for LT. In addition, patients were excluded if the date of diagnosis was unknown or in the case of concomitant liver disease. Clinical and laboratory data were collected from start of follow-up until last visit or clinical event at 6 monthly or 1-yearly intervals according to per centre patient visit intervals. Biochemical parameters collected included AST, prothrombin time (PT), international normalized ratio (INR), alanine aminotransferase (ALT), ALP, gamma-glutamyl transpeptidase (γGT), total bilirubin, albumin and platelets. Clinical data included sex, age, date of PSC diagnosis, liver histology, UDCA treatment, concomitant inflammatory bowel disease (IBD), last follow-up date, or the date of clinical outcomes. Patients with a diagnosis of IBD within the first year following a diagnosis of PSC were considered to have IBD at baseline. The primary endpoint of the current study was a combined endpoint of liver transplantation or death. For patients with a diagnosis of untreatable CCA and unknown clinical outcome we considered the last of follow-up as the date of death. This study was conducted in accordance with the protocol and the principles of the Declaration of Helsinki. The protocol was approved by the Institutional Research Board of the corresponding centre, and at each participating centre, in accordance with local regulations.

Statistical analyses

Normally distributed data are presented as mean ± SD and skewed distributed data as median and interquartile range. Statistical analyses were performed with IBM SPSS Statistics 22.0 (SPSS Inc, Chicago, IL) and SAS 9.4 (SAS institute Inc., Cary, NC). To account for missing values SAS (SAS Proc MI, MCMC method) was used to generate 10 imputed datasets of laboratory results at yearly time points between PSC diagnosis up to 5 years of follow-up. Missing data were considered to be missing at random. Rubin's rules were used for estimation of the parameters and the standard error[18-20]. The imputation model included baseline variables that were potentially predictive for outcomes in PSC (e.g. year of diagnosis, age) as well as the outcomes themselves. Only continuous biochemical variables were imputed. All analyses were performed in the original database as well as in the imputed dataset.

Start of follow-up was set at PSC diagnosis defined by the first pathological imaging result (magnetic resonance cholangiography or endoscopic retrograde cholangiography) or liver biopsy. The AOM score was calculated at yearly intervals, starting at diagnosis up to 5 years of follow-up (laboratory values used ± 3 months around time point of calculation). The AOM score was calculated using the following formula: AOM score=0.323*PSC subtype

(1=large duct PSC; 0=small duct PSC) +0.018*Age at diagnosis – 2.485*log(albumin*lower limit of normal (LLN)) + 2.451* abs (log(platelets-0.5)+ 0.347*log(aspartate aminotransferase(AST)*upper limit of normal (ULN)) + 0.393 * log(alkaline phosphatase (ALP)*ULN)) + 0.337*log(total bilirubin*ULN)). For calculation of AOM values between 1 year and 5 years following a PSC diagnosis we used the actual age at the time of laboratory assessment instead of the age at diagnosis. The association of the AOM scores with the primary endpoint was assessed in Cox-regression analyses at diagnosis and at each year up to 5 years thereafter. Discriminatory performance of the model was assessed at various time points by calculation of the C-statistic. As a next validation step, we estimated the hazard ratio of the AOM score for the combined endpoint, to assess the fit of the model (i.e. whether a model overestimates or underestimates risk)[21]. In case the log hazard ratio is equal to 1 the model has a perfect fit. A lower value indicates the model underestimates risk, while higher values suggest an overestimation of risk. In addition, the fit / potential misspecification of the AOM in our cohort at PSC diagnosis was assessed by running a Cox regression analysis including the separate variables comprising the AOM as well as the AOM score itself[22]. In this regression model the coefficient of score of the AOM was constrained to equal 1 (i.e. offsetting the score of the AOM)[22]. If the β values of the separate variables of the model are not significantly different from 0 in these analyses, the AOM gives a perfect fit. If on the other hand a β is significant different from 0, there is misclassification and the AOM can be improved by adjusting the βs of these variables. Finally, prediction accuracy (i.e. calibration) was assessed by comparing the predicted versus the observed Kaplan-Meier survival curves to assess calibration[23]. To assess prediction accuracy across different AOM score intervals, we divided patients into three risk groups based on their AOM score, using threshold points at the 20th and 80th percentiles.

A repeated linear model with a random intercept and slope per patient using an unstructured covariance matrix was performed to analyse the evolution of AOM scores over time in those with and without an endpoint at the end of follow-up. To determine an AOM threshold with the highest power to discriminate patients achieving the primary endpoint from those not achieving the primary endpoint, we performed a grid search with calculation of C-statistic between an AOM score of 0.8 and 4.0 in steps of 0.1 at each year of follow-up. The optimal threshold was subsequently included in Cox proportional hazards analyses in order to estimate the strength of association with the liver transplantation-free survival, as a baseline variable and time-dependent variable separately.

The value of ALP alone in making absolute risk predictions of transplant-free survival was assessed. Cox-proportional hazard regression analyses were used to assess the association between baseline log alkaline phosphatase and time to event. From this Cox model the baseline linear prediction equation of ALP (prognostic index), along with the baseline survival estimate $S_0(t)$, t=time, were derived. The prediction accuracy of ALP was assessed by comparing the observed versus the predicted transplant-free survival rate, for the total cohort as well as for different percentiles of risk (<20th, 20th-80th, >80th). In addition, the independent prognostic impact of IBD-phenotype was assessed using a time-fixed as well as a time-dependent covariate in Cox regression analyses, stratified for centre, and adjusted for calendar year of PSC diagnosis and age at PSC diagnosis. Finally, to compare the AOM with the MRS, the MRS was calculated with the formula: 0.0295 * age in years + 0.5373 *

ln(total bilirubin in mg/dL) - 0.8389 *serum albumin in g/dL + 0.5380*ln(AST in IU/L) + 1.2426 * (points for variceal bleeding (0=not present or 1=present)).

RESULTS

Baseline cohort characteristics

Data were obtained from 601 patients with PSC of whom 534 patients met the inclusion criteria. A total of 48 patients were excluded because we were unable to obtain a date of diagnosis and 19 patients had a follow-up < 6 months (**Supplementary Figure 1**)." A total of 13344 patient visits and a mean of 25 visits per patient were reported across the entire cohort.

The mean (SD) age was 39.2 (13.1), 66% were male, and 93% were UDCA-treated. The baseline patient characteristics are summarized in **Table 1**. The diagnosis was large duct PSC in 466 (87%), 52 (10%) had PSC with features of autoimmune hepatitis (AIH) and 16 (3%) had small-duct PSC. The year of PSC diagnosis ranged from 1977 to 2017. At baseline, 268 (60%) patients had IBD: 77% had ulcerative colitis and 20% had Crohn's disease. In total, 427 (80%) patients had complete follow-up. During the median follow-up period of 7.8 years (interquartile range, 4.0-12.6 years) a total of 232 (43%) patients reached a clinical endpoint: liver transplantation was performed in 167 patients and 65 patients died. The transplant-free survival rates were 98.3% at 1 year, 84.4% at 5 years, and 65.9% at 10 years of follow-up, as shown in **Figure 1**. The median transplant-free survival was 13.2 (11.8-14.7) years.

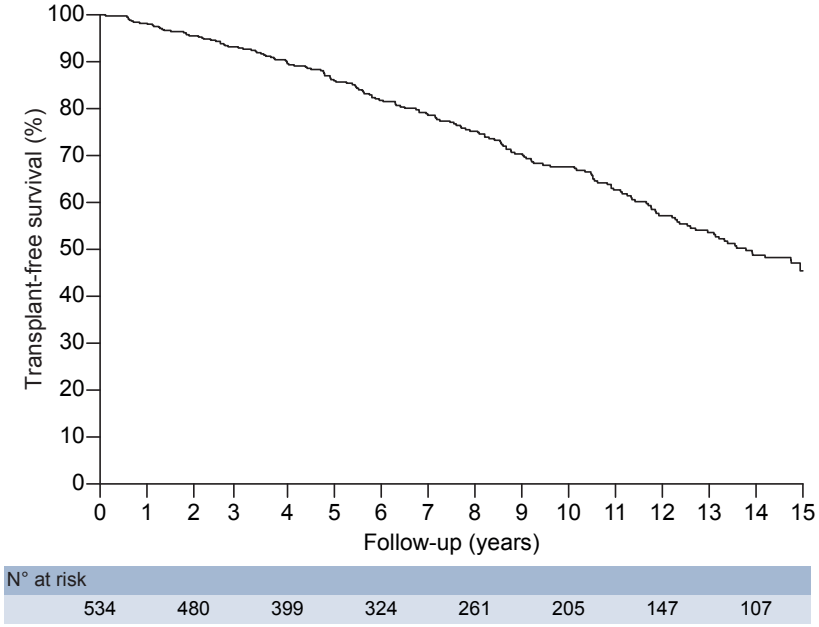


Figure 1. Kaplan-Meier estimate of transplant-free survival.

Table 1. Baseline patient characteristics

	Total cohort N=534
Age at diagnosis, y, mean (SD)	39.2 (13.1)
Male, n (%)	351 (65.7)
UDCA treated, n (%)	493 (92.3)
PSC type, n (%)	
Large duct PSC	466 (87.3)
PSC with features of AIH	52 (9.7)
Small-duct PSC	16 (2.6)
Year of diagnosis	2004 (1995-2009)
Year of diagnosis, range	1977-2017
IBD at baseline[a], n(%)	268 (60.2)
UC	206 (76.9)
CD	54 (20.1)
IBD-U	8 (3.0)
Indeterminate	0 (0)
Follow-up, years (IQR)	7.8 (4.0-12.6)
Laboratory data at diagnosis[b]	
Serum total bilirubin ×ULN	1.0 (0.52-2.30)
Serum ALP ×ULN	1.99 (1.11-3.59)
Serum γGT ×ULN	4.94 (2.11-9.94)
Serum AST ×ULN	1.79 (1.08-3.00)
Serum ALT ×ULN	2.13 (1.27-3.92)
Serum albumin ×LLN	1.17 (1.03-1.30)
Serum platelets $\times 10^3/mm^3$	258 (195-332)
Prediction model scores at diagnosis	
Amsterdam-Oxford model score	1.70 (1.30-2.17)
Mayo risk score[c]	-0.41 (-1.21-0.56)

[a] Unknown for 50 patients (9.4%).
[b] Laboratory data presented here from imputed data. Missing values in our cohort ranged from 26% to 30%.
[c] Data available for 498 patients. For 36 patients data on variceal bleeding was not available.
Abbreviations: UDCA, ursodeoxycholic acid; IBD, inflammatory bowel disease;UC, ulcerative colitis; CD, Crohn's disease; IBD-U, IBD unclassified; ULN, upper limit of normal; ALP, alkaline phosphatase; γGT, gamma-glutamyl transpeptidase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; LLN, lower limit of normal. Data presented as median (interquartile range) unless specified otherwise.

For the 65 patients who died in our cohort the cause of death was variceal bleeding in 5 patients, spontaneous bacterial peritonitis in 4, hepatorenal syndrome in 2, liver failure (unspecified) in 6, hepatocellular carcinoma in 2, signet ring cell carcinoma metastasized to the ductus choledochus in 1, pancreatic cancer in 1, colorectal cancer in 4, lung cancer in 2, renal insufficiency in 1, sepsis in 5, and surgical complications in 3. For 4 patients we could not determine the cause of death. A total of 25 patients died from CCA. Six of these patients had untreatable CCA and unknown clinical outcome. For these patients we considered the last of follow-up as the date of death.

Discriminatory performance of the Amsterdam-Oxford model and misspecification/fit
The overall discriminatory performance for death or liver transplantation of the AOM score at diagnosis calculated with C-statistic was 0.67 (95% confidence interval [CI] 0.64-0.70) and ranged to 0.75 (95% CI 0.71-0.78) at 5 years following diagnosis (**Table 2**). Assessment of the AOMs fit by univariable Cox-regression analyses, revealed the AOM had a good fit at baseline and at most time points during follow-up with a hazard ratio for clinical events ranging from 2.18 (95% CI 1.77-2.68) at diagnosis to 2.94 (95% CI 2.42-3.57) at 5 years of follow-up. Detailed assessment of the AOM fit/misspecification at PSC diagnosis indicated that only the βs for platelet count and age were significant predictors when offsetting the score of the AOM at equal 1. When adding these variables to the AOM score at baseline in the calculation of C-statistic, the C-statistic was 0.682 (95% CI 0.644-0.719).

Prediction accuracy (calibration) of the Amsterdam-Oxford model
For the total cohort, the difference between the calculated mean survival based on AOM values at PSC diagnosis and the observed Kaplan-Meier survival ranged -1.6% (96.7% predicted vs. 98.3% observed) at 1 year to 3.9 % at 5 years of follow-up (88.3% predicted vs. 84.4% observed). **Figure 2** shows further assessment of AOM calibration in different risk groups stratified by AOM score percentiles (<20th, 20-80th and > 80th percentile). Similar prediction accuracy was observed when the score was recalculated at 1 year, 3 years and 5 years after diagnosis for the 5 years following calculation, with the most accurate predictions being made in the lower-percentile and mid-percentile groups of the AOM score and underestimation of the risk of death or LT in the highest percentile (**Supplementary Figures 1A-1C**).

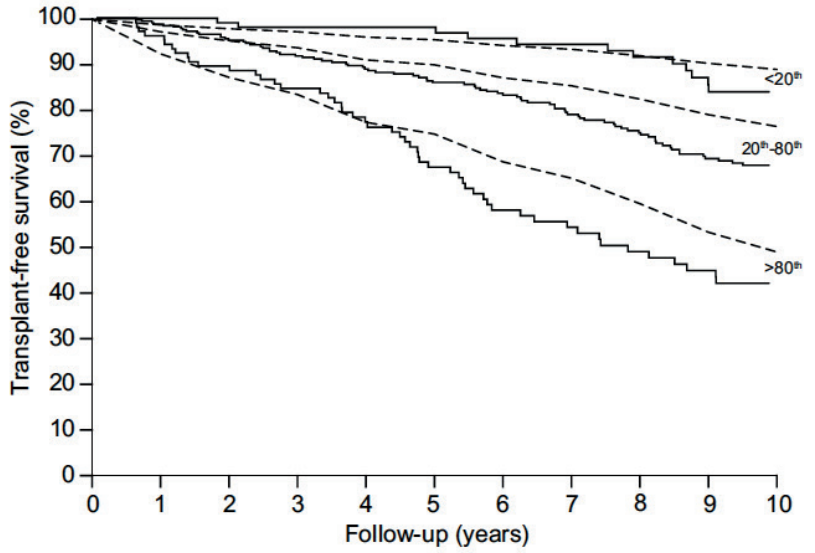
Table 2. Discriminative performance and assessment of the fit of the Amsterdam-Oxford model calculated n years after diagnosis

		Descriptives	Performance	Measure of fit
Year(s) after diagnosis	N	Median (IQR) AOM value	C-statistic (95% CI)	Hazard ratio (95% CI)[a]
0	534	1.70 (1.30-2.17)	0.6704 (0.6392-0.7015)	2.18 (1.77-2.68)
1	516	1.63 (1.24-2.17)	0.6927 (0.6615-0.7238)	2.84 (2.31-3.47)
2	480	1.67 (1.24-2.26)	0.7230 (0.6936-0.7524)	3.36 (2.75-4.11)
3	433	1.76 (1.29-2.41)	0.7484 (0.7195-0.7773)	3.70 (3.04-4.51)
4	399	1.76 (1.32-2.40)	0.7458 (0.7139-0.7777)	2.61 (2.21-3.09)
5	365	1.74 (1.31-2.44)	0.7461 (0.7114-0.7807)	2.94 (2.42-3.57)

Abbreviations: IQR, interquartile range; CI, confidence interval
[a] hazard ratios with confidence intervals that include exp(^1) (=2.72) indicate good fit. Hazard ratios greater and not including 2.72 in the confidence interval indicate over-estimation of risk by the model.

Utility of the Amsterdam-Oxford model and evolution over time
Visualization of AOM scores over time by a repeated linear model revealed that patients who were alive at the end of follow-up had consistent low AOM scores during follow-up (**Figure 3**). A subsequent grid search of AOM scores based on the C-statistic revealed that the most discriminatory threshold for death or liver transplantation ranged between 1.8

and 2.1, if calculated at baseline and during the first 5 years of follow-up (C-statistic 0.61-0.69; **Figure 4** and **Supplementary Figures 2a-2e**). In Cox-regression analyses AOM scores above 2.0 were significantly associated with clinical events (hazard ratio ranging between 2.30 (95% CI 1.75 – 2.95) at diagnosis and 4.46 (95% CI 3.19 – 6.24) at 5 years following diagnosis, **Figure 4**). At baseline a total of 174 (32.6%) patients had AOM values above 2.0. During the first 5 years following PSC diagnosis an additional 8.4% within 1 year, 13.9% within 3 years, and 25.4% within 5 years developed AOM values above the threshold of 2.0 during. Patients that reached an AOM score of 2.0 in the first 5-years of follow-up patients were at significant risk of death or liver transplantation (time-dependent HR 4.09 95% CI 2.99-5.61).



N° at risk/events						
<20th	107/0	101/1	88/2	72/4	65/7	49/12
20th-80th	321/0	288/15	240/33	206/33	161/66	128/82
>80th	106/0	91/12	71/23	46/40	35/47	28/52

Figure 2. Predicted versus observed liver transplant-free survival according to Amsterdam-Oxford model score percentiles. Figure shows prediction accuracy (calibration) of the Amsterdam-Oxford model score up to 10 years of follow-up across different percentiles of the scores (divided into 3 groups based on 20th and 80th percentile) at diagnosis.
Solid lines = actual observed transplant-free survival probabilities estimated by Kaplan-Meier analyses.
Dashed lines = the predicted mean transplant-free survival probabilities as predicted by the Amsterdam-Oxford model.

