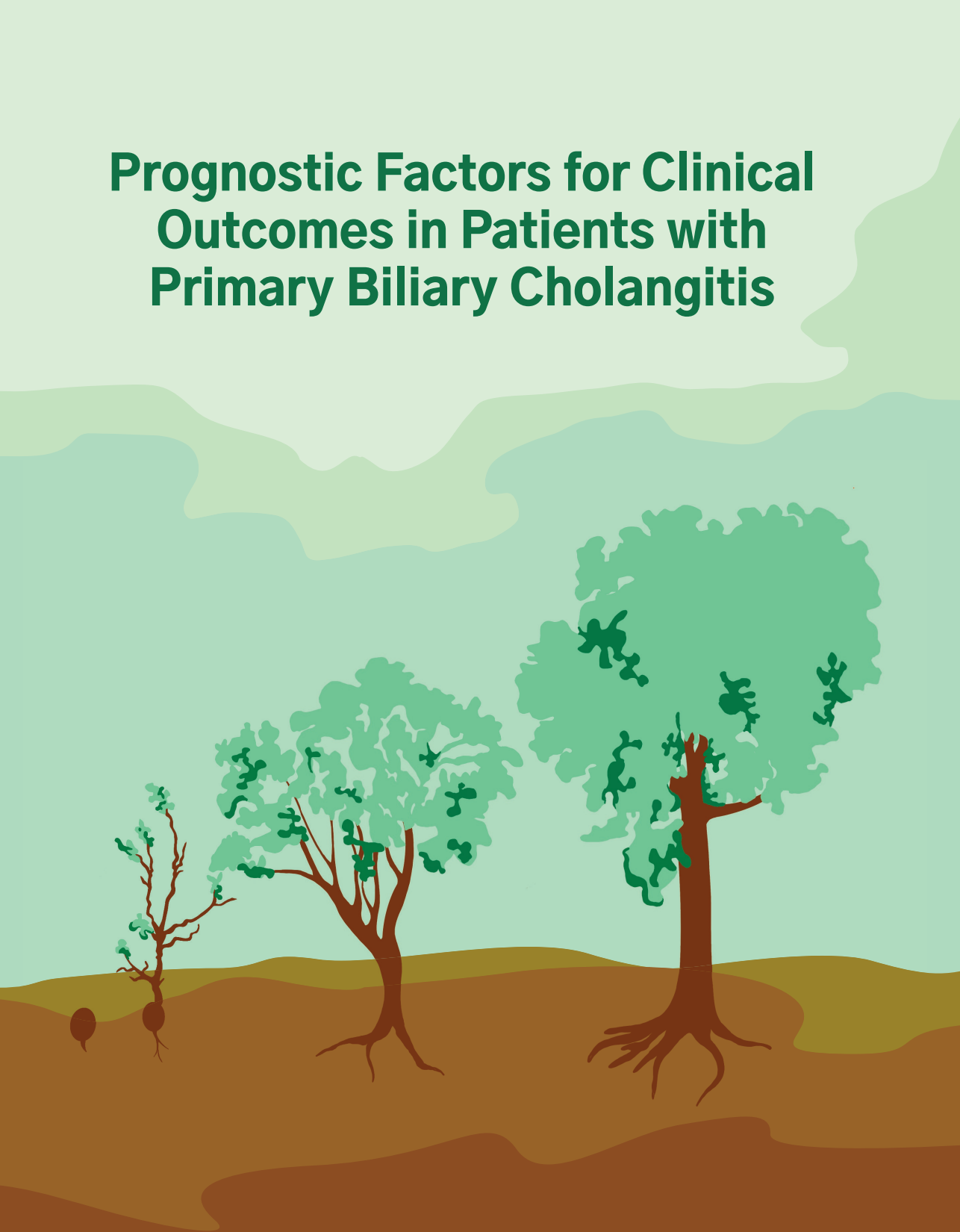
Prognostic Factors for Clinical Outcomes in Patients with Primary Biliary Cholangitis



Carla Fiorella Murillo Perez

Prognostic Factors for Clinical Outcomes in Patients with Primary Biliary Cholangitis

Carla Fiorella Murillo Perez

Copyright © 2021 by C.F. Murillo Perez

All rights reserved. No part of this thesis may be reproduced, distributed, stored in a retrieval system, or transmitted in any form or by any means, without the written permission of the author or, when appropriate, the publisher of the publications.

Cover design, layout, and printing: Proefschriftmaken.nl, Vianen, the Netherlands.

The work presented in this thesis was conducted at the Department of Gastroenterology and Hepatology, Erasmus MC University Medical Center, Rotterdam, the Netherlands and Toronto Centre for Liver Disease, Toronto General Hospital, Toronto, Canada.

The research and printing of this thesis was supported by: Department of Gastroenterology and Hepatology at Erasmus MC, Erasmus University Rotterdam, Toronto Centre for Liver Disease at Toronto General Hospital, PBC Society Canada, Intercept Pharmaceuticals, Cymabay Therapeutics, and Zambon Nederland BV.

Prognostic Factors for Clinical Outcomes in Patients with Primary Biliary Cholangitis

Prognostische factoren voor klinische resultaten bij patiënten met primaire biliaire cholangitis

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus

Prof.dr. F.A. van der Duijn Schouten

and in accordance with the decision of the Doctorate Board.

The public defence shall be held on

Friday, January 8th, 2021 at 13:30

by

Carla Fiorella Murillo Perez
born in Lima, Peru

Erasmus University Rotterdam

Doctoral Committee:

Promotors:	Prof.dr. H.J. Metselaar
	Prof.dr. H.L.A. Janssen
Other members:	Prof.dr. R.A. de Man
	Prof.dr. G.M. Hirschfield
	Prof.dr. U.H.W Beuers
Copromotor:	Dr. B.E. Hansen

TABLE OF CONTENTS

CHAPTER 1

1

General introduction and aims of thesis

Primary biliary cholangitis (PBC) is a chronic autoimmune liver disease that is characterized by immune-mediated destruction of small and medium intrahepatic bile ducts that manifests as cholestasis. Like other autoimmune diseases, it has a female predominance, co-exists with other autoimmune diseases, and presents with disease-specific autoantibodies. It has a slowly progressive course over many years that can result in ductopenia, leading to fibrosis and cirrhosis, and premature death in the absence of liver transplantation. Ursodeoxycholic acid (UDCA) has been the only therapeutic option for patients until recently, when obeticholic acid (OCA) received accelerated approval as a second-line therapy in patients who do not respond to UDCA or who lack tolerance for it.

Epidemiology

The predominant population affected by PBC are middle-aged women, with an estimated 1 in 1,000 women over the age of 40 affected globally.[1] The specific female: male ratios reported in the literature vary by study and region with a mean proportion of female patients of 92% that ranges from 76 to 100%.[2] The incidence of PBC is not limited to any particular ethnic group or geographical region and affects people from all geographical regions. However, reports on the incidence and prevalence differ according to region, with increased prevalence in Caucasian populations, particularly from Northern Europe.[3,4] Accordingly, the prevalence of PBC has been shown to be variable according to age, sex, and race in the United States (US), as demonstrated in a study from the Fibrotic Liver Disease (FOLD) Consortium that reports the highest prevalence in women, White patients, and patients aged 60-70 years.[5] A review on the epidemiology of PBC that included predominantly North American and European studies reported that the annual incidence and prevalence rates per 100,000 individuals ranged from 0.33-5.8 and 1.91-40.2, respectively.[2] While varying rates of incidence and prevalence according to region can be a result of differences in environmental factors, they may also reflect genetics and ethnicity. The prevalence of PBC is reportedly increasing over the years.[2,6] Temporal trends in the incidence of PBC have discrepancies, as some studies report stability while others report increases.[6–8]

There are several plausible explanations for the rise in prevalence and incidence rates for PBC. The increase may be a consequence of improved case ascertainment methods or increased routine testing for liver biochemistry and antimitochondrial antibody (AMA)-positivity driven by increased disease awareness, suggesting an underestimation in earlier studies.[9] Further, a true increase in prevalence can be a result of an increase in survival of PBC patients due to prompt diagnosis and enhanced care. However, one cannot exclude the possibility that environmental or behavioural changes over time may lead to increased exposure to an environmental agent that triggers the disease.[9]

Etiology: Genetics and Environment

Although the specific cause of PBC is unknown, there is consensus that it is likely triggered by a complex interaction between genetic and environmental factors. The role of genetics has been demonstrated in familial and genetic association studies. A high concordance rate of 63% (5 out of 8) for PBC has been reported in monozygotic twins.[10] Further, PBC occurs in 4-9% of family members of patients with PBC, higher than the general population.[11–14] Even in the absence of disease, first-degree relative of patients with PBC have increased AMA-positivity compared to age- and sex-matched controls, with positivity in 7-19% and 0-1%, respectively.[12,14,15]

Initially, genetic studies in PBC relied on biologically plausible genes and those selected based on associations with other autoimmune diseases.[16] From these studies, the primary genes implicated in PBC were the human leukocyte antigen (HLA) genes in the major histocompatibility complex (MHC) region, which are responsible for antigen presentation. In European and North American populations, consistent associations have been identified for HLA DRB1*0801.[17–19] In Japanese populations, associations have been identified for HLA DRB*13.[20] Genome-wide association studies have increased our ability to detect genetic variants common in a population by assessing millions of single nucleotide polymorphisms (SNP), resulting in the identification of additional PBC-associated HLA variants at the DRB1, DQA1, and DQB1 loci, as well as non-HLA loci.[16]

Supporting evidence for a critical role of environmental factors in the development of PBC is the documentation of spatial clustering of PBC cases in Northeast England and Alaska, as well as near toxic waste sites and among atomic bomb survivors from Hiroshima.[8,21–24] Research into an environmental trigger for PBC has not documented a strong correlation to any particular factor. However, there have been associations with recurrent urinary tract infections, active/past smoking, use of hormone replacement therapy, frequent use of nail polish, and hair dye.[13,25–28] Infectious agents have also been implicated, including *Escherichia coli* (E. coli), mycobacteria, *Novosphingobium* aromaticivorans, *Lactobacillus*, *Helicobacter pylori*, human retrovirus, mouse mammary tumour virus, and *Chlamydia pneumoniae*.[27,29]

Clinical Presentation and Symptoms

Like other autoimmune diseases, patients with PBC present with increased rates of co-existing autoimmune diseases that include Sjogrens/sicca syndrome, complete or incomplete CREST syndrome, rheumatoid arthritis, and thyroid disorders, of which the most common one is Sjogrens/sicca syndrome.[25,30–32]

There is variability in the type and severity of symptoms experienced by patients at diagnosis, although the proportion of patients presenting without symptoms has reportedly increased over the years.[33,34] Although an absence of symptoms generally suggests that patients are also presenting at an earlier disease stage, it is not always the case, as some have died before the development of symptoms.[35] In symptomatic patients, some common symptoms include fatigue, pruritus, jaundice, pain in the upper right quadrant, hyperlipidemia, keratoconjunctivitis, steatorrhea, and xerostomia.[32] Of these, the most commonly experienced symptoms are pruritus and fatigue[35], although still variable in prevalence and severity, since younger patients are more prone to fatigue and pruritus and non-Caucasians are more likely to have more severe pruritus.[36,37] The presence of these symptoms, particularly fatigue, imposes a great impact on quality of life even though it does not correlate with disease severity.[38–40] This emphasizes the importance of recognizing and managing health-related quality of life (HRQOL), which can be defined as 'patients' perceptions of their health status, reflecting how they feel and how much their disease affects their way of life' and is commonly measured by the PBC-40 questionnaire.[41,42] This questionnaire was specifically developed for PBC and measures six domains implicated in quality of life: fatigue, emotional, social, cognitive function, general symptoms, and itch.[42]

In early studies of untreated or largely untreated patients, the majority of asymptomatic patients would develop symptoms as the disease progressed.[43,44] One study reported 50% and 95% of patients developed symptoms after 5 and 20 years, respectively.[43] In a Japanese cohort of asymptomatic UDCA-treated patients with a mean follow-up of 5.2 years, only 15% of patients developed liver-related symptoms, in which biochemical response defined as normalization of gamma-glutamyl transpeptidase (GGT) or a reduction ≥70% at 6 months was associated with a decreased risk for symptom development.[45]

Diagnosis of PBC

The diagnosis of PBC is made when a patient fulfills two of the following criteria: i) biochemical evidence of cholestasis with an elevation in ALP for at least 6 months; ii) AMA titers above 1:40; iii) a liver biopsy with evidence of non-suppurative cholangitis and destruction of small/medium-sized bile ducts.[1] Diagnosis is mainly based on cholestatic liver biochemistry and AMA, as liver biopsies are less frequent nowadays, except when a patient lacks autoantibodies or doesn't demonstrate biochemical abnormalities.[46] In AMA-negative patients, a diagnosis can be suggested if ANA autoantibodies are detected, such as anti-gp210 or anti-sp100.[1,47]

Biochemical, serological, and histological features of PBC

Liver biochemistry

Early biochemical markers of cholestasis include elevations in ALP and GGT, supporting the inclusion of ALP in diagnostic criteria. Elevations in GGT can confirm the hepatic origin of ALP elevation and can usually be detected prior to elevations in ALP.[1] The magnitude of elevation of ALP is correlated with severity of ductopenia and inflammation, as well disease progression.[1,48] Patients can also demonstrate mild elevations in transaminases (ALT and AST) and elevations of immunoglobulins, primarily IgM. Elevations of transaminases reflects the extent of liver parenchyma inflammation and necrosis.[1,48] Later in the course of PBC, increases in conjugated bilirubin, alterations in prothrombin time, and decreases in serum albumin are observed. Hyperbilirubinemia reflects the severity of ductopenia and biliary piecemeal necrosis.[48]

Serology

The serologic hallmark for PBC is AMA given its presence in 90-95% of patients. There is no difference in biochemical, histological, and clinical features at presentation or response to treatment between AMA- positive and -negative patients.[49,50] The autoantigens of AMA correspond to the family of 2-oxo acid dehydrogenase complexes, termed M2, that are localized to the inner mitochondrial membrane.[51,52] This family of homologous enzyme complexes includes pyruvate dehydrogenase complex, branched-chain oxo acid dehydrogenase complex, and oxoglutarate dehydrogenase complex.[51] The main autoantigen of AMA is the E2 subunit of 2-oxo acid dehydrogenase complexes, for which 80-90% of sera react to E2 from pyruvate dehydrogenase, specifically its lipoic acid binding site (autoepitope).[53] Historically, the detection of AMA antibodies was predominantly performed with indirect immunofluorescence, however there has been a shift towards methods that provide greater sensitivity and specificity, as well as greater speed and automation, such as enzyme-linked immunosorbent assay (ELISA) and western immunoblots.[54,55]

Antinuclear antibodies are another class of autoantibodies that can be found in the context of PBC. They demonstrate high specificity for PBC (99%) and can be detected in 50-70% of patients.[47,56] ANAs are more frequently observed in AMA-negative patients.[49] The nuclear envelope contains the autoantigen for ANA, which yields multiple nuclear dot (ex. anti-sp100) or Rim-like/membranous patterns (ex. anti-gp210) by indirect immunofluorescence.[47]

Histology

Histologically, PBC is characterized by chronic non-suppurative inflammation of the portal sites and immune-mediated destruction of bile ducts. Although a liver biopsy is no longer essential for a diagnosis of PBC and sparely carried out due to its invasive nature, it can aid in staging of the disease. There are four histologic stages, which are mainly staged with Ludwig or Scheuer systems. According to Scheuer system, Stage I is defined by portal hepatitis with duct lesions (florid duct lesion), stage II is defined by periportal hepatitis with ductular proliferation, stage III is defined by septal fibrosis, and stage IV is defined by cirrhosis.[57] According to Ludwig staging, Stage I is defined by portal hepatitis, stage II is defined by periportal hepatitis, Stage III is defined by bridging necrosis or septal fibrosis or both, and Stage IV is defined by cirrhosis.[58] In the absence of therapy, a patient progresses histologically within 2 years, with progression rates of 62% for stage I/II and 50% for stage III.[59]

Since liver biopsies are invasive, there have been efforts to develop non-invasive methods to assess fibrosis. Non-invasive biochemical markers include AST/ALT ratio, AST to platelet ratio index (APRI), FIB-4, and ELF test. A promising non-invasive method of assessing liver fibrosis is transient elastography, which measures liver stiffness and has demonstrated high correlations with histologic fibrosis stage in PBC, with an AUROC of 0.89 for F>2 and 0.96 for F=4.[60]

Loss of self-tolerance

The trigger for loss of self-tolerance and the mechanism by which PDC-E2 becomes antigenic is not fully understood. There are various plausible mechanisms: molecular mimicry, self-alteration of PDC-E2 by xenobiotics, and the apoptotic mechanism of biliary epithelial cells that releases intact immunogenic epitopes.[14,61] There is experimental evidence for molecular mimicry between a self-antigen and an exogenous bacterial/virologic antigen for PDC-E2, as cross-reactivity was detected between human PDC-E2 and bacterial E2 from *E. coli*.[62] In the second plausible mechanism, the lipoic acid bound to E2 is replaced with a chemical xenobiotic mimic, thereby altering the host and initiating an autoimmune reaction. An experimental study tested the reactivity of more than 100 potential xenobiotics bound to PDC-E2 with the sera of PBC patients and found that nine had increased reactivity as compared to the sera of controls, as well as the native form of PDC-E2. One of the xenobiotic identified is a chemical that is largely used in cosmetics, 2-octynoic acid.[63] The last method is related to the unique apoptotic process in biliary epithelial cells, in which an intact PDC-E2 remains in the apoptotic bleb, which can also explain the targeted immune reaction to the biliary epithelial cells despite the ubiquitous distribution of PDC-E2.[64]

Pathogenesis

The specific pathogenic role of AMA, the serologic hallmark of PBC, remains to be clearly defined. Evidence suggests that it may play a role in the disease process due to its ability to inhibit the enzymatic activity of PDC and the ability of IgA AMA to undergo transcytosis in biliary epithelial cells, potentially predisposing them to apoptosis.[65,66] The loss of biliary epithelial cells is hypothesized to be carried out by autoreactive CD8+ and CD4+ T cells reacting to PDC-E2 infiltrating the portal tracts, which can also be detected at lower quantities in the peripheral blood and portal lymph nodes of patients.[24,65,67]

Another contributor to the pathogenesis of PBC is the biliary HCO_3^- umbrella hypothesis. The cholangiocyte membranes are protected by an apical alkaline barrier that is established by the secretion of bicarbonate into the bile duct lumen. This maintains bile salts in a polar state and thus unable to cross the membrane. In PBC, anion exchanger 2 (Cl^-/HCO_3^- exchanger) and type III inositoltriphosphate receptor are defective, which results in the barrier being compromised and resulting in partial protonation of bile salts.[68] Consequently, the bile salts are rendered apolar and gain the ability to cross the cholangiocyte membrane, thereby inducing apoptosis and senescence.