ALP in isolation as a predictor of transplant-free survival and the impact of IBD phenotype

In the first 5 years following a PSC diagnosis, the c-statistic for ALP alone as a predictor of transplant-free survival ranged between 0.52 and 0.63 (**Supplementary Table 1**). For the total cohort the difference between observed and predicted survival increased from -0.2% at 1 year after PSC diagnosis to -4.9% at 10 years. Further assessment of ALP calibration in different risk groups stratified according to percentiles (<20th, 20-80th and > 80th percentile), revealed the difference between observed and predicted survival between 1 year and 10 years following a PSC diagnosis ranged between -1.8% and -7.6% for the <20th percentile, between -0.5% and -2.9% for the 20th-80th percentile, and between -0.1% and -12.4% for the >20th percentile (**Supplementary Figure 3**). IBD phenotype and IBD development were not significantly associated with transplant-free survival (**Supplementary Table 2** and **Supplementary Table 3**).

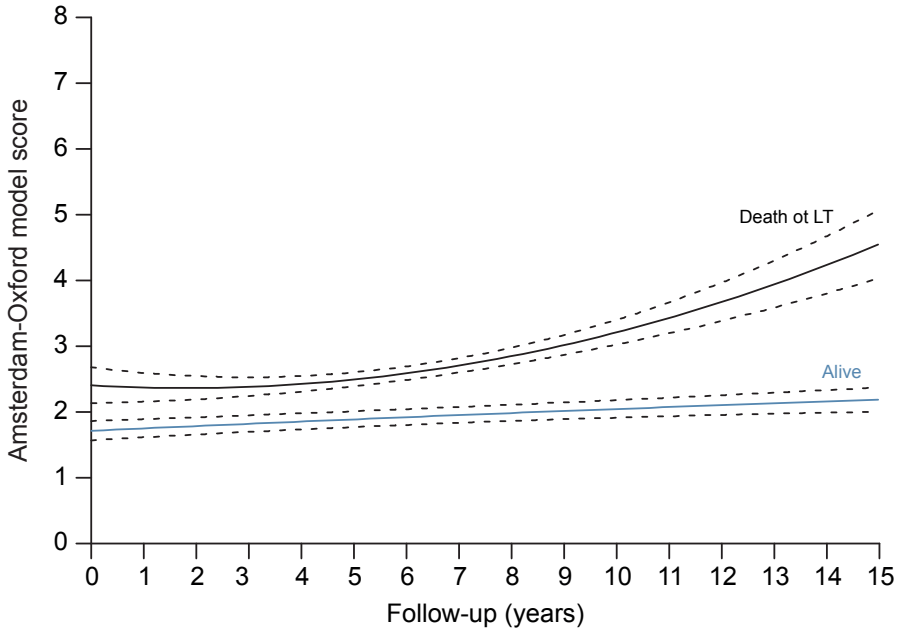


Figure 3. Evolution of AOM scores during follow-up stratified according to endpoint. Figure shows Amsterdam-Oxford model scores over time in those with a clinical event (death or liver transplantation) and those without a clinical event at the end of follow-up from a repeated linear model with a random intercept and slope per patient in an unstructured covariance matrix.
Abbreviations: LT, liver transplantation.
Solid line=predicted mean of the AOM score; dashed lines=95% confidence intervals.

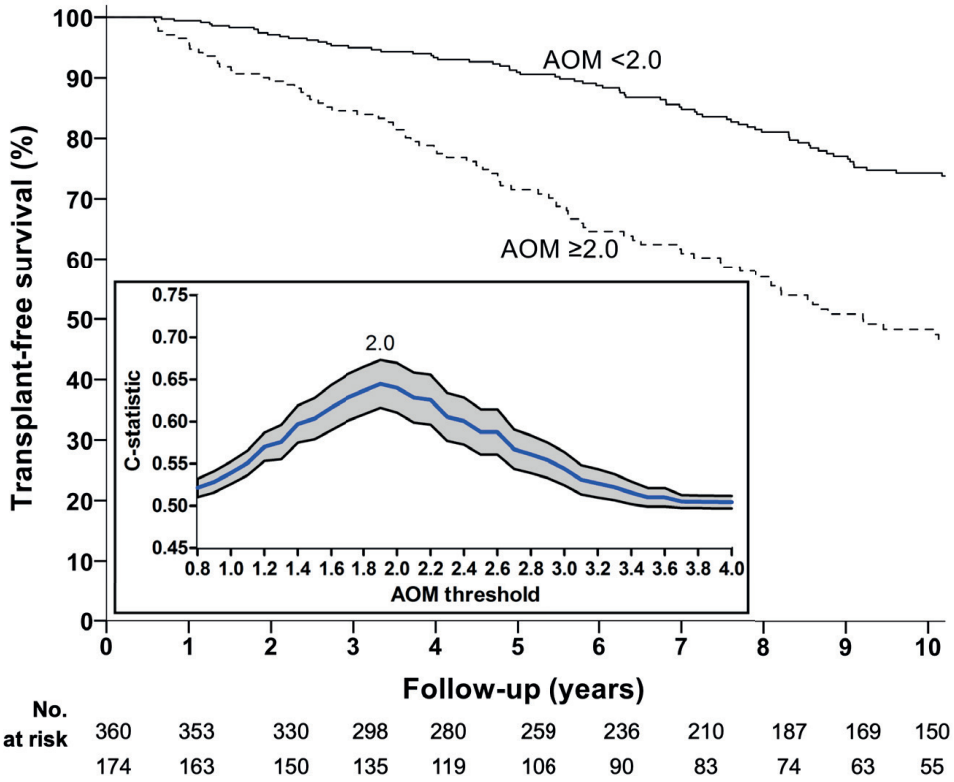


Figure 4. Kaplan Meier estimate of transplant-free survival stratified according to an Amsterdam-Oxford model threshold of 2.0 at diagnosis. Kaplan Meier estimate of transplant-free survival in patients with Amsterdam-Oxford model (AOM) scores below 2.0 (solid line) and scores equal to or above 2.0 (dashed line) at diagnosis. Panel shows C-statistics for the grid search for a threshold in AOM score between 0.8 and 4.0 in steps of 0.1 at diagnosis.

Comparison of the Mayo risk score and Amsterdam-Oxford model

The MRS could be calculated for a sub-cohort of 498 of patients at baseline. At diagnosis a total of 311 (62.4%) of the patients had a MRS value below or equal to 0 (low-risk group); 161 (32.3%) had scores above 0 but less than 2 ('intermediate risk group') and 26 (5.3%) were considered at high risk of events (MRS greater than 2). The transplant-free survival rates were significantly different between the low, intermediate and high-risk group: 99.4%, 98.1% and 92.3% at 1 year, 95.5%, 88.4% and 65.4% at 3 years, and 91.5%, 77.3% and 47.4% at 5 years of follow-up (log-rank<0.001, **Figure 5**). The discriminatory performance of the MRS calculated by C-statistic ranged from 0.73 (95% CI 0.73-0.76) at diagnosis to 0.79 (95% CI 0.76-0.82) at 5 years following PSC diagnosis. Direct comparison of discriminatory performance in patients for whom both the MRS and AOM score could be calculated at PSC diagnosis (n=498) and 1 year following diagnosis (n=482) showed higher C-statistics for the MRS than the AOM score (0.73 vs. 0.68 at diagnosis and 0.75 vs. 0.70 at 1 year following diagnosis, respectively; **Table 3**).

In terms of prediction accuracy (calibration) the MRS overestimated transplant-free survival with a difference between the calculated mean survival based on MRS scores and actual survival as observed by Kaplan Meier estimates of 5.1% at 1 year, 6.9% at 2 years, 8.9% at 3 years, and 9.6% at 4 years of follow-up. Detailed analyses of prediction accuracy in different risk-groups revealed that the difference between predicted and observed survival was most pronounced in the high-risk group (**Figure 5**).

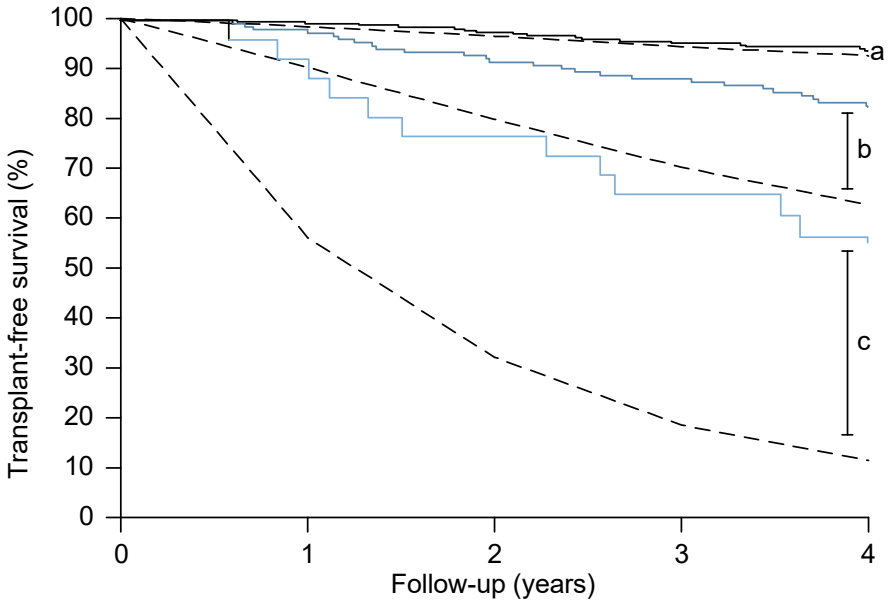


Figure 5. Predicted versus observed liver transplant-free survival according to Mayo risk score risk group. Figure shows prediction accuracy of the Mayo risk score (MRS) by comparing predicted survival based on the MRS and observed (actual) survival by Kaplan-Meier estimates at PSC diagnosis according to MRS risk groups. Solid line: observed survival by Kaplan-Meier estimates. Dashed line: predicted survival by the MRS.
[a]low-risk group (MRS value below or equal to 0); [b]intermediate risk group (scores above 0 but less than 2); [c]high risk group (MRS greater than 2).

Table 3. Direct comparison of the discriminatory performance of the Amsterdam-Oxford model and Mayo Risk Score calculated n years after diagnosis

Year(s) after diagnosis	N	Amsterdam-Oxford model C-statistic (95% CI)	Mayo Risk Score C-statistic (95% CI)
0	498	0.6816 (0.6482-0.7149)	0.7309 (0.7005-0.7613)
1	482	0.7008 (0.6685-0.7332)	0.7507 (0.7200-0.7813)
2	446	0.7334 (0.7027-0.7641)	0.7696 (0.7391-0.8001)
3	402	0.7589 (0.7291-0.7887)	0.7800 (0.7501-0.8099)
4	370	0.7592 (0.7257-0.7927)	0.7709 (0.7389-0.8029)
5	337	0.7566 (0.7199-0.7933)	0.7902 (0.7567-0.8237)

Abbreviations: IQR, interquartile range; CI, confidence interval

DISCUSSION

This long-term study of a well-characterized large cohort of PSC patients allowed us to assess the performance and utility of the AOM, a prognostic model to estimate survival for PSC patients. We confirm that the AOM has adequate discriminatory performance and satisfactory prediction accuracy when applied at PSC diagnosis. In addition, we show that the performance and accurate 5-year prediction of the AOM score remained if recalculated at different time points during follow-up, thereby extending its utility in daily clinical practice. If dichotomized, an AOM score of 2.0 had the highest discriminative power to risk stratify patients during follow-up. In patients who remained alive without liver transplantation during the observation period in our study the AOM stayed, on average, below this threshold. Although the MRS had a higher C-statistic than the AOM, the utility of this traditional prognostic tool for the individual patient may be limited due to the suboptimal prediction accuracy. The results confirm that the AOM is a valuable prognostic tool in PSC patients.

While our study was performed in tertiary referral centres, the baseline characteristics in our cohort are similar to those presented in the original AOM derivation and validation study[16]. In this population-based cohort, including patients from general hospitals and academic centres without transplant facilities, De Vries et al. reported 10-year transplant-free survival rates between 75%-80% and a median survival of 22 years. In comparison, in our cohort the 10-year transplant-free survival was 66% and the median survival was 13 years. The lower event-rate in the AOM development study may explain the AOMs slight underestimation of clinical events beyond 5 years of follow-up in the total cohort as well as in the subgroup with intermediate AOM scores (20^{th}-80^{th} percentile). As expected, the difference between predicted and observed survival was most pronounced in the subgroup of patients with highest AOM scores (>80^{th} percentile). Nonetheless, the C-statistic we found is comparable to that reported in the AOM development study (0.68).

In terms of discriminative performance, the MRS outperformed the AOM in our cohort, meaning that with the MRS more patients who experienced an event had a higher risk score and more patients without an event had a lower risk score than with the use of the AOM. However, prediction accuracy of the MRS was unsatisfactory as the score substantially overestimated the risk of liver transplantation or death. A potential explanation for the limited prediction accuracy may be the indirect clinical endpoint that was used in the development cohort of the MRS. Patients who underwent liver transplantation were considered to have died with time to death based on the expected survival in absence of transplantation. This is probably no longer accurate today. An alternative explanation may be the use of time of referral in the derivation of the MRS, which may be long after the date of diagnosis as used in our cohort. Also, the MRS was developed in specific expert centres in which it is likely that patients present with more advanced disease, especially at the time of referral. The higher event-rate in the MRS development cohort (the 5-year transplant-free survival was 65% in the MRS development cohort vs. 84% in our cohort) could thus contribute to the overestimation of events by the MRS in the current study, as well as in many centres managing PSC.

As the AOM and MRS have different prognostic qualities, it is difficult to determine which score should be used in clinical practice and how and when a score should be applied. From a clinical point of view, it is difficult to derive certainty from a single estimate of risk, especially in PSC. Rather clinicians as well as patients are likely to want to re-evaluate the risk of adverse events during follow-up. As such, accurate short-term or intermediate-term calculations of risk may be better suited. We show the AOM can be used to make such repeated estimates for patients in different categories of risk. With the inclusion of variceal bleeding -a direct clinical complication of end stage liver disease- the MRS has a high discriminative performance across different cohorts. With a current 30-day mortality of 15-20%, variceal bleeding is a strong predictor of death[24,25]. In our cohort, the MRS provides a more discriminative short-term mortality risk assessment as opposed to the AOM, albeit less accurate. This score may thus be more appropriate to estimate whether patients are at high-risk of progressive disease necessitating LT on a group level. In daily practice, however, clinicians may prefer the accurate estimates of prognosis as provided by the AOM at various time points in order to optimize the management of individual patients.

Important to consider is that the AOM was developed in an early disease stage population and that it includes ALP. ALP elevates early during the course of disease[26 27], has often been used in drug development trials as a primary endpoint, and lower ALP levels have been correlated with a favourable course of disease[28-33]. In fact, a previous study in 366 PSC patients showed that a more simplistic approach using ALP in isolation had a near identical c-statistic to that of the AOM, challenging the necessity of a complicated risk prediction model[33]. However, absolute predictions of transplant-free survival in our cohort using ALP in isolation, revealed that using ALP alone may grossly overestimate survival in different risk groups at diagnosis and when reapplied during follow-up. Moreover, the c-statistic varied between 0.52 and 0.63 during follow-up, which is clearly lower than that observed for the AOM. Therefore, the use of a prognostic score that includes more biochemical variables than ALP, especially when reapplied during follow-up and/or in a tertiary cohort, seems more appropriate.

A strength of our study is the inclusion of a well-characterized study population from multiple centres with complete follow-up in 80% of the patients. Second, we assessed a combined endpoint of LT and all-cause mortality rather than PSC-related mortality specifically. In clinical practice it is difficult to distinguish true liver-related or PSC-related death from other causes of death. In addition, one may argue that the separation between the two has limited relevance for patients and clinicians. Inherit to the retrospective nature of our study some laboratory values were missing. To overcome this problem multiple imputation techniques were used[19]. Importantly, analyses in our raw dataset revealed similar results (data not shown). While our results further validate and justify the clinical use of the AOM, the validation mainly pertains tertiary centre Caucasian patients. Further validation in a population-based setting, in particular in other ethnicities or countries, is thus warranted. Finally, the threshold of 2.0 found for the AOM should be interpreted with caution. The use of thresholds in clinical practice is widespread, but has several limitations. Categorization results in loss of predictive information and thresholds are difficult to generalize to other populations[34,35]. This study shows that with dichotomous application

of the AOM, patients may switch in risk category during follow-up. This is actually a limitation of dichotomous criteria/risk group thresholds in general. However, the AOM was not derived to be used as a dichotomous score but rather as a continuous measure of risk. The analyses in our manuscript therefore only visualize that two risk groups can be readily obtained by using the threshold of 2.0, akin to the thresholds used in the MRS.

In general, as models with a C-statistic > 0.8 are considered good prognostic models, further optimization of the risk stratification in PSC remains warranted[22]. In PSC, utility of prediction models is hampered due to heterogeneity in disease progression and outcomes as well as the lack of effective therapies. Still, reliable estimates of survival are important for patient counselling, optimization of follow-up regimens, and selection and timing of listing for LT. Once effective therapies become available, repositioning of prediction models in clinical management of PSC patients may be necessary. Cox regression analyses, with the variables comprising the AOM while offsetting the AOM score (i.e. keeping its value to equal 1), revealed a satisfactory fit of the model at PSC diagnosis. Only the βs for age and platelet count were suboptimal for our cohort, but adjustments of these betas only yielded a minor increase in terms of C-statistic (0.67 vs. 0.68). Therefore, other approaches, with for example, the addition of measures of liver fibrosis could be of value. A potential important addition to existing risk stratification models in PSC may be the inclusion of liver stiffness measurement (LSM). LSM reflects severity of fibrosis and absolute LSM values as well as longitudinal changes are strongly linked with clinical events in PSC[36]. Enhanced liver fibrosis (ELF) score, a serological measure of liver fibrosis consisting of a combination of serum concentrations of hyaluronic acid, procollagen III peptide and tissue inhibitor of metalloproteinase 1, could provide another addition to existing risk prediction models and has been shown to correlate well with LSM values and to have incremental prognostic utility to the Mayo risk[37]. In order to better assess the additive value of such measures of liver fibrosis to existing risk prediction models, large prospective multicentre collaborations with extensive datasets are necessary. New statistical techniques could further aid in the derivation of more accurate models as well. A recently introduced model by Eaton et al. using a different statistical approach with machine learning showed an impressive accuracy (C-statistic >0.9) for the prediction of endpoints[38]. This score is, however, limited by the prediction of hepatic decompensation rather than solid clinical endpoints such as LT, death or CCA. Furthermore, further validation of this score as well as the statistical technique should be awaited. Still, PSC remains a disease with a highly variable and scattered course of disease that may not be easily captured by static prognostic scores.

In conclusion, we confirm the AOM has adequate discriminatory performance and good prediction accuracy for LT-free survival, both at PSC diagnosis and other time-points during the course of disease. Hereby we extend the validity and utility of the AOM as a prognostic tool in PSC.

REFERENCES

1. Karlsen TH, Folseraas T, Thorburn D, Vesterhus M. Primary sclerosing cholangitis - a comprehensive review. *J Hepatol.* 2017;67(6):1298-1323.
2. Chapman RW, Arborgh BA, Rhodes JM, et al. Primary sclerosing cholangitis: a review of its clinical features, cholangiography, and hepatic histology. *Gut.* 1980;21(10):870-877.
3. Hirschfield GM, Karlsen TH, Lindor KD, Adams DH. Primary sclerosing cholangitis. *Lancet.* 2013;382(9904):1587-1599.
4. Williamson KD, Chapman RW. Primary sclerosing cholangitis. *Dig Dis.* 2014;32(4):438-445.
5. Boonstra K, Beuers U, Ponsioen CY. Epidemiology of primary sclerosing cholangitis and primary biliary cirrhosis: a systematic review. *J Hepatol.* 2012;56(5):1181-1188.
6. Lindor KD. Ursodiol for primary sclerosing cholangitis. Mayo Primary Sclerosing Cholangitis-Ursodeoxycholic Acid Study Group. *N Engl J Med.* 1997;336(10):691-695.
7. Lindor KD, Kowdley KV, Luketic VA, et al. High-dose ursodeoxycholic acid for the treatment of primary sclerosing cholangitis. *Hepatology.* 2009;50(3):808-814.
8. Beuers U, Spengler U, Kruis W, et al. Ursodeoxycholic acid for treatment of primary sclerosing cholangitis: a placebo-controlled trial. *Hepatology.* 1992;16(3):707-714.
9. Stiehl A, Walker S, Stiehl L, Rudolph G, Hofmann WJ, Theilmann L. Effect of ursodeoxycholic acid on liver and bile duct disease in primary sclerosing cholangitis. A 3-year pilot study with a placebo-controlled study period. *J Hepatol.* 1994;20(1):57-64.
10. Chapman RW. Medical treatment of primary sclerosing cholangitis with ursodeoxycholic acid. *Dig Liver Dis.* 2003;35(5):306-308.
11. 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(5):1464-1472.
12. Bergquist A, Ekbom A, Olsson R, et al. Hepatic and extrahepatic malignancies in primary sclerosing cholangitis. *J Hepatol.* 2002;36(3):321-327.
13. Claessen MM, Vleggaar FP, Tytgat KM, Siersema PD, van Buuren HR. High lifetime risk of cancer in primary sclerosing cholangitis. *J Hepatol.* 2009;50(1):158-164.
14. Boonstra K, Weersma RK, van Erpecum KJ, et al. Population-based epidemiology, malignancy risk, and outcome of primary sclerosing cholangitis. *Hepatology.* 2013;58(6):2045-2055.
15. Kim WR, Therneau TM, Wiesner RH, et al. A revised natural history model for primary sclerosing cholangitis. *Mayo Clin Proc.* 2000;75(7):688-694.
16. de Vries EM, Wang J, Williamson KD, et al. A novel prognostic model for transplant-free survival in primary sclerosing cholangitis. *Gut.* 2017.
17. European Association for the Study of the L. EASL Clinical Practice Guidelines: management of cholestatic liver diseases. *J Hepatol.* 2009;51(2):237-267.
18. Little RJA, Rubin DB. *Statistical analysis with missing data.* New York: John Wiley & Sons; 1987.
19. Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ.* 2009;338:b2393.
20. Rubin D. Multiple imputation aft er 18+ years. *J Am Stat Assoc.* 1996;91: 473 – 489.
21. van Houwelingen HC. Validation, calibration, revision and combination of prognostic survival models. *Stat Med.* 2000;19(24):3401-3415.
22. Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. *BMC Med Res Methodol.* 2013;13:33.

23. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology.* 2010;21(1):128-138.
24. Graham DY, Smith JL. The course of patients after variceal hemorrhage. *Gastroenterology.* 1981;80(4):800-809.
25. D'Amico G, De Franchis R. Upper digestive bleeding in cirrhosis. Post-therapeutic outcome and prognostic indicators. *Hepatology.* 2003;38(3):599-612.
26. Brensilver HL, Kaplan MM. Significance of elevated liver alkaline phosphatase in serum. *Gastroenterology.* 1975;68(6):1556-1562.
27. Kaplan MM, Righetti A. Induction of rat liver alkaline phosphatase: the mechanism of the serum elevation in bile duct obstruction. *J Clin Invest.* 1970;49(3):508-516.
28. Stanich PP, Bjornsson E, Gossard AA, Enders F, Jorgensen R, Lindor KD. Alkaline phosphatase normalization is associated with better prognosis in primary sclerosing cholangitis. *Dig Liver Dis.* 2011;43(4):309-313.
29. Al Mamari S, Djordjevic J, Halliday JS, Chapman RW. 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(2):329-334.
30. Lindstrom L, Hultcrantz R, Boberg KM, Friis-Liby I, Bergquist A. Association between reduced levels of alkaline phosphatase and survival times of patients with primary sclerosing cholangitis. *Clin Gastroenterol Hepatol.* 2013;11(7):841-846.
31. Ponsioen CY, Chapman RW, Chazouilleres O, et al. Surrogate endpoints for clinical trials in primary sclerosing cholangitis: Review and results from an International PSC Study Group consensus process. *Hepatology.* 2016;63(4):1357-1367.
32. Rupp C, Rossler 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;40(11-12):1292-1301.
33. de Vries EM, Wang J, Leeflang MM, et al. Alkaline Phosphatase at Diagnosis of Primary Sclerosing Cholangitis and One Year Later: Evaluation of Prognostic Value. *Liver Int.* 2016.
34. Altman DG, Lausen B, Sauerbrei W, Schumacher M. Dangers of using "optimal" cutpoints in the evaluation of prognostic factors. *J Natl Cancer Inst.* 1994;86(11):829-835.
35. Buettner P, Garbe C, Guggenmoos-Holzmann I. Problems in defining cutoff points of continuous prognostic factors: example of tumor thickness in primary cutaneous melanoma. *J Clin Epidemiol.* 1997;50(11):1201-1210.
36. 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(4):970-979; quiz e915-976.
37. Vesterhus M, Hov JR, Holm A, et al. Enhanced liver fibrosis score predicts transplant-free survival in primary sclerosing cholangitis. *Hepatology.* 2015;62(1):188-197.
38. Eaton JE, Vesterhus M, McCauley BM, et al. Primary Sclerosing Cholangitis Risk Estimate Tool (PREsTo) Predicts Outcomes of the Disease: A Derivation and Validation Study Using Machine Learning. *Hepatology.* 2018.

SUPPLEMENTARY DATA

Supplementary Table 1. Discriminative performance of serum alkaline phosphatase in isolation calculated n years after diagnosis

Year(s) after diagnosis	N	ALP C-statistic (95% CI)
0	534	0.6038 (0.5789 0.6287)
1	516	0.6315 (0.6054 0.6575)
2	480	0.5918 (0.5719 0.6117)
3	433	0.5872 (0.5670 0.6074)
4	399	0.6126 (0.5902 0.6351)
5	365	0.5293 (0.5215 0.5371)

Supplementary Table 2. Time fixed analyses of the impact of IBD phenotype at baseline.

	HR[a]	95% CI	p
Crohn's disease vs ulcerative colitis	0.71	0.43-1.19	0.164
IBD-U vs ulcerative colitis	1.06	0.33-3.41	0.919
No IBD vs ulcerative colitis	1.16	0.36-3.73	0.803
Crohn's disease vs no IBD	0.94	0.29-3.02	0.919
IBD-U vs no IBD	0.67	0.19-2.32	0.529
IBD-U vs Crohn's disease	1.49	0.43-5.15	0.529

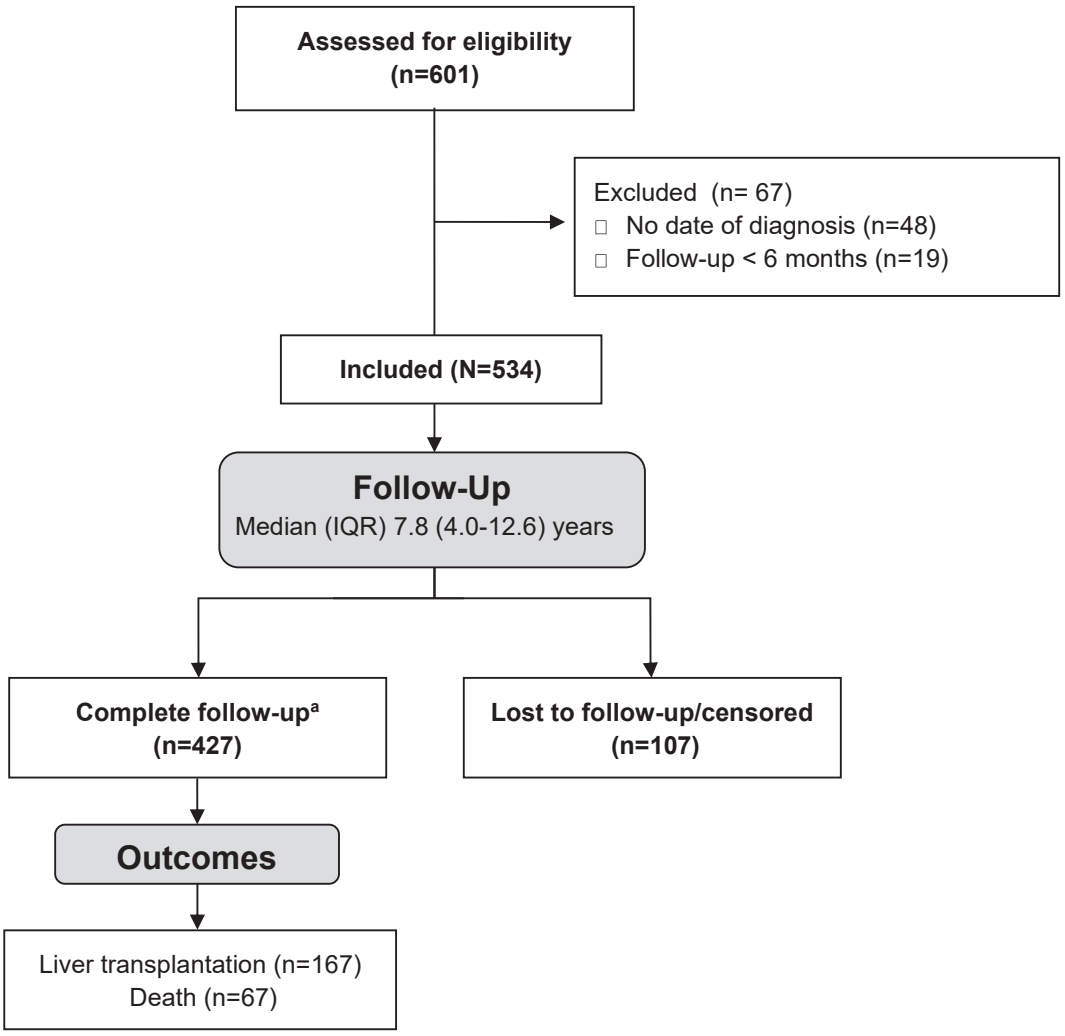
[a]Time fixed hazard ratios based on IBD phenotype at PSC diagnosis stratified by geographic region of diagnosis; adjusted for calendar year and age at diagnosis.

Supplementary Table 3. Analyses of inflammatory bowel disease phenotype defined as a time-dependent covariate.

	HR[a]	95% CI	p
Crohn's disease vs ulcerative colitis	0.76	0.49-1.20	0.250
Indeterminate vs ulcerative colitis	0.78	0.29-2.15	0.637
No IBD vs ulcerative colitis	0.91	0.68-1.23	0.556
Crohn's disease vs no IBD	0.70	0.44-1.12	0.135
IBD-U/indeterminate colitis vs no IBD	1.40	0.51-3.85	0.521
IBD-U/indeterminate Crohn's disease	1.28	0.46-3.50	0.637

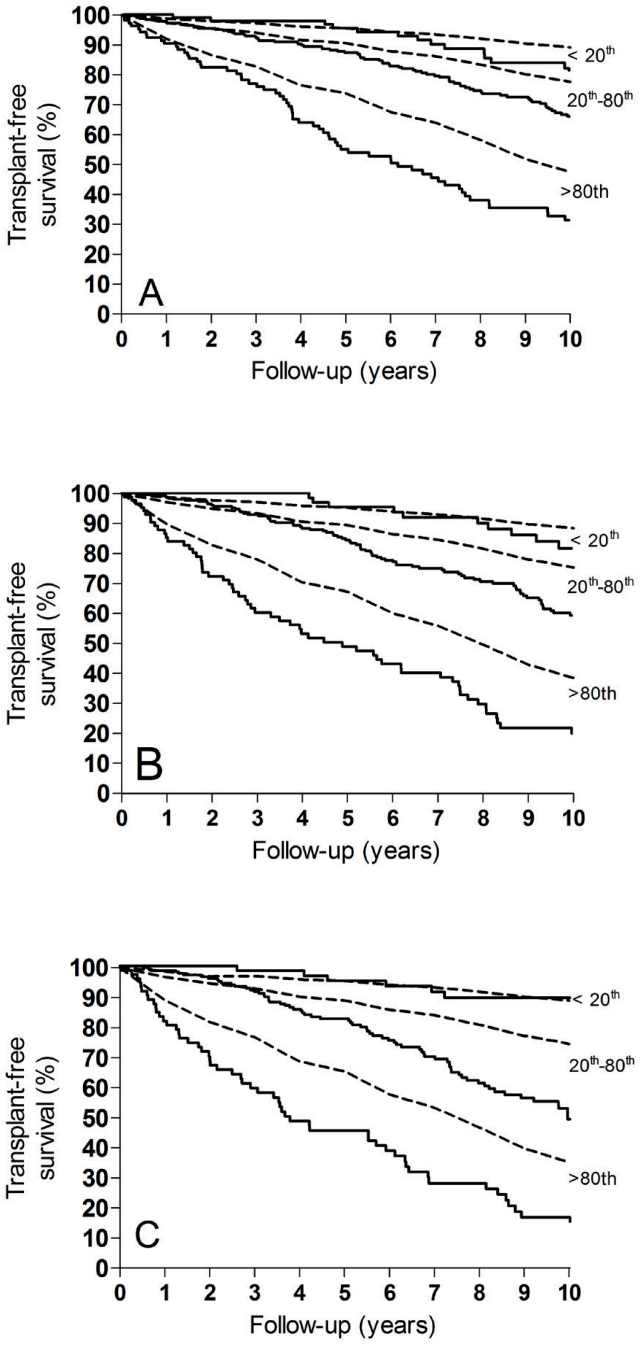
[a]Time-dependent hazard ratios based on the development of IBD phenotype during follow-up. All analyses were stratified by geographic region of diagnosis; adjusted for calendar year and age at diagnosis.

Supplementary Figure 1. Flow diagram of inclusion



[a]Defined as liver transplantation or death or clinical follow-up beyond 1 January 2016 for Padua and Rotterdam, and beyond 1 January 2018 for Ghent.

Supplementary Figure 2A-C. Predicted versus observed liver transplant-free survival according to Amsterdam-Oxford model score percentiles at various time points



Supplementary figure 2A-C shows calibration of the Amsterdam-Oxford model when re-applied at 1 year (A) and 3 years (B), and 5 years (C) after a PSC diagnosis for the subsequent 5 years of follow-up after the score was calculated. The predicted and observed survival curves are divided into different risk groups based on AOM score percentiles (<20th, 20-80th and > 20th percentile)
Solid lines = actual observed transplant-free survival probabilities estimated by Kaplan-Meier analyses. Dashed lines = the predicted mean transplant-free survival probabilities as predicted by the Amsterdam-Oxford model.