Complications

As with other liver diseases, portal hypertension is a potential complication for PBC, although it predominantly affects patients with cirrhosis. Before the introduction of UDCA, the prevalence of esophageal varices over a median 5.6 years was reported as 31% in a prospective study.[69] The development of esophageal varices was associated with a higher mortality risk in the same study, as the 1- and 3-year survival estimates were 83% and 59%, respectively. Ascites and hepatic encephalopathy are also complications that can be observed.

Hepatocellular carcinoma (HCC) is another complication that can arise. It is observed at less frequent rates compared to portal hypertension, with rates of 0.7-3.6% in patients followed for 3.6-6.8 years.[70,71] However, patients with cirrhosis are at an increased risk for HCC, as well as those with older age, male sex, history of blood transfusions, and signs of portal hypertension.[70–72] Furthermore, the development of HCC is associated with worse transplant-free survival and overall survival.[72]

Natural history

PBC is highly variable in terms of presentation, but also with regards to the disease course. Generally, patients with PBC have a diminished survival compared to age- and sex-matched individuals, which has been demonstrated in various patient populations. In the UK, it was demonstrated that untreated PBC patients had a 2.7-fold increase in adjusted mortality compared to the general population.[73] In a geographically defined cohort from Northeast England of prevalent cases from 1987 and 1994 and of whom 37% received treatment, the standard mortality ratio (SMR) was 2.87 and the 10-year survival was approximately 45%. Interestingly, patients demonstrated an increased mortality rate even when only considering liver-unrelated deaths, as the SMR was 1.73.[35] A Canadian population-based study from 1996 to 2002 reported the same SMR of 2.87 and a 10-year transplant-free survival rate of 68%, although the patient's treatment state was largely undefined.[74]

First-line Treatment

Ursodeoxycholic acid (3a, 7B-dihydroxy-5b-cholanic acid, UDCA), an endogenous bile acid that normally represents a minority (3%) of the bile acid pool, is the standard treatment for PBC and is required as life-long treatment.[75,76] It was approved by the Food and Drug Administration (FDA) in 1997 for the treatment of PBC. There are three mechanisms of action through which UDCA is thought to exert its effects.[76] First, the hydrophilic nature of UDCA protects cholangiocytes by reducing the cytotoxicity of bile and possibly reducing the concentration of hydrophobic bile acids as it becomes the predominant bile acid (40-50%). Secondly, UDCA can aid in the stimulation of hepatobiliary secretion. Third, it can inhibit the mitochondrial membrane permeability transition (MMPT) and thus prevent bile acid-induced apoptosis of hepatocytes, which it may also achieve through the stimulation of the survival pathway. The recommended dosage for UDCA is 13-15mg/kg per day. It has been demonstrated that 13-15mg/kg and 23-25mg/kg result in greater improvements in ALP and AST, compared to 5-7mg/kg, but 23-25mg/kg was not superior.[77] Furthermore, appropriate dosage can improve the rates of response in patients who were initially non-responders.[78]

The efficacy of UDCA was demonstrated in several large, randomized, double-blind, placebo-controlled trials, all of which reported that UDCA improved liver biochemistry markers, including ALP, aminotransferases, bilirubin, cholesterol, and IgM as early as 3 months from the start of treatment.[79–81] The Canadian trial demonstrated that treatment with UDCA for 2 years improved histological features, but had no impact on symptom, liver transplantation or death.[79] In the French trial, it was demonstrated that treatment with UDCA for 4 years slowed progression, as defined by hyperbilirubinemia, ascites, variceal bleeding, or encephalopathy,

and reduced the need for liver transplantation, and improved transplant-free survival.[81] The American trial demonstrated that UDCA treatment for 2 years was associated with a delay in progression, albeit had no impact on symptoms, histology, liver transplantation or survival.[80]

Research outside of the scope of clinical trials has demonstrated that UDCA delays histological progression specially in those with an early stage.[82,83] Furthermore, UDCA has shown to delay the onset of esophageal varices, as the incidence of new varices at 4 years was 16% in UDCA-treated patients, but 58% in the placebo group.[84] A meta-analysis on the impact of UDCA on pruritus or fatigue reported negative results.[83] Further, whether UDCA can have an effect on liver transplantation or patient survival has been debated, mainly due to inconsistent results. A recent study performed by the Global PBC Study Group demonstrated a lower risk for liver transplantation or death in patients receiving UDCA as compared to those who did not receive UDCA (HR=0.46, 95% CI 0.40-0.52, P<0.001), irrespective of disease stage.[85] The benefit of UDCA was also observed in those who did not achieve complete biochemical response. These results highlight the importance of using UDCA as first-line therapy for all patients with PBC.

Biochemical Response

The aim of treatment is to ultimately improve long-term clinical outcomes, for which early detection is hindered by the slow progressive nature of PBC. Therefore, the efficacy of treatment has been largely determined through an assessment of liver biochemistry, as these are the first to be altered with treatment and the fact that they are associated with clinical outcomes. There are various criteria based on liver biochemistry that have been developed in order to determine 'response' to treatment that were developed based on their association with clinical outcomes. The first study to demonstrate an association between biochemical parameters and clinical outcomes was from the Mayo clinic, in which patients with ALP <2 × ULN at 6 months were more likely to respond favourably over a 2-year period.[86] This study was followed by a study from Barcelona that proposed response be assessed at 1 year and defined by a reduction of ALP greater than 40% from baseline, or normal levels at 1 year.[87] Patients who met these criteria had transplant-free survival similar to a control population. Further, an absence of response according to Paris-I criteria (**Table 1**) was an independent predictor of liver transplantation or death.[88] Toronto response was based on the risk for progressive liver damage, as patients who did not achieve ALP<1.67×ULN at 2 years tended to have a one-stage progression in histology during extended follow-up.[89] Distinct Paris criteria were defined for patients with early stage, defined by an early histologic stage or normal albumin and bilirubin, named Paris-II (**Table 1**).[90]

While most criteria require assessment at 1 year, it has been proposed that early response identification can be done as early as 6 months, with same or higher positive and negative predictive values.[91] This is important being that patients who do not respond to UDCA will have a delay in effective therapy if response cannot determined until 1 year. To address the need for timely assessment of response, a score that employs pre-treatment parameters associated with the probability of response was developed in 2703 patients and validated in 460 patients. The goal was to predict response to treatment, which was defined as ALP<1.67×ULN at 1 year. The parameters included in the score were ALP at diagnosis, bilirubin at diagnosis, aminotransferase at diagnosis, age, the time interval between diagnosis and the start of UDCA treatment, and the absolute difference in ALP from diagnosis.[92]

The biochemical parameters included in these criteria are consistent with one another, as the majority either include ALP or bilirubin. In fact, both ALP and bilirubin have been strongly associated with long-term outcomes and deemed to be 'reasonably likely to predict clinical benefit', which has ultimately led to their inclusion as surrogate end points in clinical trials for novel therapies in PBC.[93]

Table 1. Response criteria associated with clinical outcome

Response	Time of assessment	Criteria
Rochester, 1999[86]	6 months	ALP<2×ULN and/or Mayo risk score <4.5
Barcelona, 2006[87]	1 year	>40% decrease in ALP from baseline or normal levels
Paris-I, 2008[88]	1 year	ALP ≤3×ULN, AST≤2×ULN, and bilirubin≤1mg/dL
Rotterdam, 2009[94]	1 year	Normal bilirubin and albumin given at least one was abnormal at baseline
Toronto, 2010[89]	2 years	ALP<1.67×ULN
Paris-II, 2011[90]	1 year	ALP≤1.5×ULN, AST≤1.5×ULN, and normal bilirubin
Ehime, 2011[95]	6 months	Normal GGT or ≥70% decrease in GGT

ALP, alkaline phosphatase; ULN, upper limit of normal; AST, aspartate aminotransferase; GGT, gamma-glutamyl transpeptidase.

Second-line therapies

Response to UDCA is variable and approximately 30-40% of all patients who are treated do not demonstrate an adequate response and are therefore at continued risk for disease progression and complications of PBC.[87,88] Further, even though UDCA is well tolerated by

most patients, poor tolerance has been reported in up 9% of patients.[96] Thus, there is still a need for additional therapies to become available.

Immunosuppressive therapies

Given the autoimmune nature of PBC, various immunosuppressants have been evaluated for the treatment of PBC, including methotrexate, cyclosporine, colchicine, azathioprine, and colchicine.[97–101] However, these trials did not support their use in PBC as major side effects were reported or they failed to demonstrate any benefit on liver biochemistry, histology, or survival.

Farsenoid X receptor agonists

Farsenoid X receptor (FXR) is a nuclear receptor that regulates the expression of genes essential for bile acid homeostasis. Chenodeoxycholic acid is the most potent endogenous FXR agonist.[102] Obeticholic acid (OCA) is a semi-synthetic analogue of chenodeoxycholic acid that activates FXR with 100× greater strength.[102] In 2016, OCA obtained FDA approval and became the first available therapeutic agent for PBC since the introduction of UDCA. This was prompted by the results from the PBC OCA International Study of Efficacy (POISE) phase III trial, which was a 12-month, double-blind randomized controlled trial.[103] In this trial, 216 patients with inadequate response to UDCA defined by ALP ≥1.67×ULN or abnormal bilirubin were randomized to receive OCA as adjuvant therapy or as monotherapy in patients that could not tolerate UDCA. There were three arms, placebo, 5-10mg OCA, and 10mg OCA. The endpoint of the study was a composite of ALP<1.67×ULN, a reduction in ALP of at least 15%, and normal bilirubin. The endpoint was reached by 46% and 47% of those in the 5mg and 5-10mg arms, compared to 10% in the placebo arm. Notedly, pruritus was a common side effect of OCA that was dose-dependent and could affect up to 68% of patients in the highest dose group.

The open-label extension study of OCA has demonstrated its long-term efficacy and safety, as ALP levels remains significantly lower throughout the duration of OCA, up to 4 years.[104] Pruritus and fatigue were observed in 77% and 33% of patients, respectively. The estimated survival benefit from OCA has been evaluated with the GLOBE and UK-PBC risk scores, suggesting that in comparison to placebo, patients treated with OCA (10mg) have a 26% (GLOBE) and 37% (UK-PBC) relative reduction from baseline to 1 year in the 10-year risk for liver transplantation or death.[105] The reduction in risk was nevertheless noted in patients who did not meet the primary endpoint for POISE. Whether OCA provides true survival benefit is yet to be determined in the phase IV trial (COBALT).

Fibrates

Fibrates act as ligands and exert their effects on the nuclear receptor peroxisome proliferator-activated receptor (PPAR). The receptor can exist in three isoforms: PPAR-α, PPAR-δ, PPAR-γ.[106] The first open-label study of bezafibrate, a non-selective PPAR-agonist, in PBC demonstrated that this therapeutic agent alone or adjuvant to UDCA can reduce ALP and IgM levels and improve symptoms.[107] Since then, there have been numerous studies involving bezafibrate or fenofibrate, all of which suggest similar findings.[107–114] However, the most promising findings arise from the BEZURSO trial, the first large, phase III, placebo-controlled trial of bezafibrate in combination with UDCA.[115] Patients with an incomplete response to UDCA according to Paris-II criteria were eligible for inclusion, in which 100 patients were randomized in a 1:1 ratio to receive 400mg/day of bezafibrate or placebo for 24 months.[115] The primary endpoint of the study was complete biochemical response, defined by normal bilirubin, ALP, aminotransferases, albumin, and prothrombin index, and was met by 31% of patients that received bezafibrate. Additionally, 67% achieved ALP normalization. In contrast, none of the patients in the placebo arm met the primary endpoint and only 2% achieved ALP normalization. Progression of liver stiffness was hindered in the treatment arm, as liver stiffness measures decreased 15% from baseline but increased 22% in the placebo group, all while improving pruritus and fatigue. Further research is needed to determine the impact of bezafibrate on long-term clinical outcomes. In a Japanese cohort the estimated survival benefit of bezafibrate was evaluated in 118 patients that received UDCA for at least one year and subsequent combination therapy with bezafibrate for at least another year.[116] The addition of bezafibrate was associated with a significant reduction in the GLOBE score as well as improved predicted transplant-free survival compared to pre-combination therapy. Further, in patients with normal bilirubin before the introduction of bezafibrate, combination therapy was associated with reduced risk for liver transplantation or liver-related death.[116]

Liver transplantation

In patients who reach end-stage liver disease, liver transplantation is the sole treatment option that can improve quality of life and patient survival.[117] Still, the transplantation burden for PBC has reduced in recent years as the proportion of liver transplantations attributed to PBC and the absolute number has decreased in Europe and the United States.[118–121] For example, from 1995 to 2006, there was a reduction in the absolute number of liver transplantations for PBC in the United States with an average decrease of 5.4 cases per year in spite of the increase of 249 transplants per year.[120] In a study of the European Liver Transplantation Registry, the proportion of liver transplantations for PBC decreased from 1986 to 2015 from 20% to 4%. [118]

Recurrence of PBC after liver transplantation is not uncommon and rates can range from 17% to 53%, which may be due to differences in patient population and follow-up time across studies.[119,122,123] Given the persistence of AMA after liver transplantation and the possibility for normal liver biochemistry, a diagnosis of recurrent PBC can be made when a biopsy shows evidence of histological features consistent with a florid duct lesion.[1] Although it has been shown that recurrent PBC can progress to cirrhosis in up to 15% of patients, earlier studies failed to demonstrate an impact on graft or patient survival.[124,125] A recent study from the GLOBAL PBC Study Group of 785 patients from North America and Europe showed a significant time-dependent association between recurrent PBC and graft loss (HR=2.01, 95% CI 1.16-3.51) and death (HR=1.72, 95% CI 1.11-2.65).[122] The factors found to be associated with recurrent PBC in this study were younger age at diagnosis and liver transplantation, tracolimus use, and biochemical markers of cholestasis 6 months after liver transplantation (bilirubin≥100μmol or ALP >3×ULN).

Prediction of response and clinical outcomes

Risk stratification of patients with PBC is important to determine the need for specialty care, vigilance, second-line therapies, and timing of liver transplantation, which can all be based on patient characteristics, as well as markers of disease severity. One of the first predictors of prognosis in patients with PBC that was applicable to all stages of disease was histologic stage. Multiple studies have demonstrated that an advanced histologic stage is associated with an increased risk for liver transplantation or death.[126–128] Recognition of the prognostic value of histology in PBC in conjunction with the decreased rate of biopsies has prompted the use of non-invasive markers for fibrosis. One of these markers is APRI, whose association with transplant-free survival has been demonstrated and a threshold of 0.54 was established for use at baseline and 1 year, which suggests that values that surpass this threshold are associated with worse prognosis.[128] Liver stiffness assessed by transient elastography has been associated with decompensation, liver transplantation, and death, with superiority to non-invasive biochemical markers in diagnosing significant fibrosis, severe fibrosis, or cirrhosis.[129]

Demographic factors, such as age and sex have been associated with response to UDCA and prognosis. Male sex has been proposed as being independently associated with decreased response to UDCA and increased mortality.[36,74] Meanwhile, older age has been shown to be an independent predictor for higher response to UDCA according to Paris-II criteria and increased mortality.[36,127] The impact of age, however, should be analyzed in comparison to an age- and gender-matched population since in older patients, mortality is often unrelated to PBC. In a study of asymptomatic patients, similar mortality rates between patients above 55 year old and an age- and gender-matched population were reported.[130]

Hyperbilirubinemia, a reflection of the severity of ductopenia and biliary piecemeal necrosis[131], is one of the main predictors of prognosis in untreated and treated patients. An early study of untreated patients demonstrated that there is a period of rather stable bilirubin followed by a rapid rise in bilirubin of 2.5mg/dl per year that results in death.[132] In the context of UDCA, patients who achieve normalization of bilirubin at 6 months have improved transplant-free survival compared to those without normalization.[133] Multiple studies have confirmed the predictive value of elevated bilirubin on transplant-free survival.[81,93,126,128] Bilirubin, along with albumin, is included in a three-tiered biochemical staging based on the finding that albumin and bilirubin were consistently associated with survival: early (normal bilirubin and albumin), moderately advanced (abnormal bilirubin or albumin), and advanced (abnormal bilirubin and albumin).[134] Another major liver parameter often used is ALP, which has been associated with liver transplant-free survival, with increased predictive value when combined with bilirubin.[93] Accordingly, bilirubin and ALP, are some of the most common liver parameters included in response criteria (**Table 1**).

One of the major limiting factors in the development of novel therapies for the treatment of PBC is its slowly progressive nature, which would require long-term follow-up to determine if a therapeutic agent influences clinical outcomes. Thus, surrogate endpoints such as ALP and bilirubin have been of great value, particularly for their convenience and non-invasive nature.