Supplementary Figure 3. Predicted versus observed liver transplant-free survival according to alkaline phosphatase levels at PSC diagnosis in different score percentiles

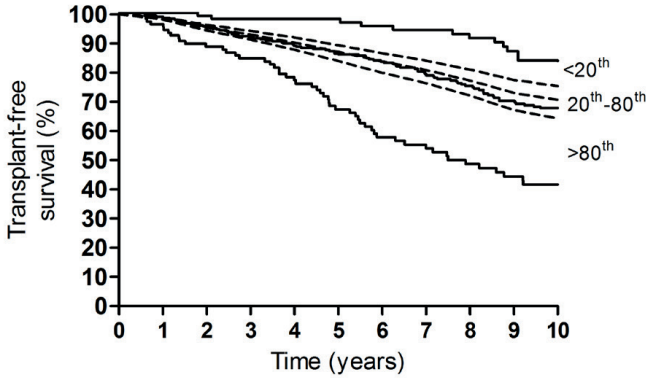
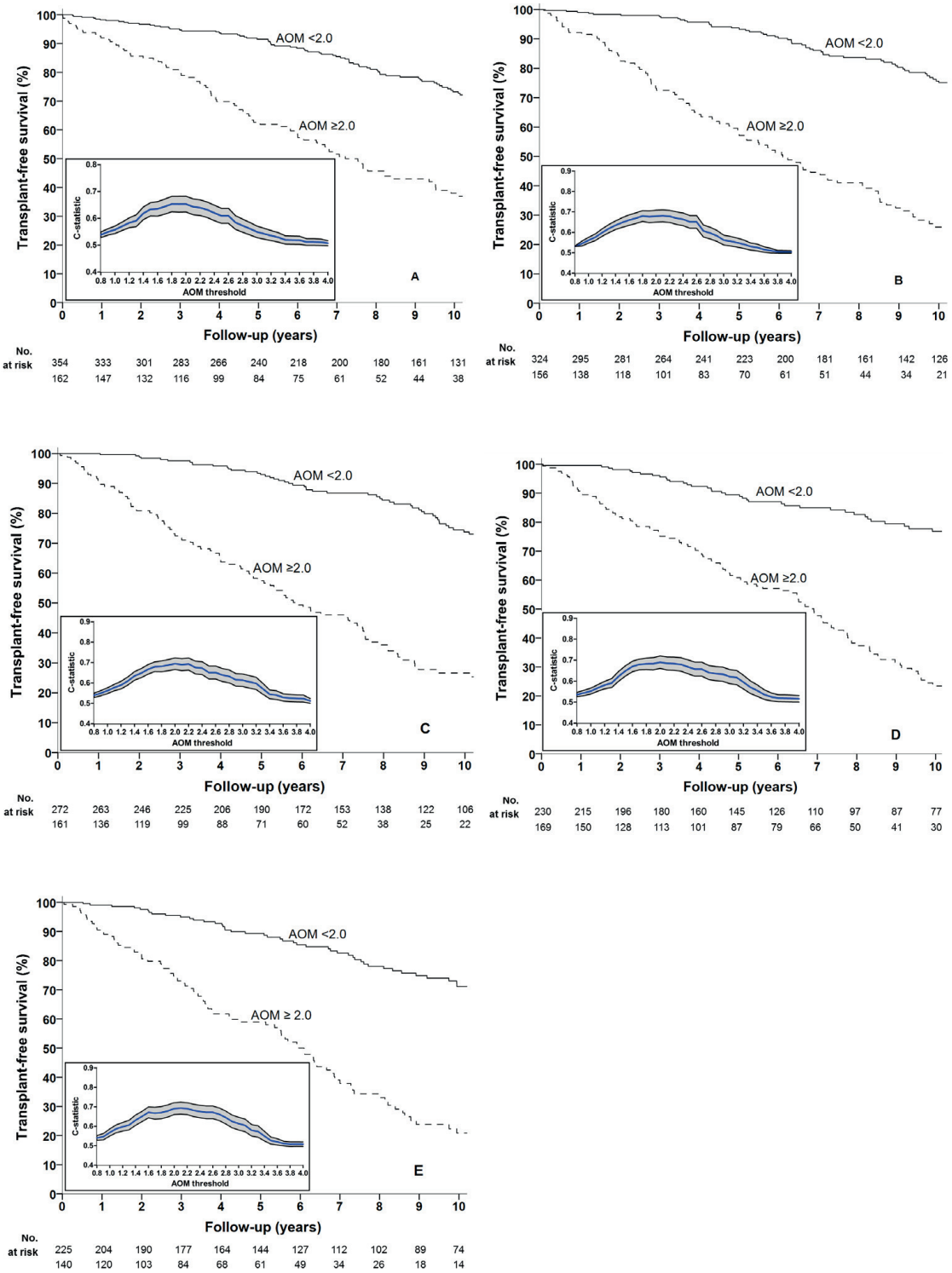


Figure shows prediction accuracy (calibration) of the alkaline phosphatase in isolation up to 10 years of follow-up across different percentiles of risk (divided into 3 groups based on 20th and 80th percentile) at diagnosis. For ALP the following linear prediction equation was obtained: 0.456825 x log(ALP x upper limit of normal). The baseline survival curve at the mean prognostic index of ALP S0(t) was 0.8861 at 5 years and 0.7367 at 10-year follow-up. The survival S(t) for any given patients was then calculated by $S(t) = S_0(t)^{\exp(\text{alkaline prognostic index})}$. As an example, for a patient with an ALP level of 2.46x the upper limit of normal, the estimated transplant-free survival would then be $0.8861^{\exp(0.456825 \times \log(2.46))} = 86.5\%$ at 5 years, and $0.7367^{\exp(0.456825 \times \log(2.46)} = 69.4\%$ at 10 years of follow-up.
Solid lines = actual observed transplant-free survival probabilities estimated by Kaplan-Meier analyses. Dashed lines = the predicted mean transplant-free survival probabilities as predicted by the Amsterdam-Oxford model.

Supplementary Figure 4.



Kaplan Meier estimate of transplant-free survival stratified according to an Amsterdam-Oxford model (AOM) threshold of 2.0 at 1 year (A), 2 years (B), 3 years (C), 4 years (D) and 5 years (E) following diagnosis. Panels show C-statistics for the grid search of AOM thresholds for AOM values of 0.8 and 4.0 in steps of 0.1 at different time points.

CHAPTER 7

7

CURRENT POLICY FOR ALLOCATION OF DONOR LIVERS IN THE NETHERLANDS ADVANTAGES PRIMARY SCLEROSING CHOLANGITIS PATIENTS ON THE LIVER TRANSPLANTATION WAITING LIST

Goet JC, Hansen BE, Tieleman M, van Hoek B, van den Berg AP, Polak WG, Dubbeld J, Porte RJ, Konijn-Janssen C, de Man RA, Metselaar HJ, de Vries AC.

Transplant International 2018; 31: 590–599.

ABSTRACT

Studies from the USA and Nordic countries indicate primary sclerosing cholangitis (PSC) patients have low mortality on the liver transplantation (LT) waiting list. However, this may vary among geographical areas. Therefore, we compared waiting list mortality and post-transplant survival between laboratory model for end-stage liver disease (LM) and MELD exception (ME)-prioritized PSC and non-PSC candidates in a nationwide study in the Netherlands. A retrospective analysis of patients waitlisted from 2006 to 2013 was conducted. A total of 852 candidates (146 PSC) were waitlisted of whom 609 (71.5%) underwent LT and 159 (18.7%) died before transplantation. None of the ME PSC patients died, and they had a higher probability of LT than LM PSC [HR obtained by considering ME as a time-dependent covariate (HR^{ME} 9.86; 95% CI 6.14–15.85)] and ME non-PSC patients (HR^{ME} 4.60; 95% CI 3.78–5.61). After liver transplantation, PSC patients alive at 3 years of follow-up had a higher probability of relisting than non-PSC patients (HR 7.94; 95% CI 1.98–31.85) but a significantly lower mortality (HR 0.51; 95% CI 0.27–0.95). In conclusion, current LT prioritization advantages PSC patients on the LT waiting list. Receiving ME points is strongly associated with timely LT.

STUDY HIGHLIGHTS

What is known:
- Data from European countries on waiting list mortality in patients with PSC after introduction of the MELD allocation system are lacking

What is new here:
- Current prioritization for liver transplantation advantages patients with PSC

INTRODUCTION

Primary sclerosing cholangitis (PSC) is a chronic, slowly progressive cholestatic liver disease characterized by intra- and extrahepatic biliary strictures which may lead to (decompensated) liver cirrhosis[1,2]. The only curative treatment for end-stage PSC is liver transplantation (LT) with an excellent survival of approximately 80% at 5 years[3-5].

Since December 2006 prioritization for liver donation in the Netherlands is performed using the Model for End-Stage Liver Disease (MELD) score, which aims to transplant patients at highest short-term mortality risk based on objective parameters[6,7]. However, allocation of donor livers using the MELD score may be less applicable for PSC patients with other complications than decompensated cirrhosis[8,9], such as recurrent episodes of cholangitis or hepatobiliary malignancies[8-12]. These complications are not associated with progressive worsening of liver function and may hinder laboratory MELD (LM) score prioritization on the LT waiting list. To counter this problem PSC patients frequently receive MELD exception (ME) points to prioritize their position on the waiting list and allow equal access to liver donation[13,14].

Recent data from the USA, however, reported that MELD score-prioritized PSC patients were less likely to die or be removed from the LT waiting list due to clinical deterioration[15,16], irrespective of ME points. These findings question the appropriateness of the current ME point system in prioritization for liver donation. Consequently, in the USA, an effort to change the exception point system has been initiated[17,18]. However, data from European countries on waiting list mortality in PSC patients after introduction of MELD are lacking. Furthermore, analyses in different cohorts are required as waiting list dynamics may vary among geographical areas, for instance due to differences in prevalence of PSC, indications for LT, deceased organ donation rate and frequency of living donor liver transplantation. This study aimed to compare waiting list mortality as well as post-transplant outcomes between PSC and non-PSC patients by current waiting list policy in the Netherlands. In addition, we aimed to determine the influence of ME points on waiting list survival.

PATIENTS AND METHODS

Population and study design

All patients aged ≥ 18 years listed for liver transplantation in the period from the introduction of MELD score prioritization in the Netherlands on December 16th, 2006 through December 31th 2013 were included. Patients were identified from the Dutch Organ Transplant Registry (NTS). Patients listed for re-transplantation, acute liver failure (high urgency-status (HU) on liver transplantation waiting list), or combined liver and kidney transplantation were excluded.

Data collection

The following clinical and laboratory data were obtained from the NTS: date of birth, sex, indications for LT, date of listing, biochemistry at listing (bilirubin, creatinine and international normalized ratio (INR)), date- and reason of delisting, and post-transplant

survival. Data were recorded until November 2016. Additional data on reason of waiting list removal and cause of death were collected from the medical records from the three liver transplant centres in the Netherlands: The University Medical Centres in Rotterdam, Groningen, and Leiden. Data from the Eurotransplant database were collected to evaluate whether MELD exception (ME) points were awarded during listing. Criteria for awarding exception points are standardized in the Eurotransplant manual[19,20]. In case of standard exceptions (SE) recipients must fulfil country and disease-specific criteria, whereas non-standard exceptions (NSE) have to be approved by a national audit group. The criteria for awarding standard exception points in most countries are: 1. at least two spontaneously occurring septic episodes within 6 months (not due to interventions, not treatable by interventions); 2. splenomegaly > 12cm; 3. body mass index-reduction > 10% within 12 months. At least two of these criteria have to be met to award SE points to PSC patients[19]. However, in the Netherlands, standard exceptions for PSC are not applied. Rather, PSC patients only receive non-standard exception points in case of recurrent infections (cholangitis/biliary sepsis, at least two episodes within 6 months (not due to interventions, not treatable by interventions) with hospitalization). These NSE are strictly enforced. The corresponding centre submits the request to the audit group which comprises auditors from all Dutch liver transplant centres who then vote on the request.

This study was conducted in accordance with the protocol and the principles of the Declaration of Helsinki. The protocol was approved by the Institutional Research Board of the corresponding centre, and at each participating centre, in accordance with local regulations.

Calculations

We calculated lab MELD score using the formula: $0.957 \times \text{Log}_e$ (creatinine mg/dL) + $0.378 \times \text{Log}_e$ (bili mg/dL) + $1.120 \times \text{Log}_e$(INR) + 0.643. Laboratory values less than 1.0 were set to 1.0 in the calculation; maximum serum creatinine in the equation was 4.0 mg/dL; lab MELD scores exceeding 40 were adjusted to 40[19].

Statistical analysis

The primary outcome was mortality on the liver transplantation waiting list, defined as the combined endpoint of death or waiting list removal due to clinical deterioration. Removal due to clinical deterioration was considered equal to death, as a fatal outcome in patients "too sick to transplant" is nearly always inevitable. Patient removed due to clinical improvement, refusal and addiction- or mental problems, as well as waiting list candidates still alive on the waiting list at the end of follow-up were censored at withdrawal from the waiting list or end of the study.

Statistical analyses were performed with IBM SPS Statistics version 22.0 (IBM Corp. Released 2013, IBM Corop, Armon, NY) and SAS software version 9.3.). Data are presented as median and interquartile range (IQR) for continuous variables. Differences in baseline characteristics were compared using the chi-square test for categorical variables, and the Wilcoxon rank-sum test for continuous variables. A value of $p<0.05$ was considered to be statistically significant.

In our study the three competing outcomes on the waiting list were LT, death and removal for other reasons. In conventional survival analysis patients are assumed to have only one type of event during follow-up. Consequently, these analyses yield less accurate estimates of waiting list survival; overestimation of the probability of death on the waiting list on one hand and underestimation of the probability of LT on the other hand[21,22]. Therefore, to determine whether there were significant differences between PSC and non-PSC patients in waiting list survival we performed competing risk analyses. This method uses cumulative incidence curves based on survival functions per event type and permits simultaneous assessment of the different outcomes[21,22].

To determine whether there were significant differences between ME and LM candidates in waiting list survival, the impact of individual covariates on the instantaneous hazard rate of events was assessed with univariate and multivariable Cox proportional hazards models. The time until patients received ME points was modelled as a time-dependent covariate. In multivariable analyses we used informal methods, keeping ME points and PSC versus non-PSC as a covariate in the model, as well as backward stepwise selection containing covariates with $p<0.20$ in univariable cox regression.

For transplanted patients, we assessed post-transplant outcomes (relisting for LT or death) using Cox Proportional Hazard analyses. For the assessment of relisting for LT we used the Landmark method[23]. In these analyses time starts at a clinically meaningful fixed time point after an intervention or initiation of therapy. As one of the main reasons for relisting for LT in PSC patients is recurrence with a median time to recurrence ranging from 3 to 5 years[24-28] we chose 3 years as a fixed time-point, but also applied the landmark method at multiple time points between 1 and 3 years of post-transplant follow-up.

RESULTS

Study population characteristics

During the study period 852 candidates (146 PSC and 706 non-PSC) were listed for LT in the Netherlands. The main indications for liver transplantation were hepatocellular carcinoma (HCC) (n=237), cholestatic liver disease/auto-immune hepatitis (n=218), alcoholic liver disease (n=142) and viral hepatitis (n=77) (**Supplementary Table 1**). Two thirds were male (68.0%); the (median (IQR)) age was 54.0 (46-61) years. PSC patients were significantly younger than non-PSC patients ($p<0.001$; **Table 1**). The median lab MELD score at listing was not significantly different between PSC patients and non-PSC patients. Bilirubin was higher in PSC patients while creatinine and INR levels were significantly higher in non-PSC patients ($p<0.001$; **Table 1**).

Table 1. Demographic and clinical characteristics of patients at the time of listing for liver transplantation

Characteristics	Total cohort N = 852	PSC patients n = 146	Non-PSC patients n = 706	P-value
Gender, male	579 (68)	106 (73)	473 (67)	0.186
Age at listing	54.0 (46-61)	46.5 (39-54)	56.0 (49-61)	<0.001
Bloodtype				0.828
O	392 (46)	66 (45)	326 (46)	
A	314 (37)	58 (40)	314 (44)	
B	102 (12)	15 (10)	102 (14)	
AB	44 (5)	7 (5)	44 (6)	
Laboratory values at listing				
Total bilirubin (μmol/L)	38 (16-87)	60 (27-136)	35 (16-76)	<0.001
Creatinine (μmol/L)	71 (58-89)	63 (52-77)	72 (60-93)	<0.001
INR	1.3 (1.1-1.5)	1.2 (1.0-1.4)	1.3 (1.1-1.5)	<0.001
MELD score: median (IQR)	13.0 (9.0-18.0)	13.5 (9.0-18.0)	13.0 (8.0-18.0)	0.532

INR, international normalized ratio; MELD, model for end-stage liver disease.
Data are presented as number and percentage for categorical data, or as median and interquartile range for continuous data. P values are calculated using the chi-square test for categorical variables and Wilcoxon rank-sum test for continuous variables. P-values illustrated in bold reflect significant findings below the cut-off of 0.05.

MELD exception points on the liver transplantation waiting list

During the study period ME points were granted to 22/146 (15.1%) PSC patients and to 228/706 (32.3%) non-PSC patients. In PSC patients all ME points awarded were NSE. In the non-PSC group that received ME points 27/228 (11.8%) patients received NSE points and 201/228 (88.2%) received SE points. Standard exceptions were mostly awarded for HCC (171 patients). Overall, PSC patients were less likely to receive ME points compared to non-PSC patient (HR 0.34; (95% confidence interval [CI]): 0.22-0.53; p<0.001). In sub-analyses, HCC patients had higher probability of receiving ME points (HR 10.1; CI 6.39-16.0; p<0.001) compared to PSC, whereas patients with alcoholic and patients with viral liver disease had a lower chance (HR 0.32; CI 0.12-0.83; p=0.020 and HR 0.23; CI 0.05-0.98; p=0.026), respectively).

Outcomes on the liver transplantation waiting list

At the end of follow-up of median 214 (IQR 62-435) days (range 8.8 years), 609 patients (71.5%) underwent LT, 159 (18.7%) died or were withdrawn due to clinical deterioration, 60 (7.0%) were withdrawn for other reasons, and 25 (2.9%) were still on the waiting list as of the November 2016 (**Supplementary Figure 1**). The causes of death or removal due to clinical deterioration are presented in **Table 2.** A total of 36 (4.2%) patients were removed because of clinical improvement and 24 (2.8%) for other reasons (refusal, addiction- or mental problems).

A total of 112/146 (76.7%) PSC patients and 397/706 (56.2%) non-PSC patients underwent LT. Six of the 146 (4.1%) PSC patients were removed because of clinical improvement and 2/146 (1.4%) for other reasons. For non-PSC patients these numbers were 30/706 (4.2%) and 22/706 (3.1%), respectively. In the PSC group a total of 18/146 (12.3 %) died or were removed due to clinical deterioration on the liver transplantation waiting list compared to 141/706 (20.0%) in the non-PSC group. None of the PSC patients died or deteriorated due to cholangitis (**Table 2**). Three of the 18 PSC patients were removed because of clinical deterioration (assumed to have died in our analyses): two patients developed cholangiocarcinoma and

one patient gallbladder carcinoma. Two of these patients died within 81- and 138 days after waitlist removal, respectively. One patient was still alive 908 days after waitlist removal. In the non-PSC group 54/141 were removed because of clinical deterioration. Data on survival after removal from the liver transplantation waiting list were available for 50/54 patients. Three patients were still alive at 118-, 370- and 708- days after waitlist removal, respectively. The other 47 patients died after waitlist removal within a median of 181 (IQR 44-400, range 2-1282) days. Most patients (33/47) died within one year after waitlist removal.