Risk scores

While response criteria are a simple way to determine prognosis, there is a loss in predictive value due to the dichotomization of continuous variables. In order to improve prognostic performance, various risk scores that culminate several variables have been developed specifically for PBC (**Table 2**). One of the earliest risk scores that employed non-invasive measurements to predict survival is the Mayo risk score. It was developed in 1989 from 312 untreated PBC patients to predict survival up to 7 years, with the intended application for selecting patients for liver transplantation and its timing.[135] The model was subsequently updated to predict short-term survival at 2 years and for use at any time during follow-up.[136]

More recent models include the GLOBE score and UK-PBC, which were developed in UDCA-treated patients. The development of the GLOBE score was conducted in globally representative cohort to predict transplant-free survival with values collected at 1 year, although it can also be used with values collected from 2-5 years.[137] The performance of the GLOBE score was superior to that of binary response criteria. The UK-PBC risk score can be used to predict the risk of liver transplantation and liver-related death at 5, 10, and 15 years.[138] The UK-PBC risk score has been validated in a cohort from the United States with excellent

discrimination.[139] Further, both scores have been validated in Chinese and Korean patient populations.[140,141] Similarly, their performance was validated in a cohort that included centers from Europe, US, and Canada for the prediction of cirrhosis-related complications, with similar and excellent prognostic performance between the scores but improved compared to published response criteria.[142]

Table 2. Risk scores developed for primary biliary cholangitis

Prognostic score	Year	Purpose	Variables
Mayo Risk score[135]	1989	Predict survival, up to 7 years and select/timing patients for liver transplantation	• Age • Total bilirubin (mg/dl) • Albumin (g/dl) • Prothrombin time • Severity of edema
Mayo Risk score[136]	1994	Predict short-term survival (2 years)	• Age • Total bilirubin (mg/dl) • Albumin (g/dl) • Prothrombin time • Severity of edema
GLOBE score[137]	2015	Predict transplant-free survival at 3, 5, 10, and 15 years	• Age at UDCA start • Bilirubin (×ULN) at 1 year • ALP (×ULN) at 1 year • Albumin (×LLN) at 1 year • Platelet count (×10^9/L) at 1 year
UK-PBC score[138]	2016	Predict liver transplantation and liver-related death at 5, 10 and 15 years	• ALP (×ULN) at 1 year • AST or ALT (×ULN) at 1 year • Bilirubin (×ULN) at 1 year • Albumin (×LLN) at baseline • Platelet count (×LLN) at baseline

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

AIMS OF THESIS

This thesis will utilize a large and globally representative cohort of patients with long-term follow-up to study patients with PBC. In **Chapters 2 and 3**, the aim is to describe temporal and spatial trends in PBC with regards to patient and disease characteristics and evaluate whether there are differences in clinical outcomes of patients according to calendar time or geographical region. **Chapters 4**, **5 and 6** aim to identify clinically relevant and important factors for risk stratification in PBC through the evaluation of individual prognostic factors as well as established risk scores that predict outcome. In **Chapter 7 and 8**, the aim is to optimize patient management, and thereby survival, through an establishment of care pathways for the need for referral and optimal biochemical treatment targets.

REFERENCES

1. 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.

2. Boonstra K, Beuers U, Ponsioen CY. Epidemiology of primary sclerosing cholangitis and primary biliary cirrhosis: a systematic review. J Hepatol 2012;56:1181-1188.

3. Metcalf J, James O. The geoepidemiology of primary biliary cirrhosis. Semin Liver Dis 1997;17:13-22.

4. Howel D, Fischbacher CM, Bhopal RS, et al. An exploratory population-based case-control study of primary biliary cirrhosis. Hepatology 2000;31:1055-1060.

5. Lu M, Li J, Haller I V, et al. Factors Associated With Prevalence and Treatment of Primary Biliary Cholangitis in United States Health Systems. Clin Gastroenterol Hepatol 2018;16:1333-1341.e6.

6. Lu M, Zhou Y, Haller I V, et al. Increasing Prevalence of Primary Biliary Cholangitis and Reduced Mortality With Treatment. Clin Gastroenterol Hepatol 2018;16:1342-1350.e1.

7. 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:e31-e38.

8. McNally RJ, James PW, Ducker S, et al. No rise in incidence but geographical heterogeneity in the occurrence of primary biliary cirrhosis in North East England. Am J Epidemiol 2014;179:492-498.

9. Prince MI, James OF. The epidemiology of primary biliary cirrhosis. Clin Liver Dis 2003;7:795-819.

10. Selmi C, Mayo MJ, Bach N, et al. Primary biliary cirrhosis in monozygotic and dizygotic twins: genetics, epigenetics, and environment. Gastroenterology 2004;127:485-492.

11. Bach N, Schaffner F. Familial primary biliary cirrhosis. J Hepatol 1994;20:698-701.

12. Hayase Y, Iwasaki S, Akisawa N, et al. Similar anti-mitochondrial antibody reactivity profiles in familial primary biliary cirrhosis. Hepatol Res 2005;33:33-38.

13. Corpechot C, Chrétien Y, Chazouillères O, et al. Demographic, lifestyle, medical and familial factors associated with primary biliary cirrhosis. J Hepatol 2010;53:162-169.

14. Lazaridis KN, Juran BD, Boe GM, et al. Increased prevalence of antimitochondrial antibodies in first-degree relatives of patients with primary biliary cirrhosis. Hepatology 2007;46:785-792.

15. Zografos TA, Gatselis N, Zachou K, et al. Primary biliary cirrhosis-specific autoantibodies in first degree relatives of Greek primary biliary cirrhosis patients. World J Gastroenterol 2012;18:4721-4728.

16. Gulamhusein AF, Juran B., Lazaridis K. Genome-Wide Association Studies in Primary Biliary Cirrhosis. Semin Liver Dis 2015;35:392-401.

17. Donaldson PT, Baragiotta A, Heneghan MA, et al. HLA class II alleles, genotypes, haplotypes, and amino acids in primary biliary cirrhosis: a large-scale study. Hepatology 2006;44:667-674.

18. Mullarkey ME, Stevens AM, McDonnell WM, et al. Human leukocyte antigen class II alleles in Caucasian women with primary biliary cirrhosis. Tissue Antigens 2005;65:199-205.

19. Begovich AB, Klitz W, Moonsamy P V, et al. Genes within the HLA class II region confer both predisposition and resistance to primary biliary cirrhosis. Tissue Antigens 1994;43:71-77.

20. Onishi S, Sakamaki T, Maeda T, et al. DNA typing of HLA class II genes; DRB1*0803 increases the susceptibility of Japanese to primary biliary cirrhosis. J Hepatol 1994;21:1053-1060.

21. Abu-Mouch S, Selmi C, Benson GD, et al. Geographic clusters of primary biliary cirrhosis. Clin Dev Immunol 2003;10:127-131.

22. Prince MI, Chetwynd A, Diggle P, et al. The geographical distribution of primary biliary cirrhosis in a well-defined cohort. Hepatology 2001;34:1083-1088.

23. Ala A, Stanca CM, Bu-Ghanim M, et al. Increased prevalence of primary biliary cirrhosis near Superfund toxic waste sites. Hepatology 2006;43:525-531.

24. Kita H, He XS, Gershwin ME. Autoimmunity and environmental factors in the pathogenesis of primary biliary cirrhosis. Ann Med 2004;36:72-80.

25. Parikh-Patel A, Gold EB, Worman H, et al. Risk factors for primary biliary cirrhosis in a cohort of patients from the United States. Hepatology 2001;33:16-21.

26. 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;42:1194-1202.

27. Mantaka A, Koulentaki M, Chlouverakis G, et al. Primary Biliary Cirrhosis in a genetically homogeneous population: Disease associations and familial occurrence rates. BMC Gastroenterol 2012;12:110.

28. Prince MI, Ducker SJ, James OF. Case-control studies of risk factors for primary biliary cirrhosis in two United Kingdom populations. Gut 2010;59:508-512.

29. Abdulkarim AS, Petrovic LM, Kim WR, et al. Primary biliary cirrhosis: an infectious disease caused by Chlamydia pneumoniae? J Hepatol 2004;40:380-384.

30. Tsianos E V, Hoofnagle JH, Fox PC, et al. Sjögren's syndrome in patients with primary biliary cirrhosis. Hepatology 1990;11:730-734.

31. Reynolds TB, Denison EK, Frankl HD, et al. Primary biliary cirrhosis with scleroderma, Raynaud's phenomenon and telangiectasia. New syndrome. Am J Med 1971;50:302-312.

32. Siegel JL, Luthra H, Donlinger J, et al. Association of primary biliary cirrhosis and rheumatoid arthritis. J Clin Rheumatol 2003;9:340-343.

33. 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.

34. Baldursdottir TR, Bergmann OM, Jonasson JG, et al. The epidemiology and natural history of primary biliary cirrhosis: a nationwide population-based study. Eur J Gastroenterol Hepatol 2012;24:824-830.

35. Prince M, Chetwynd A, Newman W, et al. Survival and symptom progression in a geographically based cohort of patients with primary biliary cirrhosis: follow-up for up to 28 years. Gastroenterology 2002;123:1044-1051.

36. 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;144:560-569.e7; quiz e13-4.

37. Levy C, Naik J, Giordano C, et al. Hispanics with primary biliary cirrhosis are more likely to have features of autoimmune hepatitis and reduced response to ursodeoxycholic acid than non-Hispanics. Clin Gastroenterol Hepatol 2014;12:1398-1405.

38. Cauch-Dudek K, Abbey S, Stewart DE, et al. Fatigue in primary biliary cirrhosis. Gut 1998;43:705-710.

39. Huet PM, Deslauriers J, Tran A, et al. Impact of fatigue on the quality of life of patients with primary biliary cirrhosis. Am J Gastroenterol 2000;95:760-767.

40. Newton JL, Davidson A, Kerr S, et al. Autonomic dysfunction in primary biliary cirrhosis correlates with fatigue severity. Eur J Gastroenterol Hepatol 2007;19:125-132.

41. Poupon RE, Chrétien Y, Chazouillères O, et al. Quality of life in patients with primary biliary cirrhosis. Hepatology 2004;40:489-494.

42. 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;54:1622-1629.

43. 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.

44. Balasubramaniam K, Grambsch PM, Wiesner RH, et al. Diminished survival in asymptomatic primary biliary cirrhosis. A prospective study. Gastroenterology 1990;98:1567-1571.

45. 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;44:630-634.

46. Zein CO, Angulo P, Lindor KD. When is liver biopsy needed in the diagnosis of primary biliary cirrhosis? Clin Gastroenterol Hepatol 2003;1:89-95.

47. Granito A, Muratori P, Muratori L, et al. Antinuclear antibodies giving the “multiple nuclear dots” or the “rim-like/membranous” patterns: diagnostic accuracy for primary biliary cirrhosis. Aliment Pharmacol Ther 2006;24:1575-1583.

48. Lindor KD, Bowlus CL, Boyer J, et al. Primary Biliary Cholangitis: 2018 Practice

Guidance from the American Association for the Study of Liver Diseases. Hepatology 2019;69:394-419.

49. Invernizzi P, Crosignani A, Battezzati PM, et al. Comparison of the clinical features and clinical course of antimitochondrial antibody-positive and -negative primary biliary cirrhosis. Hepatology 1997;25:1090-1095.

50. Kim WR, Poterucha JJ, Jorgensen RA, et al. Does antimitochondrial antibody status affect response to treatment in patients with primary biliary cirrhosis? Outcomes of ursodeoxycholic acid therapy and liver transplantation. Hepatology 1997;26:22-26.

51. Mutimer DJ, Fussey SP, Yeaman SJ, et al. Frequency of IgG and IgM autoantibodies to four specific M2 mitochondrial autoantigens in primary biliary cirrhosis. Hepatology 1989;10:403-407.

52. Flannery GR, Burroughs AK, Butler P, et al. Antimitochondrial antibodies in primary biliary cirrhosis recognize both specific peptides and shared epitopes of the M2 family of antigens. Hepatology 1989;10:370-374.

53. Van de Water J, Gershwin ME, Leung P, et al. The autoepitope of the 74-kD mitochondrial autoantigen of primary biliary cirrhosis corresponds to the functional site of dihydrolipoamide acetyltransferase. J Exp Med 1988;167:1791-1799.

54. Muratori P, Muratori L, Gershwin ME, et al. “True” antimitochondrial antibody-negative primary biliary cirrhosis, low sensitivity of the routine assays, or both? Clin Exp Immunol 2004;135:154-158.

55. Van de Water J, Cooper A, Surh CD, et al. Detection of autoantibodies to recombinant mitochondrial proteins in patients with primary biliary cirrhosis. N Engl J Med 1989;320:1377-1380.

56. Hu CJ, Song G, Huang W, et al. Identification of new autoantigens for primary biliary cirrhosis using human proteome microarrays. Mol Cell Proteomics 2012;11:669-680.

57. Scheuer P. Primary biliary cirrhosis. Proc R Soc Med 1967;60:1257-1260.

58. Ludwig J, Dickson ER, McDonald GSA. Staging of chronic nonsuppurative destructive cholangitis (Syndrome of primary biliary cirrhosis). Virchows Arch A Pathol Anat Histol 1978;379:103-112.

59. Locke GR 3rd, Therneau TM, Ludwig J, et al. Time course of histological progression in primary biliary cirrhosis. Hepatology 1996;23:52-56.

60. Gómez-Dominguez E, Mendoza J, García-Buey L, et al. Transient elastography to assess hepatic fibrosis in primary biliary cirrhosis. Aliment Pharmacol Ther 2008;27:441-447.

61. Kaplan MM, Gershwin ME. Primary biliary cirrhosis. N Engl J Med 2005;353:1261-1273.

62. Bogdanos DP, Baum H, Grasso A, et al. Microbial mimics are major targets of crossreactivity with human pyruvate dehydrogenase in primary biliary cirrhosis. J Hepatol 2004;40:31-39.

63. Amano K, Leung PS, Rieger R, et al. Chemical xenobiotics and mitochondrial autoantigens in primary biliary cirrhosis: identification of antibodies against a common

environmental, cosmetic, and food additive, 2-octynoic acid. J Immunol 2005;174:5874-5883.

64. Lleo A, Selmi C, Invernizzi P, et al. Apotopes and the biliary specificity of primary biliary cirrhosis. Hepatology 2009;49:871-879.

65. Löhr H, Fleischer B, Gerken G, et al. Autoreactive liver-infiltrating T cells in primary biliary cirrhosis recognize inner mitochondrial epitopes and the pyruvate dehydrogenase complex. J Hepatol 1993;18:322-327.

66. Matsumura S, Van De Water J, Leung P, et al. Caspase induction by IgA antimitochondrial antibody: IgA-mediated biliary injury in primary biliary cirrhosis. Hepatology 2004;39:1415-1422.

67. Kita H, Matsumura S, He XS, et al. Quantitative and functional analysis of PDC-E2-specific autoreactive cytotoxic T lymphocytes in primary biliary cirrhosis. J Clin Invest 2002;109:1231-1240.

68. van Niekerk J, Kersten R, Beuers U. Role of Bile Acids and the Biliary HCO3-Umbrella in the Pathogenesis of Primary Biliary Cholangitis. Clin Liver Dis 2018;22:457-479.

69. Gores GJ, Wiesner RH, Dickson ER, et al. Prospective evaluation of esophageal varices in primary biliary cirrhosis: development, natural history, and influence on survival. Gastroenterology 1989;96:1552-1559.

70. Shibuya A, Tanaka K, Miyakawa H, et al. Hepatocellular carcinoma and survival in patients with primary biliary cirrhosis. Hepatology 2002;35:1172-1178.

71. Suzuki A, Lymp J, Donlinger J, et al. Clinical predictors for hepatocellular carcinoma in patients with primary biliary cirrhosis. Clin Gastroenterol Hepatol 2007;5:259-264.

72. Trivedi PJ, Lammers WJ, van Buuren HR, et al. Stratification of hepatocellular carcinoma risk in primary biliary cirrhosis: A multicentre international study. Gut 2016;65:321-329.

73. Jackson H, Solaymani-Dodaran M, Card TR, et al. Influence of ursodeoxycholic acid on the mortality and malignancy associated with primary biliary cirrhosis: a population-based cohort study. Hepatology 2007;46:1131-1137.

74. 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;50:1884-1892.

75. Leuschner U, Fischer H, Kurtz W, et al. Ursodeoxycholic acid in primary biliary cirrhosis: results of a controlled double-blind trial. Gastroenterology 1989;97:1268-1274.

76. Paumgartner G, Beuers U. Ursodeoxycholic acid in cholestatic liver disease: mechanisms of action and therapeutic use revisited. Hepatology 2002;36:525-531.

77. 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;30:830-835.

78. 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;74:240-246.

79. Heathcote EJ, Cauch-Dudek K, Walker V, et al. The Canadian multicenter double-blind randomized controlled trial of ursodeoxycholic acid in primary biliary cirrhosis. Hepatology 1994;19:1149-1156.

80. Lindor KD, Dickson ER, Baldus WP, et al. Ursodeoxycholic acid in the treatment of primary biliary cirrhosis. Gastroenterology 1994;106:1284-1290.

81. Poupon RE, Poupon R, Balkau B. Ursodiol for the long-term treatment of primary biliary cirrhosis. The UDCA-PBC Study Group. N Engl J Med 1994;330:1342-1347.

82. Parés A, Caballería L, Rodés J, et al. Long-term effects of ursodeoxycholic acid in primary biliary cirrhosis: Results of a double-blind controlled multicentric trial. UDCA-Cooperative Group from the Spanish Association for the Study of the Liver. J Hepatol 2000;32:561-566.

83. Shi J, Wu C, Lin Y, et al. Long-term effects of mid-dose ursodeoxycholic acid in primary biliary cirrhosis: a meta-analysis of randomized controlled trials. Am J Gastroenterol 2006;101:1529-1538.

84. Lindor KD, Jorgensen RA, Therneau TM, et al. Ursodeoxycholic acid delays the onset of esophageal varices in primary biliary cirrhosis. Mayo Clin Proc 1997;72:1137-1140.

85. 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;71:357-365.

86. 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.

87. 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.

88. Corpechot C, Abenavoli L, Rabahi N, et al. Biochemical response to ursodeoxycholic acid and long-term prognosis in primary biliary cirrhosis. Hepatology 2008;48:871-877.

89. 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;105:2186-2194.

90. Corpechot C, Chazouillères O, Poupon R. Early primary biliary cirrhosis: Biochemical response to treatment and prediction of long-term outcome. J Hepatol 2011;55:1361-1367.

91. Zhang LN, Shi T-Y, Shi X-H, 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;58:264-272.

92. Carbone M, Nardi A, Flack S, et al. Pretreatment prediction of response to ursodeoxycholic acid in primary biliary cholangitis: development and validation of the UDCA Response Score. Lancet Gastroenterol Hepatol 2018;3:626-634.

93. 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.

94. 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;136:1281-1287.

95. 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;41:310-317.

96. Hempfling W, Dilger K, Beuers U. Systematic review: ursodeoxycholic acid--adverse effects and drug interactions. Aliment Pharmacol Ther 2003;18:963-972.

97. Combes B, Emerson SS, Flye NL, et al. Methotrexate (MTX) plus ursodeoxycholic acid (UDCA) in the treatment of primary biliary cirrhosis. Hepatology 2005;42:1184-1193.