Eighteen (14.5%) of the 124/146 (84.9%) PSC patients prioritized on lab MELD scores died; 8 (6.5%) were removed from the waiting list, 90 (72.6%) underwent LT, and 8 (6.5%) were still alive on the waiting list as of the November 2016. None of the PSC patients prioritized on (N) SE MELD scores (22/146; 15.1%) died during follow-up, and all these patients received a LT. Therefore, **Table 2** shows the causes of death or removal for the laboratory MELD prioritized PSC patients.

One hundred and fourteen patients (23.8%) in the LM non-PSC group (478/706 (67.7%)) died, 44 (9.2%) were removed from the waiting list, 309 (64.6%) received LT and 12 (2.5%) were still alive on the waiting list as of the November 2016 (**Supplementary Fig. 1**). Twenty-seven patients (12%) in the ME non-PSC group (228/706 (32.3%)) died, 8 (3.5%) patients were removed from the waiting list, 188 (82%) received a LT, and 5 (2.2%) were still alive on the waiting list as of November 2016.

Table 2. Waitlist removal due to death or clinical deterioration

	Total cohort n = 159	PSC patients n =18	Non-PSC patients n = 141[a]
End-stage liver disease/ Acute on chronic liver failure	62 (39)	8 (44)	54 (38)
Infection/ sepsis	20 (13)	6 (33)	14 (9.9)
SBP	6	2	4
Pneumonia	3	1	2
Focus unclear	11	3	8
Bleeding	11	0	11 (7.8)
Progression malignancy	45 (28)	4 (22)	41 (29)
CCA	3	3	0
HCC	39	0	39
Other	4	2[b]	2
Other (non-liver related)	12 (7.5)[c]	0	12 (8.5)[c]
Unknown	9 (5.7)	0	9 (6.4)

SBP, spontaneous bacterial peritonitis; CCA, cholangiocarcinoma; HCC, hepatocellular carcinoma.
Data are presented as number (and percentage) and represent the cumulative occurrence of endpoints in the period from waiting list acceptance until the end of the study in November 2016.
[a] 54 patients were assumed to have died after waitlist removal due to clinical deterioration. The cause of clinical deterioration could be identified for 51/54 patients: 35 patients suffered progression of HCC, 1 patient developed HCC, 6 developed end-stage liver disease, 2 patients had cancer, and 7 patients had a non-liver related cause of clinical deterioration.
[b] One patient was removed because of gallbladder carcinoma.
[c] 1 Non-Hodgkin Mantell Cell Lymphoma, 1 bladder carcinoma, 1 Alzheimer's disease, 2 heart failure/cardiac decompensation, 1 oropharyngeal cancer, 1 cardiopulmonary problems, 3 cerebral vascular accident, 1 melanoma, and 1 lung carcinoma.

Outcome on the LT waiting list: longer waiting time and low mortality for PSC patients

Although PSC patients had a significantly longer waiting time until delisting compared to non-PSC patients (HR 0.73; CI 0.61-0.88; p=0.001) they had significant better waiting list survival ($HR_{univariate}$ 0.48; CI: 0.29-0.78; p=0.003) in the cumulative incidence curves of the competing risk analyses (**Figure 1A and Figure 1B**). There were no differences in the rate of liver transplantation between PSC and non-PSC candidates (HR 0.84; CI 0.69-1.03; p=0.101; **Figure 1A and Figure 1B**).

Patients who had received MELD exception points had a higher chance of LT ((HR obtained by considering MELD exception points as a time-dependent covariate (HR^{ME}) 3.59 CI 3.01-4.28; p<0.001; **Table 3**)). In addition, ME PSC patients had a significantly higher probability of LT than had LM PSC patients (HR^{ME} 9.86 CI 6.14-15.85; p<0.001) and ME non-PSC patients (HR^{ME} in ME non-PSC patients 4.60 CI 3.78-5.61; p<0.001). The analyses revealed that the effect of age at listing was not significantly different between PSC and non-PSC patients (p-value for effect-modification 0.442).

In univariate analyses ME points (considered as a time-dependent covariate) had a numerical benefit, however not significant (p=0.069), whereas in multivariable analyses those receiving ME points had lower risk of waiting list mortality (**Table 4**). In addition, the multivariate analyses showed that the differences in waiting list survival between PSC and non-PSC patients observed in competing risk analyses are largely explained by age- and MELD scores at listing, and ME points. Older age and higher MELD scores were associated with a poorer prognosis whereas receiving ME points was associated with a better prognosis (**Table 4**). The analyses revealed that the effect of ME points and age at listing was not significantly different between PSC and non-PSC patients (p-value for effect-modification of PSC 0.944 for ME-points and 0.815 for age at listing).

Post-transplant survival is better in PSC patients, although relisting is more common

Analysis of the data of 609 transplanted patients with a mean follow-up after the first liver transplantation of 5.89 years (range: 0 days – 9.07 years) revealed no differences between PSC and non-PSC patients for the combined endpoint of death or relisting for LT (p=0.332) (**Figure 2A**). Interestingly, in sub-analyses PSC patients had a significantly lower risk of death than non-PSC patients (HR 0.51; CI 0.27-0.95; p=0.035; **Figure 2B**). The post-transplant survival rate at 1, 3 and 5 year(s) of follow-up was 91.7%, 90.5%, and 90.5% in PSC patients, while these rates were 91.2%, 83.5% and 75.6% in non-PSC patients. Proportions of patients relisted for LT did not significantly differ between the PSC and non-PSC groups (p=0.763). However, after 3 years of follow-up there was a clear distinction between PSC and non-PSC patients in this respect. The relisting rates for LT at 1, 3 and 5 year(s) were 10.7%, 14.7% and 26.8% in PSC patients, while these rates were 12.8%, 15.4% and 17.5% in non-PSC patients. An increased HR of relisting in PSC patients in patients still alive at three years of follow-up, and over the period 1- and 3 years post-transplant (HR hazard of relisting 7.94; CI 1.98-31.85; p=0.003) as compared to non-PSC patients was observed according to the landmark method[23] (**Figure 2C**).

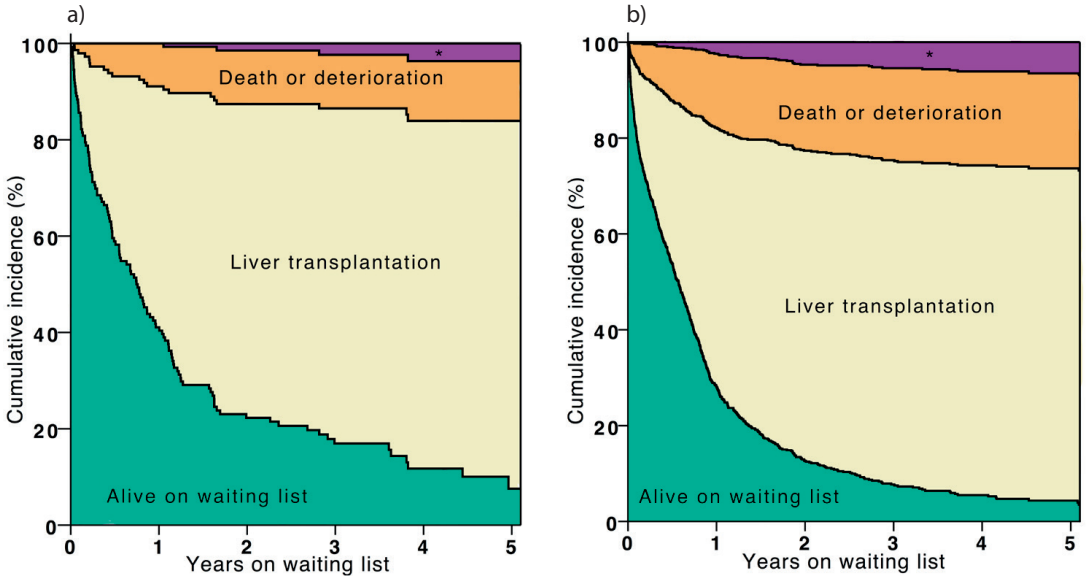


Figure 1. Competing risk analyses with cumulative incidence curves comparing the outcomes on the liver transplantation waiting list (removal, death or clinical deterioration, LT, and still alive) in PSC patients (a) and non-PSC patients (b). The cumulative incidence curves show that although PSC patients had a longer waiting time on the LT waiting list, they had better waiting list survival compared to non-PSC patients in univariable analyses. The transplantation rate between both groups was equal.
*other reasons for removal from the waiting list including clinical improvement, patients' refusal and addiction- or mental problems.

Table 3. The association of time-dependent MELD exception points with liver transplantation

	Univariable analysis				Multivariable analysis			
	HR	95% CI		P-value	HR	95% CI		P-value
Male sex	0.97	0.82	1.15	0.725				
Age at listing	1.01	1.00	1.02	0.046	1.02	1.01	1.02	<0.001
MELD score at listing	1.10	1.08	1.12	<0.001	1.10	1.09	1.12	<0.001
PSC vs. non-PSC	0.84	0.69	1.03	0.101				
Without exception points					0.95	0.74	1.21	0.654
With exception points[a]					2.27	1.45	3.56	<0.001
ME points vs. LM[ab]	3.59	3.01	4.28	<0.001				
PSC	6.87	4.24	11.13	<0.001	9.86	6.14	15.85	<0.001
Non-PSC	3.38	2.78	4.10	<0.001	4.60	3.78	5.61	<0.001

HR, hazard ratio; MELD, model for end-stage liver disease; PSC, primary sclerosing cholangitis; ME, MELD exception; LM, laboratory MELD.
[a] These hazard ratios were obtained by considering MELD exception points as a time-dependent covariate in univariable and multivariable analyses.
[b] The effect of receiving MELD exception points on the probability of liver transplantation was significantly different between PSC and non-PSC patients (interaction, p=0.003). PSC patients that received MELD exception points during follow-up were more likely to receive liver transplantation than non-PSC patients that received ME points.

Table 4. The association of time-dependent MELD exception points with death or clinical deterioration

	Univariable analysis				Multivariable analysis			
	HR	95% CI		P-value	HR	95% CI		P-value
Male sex	1.11	0.80	1.54	0.538				
Age at listing	1.04	1.02	1.06	<0.001	1.05	1.04	1.07	<0.001
PSC vs. non-PSC	0.48	0.29	0.78	0.003	0.72	0.43	1.21	0.211
MELD score at listing	1.11	1.09	1.13	<0.001	1.15	1.13	1.18	<0.001
MELD Exception points[a]	0.67	0.44	1.03	0.069	0.43	0.28	0.68	<0.001

HR, hazard ratio; MELD, model for end-stage liver disease; PSC, primary sclerosing cholangitis.
[a] These hazard ratios were obtained by considering MELD exception points as a time-dependent covariate in univariable and multivariable analyses

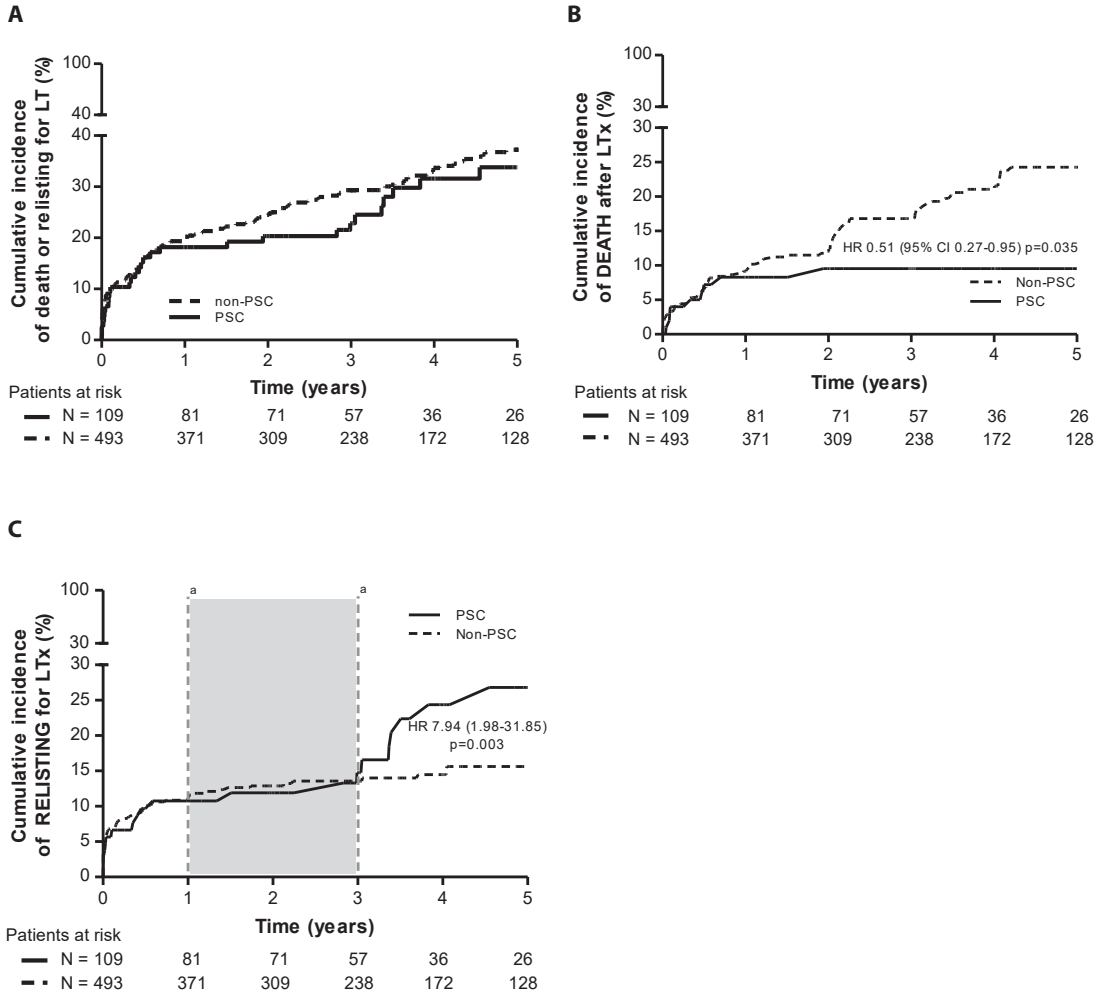


Figure 2. Cumulative incidence of post-transplant relisting and death Kaplan Meier estimates of posttransplant outcomes, stratified according to main indication for liver transplantation (PSC versus non- PSC). The solid line shows values for PSC patients and the dotted line for non-PSC patients. (a) Cumulative incidence of relisting for liver transplantation or death, whichever came first. There were no differences between PSC and non-PSC patients (p=0.301). (b) PSC patients alive after 3 years post-transplant followup had higher probability of relisting for LT compared to non-PSC patients. (c) Overall, PSC patients had a lower risk of post-transplant death compared to non-PSC patients.
[a] Grey area represents interval in which the landmark method was applied.

DISCUSSION

The results of this nationwide study in the Netherlands demonstrate that under current policy for liver donation prioritization the waiting time for PSC patients is longer than that of patients with other indications for liver donation. However, this does not result in increased waiting list mortality or a lower probability of liver transplantation. Although receiving ME points on the LT waiting list during follow-up is associated with better survival and higher probability of LT across all indications for liver donation, this finding is most pronounced in PSC patients. PSC patients who have received MELD exception points have a higher probability of liver transplantation than non-PSC patients and no mortality during waiting for LT was observed in these patients. Lastly, our study suggests that post-transplant PSC patients have better post-transplant survival than non-PSC patients, although they are more often relisted for liver transplantation.

Our study results are in accordance with those of Freeman et al. (2004)[29], who reported a lower risk of death or removal from the LT waiting list in PSC patients compared to other indications for LT after the introduction of MELD allocation in the USA. Moreover, our findings match those from a Scandinavian study that found an equal probability of LT for PSC and non-PSC patients and lower waiting list mortality in PSC patients[30]. These results are not directly applicable to the Netherlands or the USA, as Scandinavian countries do not use the MELD score for liver allocation. Lastly, also consistent with our findings, a recent nationwide study from the USA in more than 79.000 patients reported that MELD score-allocated PSC patients were less likely to die or be removed from the LT waiting list due to clinical deterioration compared to non-PSC patients, irrespective of MELD exception points[15].

The MELD score comprises laboratory parameters that may not reflect PSC disease severity[6,7,31]. As such, as observed in the current study, time on the LT waiting list may be longer for PSC patients, thereby increasing the risk of development of PSC-associated complications[14,15]. In this regard cholangiocarcinoma (CCA), which develops in 6-36% of PSC patients, is an important complication[9,12,32,33]. Nonetheless, only 3 (2%) patients were withdrawn because of biliary tract cancer (2 CCA and 1 gallbladder carcinoma) in our study. This can be explained by CCA being a contra-indication for liver transplantation during the study period[19]. Interestingly, but in keeping with Goldberg et al. (2012)[16], none of the PSC patients died or deteriorated due to fulminant cholangitis; one of the PSC-associated complications suggested to affect waiting list mortality and for which standard ME points can be granted[19,20].

Although the exact reasons for relisting after the first liver transplantation in the current study are unknown, one might speculate that the observed higher probability of relisting in PSC patients was due to recurrent disease. While the 5-year post-transplant survival of PSC patients exceeds 80%[3-5], approximately 20% of PSC patients will develop recurrent disease within a median time of 3-5 years[24-28]). This is associated with increased risk of graft loss and mortality[27,28,34]. Interestingly, recently published data from the United Network for Organ Sharing (UNOS) by Henson et al. (2016)[35] indicates that PSC patients with a late re-transplantation for recurrent disease have an excellent 5-year graft survival of

approximately 75.7%. Based on these - and our results, PSC patients have a high 'transplant benefit'; meaning a high post-transplant survival in addition to post-acceptance survival. From an economic and ethical perspective this is an important consideration in the prioritization for liver donation. The high transplant benefit in PSC patients may warrant currently observed waiting list advantage.