98. Wiesner RH, Ludwig J, Lindor KD, et al. A controlled trial of cyclosporine in the treatment of primary biliary cirrhosis. N Engl J Med 1990;322:1419-1424.

99. Almasio PL, Floreani A, Chiaramonte M, et al. Multicentre randomized placebo-controlled trial of ursodeoxycholic acid with or without colchicine in symptomatic primary biliary cirrhosis. Aliment Pharmacol Ther 2000;14:1645-1652.

100. Gong Y, Christensen E, Gluud C. Azathioprine for primary biliary cirrhosis. Cochrane Database Syst Rev 2007:CD006000.

101. Leuschner M, Güldütuna S, You T, et al. Ursodeoxycholic acid and prednisolone versus ursodeoxycholic acid and placebo in the treatment of early stages of primary biliary cirrhosis. J Hepatol 1996;25:49-57.

102. Pellicciari R, Fiorucci S, Camaioni E, et al. 6alpha-ethyl-chenodeoxycholic acid (6-ECDCA), a potent and selective FXR agonist endowed with anticholestatic activity. J Med Chem 2002;45:3569-3572.

103. Nevens F, Andreone P, Mazzella G, et al. A placebo-controlled trial of obeticholic acid in primary biliary cholangitis. N Engl J Med 2016;375:631-643.

104. Trauner M, Nevens F, Shiffman ML, et al. Long-term efficacy and safety of obeticholic acid for patients with primary biliary cholangitis: 3-year results of an international open-label extension study. Lancet Gastroenterol Hepatol 2019;4:445-453.

105. Carbone M, Harms MH, Lammers WJ, et al. Clinical application of the GLOBE and United Kingdom-primary biliary cholangitis risk scores in a trial cohort of patients with primary biliary cholangitis. Hepatol Commun 2018;2:683-692.

106. Tyagi S, Gupta P, Saini AS, et al. The peroxisome proliferator-activated receptor: A family of nuclear receptors role in various diseases. J Adv Pharm Technol Res 2011;2:236-240.

107. Iwasaki S, Tsuda K, Ueta H, et al. Bezafibrate may have a beneficial effect in pre-cirrhotic primary biliary cirrhosis. Hepatol Res 1999;16:12-18.

108. Lens S, Leoz M, Nazal L, et al. Bezafibrate normalizes alkaline phosphatase in primary biliary cirrhosis patients with incomplete response to ursodeoxycholic acid.

Liver Int 2014;34:197-203.

109. Tanaka A, Hirohara J, Nakanuma Y, et al. Biochemical responses to bezafibrate improve long-term outcome in asymptomatic patients with primary biliary cirrhosis refractory to UDCA. J Gastroenterol 2015;50:675-682.

110. Honda A, Ikegami T, Nakamuta M, et al. Anticholestatic effects of bezafibrate in patients with primary biliary cirrhosis treated with ursodeoxycholic acid. Hepatology 2013;57:1931-1941.

111. Grigorian AY, Mardini HE, Corpechot C, et al. Fenofibrate is effective adjunctive therapy in the treatment of primary biliary cirrhosis: A meta-analysis. Clin Res Hepatol Gastroenterol 2015;39:296-306.

112. Cheung AC, Lapointe-Shaw L, Kowgier M, et al. Combined ursodeoxycholic acid (UDCA) and fenofibrate in primary biliary cholangitis patients with incomplete UDCA response may improve outcomes. Aliment Pharmacol Ther 2016;43:283-293.

113. Zhang Y, Li S, He L, et al. Combination therapy of fenofibrate and ursodeoxycholic acid in patients with primary biliary cirrhosis who respond incompletely to UDCA monotherapy: A meta-analysis. Drug Des Devel Ther 2015;9:2757-2766.

114. Hegade VS, Khanna A, Walker LJ, et al. Long-Term Fenofibrate Treatment in Primary Biliary Cholangitis Improves Biochemistry but Not the UK-PBC Risk Score. Dig Dis Sci 2016;61:3037-3044.

115. Corpechot C, Chazouillères O, Rousseau A, et al. A placebo-controlled trial of bezafibrate in primary biliary cholangitis. N Engl J Med 2018;378:2171-2181.

116. Honda A, Tanaka A, Kaneko T, et al. Bezafibrate Improves GLOBE and UK-PBC Scores and Long-Term Outcomes in Patients With Primary Biliary Cholangitis. Hepatology 2019;70:2035-2046.

117. Markus BH, Dickson ER, Grambsch PM, et al. Efficacy of liver transplantation in patients with primary biliary cirrhosis. N Engl J Med 1989;320:1709-1713.

118. 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;49:285-295.

119. Liermann Garcia RF, Evangelista Garcia C, McMaster P, et al. Transplantation for primary biliary cirrhosis: retrospective analysis of 400 patients in a single center. Hepatology 2001;33:22-27.

120. Lee J, Belanger A, Doucette JT, et al. Transplantation trends in primary biliary cirrhosis. Clin Gastroenterol Hepatol 2007;5:1313-1315.

121. Kuiper EMM, Hansen BE, Metselaar HJ, et al. Trends in liver transplantation for primary biliary cirrhosis in the Netherlands 1988-2008. BMC Gastroenterol 2010;10:144.

122. Montano-Loza AJ, Hansen BE, Corpechot C, et al. Factors Associated With Recurrence of Primary Biliary Cholangitis After Liver Transplantation and Effects on Graft and Patient Survival. Gastroenterology 2019;156:96-107.e1.

123. Bosch A, Dumortier J, Maucort-Boulch D, et al. Preventive administration of UDCA

after liver transplantation for primary biliary cirrhosis is associated with a lower risk of disease recurrence. J Hepatol 2015;63:1449-1458.

124. Bosch A, Dumortier J, Maucort-Boulch D, et al. Preventive administration of UDCA after liver transplantation for primary biliary cirrhosis is associated with a lower risk of disease recurrence. J Hepatol 2015;63:1449-1458.

125. Charatcharoenwitthaya P, Pimentel S, Talwalkar JA, et al. Long-term survival and impact of ursodeoxycholic acid treatment for recurrent primary biliary cirrhosis after liver transplantation. Liver Transplant 2007;13:1236-1245.

126. 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;78:236-246.

127. 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;20:707-713.

128. 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. J Hepatol 2014;60:1249-1258.

129. 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;56:198-208.

130. Kubota J, Ikeda F, Terada R, et al. Mortality rate of patients with asymptomatic primary biliary cirrhosis diagnosed at age 55 years or older is similar to that of the general population. J Gastroenterol 2009;44:1000-1006.

131. Lindor KD, Gershwin ME, Poupon R, et al. Primary biliary cirrhosis. Hepatology 2009;50:291-308.

132. Shapiro JM, Smith H, Schaffner F. Serum bilirubin: a prognostic factor in primary biliary cirrhosis. Gut 1979;20:137-140.

133. 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.

134. 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.

135. Dickson ER, Grambsch PM, Fleming TR, et al. Prognosis in primary biliary cirrhosis: model for decision making. Hepatology 1989;10:1-7.

136. 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.

137. 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.

138. 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.

139. Cheung AC, Gulamhusein AF, Juran BD, et al. External validation of the United Kingdom-primary biliary cholangitis risk scores of patients with primary biliary cholangitis treated with ursodeoxycholic acid. Hepatol Commun 2018;2:676-682.

140. Yoo JJ, Cho EJ, Lee B, et al. Prognostic Value of Biochemical Response Models for Primary Biliary Cholangitis and the Additional Role of the Neutrophil-to-Lymphocyte Ratio. Gut Liver 2018;12:714-721.

141. 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;45:733-743.

142. Efe C, Taşçilar K, Henriksson I, et al. Validation of Risk Scoring Systems in Ursodeoxycholic Acid-Treated Patients With Primary Biliary Cholangitis. Am J Gastroenterol 2019;114:1101-1108.

143. Huang C, Han W, Wang C, et al. Early Prognostic Utility of Gp210 Antibody-Positive Rate in Primary Biliary Cholangitis: A Meta-Analysis. Dis Markers 2019;2019.

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, Battezzati PM, Floreani A, Pares A, Nevens F, Kowdley KV, Bruns T, Dalekos GN, Thorburn D, Hirschfield G, LaRusso NF, Lindor KD, Zachou K, Poupon R, Trivedi PJ, Verhelst X, Janssen HLA, Hansen BE

Hepatology, 2018

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.

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.*[10], 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 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, the characteristics of which 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] 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 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 nontreated 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 dichotomized according to Ludwig et al.'s[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 was also retrieved.

In patients who 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-squared tests for categorical variables and analyses of variance for continuous data. $P < 0.05$ was considered significant for all statistical analyses. Significant results were further analyzed 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 of the PBC population was also compared within each decade to an age- and gender-matched Dutch population. These outcomes were also estimated by Cox proportional hazards' modeling (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 4,805 PBC patients, diagnosed between 1970 and 2014, were included and divided into five cohorts according to their year of diagnosis (**Table 1, Table S1**): 143 patients were diagnosed from 1970 to 1979, 858 patients from 1980 to 1989, 1,754 patients from 1990 to 1999, 1,815 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 times 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 subgroup of patients (n = 3,518) with a maximum 2-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 (**Table S2**).

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 (**Figure S1A-C**). The effect of calendar time on the increase in age at diagnosis remained significant (P < 0.001) after correcting for sex (**Table S3**). 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 that of 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**).

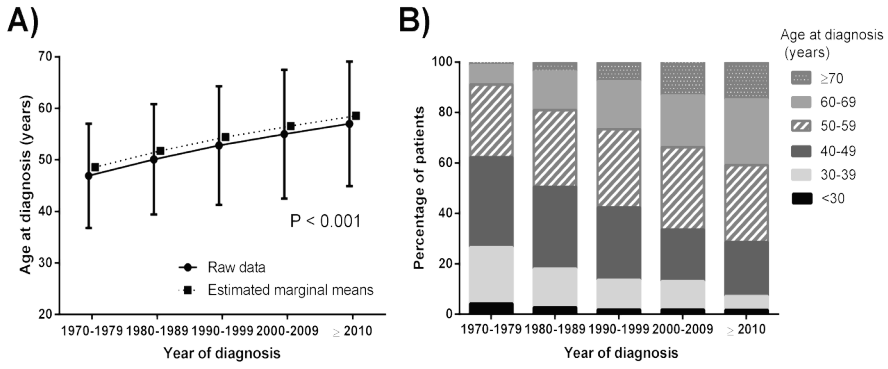


Figure 1. Age at diagnosis of PBC patients across different decades. A) Mean age (± standard deviation) at diagnosis (dots) and estimated marginal means (squares) obtained after adjusting for sex. B) Distribution of age groups over calendar time.

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 (×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).
Primary biliary cholangitis, PBC; 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%).

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 (times the 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 <2 times 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**).

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 2,831 patients who underwent liver biopsy at diagnosis, 2,217 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 through 1989, 948 biopsies from 1990 through 1999, and 943 from 2000 through 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).

Trends in UDCA response rates

The proportion of patients who ever received UDCA increased across the investigated decades (P < 0.001) (**Table 1**). In patients who 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).

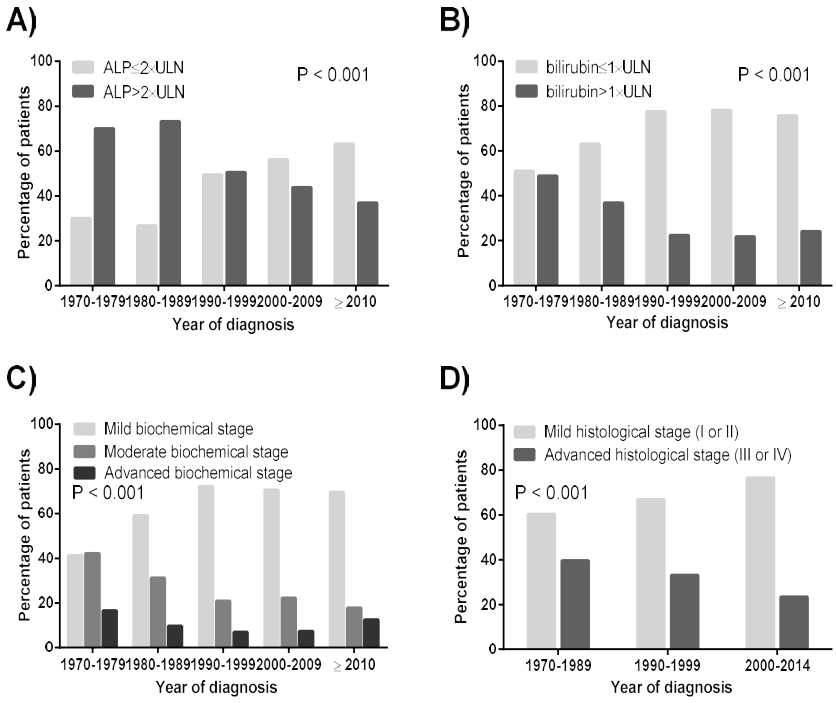


Figure 2. Study entry characteristics associated with disease severity of patients diagnosed in different decades. A) Percentage of patients with alkaline phosphatase (ALP) above or below 2 times the upper limit of normal (×ULN). B) Percentage of patients with bilirubin above or below 1×ULN. C) Percentage 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).

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, Table S4**). Importantly, this trend remained true in patients who did not meet the individual criteria at baseline (**Table S5**). In a multivariable logistic regression, calendar time was not a significant predictor for UDCA response according to Paris-I criteria (**Table 2**). 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).

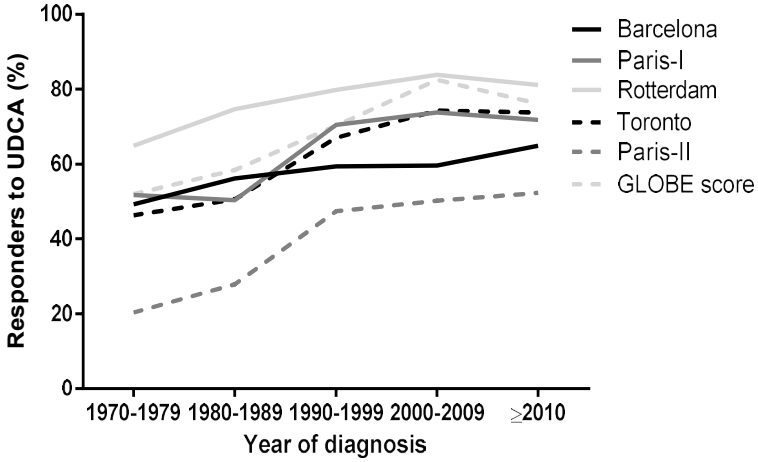


Figure 3. Response rates to ursodeoxycholic acid (UDCA) therapy over calendar time. Response was determined according to various published criteria: Barcelona, Paris-I, Rotterdam, Toronto, Paris-II, and the GLOBE score.[7-9,26,34,35] Response rates according to all criteria were significantly different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

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.

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).

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 through 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, Table S6**). 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).

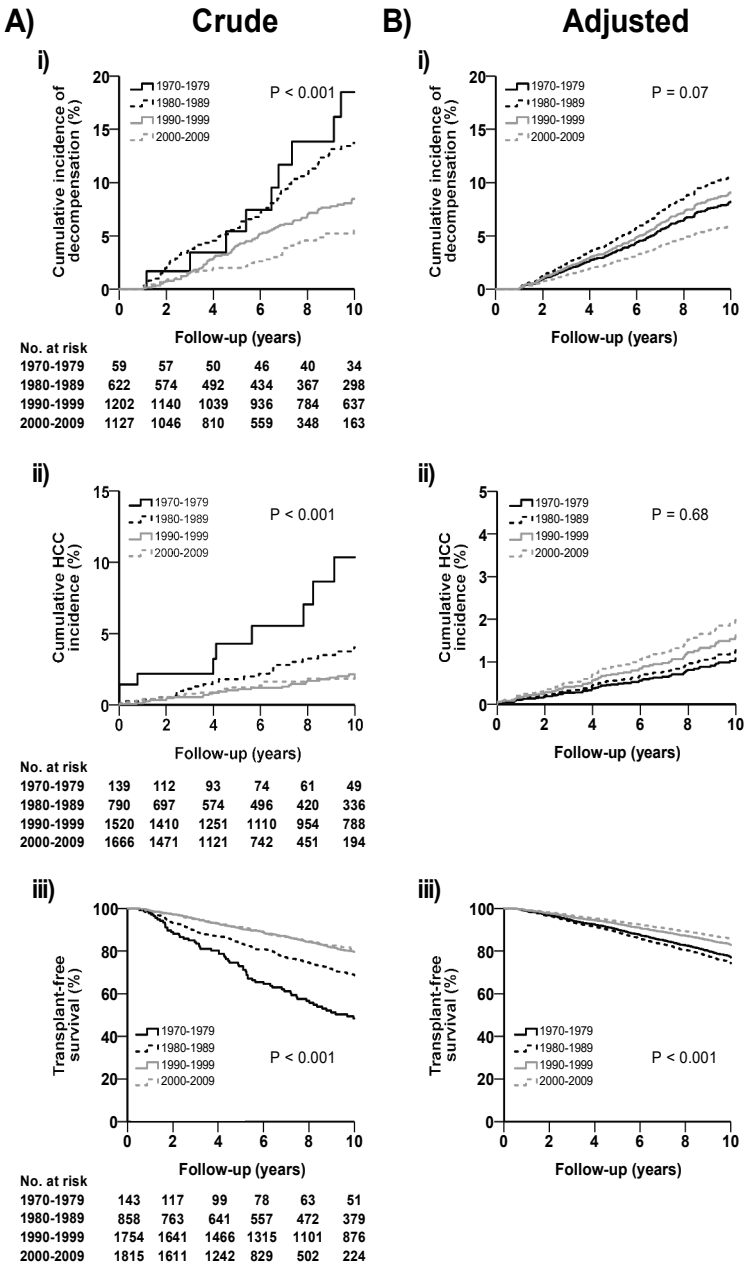


Figure 4. Time-to-event analyses of decompensation, hepatocellular carcinoma (HCC), and liver transplantation or death over calendar time. A) Kaplan-Meier (crude) and B) Multivariable Cox regression (adjusted) estimates of i) cumulative incidence of decompensation, ii) cumulative incidence of hepatocellular 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. 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 enzyme-linked immunosorbent assay-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 was reported in a 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 North 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) (39). 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/day has been emphasized in a study that found 40% of UDCA nonresponders 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 documented. In a Canadian population-based study of patients diagnosed between 1996 and 2002, Myers *et al.* 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 6 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 of PBC patients was compared to that of 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 subgroup of patients with a maximum 2-year difference. The same trends emerged in the subgroup 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 were 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, and 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;353:1261-1273.
2. Poupon R. Primary biliary cirrhosis: a 2010 update. J Hepatol 2010;52:745-758.
3. Lindor KD, Gershwin ME, Poupon R, Kaplan M, Bergasa NV, Heathcote EJ. Primary biliary cirrhosis. Hepatology 2009;50:291-308.
4. Pratt DS. Primary Biliary Cholangitis—A New Name and a New Treatment. N Engl J Med 2016;375:685-687.
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;113:884-890.
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;122:652-658.
7. 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.
8. Corpechot C, Abenavoli L, Rabahi N, Chrétien Y, Andréani T, Johanet C, et al. Biochemical response to ursodeoxycholic acid and long-term prognosis in primary biliary cirrhosis. Hepatology 2008;48:871-877.
9. Kuiper EM, Hansen BE, de Vries RA, den Ouden–Muller JW, Van Ditzhuijsen TJ, Haagsma EB, et al. Improved prognosis of patients with primary biliary cirrhosis that have a biochemical response to ursodeoxycholic acid. Gastroenterology 2009;136:1281-1287.
10. Boonstra K, Beuers U, Ponsioen CY. Epidemiology of primary sclerosing cholangitis and primary biliary cirrhosis: a systematic review. J Hepatol 2012;56:1181-1188.
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;26:830-836.
12. Boonstra K, Kunst AE, Stadhouders PH, Tuynman HA, Poen AC, van Nieuwkerk KM, et al. Rising incidence and prevalence of primary biliary cirrhosis: a large population-based study. Liver Int 2014;34:e31-e38.
13. Eriksson S, Lindgren S. The prevalence and clinical spectrum of primary biliary cirrhosis in a defined population. Scand J Gastroenterol 1984;19:971-976.
14. Myszor M, James OF. The epidemiology of primary biliary cirrhosis in north-east England: an increasingly common disease? Q J Med 1990;75:377-385.
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;30:367-371.
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;30:390-394.