Strengths of our study are its nationwide coverage and long-term follow-up period from 2006 through 2016. Furthermore, we used competing risk analyses. Whereas normal survival analyses would have provided an overestimation of the risk of death or clinical deterioration and an underestimation of the probability of LT, our analyses provide a reliable overview of LT waiting list survival. Moreover, in addition to an in-depth analysis of the influence of ME points on waiting list survival we assessed "transplant benefit" (the combination of post-acceptance and post-transplant survival) of the current allocation system. As such we provide a comprehensive overview of LT waiting list dynamics of PSC patients in the Netherlands.

However, some limitations need to be considered. First, the considerable proportion of HCC patients may have influenced the results, as early-stage HCC patients receive standard ME points. Still, when we excluded the HCC group from analysis, we found no differences in granting ME points in PSC versus non-PSC patients, and the results of all other analyses remained unchanged. Secondly, our study only indicates that the advantage obtained from current ME policy is greater than appropriate. To obtain a definite answer on the appropriateness of ME priority a study to determine the outcomes of all patients in the counterfactual case they had not been given ME priority would be needed. However, the current data do not allow for such a comparative analysis. In addition, to study this in a prospective study would raise ethical issues. Third, in our study removal due to clinical deterioration was considered equal to death. However, studies using the United States Organ Procurement and Transplantation Network (OPTN) have shown that removal for medical deterioration is not always a reliable indicator of death[36]. While in the US a marked variability in the use of removal codes among the different OPTN regions may significantly impact the estimates of deaths, this variation is negligible in the Netherlands as LT is centralized in three centres. Moreover, in our study most patients that deteriorated died within one year after removal (2/3 PSC and 33/50 non-PSC). Therefore, it is reasonable to assume that the impact of considering clinical deterioration equal to death has a negligible impact on our results. Finally, this nationwide study may be difficult to generalize. However, several European countries use the same standard MELD exceptions that are common to the Eurotransplant system. Consequently, variability mainly concerns the non-standard ME points. Furthermore, in the US the PSC aspect of exception points and other challenges regarding the ME system (including lack of standardization, geographical differences in the approval of exceptions, and limited evidence base to support certain exceptions[17,37,38]) have already triggered a broader effort to change the exception point system[17,18]. Our data warrant reconsideration of the ME system in Europe, similar to initiatives in the US.

In conclusion, this nationwide study in the Netherlands confirms previously reported challenges in granting equal access to donor livers across patients with various end-stage liver diseases. Despite a longer waiting time, current MELD score prioritization does not

result in increased waiting list mortality or a lower probability of liver transplantation in PSC patients, while the MELD exception points system advantages PSC patients on the liver transplantation waiting list in the Netherlands. These findings need to be weighed against higher transplant benefit in PSC patients during the continuous process of reassessment and adjustment of liver transplantation prioritization.

REFERENCES

1. Chapman RW, Arborgh BA, Rhodes JM, et al. Primary sclerosing cholangitis: a review of its clinical features, cholangiography, and hepatic histology. *Gut*. 1980;21(10):870-877.
2. Hirschfield GM, Karlsen TH, Lindor KD, Adams DH. Primary sclerosing cholangitis. *Lancet*. 2013;382(9904):1587-1599.
3. European Association for the Study of the L. EASL Clinical Practice Guidelines: management of cholestatic liver diseases. *J Hepatol*. 2009;51(2):237-267.
4. Roberts MS, Angus DC, Bryce CL, Valenta Z, Weissfeld L. Survival after liver transplantation in the United States: a disease-specific analysis of the UNOS database. *Liver Transpl*. 2004;10(7):886-897.
5. Singal AK, Guturu P, Hmoud B, Kuo YF, Salameh H, Wiesner RH. Evolving frequency and outcomes of liver transplantation based on etiology of liver disease. *Transplantation*. 2013;95(5):755-760.
6. Malinchoc M, Kamath PS, Gordon FD, Peine CJ, Rank J, ter Borg PC. A model to predict poor survival in patients undergoing transjugular intrahepatic portosystemic shunts. *Hepatology*. 2000;31(4):864-871.
7. Wiesner R, Edwards E, Freeman R, et al. Model for end-stage liver disease (MELD) and allocation of donor livers. *Gastroenterology*. 2003;124(1):91-96.
8. Folseraas T, Boberg KM. Cancer Risk and Surveillance in Primary Sclerosing Cholangitis. *Clin Liver Dis*. 2016;20(1):79-98.
9. Broome U, Olsson R, Loof L, et al. Natural history and prognostic factors in 305 Swedish patients with primary sclerosing cholangitis. *Gut*. 1996;38(4):610-615.
10. Boonstra K, Weersma RK, van Erpecum KJ, et al. Population-based epidemiology, malignancy risk, and outcome of primary sclerosing cholangitis. *Hepatology*. 2013;58(6):2045-2055.
11. Claessen MM, Vleggaar FP, Tytgat KM, Siersema PD, van Buuren HR. High lifetime risk of cancer in primary sclerosing cholangitis. *J Hepatol*. 2009;50(1):158-164.
12. Bergquist A, Ekbom A, Olsson R, et al. Hepatic and extrahepatic malignancies in primary sclerosing cholangitis. *J Hepatol*. 2002;36(3):321-327.
13. Goldberg D, Bittermann T, Makar G. Lack of standardization in exception points for patients with primary sclerosing cholangitis and bacterial cholangitis. *Am J Transplant*. 2012;12(6):1603-1609.
14. Goldberg DS, French B, Thomasson A, Reddy KR, Halpern SD. Current trends in living donor liver transplantation for primary sclerosing cholangitis. *Transplantation*. 2011;91(10):1148-1152.
15. Goldberg D, French B, Thomasson A, Reddy KR, Halpern SD. Waitlist survival of patients with primary sclerosing cholangitis in the model for end-stage liver disease era. *Liver Transpl*. 2011;17(11):1355-1363.
16. Goldberg DS, Camp A, Martinez-Camacho A, Forman L, Fortune B, Reddy KR. Risk of waitlist mortality in patients with primary sclerosing cholangitis and bacterial cholangitis. *Liver Transpl*. 2013;19(3):250-258.
17. Committee OaULaIOT. OPTN/UNOS Public Comment Proposal to Establish a National Liver Review Board. https://optn.transplant.hrsa.gov/media/2085/liver_pcproposal_review_board_policy_201701.pdf. Published 2017. Accessed 8-8-2017, 2017.

18. Committee LIOT. Adult MELD exception review guidance. Organ Procurement and Transplantation Network. https://optn.transplant.hrsa.gov/resources/by-organ/liver-intestine/guidance-on-meld-peld-exception-review/#References. Published 2015. Updated June 2, 2015. Accessed 8-8-2017, 2017.
19. Eurotransplant Manual, Chapter 5 Liver Allocation System (ELAS). Eurotransplant. https://www.eurotransplant.org/cms/mediaobject.php?file=chapter5_elas21.pdf. Published 2014. Updated January 13. Accessed July 11, 2014.
20. Eurotransplant Manual, Chapter 5 Liver Allocation System (ELAS). https://www.eurotransplant.org/cms/mediaobject.php?file=H5+ELAS+MELD+Oct+20161.pdf. Published 2016. Accessed January 1, 2017.
21. Kim WR, Therneau TM, Benson JT, et al. Deaths on the liver transplant waiting list: an analysis of competing risks. *Hepatology.* 2006;43(2):345-351.
22. Satagopan JM, Ben-Porat L, Berwick M, Robson M, Kutler D, Auerbach AD. A note on competing risks in survival data analysis. *Br J Cancer.* 2004;91(7):1229-1235.
23. Anderson JR, Cain KC, Gelber RD. Analysis of survival by tumor response. *J Clin Oncol.* 1983;1(11):710-719.
24. Tamura S, Sugawara Y, Kaneko J, Matsui Y, Togashi J, Makuuchi M. Recurrence of primary sclerosing cholangitis after living donor liver transplantation. *Liver Int.* 2007;27(1):86-94.
25. Gautam M, Cheruvattath R, Balan V. Recurrence of autoimmune liver disease after liver transplantation: a systematic review. *Liver Transpl.* 2006;12(12):1813-1824.
26. Charatcharoenwitthaya P, Lindor KD. Recurrence of primary sclerosing cholangitis: what do we learn from several transplant centers? *Liver Transpl.* 2008;14(2):130-132.
27. Campsen J, Zimmerman MA, Trotter JF, et al. Clinically recurrent primary sclerosing cholangitis following liver transplantation: a time course. *Liver Transpl.* 2008;14(2):181-185.
28. Alabraba E NP, Gunson B, et al. A re-evaluation of the risk factors for the recurrence of primary sclerosing cholangitis in liver allografts. *Liver Transplant.* 2009;15:330e340.
29. Freeman RB, Wiesner RH, Edwards E, et al. Results of the first year of the new liver allocation plan. *Liver Transpl.* 2004;10(1):7-15.
30. Brandsaeter B, Broome U, Isoniemi H, et al. Liver transplantation for primary sclerosing cholangitis in the Nordic countries: outcome after acceptance to the waiting list. *Liver Transpl.* 2003;9(9):961-969.
31. Kamath PS, Wiesner RH, Malinchoc M, et al. A model to predict survival in patients with end-stage liver disease. *Hepatology.* 2001;33(2):464-470.
32. Miros M, Kerlin P, Walker N, Harper J, Lynch S, Strong R. Predicting cholangiocarcinoma in patients with primary sclerosing cholangitis before transplantation. *Gut.* 1991;32(11):1369-1373.
33. Nashan B, Schlitt HJ, Tusch G, et al. Biliary malignancies in primary sclerosing cholangitis: timing for liver transplantation. *Hepatology.* 1996;23(5):1105-1111.
34. Alexander J, Lord JD, Yeh MM, Cuevas C, Bakthavatsalam R, Kowdley KV. Risk factors for recurrence of primary sclerosing cholangitis after liver transplantation. *Liver Transpl.* 2008;14(2):245-251.
35. Henson JB, Patel YA, King LY, Zheng J, Chow SC, Muir AJ. Outcomes of Liver Retransplantation in Patients with Primary Sclerosing Cholangitis. *Liver Transpl.* 2016.
36. Voigt MD, Hunsicker LG, Snyder JJ, Israni AK, Kasiske BL. Regional variability in liver waiting list removals causes false ascertainment of waiting list deaths. *Am J Transplant.* 2013;13(2):369-375.

37. Argo CK, Stukenborg GJ, Schmitt TM, Kumer SC, Berg CL, Northup PG. Regional variability in symptom-based MELD exceptions: a response to organ shortage? *Am J Transplant*. 2011;11(11):2353-2361.
38. Rodriguez-Luna H, Vargas HE, Moss A, Reddy KS, Freeman RB, Mulligan D. Regional variations in peer reviewed liver allocation under the MELD system. *Am J Transplant*. 2005;5(9):2244-2247.

SUPPLEMENTARY DATA

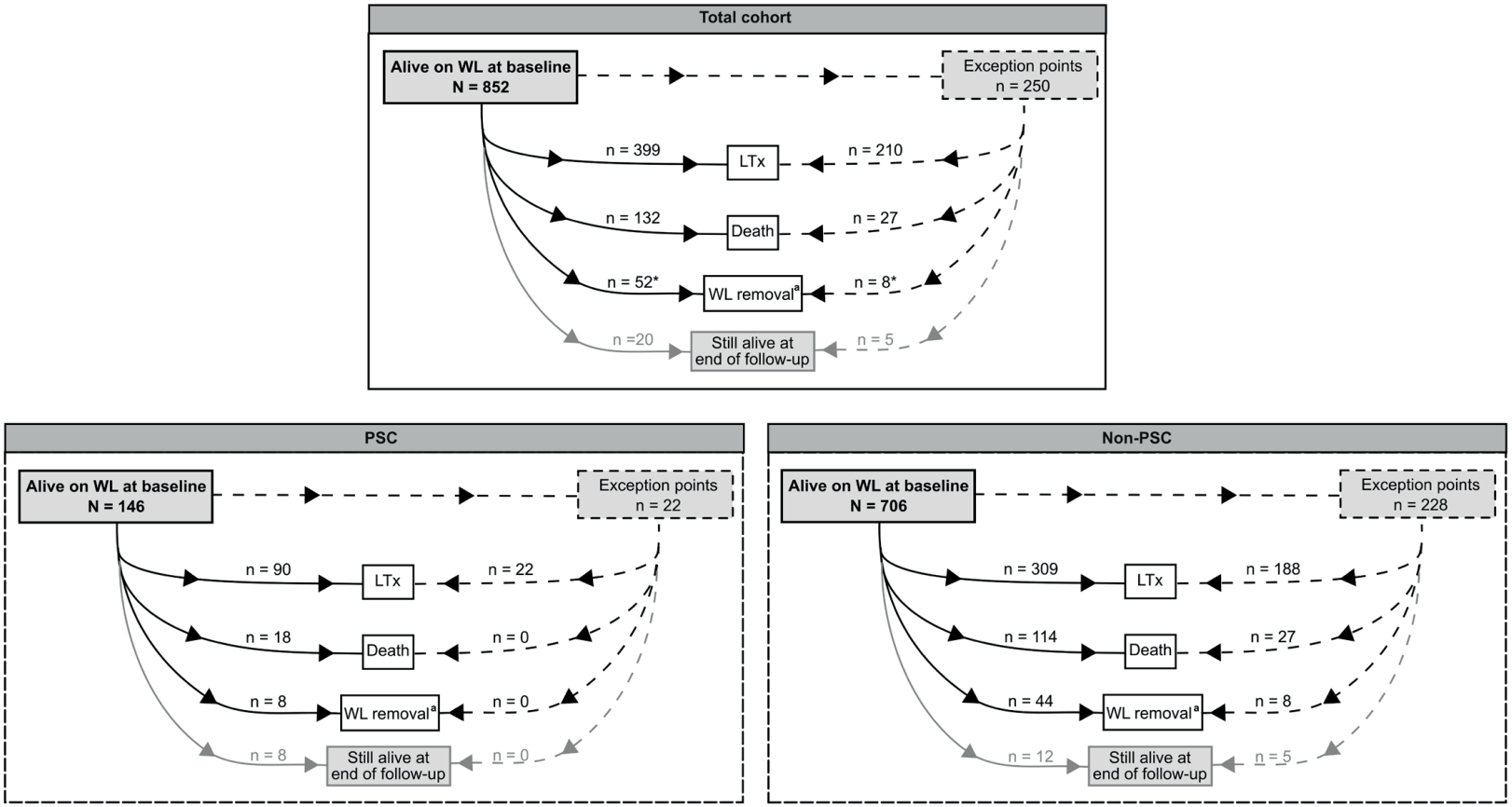
Supplementary Table 1. Main indications for liver transplantation

	Frequency, n	Percentage, %
Hepatocellular carcinoma[a]	237	28
Cholestatic liver disease / auto-immune hepatitis	218	26
PSC	146	17
PBC	29	3.4
AIH	30	3.5
Other	13	1.5
Alcoholic liver disease	142	17
Viral hepatitis	77	9.0
HBV	24	2.8
HCV	53	6.2
Metabolic liver disease	93	11
NASH	45	5.3
Alpha-1-antityripsin deficiency	13	1.5
Haemochromatosis	11	1.3
Other	24	2.8
Other[b]	36	4.2
Cryptogenic	49	5.8
Total	852	100

PSC, Primary sclerosing cholangitis; PBC, primary biliary cholangitis; AIH, auto-immune hepatitis; HBV, hepatitis B virus; HCV, hepatitis C virus; NASH, Non-alcoholic steatohepatitis..

[a] Underlying disease in HCC was (indication (frequency)): alcoholic (65), HBV (34), HCV (85), PSC (5), PBC (4), cholestatic other and AIH (4), metabolic (23) (NASH in 16 patients), other (2), cryptogenic (11), and unknown in 4 patients.

[b] 'Other' includes 'Toxic-drugs related', 'Polycystic liver disease', 'Vascular liver disease' and 'benign liver tumors'.



Supplementary Figure 1. Flow diagram of outcomes. Patients granted MELD Exception (ME) points in the period from waiting list acceptance (cumulative) until the end of the study in November 2016 are included in the group 'Exception points'. Numbers represent cumulative occurrence of endpoints in the period from waiting list acceptance until the end of the study in November 2016. [a] Removal for clinical improvement, patients' refusal and addiction- or mental problems.

CHAPTER 8

8

GENERAL DISCUSSION AND CONCLUSIONS

BACKGROUND

Primary biliary cholangitis

PBC is a chronic cholestatic liver disease[1,2]. While the course of the disease is beneficially altered in most patients due to the initiation of ursodeoxycholic acid (UDCA) therapy, a subgroup of patients has an inadequate treatment response and may develop liver failure. Therefore, the development of second-line therapies is of utmost importance. Due to PBC being a rare, slowly progressive, disease the conduction of clinical trials within a feasible time-frame has been challenging. A landmark paper by the Global PBC Study Group has changed this landscape with the identification of alkaline phosphatase and bilirubin as independent (surrogate) markers of long-term outcome[3]. These markers now allow the assessment of the effect of new drug therapies within a shorter and feasible time-frame. Another key development has been the derivation of continuous risk scores that estimate the risk of future events in patients receiving UDCA therapy[4,5]. In the setting of the development of new drug therapies for PBC, these risk stratification tools are critical to identify patients in need of second-line therapies or patients who can be safely continued on UDCA monotherapy[4,5].

Primary sclerosing cholangitis

PSC is a chronic, slowly progressive cholestatic liver disease characterized by intrahepatic and extrahepatic biliary strictures which may lead to (decompensated) liver cirrhosis[6-8]. In contrast to PBC, PSC lacks approved therapies that are able to halt the disease progression. The search for therapies has been hindered by uncertainties regarding the pathogenesis of the disease as well as possible pathways that detrimentally affect disease progression. In addition, studying the impact of new drugs on hard clinical endpoints within the time-span of a clinical trial is demanding as PSC is a heterogeneous disease abounding in different phenotypic appearances. Therefore, there is a great need for surrogate markers of disease progression that can be used in clinical trials. While surrogate markers are not yet available for PSC, collaboration of various PSC consortia in past years has boosted research in PSC. With the derivation of various prognostic models for PSC, namely the Amsterdam-Oxford score, UK-PSC score and Primary Sclerosing Cholangitis Risk Estimate Tool (PREsTo) clinicians are now able to quantify the risk of future events for the individual patient, allowing more individualized risk prediction[9-11].

RISK STRATIFICATION IN PRIMARY BILIARY CHOLANGITIS

In recent years, a number of studies have demonstrated an increased incidence and prevalence[12-20]. In addition, the clinical presentation of PBC has changed. Whereas most patients presented with an advanced histological stage in earlier decades, nowadays most patients present during an early (asymptomatic) stage[21,22]. **Chapter 2** evaluated these temporal trends in patient- and disease characteristics over a 44-year period in 4,805 European and American PBC patients diagnosed between 1970 and 2014. We found the mean age at diagnosis increased incrementally by 2-3 years per decade from 47 years in the 70s to 57 years from those diagnosed in 2010 onwards. Furthermore, more patients presented with milder biochemical disease severity in recent years (near 40% in 1970 and

more than 70% nowadays) as well as a milder histological disease stage. Correspondingly, response to UDCA according to Paris-I criteria has increased and decompensation rates have decreased. Our results demonstrate that in recent decades, PBC patients may present at an older age at diagnosis along with a reduced disease severity. A possible explanation for these findings may be an increase in routine testing of liver function.

In **chapter 3**, we assessed the subpopulation of patients that present with an early biochemical disease stage, as defined by normal serum bilirubin and albumin concentrations according to Rotterdam criteria[23]. In 1615 patients with early stage PBC, we found that ~46% at 5 years following initiation of UDCA treatment or initial evaluation transitioned from early to moderate stage (abnormal bilirubin or albumin). Thereafter, 16% transitioned to advanced disease stage (abnormal bilirubin and albumin) within 5 years after the initial transition from early to moderate disease stage. Transitions from early to moderate PBC and from moderate to advanced PBC were associated with higher probabilities of a clinical event. These findings stress the importance of surveillance in patients with an early biochemical disease stage.

During the last decades UDCA was the only licensed treatment option available[24-28]. However, the treatment landscape evolved in 2016 with obeticholic acid (OCA) being granted approval as adjuvant therapy by the FDA and EMA for patients having incomplete response to UDCA alone[29]. In the near future, it is likely that other agents, including fibric acid derivatives and other nuclear bile acid therapies will also gain approval as second-line treatment options[30]. In this setting, reliable assessment of response to UDCA and estimates of transplant-free survival are of critical importance to identify individuals who may benefit from these second-line treatment options; and reciprocally, those for whom UDCA monotherapy can be continued safely. However, the proposed assessment at one year of UDCA treatment poses a limitation in daily clinical practice since in daily clinical practice data with respect to the initial treatment response or the initial estimate of transplant-free survival may not, or no longer, be available. Moreover, (repeated) risk stratification after more prolonged treatment can be considered at least of equal key relevance in patient management. Therefore, the ability of a risk stratification tool to be used at any point following treatment initiation, even after prolonged treatment, is of relevance to patients and important for effective clinical care delivery. **Chapter 4** therefore assessed the prognostic performance and applicability of the GLOBE score during prolonged follow-up using data from a large internationally representative cohort of more than 3500 UDCA-treated PBC patients. We show that irrespective of the time of assessment, the GLOBE score is a reliable tool to predict prognosis in UDCA treated PBC up to 10 years of follow-up. A reliable estimate of transplant-free survival can be obtained by calculating the GLOBE score based on age at initiation of UDCA as well as age as using the actual age at the time of assessment. This study therefore markedly increases the utility of the GLOBE score as a practical clinical tool to estimate prognosis.

Over the past years, several risk scores based on prediction models have been proposed. There is scarce data on the comparative value of these scores to predict outcome in PBC. **Chapter 5** compares the predictive performance of the Mayo Risk Score (MRS) and Model for End-stage Liver Disease (MELD) with newly the introduced GLOBE score[4] and UK-PBC

risk score[5]. We demonstrate that all prognostic scores developed for PBC have comparable performance in the prediction of clinical outcomes. This is true for various time points during UDCA treatment as well as in sub-groups stratified according to biochemical and fibrosis disease stage. This suggests that implementation ought to be based on clinical context.

RISK STRATIFICATION IN PRIMARY SCLEROSING CHOLANGITIS

Recently the Amsterdam-Oxford model (AOM) was introduced, a prognostic model developed in a population-based cohort to predict the long-term risk of PSC-related death and/or liver transplantation[9]. **Chapter 6** assessed the predictive performance and applicability of the AOM in clinical practice in a large cohort of PSC patients from three tertiary centres in Europe. We show that the C-statistic for the AOM remains around 0.70 in the first 5 years following PSC diagnosis, thereby further validating the AOM as a valuable risk stratification tool in PSC and extending its utility.

Once PSC patients have progressed to end stage liver disease, prioritization for liver transplantation on the liver transplantation waiting list in The Netherlands is performed using the MELD score[31,32]. However, PSC specific complications such as recurrent episodes of cholangitis or hepatobiliary malignancy may not be adequately reflected by the MELD score[33-37]. These complications may hinder laboratory MELD (LM) score prioritization on the liver transplantation waiting list. To counter this problem PSC patients frequently receive MELD exception (ME) points[38,39]. In **Chapter 7** we showed that between 2006 and 2013 none of the PSC patients waitlisted who received ME died, and that, despite a longer waiting time, they had a higher probability of LT than LM PSC and non-PSC patients who received ME. In addition, post-LT survival was higher for PSC patients. These results indicate that current MELD prioritization does not result in increased waiting list mortality or a lower probability of LT in PSC patients, while the MELD exception points system advantages PSC patients on the liver transplantation waiting list in the Netherlands. These findings may warrant reconsideration of the liver transplant prioritization policy for PSC patients, but should, however, be weighed against higher transplant benefit in PSC patients.

CONCLUSIONS

Nowadays, PBC patients may present at an older age at diagnosis along with an early biochemical or histological disease severity. Within the subgroup of patients that present with an early disease as defined by normal bilirubin and albumin, the percentage of biochemical transition to a moderate or advanced disease stage may be substantial. These transitions are associated with worse survival. Therefore, early stage PBC patients should be offered structured life-long (biochemical) follow-up and if needed adjustment of treatment. During prolonged treatment with UDCA, the GLOBE score can be used to adequately estimate prognosis. Changes in GLOBE score over time, especially when resulting in sustained GLOBE scores above the set threshold, should prompt reconsideration of optimal management and possible use of second-line therapies. When assessing a patient's risk after initiation or during the course of UDCA therapy, the various risk prediction scores have equal performance. Choice of one score over the other should therefore be based on clinical context. In PSC, the Amsterdam-Oxford score can aid physicians in estimating LT-free survival. In doing so, it has an adequate predictive performance at PSC diagnosis, but also 5 years thereafter. Waiting list survival for PSC patients under current MELD score prioritization is excellent, but may warrant reconsideration of transplant prioritization policy.

FUTURE DIRECTIONS

The currently changing clinical scenario in PBC in which new drugs are coming available, all with different profiles of efficacy and tolerability, poses a new challenge in the risk stratification of PBC patients. So what is next? Risk stratification and prediction models are in constant development. Various factors could provide an important addition to existing prediction models. Liver stiffness measurement might be a promising candidate, but requires further validation, and has practical implications and limitations. In addition, the inclusion of symptoms or comorbidities, including for example non-alcoholic steatohepatitis or cardiovascular disease, could aid in even more accurate risk predictions and requires further exploration. Finally, new biomarkers and measurements techniques are emerging, the knowledge of genetic associations and environmental triggers is increasing, and new insights in disease and host factors are being gained. These developments will likely further improve our ability to identify patients at risk, provide individual estimates of survival or risk of events, and aid in the prediction of response to various treatments.

Similar to other biochemical response criteria, both the GLOBE and UK-PBC score require further validation in evaluating the accuracy of estimated survival in patients treated with newly developed drugs. Likewise, the utility of these scores as a surrogate marker of treatment success for novel therapies awaits confirmation. Such measurement neglects the markers of fibrosis. Efforts are currently ongoing to assess the benefits of using the GLOBE and UK-PBC risk scores as response criteria to predict risk reduction and benefit induced by a second-line choleretic drug, namely obeticholic acid[40].

A pragmatic approach towards future sophistication of prognostic tools is to expand existing, mainly retrospective databases by adding prospectively collected data, while continuing the recruitment of incident cases. The capture of additional data can allow the modelling of outcomes using sets of variables measured at different time points after starting treatment; enabling the development of models incorporating repeated measurements. Such a dynamic prediction model allows for an update of survival probabilities at new observations of biochemistry or clinical events, and thereby could provide even more individualized predictions than continuous models.

In the near future, gastroenterologists and hepatologists will likely be faced with multiple therapeutic options for PBC, emphasising the clinical importance of an accurate estimation of prognoses and selection of the appropriate therapy for the individual patient. An important question in the applicability of a prediction model is whether the predictions are generalizable. This will depend on the setting and assumptions under which they were created and validated; in a different setting the performance may be less good and may even need new validation. Applied in a setting similar to the conditions and assumptions a model was created in, the prediction performs best.

References

1. Kaplan MM, Gershwin ME. Primary biliary cirrhosis. *N Engl J Med.* 2005;353(12):1261-1273.
2. Poupon R. Primary biliary cirrhosis: a 2010 update. *J Hepatol.* 2010;52(5):745-758.
3. Lammers WJ, van Buuren HR, Hirschfield GM, 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.* 2014;147(6):1338-1349.
4. Lammers WJ, Hirschfield GM, Corpechot C, et al. Development and Validation of a Scoring System to Predict Outcomes of Patients With Primary Biliary Cirrhosis Receiving Ursodeoxycholic Acid Therapy. *Gastroenterology.* 2015;149(7):1804-1812 e1804.
5. Carbone M, Sharp SJ, Flack S, et al. The UK-PBC risk scores: Derivation and validation of a scoring system for long-term prediction of end-stage liver disease in primary biliary cirrhosis. *Hepatology.* 2015:doi: 10.1002/hep.28017.
6. Chapman RW, Arborgh BA, Rhodes JM, et al. Primary sclerosing cholangitis: a review of its clinical features, cholangiography, and hepatic histology. *Gut.* 1980;21(10):870-877.
7. Hirschfield GM, Karlsen TH, Lindor KD, Adams DH. Primary sclerosing cholangitis. *Lancet.* 2013;382(9904):1587-1599.
8. Williamson KD, Chapman RW. Primary sclerosing cholangitis. *Dig Dis.* 2014;32(4):438-445.
9. de Vries EM, Wang J, Williamson KD, et al. A novel prognostic model for transplant-free survival in primary sclerosing cholangitis. *Gut.* 2017.
10. Goode EC, Clark AB, Mells GF, et al. Factors Associated With Outcomes of Patients With Primary Sclerosing Cholangitis and Development and Validation of a Risk Scoring System. *Hepatology.* 2019;69(5):2120-2135.
11. Eaton JE, Vesterhus M, McCauley BM, et al. Primary Sclerosing Cholangitis Risk Estimate Tool (PREsTo) Predicts Outcomes in PSC: A Derivation & Validation Study Using Machine Learning. *Hepatology.* 2018.
12. Boonstra K, Beuers U, Ponsioen CY. Epidemiology of primary sclerosing cholangitis and primary biliary cirrhosis: a systematic review. *J Hepatol.* 2012;56(5):1181-1188.
13. Metcalf JV, Bhopal RS, Gray J, Howel D, James OF. Incidence and prevalence of primary biliary cirrhosis in the city of Newcastle upon Tyne, England. *Int J Epidemiol.* 1997;26(4):830-836.
14. Boonstra K, Kunst AE, Stadhouders PH, et al. Rising incidence and prevalence of primary biliary cirrhosis: a large population-based study. *Liver Int.* 2014;34(6):e31-38.
15. Eriksson S, Lindgren S. The prevalence and clinical spectrum of primary biliary cirrhosis in a defined population. *Scand J Gastroenterol.* 1984;19(7):971-976.
16. Myszor M, James OF. The epidemiology of primary biliary cirrhosis in north-east England: an increasingly common disease? *Q J Med.* 1990;75(276):377-385.
17. Remmel T, Remmel H, Uibo R, Salupere V. Primary biliary cirrhosis in Estonia. With special reference to incidence, prevalence, clinical features, and outcome. *Scand J Gastroenterol.* 1995;30(4):367-371.
18. James OF, Bhopal R, Howel D, Gray J, Burt AD, Metcalf JV. Primary biliary cirrhosis once rare, now common in the United Kingdom? *Hepatology.* 1999;30(2):390-394.
19. Pla X, Vergara M, Gil M, et al. Incidence, prevalence and clinical course of primary biliary cirrhosis in a Spanish community. *Eur J Gastroenterol Hepatol.* 2007;19(10):859-864.
20. Myers RP, Shaheen AA, Fong A, et al. Epidemiology and natural history of primary biliary cirrhosis in a Canadian health region: a population-based study. *Hepatology.* 2009;50(6):1884-1892.

21. Christensen E, Crowe J, Doniach D, et al. Clinical pattern and course of disease in primary biliary cirrhosis based on an analysis of 236 patients. *Gastroenterology.* 1980;78(2):236-246.
22. Locke GR, 3rd, Therneau TM, Ludwig J, Dickson ER, Lindor KD. Time course of histological progression in primary biliary cirrhosis. *Hepatology.* 1996;23(1):52-56.
23. ter Borg PC, Schalm SW, Hansen BE, van Buuren HR, Dutch PBCSG. Prognosis of ursodeoxycholic Acid-treated patients with primary biliary cirrhosis. Results of a 10-yr cohort study involving 297 patients. *Am J Gastroenterol.* 2006;101(9):2044-2050.
24. Lindor KD, Gershwin ME, Poupon R, et al. Primary biliary cirrhosis. *Hepatology.* 2009;50(1):291-308.
25. Poupon RE, Lindor KD, Cauch-Dudek K, Dickson ER, Poupon R, Heathcote EJ. Combined analysis of randomized controlled trials of ursodeoxycholic acid in primary biliary cirrhosis. *Gastroenterology.* 1997;113(3):884-890.
26. Shi J, Wu C, Lin Y, Chen YX, Zhu L, Xie WF. Long-term effects of mid-dose ursodeoxycholic acid in primary biliary cirrhosis: a meta-analysis of randomized controlled trials. *Am J Gastroenterol.* 2006;101(7):1529-1538.
27. Corpechot C, Carrat F, Bahr A, Chretien Y, Poupon RE, Poupon R. The effect of ursodeoxycholic acid therapy on the natural course of primary biliary cirrhosis. *Gastroenterology.* 2005;128(2):297-303.
28. Liver EAftSot. EASL Clinical Practice Guidelines: The diagnosis and management of patients with primary biliary cholangitis. *J Hepatol.* 2017.
29. Nevens F, Andreone P, Mazzella G, et al. A Placebo-Controlled Trial of Obeticholic Acid in Primary Biliary Cholangitis. *N Engl J Med.* 2016;375(7):631-643.
30. Corpechot C, Chazouilleres O, Rousseau A, et al. A Placebo-Controlled Trial of Bezafibrate in Primary Biliary Cholangitis. *N Engl J Med.* 2018;378(23):2171-2181.
31. Malinchoc M, Kamath PS, Gordon FD, Peine CJ, Rank J, ter Borg PC. A model to predict poor survival in patients undergoing transjugular intrahepatic portosystemic shunts. *Hepatology.* 2000;31(4):864-871.
32. Wiesner R, Edwards E, Freeman R, et al. Model for end-stage liver disease (MELD) and allocation of donor livers. *Gastroenterology.* 2003;124(1):91-96.
33. Boonstra K, Weersma RK, van Erpecum KJ, et al. Population-based epidemiology, malignancy risk, and outcome of primary sclerosing cholangitis. *Hepatology.* 2013;58(6):2045-2055.
34. Claessen MM, Vleggaar FP, Tytgat KM, Siersema PD, van Buuren HR. High lifetime risk of cancer in primary sclerosing cholangitis. *J Hepatol.* 2009;50(1):158-164.
35. Bergquist A, Ekbom A, Olsson R, et al. Hepatic and extrahepatic malignancies in primary sclerosing cholangitis. *J Hepatol.* 2002;36(3):321-327.
36. Folseraas T, Boberg KM. Cancer Risk and Surveillance in Primary Sclerosing Cholangitis. *Clin Liver Dis.* 2016;20(1):79-98.
37. Broome U, Olsson R, Loof L, et al. Natural history and prognostic factors in 305 Swedish patients with primary sclerosing cholangitis. *Gut.* 1996;38(4):610-615.
38. Goldberg D, Bittermann T, Makar G. Lack of standardization in exception points for patients with primary sclerosing cholangitis and bacterial cholangitis. *Am J Transplant.* 2012;12(6):1603-1609.
39. Goldberg DS, French B, Thomasson A, Reddy KR, Halpern SD. Current trends in living donor liver transplantation for primary sclerosing cholangitis. *Transplantation.* 2011;91(10):1148-1152.
40. Harms M.H. CM, Hansen, B.E., Mells, G., Pencek R., Smoot-Malecha E., MacConell L. Risk reduction with obeticholic acid in patients not achieving the POISE primary endpoint. *Journal of Hepatology.* 2017;66(1):Supplement, S1-S876.