17. Pla X, Vergara M, Gil M, Dalmau B, Cisteró B, Bella RM, et al. Incidence, prevalence and clinical course of primary biliary cirrhosis in a Spanish community. Eur J Gastroenterol Hepatol 2007;19:859-864.

18. Myers RP, Shaheen AA, Fong A, Burak KW, Wan A, Swain MG, et al. Epidemiology and natural history of primary biliary cirrhosis in a Canadian health region: A population-based study. Hepatology 2009;50:1884-1892.

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;179:492-498.

20. Baldursdottir TR, Bergmann OM, Jonasson JG, Ludviksson BR, Axelsson TA, Björnsson ES. The epidemiology and natural history of primary biliary cirrhosis: a nationwide population-based study. Eur J Gastroenterol Hepatol 2012;24:824-830.

21. Gross RG, Odin JA. Recent advances in the epidemiology of primary biliary cirrhosis. Clin Liver Dis 2008;12:289-303.

22. Lee J, Belanger A, Doucette JT, Stanca C, Friedman S, Bach N. Transplantation trends in primary biliary cirrhosis. Clin Gastroenterol Hepatol 2007;5:1313-1315.

23. Hirschfield GM, Mason A, Luketic V, Lindor K, Gordon SC, Mayo M, et al. Efficacy of obeticholic acid in patients with primary biliary cirrhosis and inadequate response to ursodeoxycholic acid. Gastroenterology 2015;148:751-761.

24. Christensen E, Crowe J, Doniach D, Popper H, Ranek L, Rodés J, et al. Clinical pattern and course of disease in primary biliary cirrhosis based on an analysis of 236 patients. Gastroenterology 1980;78:236-246.

25. Locke GR 3rd, Therneau TM, Ludwig J, Dickson ER, Lindor KD. Time course of histological progression in primary biliary cirrhosis. Hepatology 1996;23:52-56.

26. Lammers WJ, Hirschfield GM, Corpechot C, Nevens F, Lindor KD, Janssen HL, 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.

27. Lammers WJ, Van Buuren HR, Hirschfield GM, Janssen HL, Invernizzi P, Mason AL, 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.

28. Nguyen DL, Juran BD, Lazaridis KN. Primary Biliary Cirrhosis. Best Pract Res Clin Gastroenterol 2010;24:647-654.

29. European Association for the Study of the Liver. EASL Clinical Practice Guidelines: The diagnosis and management of patients with primary biliary cholangitis. J Hepatol 2017;67:145-172.

30. Ludwig J, Dickson ER, McDonald GA. Staging of chronic nonsuppurative destructive cholangitis (syndrome of primary biliary cirrhosis). Virchows Arch A Pathol Anat Histol 1978;379:103-112.

31. Scheuer P. Primary biliary cirrhosis. Proc R Soc Med 1967;60:1257-1260.

32. Ter Borg PC, Schalm SW, Hansen BE, Van Buuren HR. 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.

33. **Trivedi PJ**, **Bruns T**, Cheung A, Li KK, Kittler C, Kumagi T, et al. Optimising risk stratification in primary biliary cirrhosis: AST/platelet ratio index predicts outcome independent of ursodeoxycholic acid response. J Hepatol 2014;60:1249-1258.

34. Kumagi T, Guindi M, Fischer SE, Arenovich T, Abdalian R, Coltescu C, et al. Baseline ductopenia and treatment response predict long-term histological progression in primary biliary cirrhosis. Am J Gastroenterol 2010;105:2186-2194.

35. Corpechot C, Chazouillères O, Poupon R. Early primary biliary cirrhosis: biochemical response to treatment and prediction of long-term outcome. J Hepatol 2011;55:1361-1367.

36. Bergquist A, Said K, Broomé U. Changes over a 20-year period in the clinical presentation of primary sclerosing cholangitis in Sweden. Scand J Gastroenterol 2007;42:88-93.

37. Oertelt S, Rieger R, Selmi C, Invernizzi P, Ansari AA, Coppel RL, Podda M, Leung PS, Gershwin ME. A sensitive bead assay for antimitochondrial antibodies: Chipping away at AMA-negative primary biliary cirrhosis. Hepatology 2007;45:659-665.

38. **Carbone M**, **Mells GF**, Pells G, Dawwas MF, Newton JL, Heneghan MA, et al. Sex and age are determinants of the clinical phenotype of primary biliary cirrhosis and response to ursodeoxycholic acid. Gastroenterology 2013;144:560-569.

39. United Nations, Department of Economic and Social Affairs, Population Division (2017). World Population Prospects: The 2017 Revision, custom data acquired via website.

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;20:707-713.

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;123:1044-1051.

42. Lammers WJ, Leeman M, Ponsioen CI, Boonstra K, van Erpecum KJ, Wolfhagen FH, et al. How the concept of biochemical response influenced the management of primary biliary cholangitis over time. Neth J Med 2016;74:240-246.

43. Angulo P, Dickson ER, Therneau TM, Jorgensen RA, Smith C, DeSotel CK, et al. Comparison of three doses of ursodeoxycholic acid in the treatment of primary biliary cirrhosis: a randomized trial. J Hepatol 1999;30:830-835.

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;53:865-870.

45. European Association for the Study of the Liver. EASL Clinical Practice Guidelines: management of cholestatic liver diseases. J Hepatol 2009;51:237-267.

46. Corpechot C, Carrat F, Bahr A, Chrétien Y, Poupon RE, Poupon R. The effect of ursodeoxycholic acid therapy on the natural course of primary biliary cirrhosis. Gastroenterology 2005;128:297-303.

47. Gershwin ME, Selmi C, Worman HJ, Gold EB, Watnik M, Utts J, et al. Risk factors and comorbidities in primary biliary cirrhosis: A controlled interview-based study of 1032 patients. Hepatology 2005;42:1194-1202.

Author names in bold designate shared co-first authorship.

SUPPLEMENTARY DATA

Table S1. 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.

Table S2. 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 (×10^9/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
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.

Figure S1. 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.

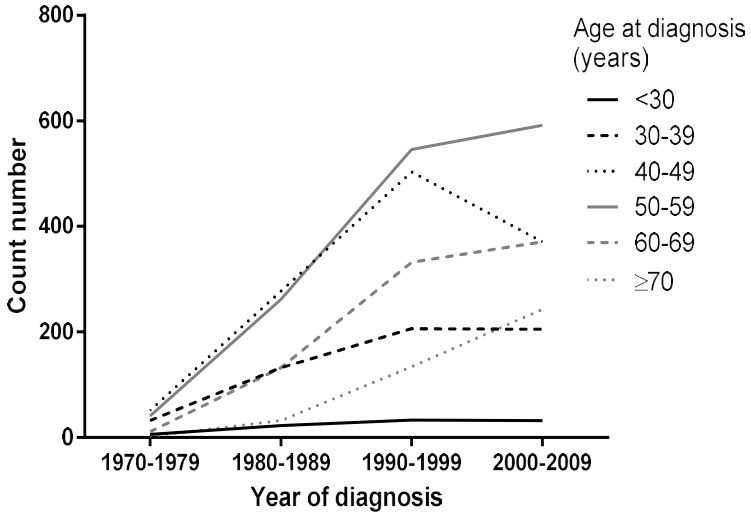


Figure S2. Absolute number of patients according to age at diagnosis and over calendar time.

Table S3. 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.

Table S4. 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]

Table S5. 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.

Table S6. 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.

CHAPTER 3

3

The impact of geographical region on outcomes of patients with primary biliary cholangitis from Western Europe

Murillo Perez CF, Gerussi A, Trivedi PJ, Corpechot C, van der Meer AJ, Battezzati PM, Lindor KD, Nevens F, Kowdley KV, Bruns T, Cazzagon N, Floreani A, Tanaka A, Ma X, Mason AL, Gulamhusein A, Ponsioen CY, Carbone M, Lleo A, Mayo MJ, Dalekos GN, Gatselis NK, Thorburn D, Verhelst X, Parés A, Janssen HLA, Hirschfield GM, Hansen BE, Invernizzi P, Lammers WJ

ABSTRACT

Background and Aims: The incidence and prevalence of primary biliary cholangitis (PBC) varies according to geographical region. The aim of this study was to assess whether clinical outcomes of patients with PBC within Western Europe also vary according to geographical region.

Methods: Patients from the Global PBC database from European centers diagnosed from 1990 onwards that were ursodeoxycholic acid (UDCA)-treated were included in the study. Patients with a time lag greater than 1 year from diagnosis to the start of follow-up were excluded. Differences in baseline characteristics were studied according to North/South and East/West, while 10-year outcomes (transplant-free survival and decompensation) after 1 year of UDCA were studied with latitude and longitude. Multivariable Cox regression analysis were adjusted for biochemical markers of disease severity and cirrhosis as a time-dependent covariate, as well as age, sex, and diagnosis year.

Results: A total of 1878 patients were included in the study. There were no differences in patient age or sex, with an overall mean age of 54 years and 89% female patients. Those in North Europe were more often of a moderately advanced or advanced Rotterdam biochemical stage (28.4%) compared to those in South Europe (20.6%). Additionally, they exhibited higher median alkaline phosphatase (ALP [2.0×ULN vs. 1.4×ULN]) and transaminases. In multivariable Cox regression analysis, there was a significant interaction between latitude and longitude in the prediction of decompensation ($P<0.001$) and a trend for transplant-free survival. An increased risk for poor outcomes was observed for the Northwestern area, pertaining to the UK centers, for decompensation (HR 1.6-2.3) and transplant-free survival (HR 1.2-1.3) as compared to the reference (Paris).

Conclusion: Patients from Northwestern Europe with PBC and of similar disease severity have increased risk of decompensation and liver transplantation or death. This suggests that geographical region may have an impact on patient outcome.

INTRODUCTION

Primary biliary cholangitis, a slowly progressive chronic cholestatic liver disease, can affect people from all geographical regions and of various ethnicities. The incidence and prevalence rates of PBC vary widely by geographical region, with the highest incidence in Caucasian populations from Northern Europe and lowest in the Indian subcontinent and Africa.[1] In a systematic review, the incidence rates of PBC varied from 0.33-05.8 per 100,000 per year, while the prevalence rates varied from 1.91-40.2 per 100,000.[2] Respective differences in incidence and prevalence can potentially be attributed to case ascertainment, genetics, but also environmental factors. The involvement of environmental factors in the etiology of PBC is supported by reports of spatial clustering of cases, as reported in Northeast England and Alaska.[3–5] Furthermore, an increased prevalence of PBC patients has been reported near toxic waste sites and among atomic bomb survivors from Hiroshima.[6,7]

Based on these findings, it is plausible that geographical region may also affect disease severity and clinical outcomes. However, there has been no formal assessment on the differences in presentation and outcomes of patients with PBC according to geographical region. Thus, whether regions with a higher prevalence for PBC, such as Northern Europe, also have increased rates of clinical outcomes is unknown. Elucidating such potential differences can inform whether homogeneous clinical practice guidelines for patients with PBC across all geographical regions is applicable. The aim of this study is to determine whether there is an association between geographical region and clinical outcomes of patients with PBC within Western Europe.

METHODS

Population and study design

Patients from Western European centers in the Global PBC database diagnosed with PBC from 1990 onwards that were ursodeoxycholic acid (UDCA)-treated were included in the study. The cities covered are Paris, Birmingham, London, Ghent, Leuven, Jena, Larissa, Barcelona, Milan, Padova, and a nation-wide cohort from the Netherlands of which patients are predominantly from Rotterdam. Patients with a short follow-up (<6 months), UDCA discontinuation, unknown dates of clinical events, autoimmune hepatitis overlap syndrome, or another concomitant liver disease were excluded from the study. Additionally, those with a time lag greater than 1 year from diagnosis to the start of follow-up were excluded. Baseline characteristics according to geographical region were assessed categorically (North/South and East/West). Meanwhile, the impact of geographical region on response to UDCA and clinical outcomes (hepatic decompensation and transplant-free survival) was assessed as longitude and latitude. This study was conducted in accordance with the 1975 Declaration of Helsinki. The protocol was approved by the institutional research board at all participating centers as per local regulations.

Data collection

Baseline (study entry) was defined as the date of UDCA initiation. At study entry, the following data were collected: sex, age at diagnosis, anti-mitochondrial (AMA) antibody serological status, biochemical disease stage according to Rotterdam criteria[8], and histologic stage, staged according to Ludwig or Scheuer's criteria.[9,10] Cirrhosis was defined as histologic stage 4 or evidence for cirrhosis. The following laboratory parameters were collected every 6-12 months: total bilirubin, ALP, albumin, aspartate aminotransferase (AST), alanine aminotransferase (ALT), and platelet count. The longitude and latitude correspond to the coordinates of each city and Rotterdam for the nation-wide cohort of the Netherlands as it represents the majority of patients. A center survey was distributed to the centers to gain insight into their treatment regimen, hepatocellular carcinoma (HCC) surveillance, and liver transplantation allocation.

Statistical Analyses

Centers were allocated into North/South and East/West; categories that are not mutually exclusive as described in **Supplementary Table 1**. Centers included in the South category were Mediterranean, while Paris was as it is in the North of France. Categorization into East and West was more arbitrary. Patient and disease characteristics were reported according to North/South and East/West and compared with a Chi-square test for categorical variables and

Student's t-test or Mann-Whitney U test for continuous variables. The association of longitude, latitude, and their interaction with response to UDCA was assessed with univariate and multivariable logistic regression analyses. In addition to demographics, the analyses were adjusted for various markers of disease severity including Rotterdam disease stage, bilirubin, ALP, and cirrhosis, defined as cirrhosis from 2 years prior and up to 1 year after UDCA treatment. The predicted probability of response to UDCA was calculated and illustrated according to latitude and longitude.

The clinical outcomes assessed were cumulative hepatic decompensation incidence (ascites, variceal bleeding, and hepatic encephalopathy) and transplant-free survival. Due to the differences in follow-up across centers, clinical outcomes were assessed at 10 years. Patients lost to follow-up after 10 years or without an event in the first 10 years were censored at that time. Furthermore, decompensation events in the first six months were excluded.

The influence of latitude, longitude, and their interaction on clinical outcomes was assessed with Cox regression analyses. In order to adjust for the influence of disease severity in the differences in outcomes according to geographical region, various biochemical variables and cirrhosis as a time-dependent variable were included in the multivariable Cox regression analyses. The hazard ratio was calculated by utilizing the Beta coefficients from the models and illustrated according to latitude and longitude with Paris as the reference.

Logistic regression and Cox regression analyses were carried out with the imputed database. Multiple imputation was conducted with by the Markov chain Monte Carlo (MCMC) method for missing data with SAS version 9.4 (SAS Institute Inc., Cary, NC, USA). Ten imputed data sets for missing biochemical values were generated to reduce sampling variability of laboratory results at baseline and yearly thereafter up to 10 years (SAS Proc MI, MCMC method).[11] The imputation was performed based on the assumption that data were missing at random, in which variables predicting outcomes and outcomes themselves were included in the imputation model. Rubin's rules were used to estimate the parameter and standard error.[12,13] Laboratory data included in the multivariable model that were not normally distributed were transformed with the natural log (LN).

A *P*-value less than 0.05 was considered statistically significant. All analyses were two-sided and were performed using IBM SPSS Statistics for Windows, version 25.0 (IBM Corp., Armonk, NY, USA).

RESULTS

Patient population characteristics according to geographical region

In this study, a total of 1878 patients from Europe who are representative of a PBC population were included (**Table 1**). There were 391 patients excluded due to the time lag greater than 1 year from diagnosis to the start of UDCA. The mean age at diagnosis was 54.4 years (SD 12.4) and 1677 (89.3%) were female. The median follow-up in this population was 7.7 years (IQR 4.0-11.9). In the first 10 years of follow-up, there were 219 clinical events, of which 77 were liver transplantations and 142 were deaths.