CHAPTER 9

9

DUTCH SUMMARY
(NEDERLANDSE SAMENVATTING)

ACHTERGROND

Primaire biliaire cholangitis

Primaire biliaire cholangitis (PBC) is een chronische cholestatische leverziekte[1,2]. De meeste patiënten met PBC hebben baat bij het gebruik van ursodeoxycholzuur (UDCA). Echter, bij een deel van de patiënten heeft UDCA onvoldoende effect. Deze patiënten kunnen leverfalen ontwikkelen. De ontwikkeling van tweedelijns therapieën is dan ook essentieel. Omdat PBC een zeldzame en langzaam progressieve ziekte is, is het opzetten en uitvoeren van studies naar deze nieuwe therapieën lastig. Hier is verandering in gekomen door een studie die heeft aangetoond dat alkalische fosfatase (ALP) en bilirubine onafhankelijk geassocieerd zijn met lange-termijn uitkomsten zoals levertransplantatie en overlijden bij patiënten met PBC. Deze biochemische surrogaatmarkers kunnen gebruikt worden als surrogaateindpunten in studies die het effect van nieuwe therapieën onderzoeken[3]. Deze studies kunnen nu in een kort tijdsbestek opgezet en uitgevoerd worden. Een andere belangrijke recente ontwikkeling is de mogelijkheid om het risico op belangrijke klinische eindpunten als levertransplantatie en overlijden te voorspellen met predictiemodellen[4,5]. Deze predictiemodellen spelen een belangrijke rol in de ontwikkeling van nieuwe therapieën omdat ze de clinicus in staat stellen de transplantatievrije overleving voor individuele patiënten te voorspellen. Ook spelen deze scores een rol in het identificeren van patiënten die mogelijk tweedelijns therapie nodig hebben, of juist vast te stellen dat toevoegen van nieuwe medicatie niet nodig is[4,5].

Primaire scleroserende cholangitis

Primaire scleroserende cholangitis (PSC) is een chronische cholestatische leverziekte die gekarakteriseerd wordt door ontsteking van zowel de intra- als extrahepatische galwegen wat kan lijden tot (gedecompenseerde) levercirrose[6-8]. In tegenstelling tot PBC, zijn er voor PSC geen therapieën die de progressie van de ziekte kunnen afremmen. De ontwikkeling van nieuwe therapieën voor PSC kent een aantal uitdagingen. Ten eerste is de pathogenese niet volledig opgehelderd. Ten tweede ontwikkelen klinisch relevante eindpunten zich vaak niet binnen het tijdsbestek van een studie, waardoor een eventueel effect van nieuwe therapie niet goed te onderzoeken is. Ten derde is PSC een heterogene ziekte met verschillende fenotypen en verschillende klinische relevante eindpunten. Hoewel er voor PSC nog geen surrogaateindpunten zijn geïdentificeerd, zijn er de afgelopen jaren wel belangrijke ontwikkelingen geweest in de risicostratificatie van PSC-patiënten. Dit komt met name door samenwerking tussen verschillende centra. Belangrijke prognostische modellen in PSC zijn de Amsterdam-Oxford score, UK-PSC score en de Primary Sclerosing Cholangitis Risk Estimate Tool (PREsTo). Deze scores helpen clinici om voor individuele patiënten het risico op klinisch relevante uitkomsten in te schatten[9-11].

RISICOSTRATIFICATIE VAN PATIENTEN MET PRIMAIRE BILIAIRE CHOLANGITIS

In de afgelopen jaren hebben verscheidene studies aangetoond dat er een toegenomen incidentie en prevalentie is van PBC[12-20]. Daarnaast suggereren studies dat de klinische presentatie van PBC is veranderd. In het verleden presenteerden patiënten zich veelal met

vergevorderde ziekte waarbij al cirrose was opgetreden. Tegenwoordig worden patiënten gediagnosticeerd in een vroeg en vaak asymptomatisch stadium van de ziekte[21,22]. **Hoofdstuk 2** richt zich op deze veranderingen in patiënt- en ziekte karakteristieken in 4805 PBC patiënten die gediagnosticeerd zijn tussen 1970 en 2014. Onze analyses laten zien dat de gemiddelde leeftijd waarop patiënten gediagnosticeerd worden toeneemt. Daarbij presenteren patiënten zich vaker in een vroeger stadium van de ziekte. Een mogelijke verklaring is dat tegenwoordig meer routine bloedonderzoek wordt gedaan. Hierdoor worden mogelijk meer patiënten gediagnosticeerd met PBC waarbij in het verleden wellicht nooit aanleiding zou zijn geweest om nader onderzoek te doen.

In **hoofdstuk 3** richten we ons verder op patiënten met een milde biochemische ziekte volgens de Rotterdam criteria[23]. In een cohort van 1615 patiënten met een vroeg stadium PBC laten we zien dat binnen 5 jaar na start van UDCA ~46% progressie van de ziekte heeft naar een matig-ernstig stadium van de ziekte. Binnen deze populatie heeft vervolgens 16% progressie naar gevorderde ziekte in de daaropvolgende 5 jaar. Deze veranderingen in biochemisch stadium zijn geassocieerd met de noodzaak tot levertransplantatie en met overlijden. Deze bevindingen benadrukken het belang van goede follow-up van PBC-patiënten. Ook als zij zich presenteren met een vroeg biochemisch stadium van de ziekte bij start van UDCA.

In de afgelopen decennia was UDCA het enige goedgekeurde medicijn voor PBC[24-28]. In 2016 is dit veranderd met de komst van obitocholzuur (OCA) wat ingezet kan worden bij patiënten die onvoldoende responderen op UDCA[29]. De verwachting is dat in de nabije toekomst meer nieuwe tweedelijns therapieën beschikbaar komen[30]. Bij veranderingen in de behandelmogelijkheden van patiënten met PBC is het essentieel dat de clinicus betrouwbaar patiënten met (on)voldoende response kan identificeren; niet alleen kort na start van therapie, maar ook gedurende de daaropvolgende jaren. **Hoofdstuk 4** is een validatiestudie van het gebruik van de GLOBE score gedurende follow-up voor PBC-patiënten die behandeld worden met UDCA. We laten zien dat de GLOBE score betrouwbaar de transplantatie-vrije overleving voorspelt op verschillende tijdspunten in de eerste 10 jaar na start van therapie. Hierbij kan zowel de leeftijd bij start van UDCA gebruikt worden als de leeftijd ten tijde van berekening van de score. Met deze studie wordt de bruikbaarheid van de GLOBE score uitgebreid.

De afgelopen jaren zijn er verschillende predictiemodellen geïntroduceerd voor patiënten met PBC. In **hoofdstuk 5** hebben we de waarde en betrouwbaarheid onderzocht van verscheidene risicoscores die ontwikkeld zijn voor PBC, waaronder de Mayo Risk Score (MRS), Model for End-stage Liver Disease (MELD), GLOBE score[4] en UK-PBC risk score[5]. We laten zien dat de scores vergelijkbare betrouwbaarheid hebben in het voorspellen van klinisch relevante uitkomsten. Welke risicoscore gebruikt wordt in de kliniek is daarom afhankelijk van de klinische context.

RISICOSTRATIFICATIE VAN PATIENTEN MET PRIMAIRE SCLEROSERENDE CHOLANGITIS

In **hoofdstuk 6** onderzochten we de betrouwbaarheid van het Amsterdam-Oxford model (AOM) bij het voorspellen van lange-termijn uitkomsten van patiënten met PSC[9]. Deze studie keek niet alleen naar de betrouwbaarheid van het model ten tijde van diagnose, het moment waar dit model voor gevalideerd is, maar ook naar de betrouwbaarheid wanneer de score op andere momenten werd berekend. We laten zien dat de C-statistiek, als maat voor accuraatheid van de voorspelling, ongeveer 0.70 blijft in de eerste 5 jaar na het stellen van de diagnose PSC. Deze bevindingen rechtvaardigen het gebruik van de AOM in de eerste 5 jaar na diagnose. Hiermee valideert deze studie de AOM voor een langere periode na de diagnose PSC en daarmee vergroot het de bruikbaarheid van dit predictiemodel in de klinische praktijk.

Wanneer patiënten eenmaal eindstadium leverfalen hebben ontwikkeld komen zij in aanmerking voor levertransplantatie. In Nederland wordt de prioritering op de wachtlijst voor levertransplantatie bepaald door de MELD score[31,32]. PSC heeft echter specifieke complicaties zoals recidiverende cholangitis en een hoog risico op het ontstaan van hepatobiliaire maligniteiten. Deze complicaties komen niet goed tot uiting in de MELD[33-37]. Om gelijke toegang tot levertransplantatie te waarborgen krijgen patiënten met PSC regelmatig MELD exceptie (ME) punten[38,39]. In **hoofdstuk 7** laten we zien dat tussen 2006 en 2013 geen van de patiënten met PSC op de wachtlijst die ME punten kreeg overleden is. PSC patiënten met ME punten moesten wel langer wachten op een levertransplantatie. Ze hadden echter wel een hogere kans op levertransplantatie dan patiënten met PSC die enkel op basis van de MELD score waren geprioriteerd en patiënten met een andere oorzaak voor eindstadium leverfalen. Na levertransplantatie hadden patiënten met PSC een betere overleving dan patiënten met een andere oorzaak voor het leverfalen. Deze bevindingen laten zien dat er opnieuw discussie nodig is over hoe patiënten te prioriteren op de wachtlijst voor levertransplantatie.

CONCLUSIE

Tegenwoordig presenteren patiënten met PBC zich op oudere leeftijd en met een gunstiger biochemisch profiel dan in het verleden. Binnen de subgroep van patiënten die zich presenteert met een vroeg biochemisch stadium van de ziekte, zal een substantieel deel alsnog progressie laten zien naar een matig-ernstig of ernstig stadium. Deze progressie is geassocieerd met een slechtere overleving. Patiënten met een nog gunstig biochemisch profiel bij start van behandeling moeten daarom alsnog onderworpen worden aan een goed gestructureerde follow-up met zo nodig aanpassing van medicatie. Gedurende langdurige behandeling met UDCA kan de noodzaak tot aanpassing van medicatie goed worden ingeschat met de GLOBE score. Veranderingen in de GLOBE score gedurende behandeling, met name stijging in deze score, kunnen aanleiding zijn om behandeling opnieuw te evalueren en eventueel tweedelijns therapie te starten. Naast de GLOBE score kunnen ook andere scores, zoals de UK-PBC en MRS betrouwbaar toegepast worden. Welke score men gebruikt hangt dan af van de klinische context. Bij patiënten met PSC

kan de transplantatievrije overleving ingeschat worden met de Amsterdam-Oxford score. Dit kan met redelijke betrouwbaarheid ten tijde van de diagnose van PSC, maar ook in de eerste 5 jaar daarna. Patiënten met PSC die uiteindelijk toch progressie van ziekte hebben, en op basis van de MELD score op de wachtlijst komen voor levertransplantatie worden geplaatst, hebben een goede kans om tijdig getransplanteerd te worden.

REFERENCES

1. Kaplan MM, Gershwin ME. Primary biliary cirrhosis. *N Engl J Med.* 2005;353(12):1261-1273.
2. Poupon R. Primary biliary cirrhosis: a 2010 update. *J Hepatol.* 2010;52(5):745-758.
3. Lammers WJ, van Buuren HR, Hirschfield GM, 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.* 2014;147(6):1338-1349.
4. Lammers WJ, Hirschfield GM, Corpechot C, et al. Development and Validation of a Scoring System to Predict Outcomes of Patients With Primary Biliary Cirrhosis Receiving Ursodeoxycholic Acid Therapy. *Gastroenterology.* 2015;149(7):1804-1812 e1804.
5. Carbone M, Sharp SJ, Flack S, et al. The UK-PBC risk scores: Derivation and validation of a scoring system for long-term prediction of end-stage liver disease in primary biliary cirrhosis. *Hepatology.* 2015:doi: 10.1002/hep.28017.
6. Chapman RW, Arborgh BA, Rhodes JM, et al. Primary sclerosing cholangitis: a review of its clinical features, cholangiography, and hepatic histology. *Gut.* 1980;21(10):870-877.
7. Hirschfield GM, Karlsen TH, Lindor KD, Adams DH. Primary sclerosing cholangitis. *Lancet.* 2013;382(9904):1587-1599.
8. Williamson KD, Chapman RW. Primary sclerosing cholangitis. *Dig Dis.* 2014;32(4):438-445.
9. de Vries EM, Wang J, Williamson KD, et al. A novel prognostic model for transplant-free survival in primary sclerosing cholangitis. *Gut.* 2017.
10. Goode EC, Clark AB, Mells GF, et al. Factors Associated With Outcomes of Patients With Primary Sclerosing Cholangitis and Development and Validation of a Risk Scoring System. *Hepatology.* 2019;69(5):2120-2135.
11. Eaton JE, Vesterhus M, McCauley BM, et al. Primary Sclerosing Cholangitis Risk Estimate Tool (PREsTo) Predicts Outcomes in PSC: A Derivation & Validation Study Using Machine Learning. *Hepatology.* 2018.
12. Boonstra K, Beuers U, Ponsioen CY. Epidemiology of primary sclerosing cholangitis and primary biliary cirrhosis: a systematic review. *J Hepatol.* 2012;56(5):1181-1188.
13. Metcalf JV, Bhopal RS, Gray J, Howel D, James OF. Incidence and prevalence of primary biliary cirrhosis in the city of Newcastle upon Tyne, England. *Int J Epidemiol.* 1997;26(4):830-836.
14. Boonstra K, Kunst AE, Stadhouders PH, et al. Rising incidence and prevalence of primary biliary cirrhosis: a large population-based study. *Liver Int.* 2014;34(6):e31-38.
15. Eriksson S, Lindgren S. The prevalence and clinical spectrum of primary biliary cirrhosis in a defined population. *Scand J Gastroenterol.* 1984;19(7):971-976.
16. Myszor M, James OF. The epidemiology of primary biliary cirrhosis in north-east England: an increasingly common disease? *Q J Med.* 1990;75(276):377-385.
17. Remmel T, Remmel H, Uibo R, Salupere V. Primary biliary cirrhosis in Estonia. With special reference to incidence, prevalence, clinical features, and outcome. *Scand J Gastroenterol.* 1995;30(4):367-371.
18. James OF, Bhopal R, Howel D, Gray J, Burt AD, Metcalf JV. Primary biliary cirrhosis once rare, now common in the United Kingdom? *Hepatology.* 1999;30(2):390-394.
19. Pla X, Vergara M, Gil M, et al. Incidence, prevalence and clinical course of primary biliary cirrhosis in a Spanish community. *Eur J Gastroenterol Hepatol.* 2007;19(10):859-864.
20. Myers RP, Shaheen AA, Fong A, et al. Epidemiology and natural history of primary biliary cirrhosis in a Canadian health region: a population-based study. *Hepatology.* 2009;50(6):1884-1892.

21. Christensen E, Crowe J, Doniach D, et al. Clinical pattern and course of disease in primary biliary cirrhosis based on an analysis of 236 patients. *Gastroenterology.* 1980;78(2):236-246.
22. Locke GR, 3rd, Therneau TM, Ludwig J, Dickson ER, Lindor KD. Time course of histological progression in primary biliary cirrhosis. *Hepatology.* 1996;23(1):52-56.
23. ter Borg PC, Schalm SW, Hansen BE, van Buuren HR, Dutch PBCSG. Prognosis of ursodeoxycholic Acid-treated patients with primary biliary cirrhosis. Results of a 10-yr cohort study involving 297 patients. *Am J Gastroenterol.* 2006;101(9):2044-2050.
24. Lindor KD, Gershwin ME, Poupon R, et al. Primary biliary cirrhosis. *Hepatology.* 2009;50(1):291-308.
25. Poupon RE, Lindor KD, Cauch-Dudek K, Dickson ER, Poupon R, Heathcote EJ. Combined analysis of randomized controlled trials of ursodeoxycholic acid in primary biliary cirrhosis. *Gastroenterology.* 1997;113(3):884-890.
26. Shi J, Wu C, Lin Y, Chen YX, Zhu L, Xie WF. Long-term effects of mid-dose ursodeoxycholic acid in primary biliary cirrhosis: a meta-analysis of randomized controlled trials. *Am J Gastroenterol.* 2006;101(7):1529-1538.
27. Corpechot C, Carrat F, Bahr A, Chretien Y, Poupon RE, Poupon R. The effect of ursodeoxycholic acid therapy on the natural course of primary biliary cirrhosis. *Gastroenterology.* 2005;128(2):297-303.
28. Liver EAftSot. EASL Clinical Practice Guidelines: The diagnosis and management of patients with primary biliary cholangitis. *J Hepatol.* 2017.
29. Nevens F, Andreone P, Mazzella G, et al. A Placebo-Controlled Trial of Obeticholic Acid in Primary Biliary Cholangitis. *N Engl J Med.* 2016;375(7):631-643.
30. Corpechot C, Chazouilleres O, Rousseau A, et al. A Placebo-Controlled Trial of Bezafibrate in Primary Biliary Cholangitis. *N Engl J Med.* 2018;378(23):2171-2181.
31. Malinchoc M, Kamath PS, Gordon FD, Peine CJ, Rank J, ter Borg PC. A model to predict poor survival in patients undergoing transjugular intrahepatic portosystemic shunts. *Hepatology.* 2000;31(4):864-871.
32. Wiesner R, Edwards E, Freeman R, et al. Model for end-stage liver disease (MELD) and allocation of donor livers. *Gastroenterology.* 2003;124(1):91-96.
33. Boonstra K, Weersma RK, van Erpecum KJ, et al. Population-based epidemiology, malignancy risk, and outcome of primary sclerosing cholangitis. *Hepatology.* 2013;58(6):2045-2055.
34. Claessen MM, Vleggaar FP, Tytgat KM, Siersema PD, van Buuren HR. High lifetime risk of cancer in primary sclerosing cholangitis. *J Hepatol.* 2009;50(1):158-164.
35. Bergquist A, Ekbom A, Olsson R, et al. Hepatic and extrahepatic malignancies in primary sclerosing cholangitis. *J Hepatol.* 2002;36(3):321-327.
36. Folseraas T, Boberg KM. Cancer Risk and Surveillance in Primary Sclerosing Cholangitis. *Clin Liver Dis.* 2016;20(1):79-98.
37. Broome U, Olsson R, Loof L, et al. Natural history and prognostic factors in 305 Swedish patients with primary sclerosing cholangitis. *Gut.* 1996;38(4):610-615.
38. Goldberg D, Bittermann T, Makar G. Lack of standardization in exception points for patients with primary sclerosing cholangitis and bacterial cholangitis. *Am J Transplant.* 2012;12(6):1603-1609.
39. Goldberg DS, French B, Thomasson A, Reddy KR, Halpern SD. Current trends in living donor liver transplantation for primary sclerosing cholangitis. *Transplantation.* 2011;91(10):1148-1152.