Table 1. Baseline characteristics of cohort

	Whole cohort N=1878
Female	1677 (89.3)
Age at the start of follow-up, years	54.4 (12.4)
Age at the start of follow-up <50	684 (36.4)
Year of diagnosis	2002 (1997-2006)
AMA-positive	1687 (90.5)
UDCA dosage, mg/kg	12.3 (9.6-14.7)
Biochemical stage	
Mild	852 (74.4)
Moderately advanced	235 (20.5)
Advanced	58 (5.1)
Histological stage	
I or II	672 (74.0)
III or IV	236 (26.0)
Laboratory parameters	
Total bilirubin (×ULN)	0.60 (0.44-1.65)
ALP (×ULN)	1.78 (1.18-3.10)
AST (×ULN)	1.37 (0.94-2.12)
ALT (×ULN)	1.58 (0.98-2.53)
Albumin (×LLN)	1.16 (1.08-1.26)
Platelet count (×10^9/L)	255 (207-304)

AMA, anti-mitochondrial antibody; UDCA, ursodeoxycholic acid; ULN, upper limit of normal; ALP, alkaline phosphatase, AST, aspartate aminotransferase; ALT, alanine aminotransferase; LLN, lower limit of normal.

Data presented as n (%) or median (25^{th} percentile-75^{th} percentile).

AMA status was available for 1865 (99.3%). UDCA dosage was available for 1037 (55.2%).

Biochemical disease stage established according to Rotterdam criteria[8] and available for 1145 patients (61.0%).

Histological staging according to Ludwig's[9] or Scheuer's[10] criteria was available for 908 (48.3%) as not all patients were biopsied at baseline.

Availability of laboratory parameters was as follows:

Bilirubin: n=1429 (76.1%); ALP: n=1496 (79.7%); AST: n=1484 (79.0%); ALT: n=1437 (76.5%); albumin: n=1193 (63.5%); platelet count: n=1182 (62.9%).

Of those included in the study, 1199 (63.8%) patients were allocated into the North region, while 679 (36.2%) patients were allocated into the South. In comparison to the South region, an increased proportion of patients in the North were AMA-positive and of moderately advanced or advanced biochemical disease stage (**Table 2**). They also presented with higher ALP, AST, ALT, albeit higher platelet count. In contrast, 612 (32.6%) were allocated into the East region and 1266 (67.4%) into the West. A higher proportion of patients from the West were AMA-positive and presented with higher ALP, AST, ALT, albeit lower total bilirubin and higher platelet count (**Table 3**). There were no differences detected in the age at the start of follow-up or proportion of female patients according to geographical region. For descriptive purposes, a center-level evaluation of the treatment, HCC, and liver transplantation regimens according to calendar year and geographical region are depicted in **Supplementary Figures 1-3**.

Table 2. Baseline characteristics of patients with PBC from North and South Europe

N=1878	North Europe n=1199	South Europe n=679	P
Female, n (%)	1061 (88.5)	616 (90.7)	0.13
Age at the study entry, mean (SD)	54.4 (12.0)	54.3 (13.1)	0.89
Age at study entry <50, n (%)	426 (35.5)	258 (38.0)	0.29
Year of diagnosis, median (IQR)	2002 (1997-2007)	2002 (1997-2006)	0.97
AMA-positive, n (%)	1092 (91.6)	595 (88.4)	0.02
UDCA dosage, mg/kg	12.0 (9.5-14.2)	12.3 (9.2-15.0)	0.28
Biochemical stage, n (%)			0.008
Mild	528 (71.6)	324 (79.4)	
Moderately advanced	164 (22.3)	71 (17.4)	
Advanced	45 (6.1)	13 (3.2)	
Histological stage, n (%)			0.19
I or II	342 (73.5)	204 (69.2)	
III or IV	123 (26.5)	91 (30.8)	
Laboratory parameters, median (IQR)			
Total bilirubin (×ULN)	0.59 (0.41-0.95)	0.61 (0.48-0.83)	0.14
ALP (×ULN)	2.00 (1.30-3.49)	1.41 (1.00-2.47)	<0.001
AST (×ULN)	1.44 (0.98-2.20)	1.26 (0.89-1.94)	<0.001
ALT (×ULN)	1.68 (1.04-2.66)	1.41 (0.91-2.35)	<0.001
Albumin (×LLN)	1.14 (1.06-1.23)	1.19 (1.11-1.29)	<0.001
Platelet count (×10^9/L)	265 (217-322)	235 (197-283)	<0.001

AMA, anti-mitochondrial antibody; UDCA, ursodeoxycholic acid; ULN, upper limit of normal; ALP, alkaline phosphatase, AST, aspartate aminotransferase; ALT, alanine aminotransferase; LLN, lower limit of normal.

AMA status was available for 1192 (99.4%) from North and 673 (99.1%) from South. UDCA dosage was available for 596 (49.7%) from North and 330 from South (48.6%).

Biochemical disease stage established according to Rotterdam criteria[8] and available for 737 patients (61.5%) from North and 408 patients (60.0%) from South.

Histological staging according to Ludwig's[9] or Scheuer's[10] criteria was available for 465 (38.8%) from the North and 295 (43.4%) from the South, as not all patients were biopsied at baseline.

Availability of laboratory parameters was as follows:
North Europe: bilirubin: n=937 (78.1%); ALP: n=985 (82.2%); AST: n=959 (80.0%); ALT: n=913 (76.1%); albumin: n=760 (63.4%); platelet count: n=730 (60.9%). South Europe: bilirubin: n=492 (72.5%); ALP:

n=511 (75.3%); AST: n=525 (77.3%); ALT: n=524 (77.2%); albumin: n=433 (63.8%); platelet count: n=452 (66.6%).

Table 3. Baseline characteristics of patients with PBC from East and West Europe

N=1878	East Europe n=612	West Europe n=1266	P
Female, n (%)	554 (90.5)	1123 (88.7)	0.23
Age at the study entry, mean (SD)	54.5 (12.7)	54.3 (12.2)	0.73
Age at study entry <50, n (%)	222 (36.3)	462 (36.5)	0.93
Year of diagnosis, median (IQR)	2003 (1998-2008)	2001 (1996-2006)	<0.001
AMA-positive, n (%)	534 (88.0)	1153 (91.7)	0.011
UDCA dosage, mg/kg	12.2 (9.1-15.0)	12.0 (9.6-14.3)	0.41
Biochemical stage, n (%)			0.68
Mild	271 (75.7)	581 (73.8)	
Moderately advanced	68 (19.0)	167 (21.2)	
Advanced	19 (5.3)	39 (5.0)	
Histological stage, n (%)			0.001
I or II	93 (60.8)	453 (74.6)	
III or IV	60 (39.2)	154 (25.4)	
Laboratory parameters, median (IQR)			
Total bilirubin (×ULN)	0.64 (0.48-0.87)	0.59 (0.42-0.93)	0.015
ALP (×ULN)	1.32 (0.91-2.15)	2.01 (1.32-3.48)	<0.001
AST (×ULN)	1.25 (0.87-1.96)	1.42 (0.98-2.16)	0.001
ALT (×ULN)	1.37 (0.85-2.28)	1.68 (1.05-2.63)	<0.001
Albumin (×LLN)	1.20 (1.09-1.31)	1.14 (1.06-1.23)	<0.001
Platelet count (×10^9/L)	236 (196-286)	262 (215-316)	<0.001

AMA, anti-mitochondrial antibody; UDCA, ursodeoxycholic acid; ULN, upper limit of normal; ALP, alkaline phosphatase, AST, aspartate aminotransferase; ALT, alanine aminotransferase; LLN, lower limit of normal.

AMA status was available for 607 (99.2%) from East and 1258 (99.4%) from West. UDCA dosage was available for 319 (52.1%) from East and 607 from West (47.9%).

Biochemical disease stage established according to Rotterdam criteria[8] and available for 358 patients (58.5%) from East and 787 patients (62.2%) from West.

Histological staging according to Ludwig's[9] or Scheuer's[10] criteria was available for 153 (25.0%) from the North and 607 (47.9%) from the South, as not all patients were biopsied at baseline.

Availability of laboratory parameters was as follows:
East Europe: bilirubin: n=423 (69.1%); ALP: n=411 (67.2%); AST: n=455 (74.3%); ALT: n=454 (74.2%); albumin: n=382 (62.4%); platelet count: n=424 (69.3%). West Europe: bilirubin: n=1006 (79.5%); ALP: n=1055 (83.3%); AST: n=1029 (81.3%); ALT: n=983 (77.6%); albumin: n=811 (64.1%); platelet count: n=758 (59.9%).

Response to ursodeoxycholic acid according to geographical region

In univariate analysis, there was a significant interaction between latitude and longitude in their association with response to UDCA according to Paris-I criteria (P=0.04). Predicted probability of response to UDCA was lower in the northwestern region (London and Birmingham) (**Figure 1A**). The interaction remained significant in multivariable analysis adjusted for age, sex, diagnosis year, cirrhosis in the first year, ALP, bilirubin and Rotterdam disease stage (P=0.01) (**Figure 1B**). The same trend was observed for Barcelona criteria in univariate and

multivariable analysis. Further, in a sub-group of patients that excluded those with early cirrhosis, the trend for decreased response rates in the Northwestern region remained (Paris-I, P=0.02; Barcelona, P=0.006). The interaction between latitude and longitude was not found to be significant for Rotterdam, Paris-II, Toronto, GLOBE criteria of response.

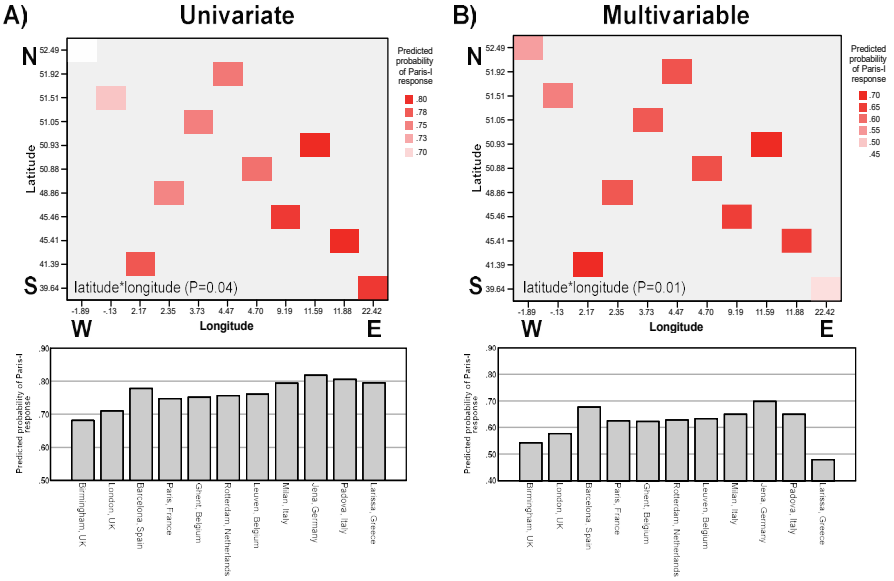


Figure 1. Predicted probability of Paris-I response according to geographical region. Probabilities for the multivariable analysis are based on the following fixed patient and disease characteristics: female, mild Rotterdam disease stage, no cirrhosis, diagnosis year 2000, age 50, ALP=3.5×ULN, bilirubin=1.2×ULN.

Clinical outcomes according to geographical region

The impact of geographical region on hepatic decompensation after 1 year of UDCA was assessed with longitude and latitude, in which there was a significant interaction between these in univariate analysis (P<0.001, **Table 3**). The hazard ratio for hepatic decompensation according to region was relatively stable throughout with an increased risk for decompensation in the Northwestern regions, pertaining to cities in the UK (London and Birmingham), with Paris as the reference (HR 2.0-3.4) (**Figure 2A**). In multivariable analysis, while adjusting for age, sex, year of diagnosis, Rotterdam disease stage, APRI[14], GLOBE response[14], and cirrhosis as a time-dependent covariate, the increased risk for decompensation in the Northwestern regions remained (HR 1.6-2.3) (**Figure 2B**).

Table 3. Univariate and multivariable analysis of latitude and longitude for decompensation and transplant-free survival after 1 year of ursodeoxycholic acid

Decompensation	Univariate			Multivariable[a]		
	B	Standard error	P value	B	Standard error	P value
Latitude	.072	.028	.011	.050	.038	.193
Longitude	.855	.127	<0.001	.578	.150	<0.001
Latitude*longitude	-.021	.003	<0.001	-.015	.003	<0.001
Transplant-free survival	**Univariate**			**Multivariable[b]**		
	B	Standard error	P value	B	Standard error	P value
Latitude	.069	.025	.005	.030	.028	.292
Longitude	.302	.104	.004	.146	.115	.203
Latitude*longitude	-.007	.002	.002	-.004	.003	.148

[a]Adjusted for age, sex, year of diagnosis, Rotterdam disease stage, APRI, GLOBE response, and cirrhosis as a time-dependent covariate.

[b]Adjusted for age, sex, diagnosis year, Rotterdam disease stage, GLOBE response, and cirrhosis as a time-dependent covariate.

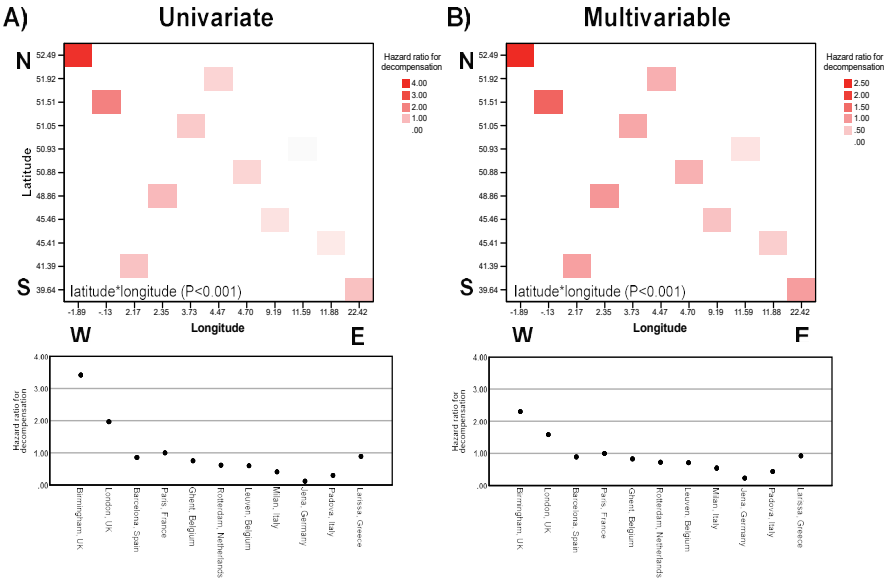


Figure 2. Hazard ratio for decompensation according to geographical region (latitude and longitude) for univariate and multivariable Cox regression analyses after 1 year of ursodeoxycholic acid. The reference is Paris, France.

In univariate analysis, there was a significant interaction between latitude and longitude for transplant-free survival after 1 year of UDCA (P=0.04) (**Table 3**). The Northwestern regions also had an increased risk for liver transplantation or death (HR 1.4-1.7) (**Figure 3A**). In multivariable analyses while adjusting for age, sex, diagnosis year, Rotterdam disease stage, GLOBE response, and cirrhosis as a time-dependent covariate, the trend remained but was no longer statistically significant (HR 1.2-1.3). Exclusion of the Dutch cohort from the population did not alter the findings.

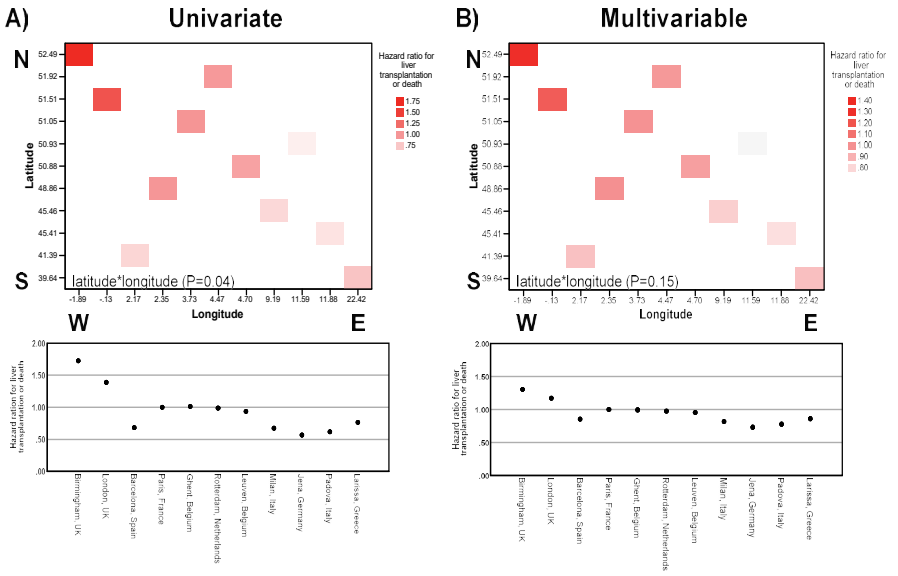


Figure 3. Hazard ratio for liver transplantation or death according to geographical region (latitude and longitude) for univariate and multivariable Cox regression analyses after 1 year of ursodeoxycholic acid. The reference is Paris, France.

DISCUSSION

This study assessed whether patient and disease characteristics, as well as clinical outcomes differ according to geographical region in patients with PBC from Western Europe. While there was no difference in patient age or sex according to region, we demonstrate a lower probability of response (Paris-I and Barcelona) in the Northwestern region that pertains to London and Birmingham (UK). Further, we demonstrate that this region also has an increased risk for decompensation, liver transplantation, and death, even when considering various markers of disease severity such as liver biochemistry and cirrhosis.

PBC mainly affects middle-aged women such that a systematic review of population-based epidemiological studies on PBC report 92% (~9:1) as the mean proportion of female patients.[2]

However, a population-based study of European regions Lombardia (Italy) and Denmark have reported an increased proportion of male patients of 2.3:1 and 4.2:1, respectively.[15] In contrast, we report a 9:1 ratio that was not significantly different according to region. This suggests that there may be differences in certain regions of Europe that are not captured in this study due to the fact that this is a cohort study.

There is evidence that demonstrates that there is a wide range in the incidence and prevalence of PBC according to region.[2] Therefore, it is plausible that region can also have an impact on the disease severity and clinical outcomes of patients. The decreased probability of response and increased risk for clinical outcomes in the Northwestern region (UK) is in line with studies that highlight an increased incidence and prevalence of PBC in the Northeast England.[16,17] While the association of the interaction between latitude and longitude was not significant for transplant-free survival, there was a clear trend present. The lack of significance may be due to the fact that liver transplantation and death are longer-term outcomes, as compared to decompensation. Nonetheless, these findings may be driven by an environmental factor that cannot be readily identified at this moment. One potential contributing factor is the concentration of serum Vitamin D (25-dihydroxyvitamin D), for which the main source is sun exposure. It is estimated from a cross-sectional National Diet and Nutrition survey that 8.4% of people in UK people aged 14-64 have vitamin D deficiency in the summer and 39.3% in the winter, with a latitudinal pattern in which Northern regions have lower vitamin D.[18,19] It has been demonstrated that Vitamin D concentrations are associated with advanced histological stage, response to UDCA (Paris-I and Barcelona), and liver transplantation and death in patients with PBC.[20,21] Conversely, these findings may result from center-effects of referral centers. Still, the possibility that these findings suggest a true environmental effect cannot be excluded.

Another potential contributing factor could be obesity rates in the UK and its association with fatty liver disease. In a European health report from 2018, the UK ranks third with the most patients with self-reported obsesity at 20% of the population.[22] The role of body mass index (BMI) on the observed differences in clinical outcomes according to region could not be evaluated at this time as it was not available in our dataset.

There are various limitations to this study. First, although various disease severity covariates were considered in the analysis, it does not mean that all potential confounders were considered. There is the possibility that some confounding factors were not included, some of which are not readily available in the Global dataset or are difficult to capture, such as BMI and UDCA dosage (mg/kg). Second, this was not a population-based study, but rather a cohort study of centers included in the Global PBC Study Group. Thus, there was a select number of centers included in this analysis that did not allow for an extensive analysis into the impact of

geographical region in other parts of Europe. Further research is warranted to report how these may differ to other part of Europe and to further explore the potential contribution of an environmental factor such as Vitamin D.

In conclusion, patients from the Northwestern part of the region evaluated, pertaining to UK centers, demonstrated and increased risk for decompensation, liver transplantation, and death while considering markers of disease severity. While these findings raise the possibility of geographical region influencing a patient's outcome, likely by an environmental factor, they also highlight the need for further research to validate these findings and evaluate other regions within Europe and beyond. These studies can inform whether particular interest should be provided to patients in these regions and possibly elucidate the role of an environmental agent.

REFERENCES

1. Metcalf J, James O. The geoepidemiology of primary biliary cirrhosis. Semin Liver Dis 1997;17:13-22.
2. Boonstra K, Beuers U, Ponsioen CY. Epidemiology of primary sclerosing cholangitis and primary biliary cirrhosis: a systematic review. J Hepatol 2012;56:1181-1188.
3. Abu-Mouch S, Selmi C, Benson GD, et al. Geographic clusters of primary biliary cirrhosis. Clin Dev Immunol 2003;10:127-131.
4. McNally RJ, James PW, Ducker S, et al. No rise in incidence but geographical heterogeneity in the occurrence of primary biliary cirrhosis in North East England. Am J Epidemiol 2014;179:492-498.
5. Prince MI, Chetwynd A, Diggle P, et al. The geographical distribution of primary biliary cirrhosis in a well-defined cohort. Hepatology 2001;34:1083-1088.
6. Ala A, Stanca CM, Bu-Ghanim M, et al. Increased prevalence of primary biliary cirrhosis near Superfund toxic waste sites. Hepatology 2006;43:525-531.
7. Kita H, He XS, Gershwin ME. Autoimmunity and environmental factors in the pathogenesis of primary biliary cirrhosis. Ann Med 2004;36:72-80.
8. 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.
9. Ludwig J, Dickson ER, McDonald GSA. Staging of chronic nonsuppurative destructive cholangitis (Syndrome of primary biliary cirrhosis). Virchows Arch A Pathol Anat Histol 1978;379:103-112.
10. Scheuer P. Primary biliary cirrhosis. Proc R Soc Med 1967;60:1257-1260.
11. 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.
12. Rubin DB. Multiple imputation after 18+ years. J Am Statitical Assoc 1996;91:473-489.
13. Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York: Wiley, 1987.
14. 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.
15. 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.
16. McNally RJQ, James PW, Ducker S, et al. No Rise in Incidence but Geographical Heterogeneity in the Occurrence of Primary Biliary Cirrhosis in North East England. Am J Epidemiol 2014;179:492-498.
17. James OF, Bhopal R, Howel D, et al. Primary biliary cirrhosis once rare, now common in the United Kingdom? Hepatology 1999;30:390-394.
18. Bates B, Lennox A, Prentice A, et al. National Diet and Nutrition Survey: Results from Years 1, 2, 3 and 4 (Combined) of the Rolling Programme (2008/2009-2011/2012): A Survey Carried out on Behalf of Public Health England and the Food Standards Agency., 2014.
19. Hyppönen E, Power C. Hypovitaminosis D in British adults at age 45 y: nationwide

cohort study of dietary and lifestyle predictors. Am J Clin Nutr 2007;85:860-868.

20. Guo GY, Shi YQ, Wang L, et al. Serum vitamin D level is associated with disease severity and response to ursodeoxycholic acid in primary biliary cirrhosis. Aliment Pharmacol Ther 2015;42:221-230.

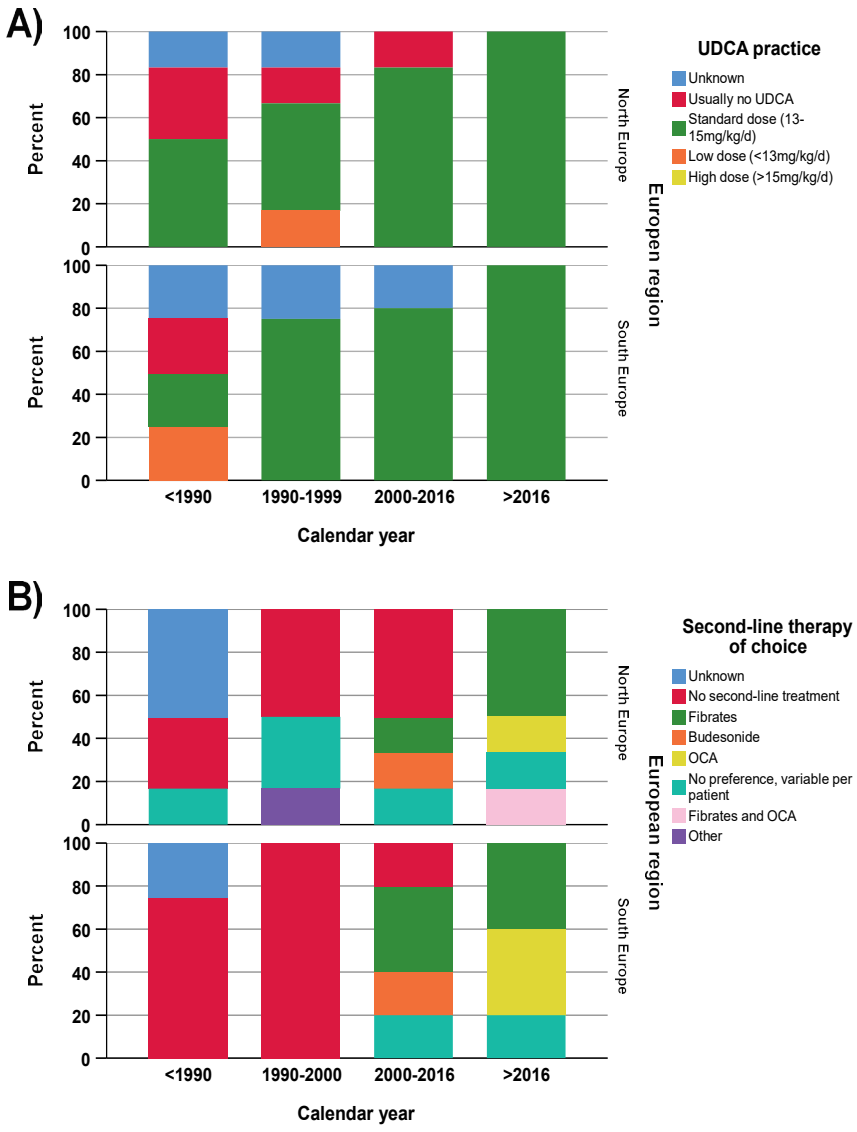
21. Agmon-Levin N, Kopilov R, Selmi C, et al. Vitamin D in primary biliary cirrhosis, a plausible marker of advanced disease. Immunol Res 2015;61:141-146.

22. OECD/ European Union. Obesity among adults. In: Health at a Glance: Europe 2018: State of Health in the EU Cycle. Paris/European Union, Brussels: OECD Publishing, 2018:126-127. https://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-europe-2018/obesity-among-adults_health_glance_eur-2018-27-en.

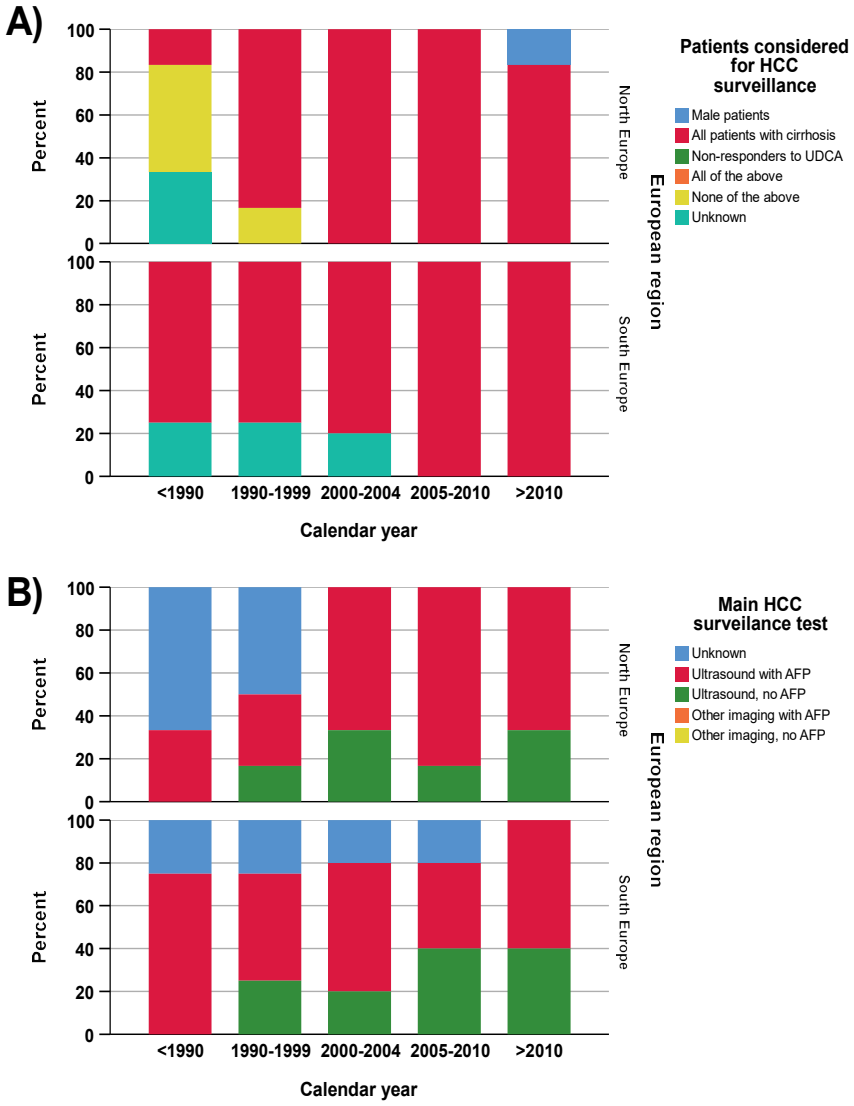
SUPPLEMENTARY DATA

Supplementary Table 1. Inclusion and allocation of Western European centers in study according to North/South and East/West

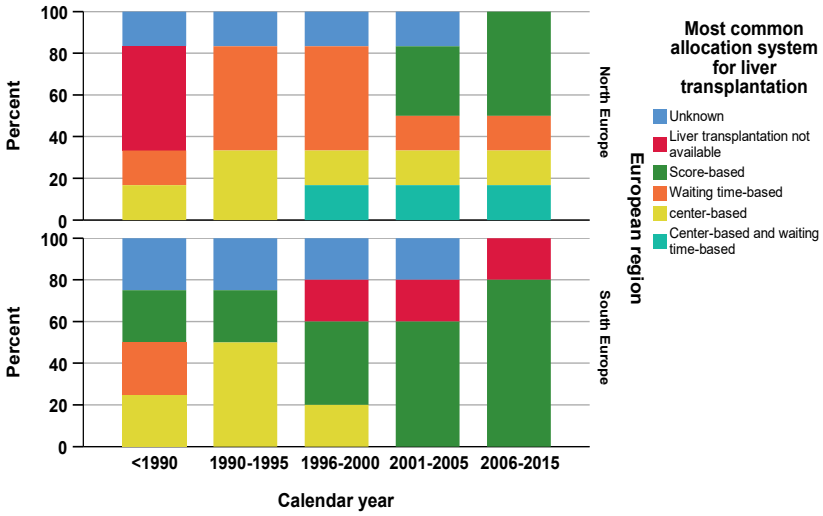
	n	North	South	East	West
the Netherlands (nation-wide cohort)	609	✓			✓
Paris, France	122	✓			✓
Birmingham, UK	214	✓			
London, UK	17	✓			✓
Ghent, Belgium	26	✓			✓
Leuven, Belgium	108	✓			✓
Jena, Germany	103	✓		✓	
Larissa, Greece	205		✓	✓	
Barcelona, Spain	170		✓		✓
Milan, Italy	78		✓	✓	
Padova, Italy	226		✓	✓	



Supplementary Figure 1. UDCA practice and second-line therapy preferability according to geographical region and calendar year for European centers in the GLOBAL PBC Study Group database.



Supplementary Figure 2. Patient population considered for hepatocellular carcinoma (HCC) surveillance and modality of HCC surveillance according to geographical region and calendar year for European centers in the GLOBAL PBC Study Group database.



Supplementary Figure 3. Liver transplantation allocation according to geographical region and calendar year for European centers in the GLOBAL PBC Study Group database.

CHAPTER 4

4

Effects of age and sex of response to ursodeoxycholic acid and transplant-free survival in patients with primary biliary cholangitis

Cheung AC*, Lammers WJ*, Murillo Perez CF, van Buuren HR, Gulamhusein A, Trivedi PJ, Lazaridis KN, Ponsioen CY, Floreani A, Hirschfield GM, Corpechot C, Mayo MJ, Invernizzi P, Battezzati PM, Parés A, Nevens F, Thorburn D, Mason AL, Carbone M, Kowdley KV, Bruns T, Dalekos GN, Gatselis NK, Verhelst X, Lindor KD, Lleo A, Poupon R, Janssen HLA, Hansen BE

**shared first co-authorship*

Clinical Gastroenterology & Hepatology, 2019

ABSTRACT

Background & Aims: Primary biliary cholangitis (PBC) predominantly affects middle-aged women; there are few data on disease phenotypes and outcomes of PBC in men and younger patients. We investigated whether differences in sex and/or age at the start of ursodeoxycholic acid (UDCA) treatment are associated with response to therapy, based on biochemical markers, or differences in transplant-free survival.

Methods: We performed a longitudinal retrospective study of 4355 adults in the Global PBC Study cohort, collected from 17 centers across Europe and North America. Patients received a diagnosis of PBC from 1961 through 2014. We evaluated the effects of sex and age on response to UDCA treatment (based on GLOBE score) and transplant-free survival using logistic regression and Cox regression analyses, respectively.

Results: Male patients were older at the start of treatment (58.3±12.1 years vs 54.3±11.6 years for women; *P*<.0001) and had higher levels of bilirubin and lower circulating platelet counts (*P*<.0001). Younger patients (45 years or younger) had increased serum levels of transaminase than older patients (older than 45 years). Patients older than 45 years at time of treatment initiation had increased odds of a biochemical response to UDCA therapy, based on GLOBE score, compared to younger patients. The greatest odds of response to UDCA were observed in patients older than 65 years (odds ratio compared to younger patients 45 years or younger, 5.48; 95% CI, 3.92–7.67; *P*<.0001). Risk of liver transplant or death (compared to a general population matched for age, sex, and birth year) decreased significantly with advancing age: hazard ratio for patients 35 years or younger, 14.59 (95% CI, 9.66–22.02) vs hazard ratio for patients older than 65 years, 1.39 (95% CI, 1.23–1.57) (*P*<.0001). On multivariable analysis, sex was not independently associated with response or transplant-free survival.

Conclusion: In longitudinal analysis of 4355 adults in the Global PBC Study, we associated patient age, but not sex, with response to UDCA treatment and transplant-free survival. Younger age at time of treatment initiation is associated with increased risk of treatment failure, liver transplant, and death.

INTRODUCTION

Primary biliary cholangitis (PBC) is a chronic autoimmune cholestatic liver disease in which patient outcome is largely dictated by the development of cirrhosis and portal hypertension.[1,2] Between 83 to 95% of patients are women, most often presenting between 40 and 60 years of age.[3]

Several studies have demonstrated that the clinical impact of PBC differs according to sex and age group.[3,4] Compared with male patients, female patients are more frequently symptomatic, with an increased burden of pruritus[5,6] and greater scores in the fatigue domain of the PBC-40 quality of life questionnaire.[7] In contrast, male PBC patients are more likely to present with advanced disease[5-7], harbor an increased risk for hepatocellular carcinoma[8], and appear to have significantly worse transplant-free survival.[4,9] Male sex has also recently been identified as a risk factor for non-response to ursodeoxycholic acid (UDCA) independent of age at presentation, presence of portal hypertension, and biochemical indices of disease severity[7], alluding to the possibility of a more rapidly progressive disease course. Age appears to add another layer of complexity to clinical phenotypes, because a study conducted by the UK-PBC consortium recognized that younger patients are affected by more severe pruritus and fatigue. Furthermore, there was a positive correlation between older age at presentation and response to UDCA in female patients; with a lesser impact evident in patients of male sex.[7]

The aim of this study was to validate the prognostic impact of presenting age and sex on treatment responses and transplant-free survival using a large, internationally representative cohort of patients with PBC.

PATIENTS AND METHODS

Subjects and study design

This was a longitudinal study of treatment response and transplant-free survival according to age and sex in a well-defined cohort from the Global PBC Study Group. Demographic, clinical, and outcome data were collected from 17 centers across Europe and North America. Patients with a short follow-up (<6 months), overlap syndrome, or another concomitant liver disease were excluded. This study included adult patients (≥18 years of age) diagnosed between 1961 and 2014 with PBC as defined by published criteria[1,10] and who were treated with UDCA.

Baseline was defined as the date of UDCA initiation. The primary endpoints were biochemical response as per the GLOBE score criteria[11] (GLOBE score below the age-specific threshold) and liver transplant-free survival. Secondary endpoints included biochemical response according to the following response criteria: Barcelona, Paris-I, Rotterdam, Toronto, and Paris-II. Patients who did not meet clinical endpoints (liver transplant or death) were censored at their last date of available follow-up. The protocol was reviewed and approved by all local Institutional Review Boards across the 17 centers.

Statistical analysis

Continuous data are expressed as mean ± standard deviation (SD) or median and interquartile range (IQR) as appropriate, and categorical data as proportions. Unpaired t-test, Mann-Whitney U test, or analysis of variance was used to determine whether there was a significant difference for continuous data, and differences in categorical data were analyzed by using the Chi-square test. Life table analysis was conducted to assess the effect of age on transplant-free survival in PBC patients when compared with a control Dutch population. PBC patients were stratified into various age groups, after which they were independently analyzed relative to the control population matched for age, sex, and birth year. Unadjusted differences in transplant-free survival between male and female patients were assessed using Kaplan-Meier estimates and compared by using the log-rank test. Univariable and multivariable associations were computed using a logistic regression for biochemical response (odds ratio [OR] and 95% confidence interval [CI]) and Cox proportional hazards regression for transplant-free survival (hazard ratio [HR] and 95% CI). Univariable analysis included sex, age at UDCA initiation, year of diagnosis, histologic stage at baseline as defined by the criteria of Scheuer[12] and Ludwig et al[3], biochemical stage at baseline as defined by ter Borg et al.[14] (mild: normal bilirubin and albumin, moderately advanced: either abnormal bilirubin or albumin, advanced: abnormal bilirubin and albumin), biochemical response, and surrogates of portal hypertension (platelet count $<150\times10^9$/L).[15,16] Age at UDCA initiation was analyzed as a continuous and categorical variable (grouped as $\leq$35, 36-45, 46-55, 56-65, and >65 years) to allow for an equitable

distribution during analysis. To account for the lack of an adequate threshold of response for the age group ≤35, age was analyzed in the following age groups for response to UDCA (≤45, 46-55, 56-65, and >65 years). The association of age with response was further assessed with a restricted cubic spline function with three knots. To determine whether age was an independent determinant of response in various subgroups, patients were categorized into 2 groups, according to whether their GLOBE score at baseline was below or above the age-specific threshold (GLOBE score status at baseline). An interaction between age and sex, and age and GLOBE score status at baseline were included in the analysis.

All analyses were performed using multiple imputation by Markov chain Monte Carlo method for missing data (Proc MI in SAS version 9.3). Multiple imputation was based on the assumption that data were missing at random, with 10 imputed datasets created from iterations to reduce sampling variability. Rubin's rules were used for the estimation of parameters of interest and standard error.[17,18] The variables included in the process of imputation were: aspartate aminotransferase (AST), alanine aminotransferase (ALT), alkaline phosphatase (ALP), total bilirubin, albumin, and platelet count.

A two-sided *P*-value of <.05 was considered statistically significant. Statistical analyses were performed using IBM SPSS Statistics for Windows, version 24.0 (IBM Corp., Armonk, NY, USA) and SAS version 9.3 (SAS Institute Inc., Cary, NC, USA).

RESULTS

Patient characteristics

For the current study, a total of 4355 UDCA-treated PBC patients were analyzed after excluding those <18 years of age (n=3) and with no (n=647) or unknown treatment status (n=96). Four hundred forty-six (10%) were male, and 3909 (90%) were female with a median follow-up of 7.7 years (IQR 3.9-12.0); 576 patients died (276 deaths were liver-related) and 330 patients underwent transplantation.

Clinical differences at baseline between sexes and age groups

At the time of UDCA initiation, male patients were older than female patients (58.3 ± 12.1 years vs 54.3 ± 11.6 years, *P*<.0001), exhibited greater median serum bilirubin values (0.82 ×ULN [IQR 0.59-1.49] vs 0.62 ×ULN [IQR 0.44-1.00], *P*<.0001) and were more often thrombocytopenic (platelet count <150×10^9/L, 21% vs 14%, *P*=.001) (**Table 1**). Concurrently, patients presenting at a younger age more often manifest an earlier disease stage, both biochemically and histologically, albeit with significantly greater serum transaminases and ALP values than older patients (*P*<.0001) (**Table 2**). The ALP levels in age groups 36-45 and 46-55 were significantly lower from that of 56-65 and >65 (P<.001). Furthermore, there was

significantly higher ALP in the group 56-65 compared to the age group >65 (P=.001). Younger patients were more likely to present with an ALP>4×ULN: 27%, 31%, 27%, 18%, and 14%, in order from youngest to oldest (P<.001).

Table 1. Baseline characteristics of the total cohort of PBC patients and according to sex

Parameter	All patients n=4355	Male n=446	Female n=3909	P-value
Age at beginning of follow-up, mean±SD	54.7 ± 11.7	58.3 ± 12.1	54.3 ± 11.6	<.0001
AMA-positive, no. (%)[a]	3849 (90.7)	410 (92.8)	3439 (90.5)	.12
Year of diagnosis, no. (%)				.44
<1990	816 (18.7)	79 (17.7)	737 (18.9)	
1990-2000	1678 (38.5)	161 (36.1)	1517 (38.8)	
2000-2010	1655 (38.0)	181 (40.6)	1474 (37.7)	
>2010	206 (4.7)	25 (5.6)	181 (4.6)	
Biochemical disease stage, no. (%)[b]				<.0001
Early	1731 (67.9)	146 (52.9)	1585 (69.8)	
Moderate	618 (24.3)	91 (33.0)	527 (23.2)	
Advanced	199 (7.8)	39 (14.1)	160 (7.0)	
Histological disease stage, no. (%)[c]				.84
Early stage disease (F1-2)	1225 (68.3)	121 (67.6)	1104 (68.4)	
Late stage disease (F3-4)	569 (31.7)	58 (32.4)	511 (31.6)	
Portal hypertension, no. (%)[d]	368 (14.8)	59 (21.2)	309 (13.9)	.001
Laboratory parameters, median (IQR)[e]				
AST (×ULN)	1.43 (0.94-2.23)	1.40 (0.92-2.13)	1.45 (0.94-2.23)	.42
ALT (×ULN)	1.64 (1.00-2.60)	1.65 (1.00-2.61)	1.64 (1.00-2.60)	.96
ALP (×ULN)	2.07 (1.30-3.71)	2.00 (1.30-3.40)	2.10 (1.30-3.74)	.26
Albumin (×LLN)	1.14 (1.06-1.23)	1.14 (1.03-1.24)	1.14 (1.06-1.23)	.05
Total bilirubin (×ULN)	0.65 (0.45-1.04)	0.82 (0.59-1.49)	0.62 (0.44-1.00)	<.0001
Platelets (x10^9/L)	244 (186-297)	216 (162-262)	248 (190-300)	<.0001

PBC, primary biliary cholangitis; SD, standard deviation; AMA, anti-mitochondrial antibody; IQR, interquartile range; AST, aspartate aminotransferase; ULN, upper limit of normal; ALT, alanine aminotransferase; ALP, alkaline phosphatase; LLN, lower limit of normal

[a] AMA status was unavailable for 112 patients (4 males, 108 females).

[b] Biochemical disease stage defined as per ter Borg et al.[14] (early: normal serum bilirubin and albumin levels, moderate: abnormal serum bilirubin or albumin levels, advanced: abnormal serum bilirubin and albumin levels). Insufficient data for determination in 41.5% (n=1807, all patients); 38.1% (n=170, males) and 41.9% (n=1637, females).

[c] Baseline biopsy was performed in 51.5% (2244 patients; 232 males and 2012 females). Baseline histological disease stage was unavailable in 20.1% (n=450, all patients), 22.8% of male patients (n=53) and 19.7% of female patients (n=397).

[d] Portal hypertension defined as a platelet count <150x10^9/L. Platelet count was available for 57.2% (2494 patients; 278 males and 2216 females).

[e] Due to differences in normal thresholds between centres, laboratory values are listed as factors of the upper and lower limit of normal.

Table 2. Baseline characteristics according to age at UDCA initiation

Parameter	≤35 n=199	36-45 n=727	46-55 n=1305	56-65 n=1234	>65 n=890	P-value
Male sex	15 (7.5)	55 (7.6)	97 (7.4)	140 (11.3)	139 (15.6)	<.0001
AMA-positive	172 (87.8)	634 (90.7)	1139 (89.9)	1094 (90.6)	810 (92.8)	.11
Diagnosis year						<.0001
<1990	28 (14.0)	112 (15.4)	239 (18.3)	268 (21.7)	169 (19.0)	
1990-1999	64 (32.2)	330 (45.4)	530 (40.6)	460 (37.3)	294 (33.0)	
2000-2010	101 (50.8)	259 (35.6)	482 (36.9)	444 (36.0)	369 (41.5)	
>2010	6 (3.0)	26 (3.6)	54 (4.1)	62 (5.0)	58 (6.5)	
Biochemical Disease stage[a]						.003
Early	71 (67.6)	238 (63.1)	496 (69.5)	549 (71.6)	377 (64.4)	
Moderate	30 (28.6)	116 (30.8)	164 (23.0)	159 (20.7)	149 (25.5)	
Advanced	4 (3.8)	23 (6.1)	54 (7.6)	59 (7.7)	59 (10.1)	
Histological Disease stage[b]						.03
Early (F1-2)	74 (77.1)	250 (69.6)	443 (70.8)	288 (64.4)	170 (63.9)	
Late (F3-4)	22 (22.9)	109 (30.4)	183 (29.2)	159 (35.6)	96 (36.1)	
Portal hypertension[c]	8 (7.4)	29 (7.5)	81 (11.6)	117 (15.8)	133 (23.8)	<.0001
Laboratory values						
AST (×ULN)	1.67 (1.03-2.75)	1.89 (1.15-2.80)	1.53 (1.00-2.40)	1.30 (0.90-2.00)	1.23 (0.83-1.87)	<.0001
ALT (×ULN)	2.25 (1.43-3.89)	2.46 (1.37-3.80)	1.87 (1.20-2.86)	1.49 (1.00-2.20)	1.20 (0.79-1.84)	<.0001
ALP (×ULN)	2.01 (1.12-4.22)	2.55 (1.49-4.83)	2.33 (1.45-4.13)	2.00 (1.30-3.31)	1.15 (1.20-2.76)	<.0001
Albumin (×LLN)	1.17 (1.09-1.27)	1.16 (1.06-1.25)	1.16 (1.08-1.26)	1.14 (1.06-1.24)	1.11 (1.01-1.20)	<.0001
Total bilirubin (×ULN)	0.64 (0.42-1.06)	0.71 (0.48-1.23)	0.60 (0.41-1.00)	0.67 (0.48-1.00)	0.67 (0.48-1.05)	.003
Platelet count ($x10^9$/L)	261 (224-302)	277 (222-331)	253 (202-304)	233 (177-284)	216 (151-272)	<.0001

UDCA, ursodeoxycholic acid; AMA, anti-mitochondrial antibody; IQR, interquartile range; AST, aspartate aminotransferase; ULN, upper limit of normal; ALT, alanine aminotransferase; ALP, alkaline phosphatase; LLN, lower limit of normal.
All data presented as no. (%), expect laboratory values, which are expressed as median (IQR).
[a] Biochemical disease stage defined as per ter Borg et al.[14] - early: normal serum bilirubin and albumin levels, moderate: abnormal serum bilirubin or albumin levels, advanced: abnormal serum bilirubin and albumin levels. Insufficient data for determination in 41.5% (94, 350, 591, 467, and 305 for each age group).
[b] Baseline histological disease stage was not available in 23% (n=28), 21% (n=95), 15% (n=114), 23% (n=135), 23% (n=266) in each respective age group (listed from youngest to oldest).
[c] Portal hypertension defined as a platelet count <$150x10^9$/L. Platelet count was unavailable for 57.3% (91, 339, 608, 493, 330 for each respective age group listed from youngest to oldest).

Effect of age on biochemical response to ursodeoxycholic acid

Across the cohort in its entirety, laboratory data after 1 year of UDCA therapy were available for 4200 patients (96%). On univariable analysis, older age at UDCA initiation was associated with higher likelihood of achieving biochemical response according to the GLOBE score (per 10-year increase in age: OR 1.25, 95% CI 1.18-1.32, *P*<.0001). After adjusting for additional baseline factors, older patients appeared to have significantly better response than younger

patients (**Table 3**). The same relationship was observed when age was analyzed as a categorical variable in multivariable analysis: ≤45 years (reference group), 46-55 years (OR 2.67, 95% CI 2.06-3.46, *P*<.0001), 56-65 years (OR 4.91, 95% CI 3.68-6.56 *P*<.0001), >65 years (OR 5.48, 95% CI 3.92-7.67, *P*<.0001). When analyzing the effect of age on biochemical response to UDCA in male and female patients separately, age had a similar effect in both (**Supplementary Figure 1**). In independent multivariable logistic regressions, the OR per 10-year increase in age was comparable for male and female patients (male: OR 1.65, 95% CI 1.27-2.12, P<.001; female: OR 1.49, 95% CI 1.36-1.64, P<.001). Furthermore, the interaction term between age and sex was not significantly different (*P*=.66) and there was no evidence of an additive interaction. Age was also a significant predictor for the other response criteria except Rotterdam criteria (**Supplementary Table 1**). Older patients had a higher probability of response than younger patients, irrespective of whether their GLOBE score at baseline was below or above the age-specific thresholds (**Figure 1**). The effect of age was also assessed with a restricted cubic spline function, which suggested the positive effect of age is less pronounced after the age of 65 (P=.004) (**Supplementary Figure 2**).

Because younger patients were more likely to have elevated ALT and AST levels, we assessed whether these biochemical markers were independently associated with response to UDCA in separate multivariable models while adjusting for center, sex, age, year of diagnosis, response at baseline and log bilirubin. AST but not ALT was an independent predictor of response (ALT [log]: OR 1.15, 95% CI 0.84-1.58, P=.39; AST [log]: OR 0.57, 95% CI 0.39-0.84, P=.004).

Effect of sex on biochemical response to ursodeoxycholic acid

Overall, male patients had significantly lower biochemical response compared to female patients (62% [n=263] vs 72% [n=2732], *P*<.0001) and male sex was associated with lower response (OR 0.62, 95% CI 0.50-0.76, *P*<.0001). However, after adjusting for parameters corresponding to disease severity (baseline bilirubin, ALP, albumin, and platelet count), age at UDCA initiation, year of diagnosis, GLOBE score status at baseline, and center, male sex was no longer an independent predictor of response (OR 0.77, 95% CI 0.57-1.04, *P*=.09) (**Table 3**). Comparable results were found for the other response criteria (**Supplementary Table 2**).

Table 3. Multivariable logistic regression for factors affecting biochemical response to UDCA according to the GLOBE score criteria

Baseline variable	Entire cohort (n=4200)		
	OR	95% CI	P-value
Male sex	0.77	0.57-1.04	.09
Age at UDCA initiation (per 10-year increase)	1.51	1.37-1.64	<.0001
Year of diagnosis (per 10-year increase)	1.11	0.96-1.27	0.16
Baseline log bilirubin (×ULN)	0.06	0.03-0.09	<.0001
Baseline log ALP (×ULN)	0.28	0.19-0.40	<.0001
Baseline albumin (×LLN) (per 0.5 increase)	3.75	2.57-5.49	<.0001
Baseline platelet count (per 50×10^9/L increase)	1.49	1.42-1.57	<.0001
GLOBE score below threshold at baseline	3.76	2.85-4.95	<.0001

UDCA, ursodeoxycholic acid; OR, odds ratio; CI, confidence interval; ULN, upper limit of normal; ALP, alkaline phosphatase; AST, aspartate aminotransferase.

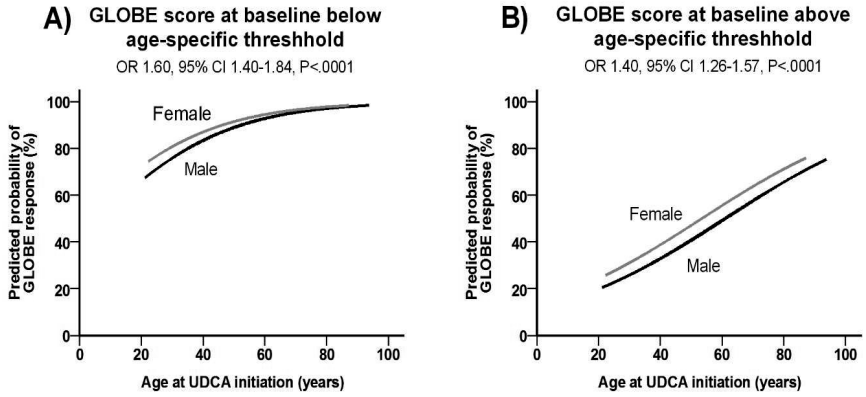


Figure 1. Predicted probability of GLOBE response according to age at the start of ursodeoxycholic acid (UDCA) treatment stratified by GLOBE status at baseline. Predicted probability of GLOBE response according to age at UDCA initiation in (A) patients whose GLOBE score at baseline is below the age-specific threshold (n=2621) and (B) patients whose GLOBE score at baseline is above the age-specific threshold (n=1579). Predicted probabilities obtained from a logistic regression correspond to a PBC patient diagnosed in 2000 with median laboratory values after adjusting for diagnosis year, sex, bilirubin, albumin, and platelet count. Odds ratios (ORs) and 95% confidence intervals (CIs) are with respect to 10-year increase in age.

Transplant-free survival among different age groups

The 10-year transplant-free survival rate decreased with age in the corresponding age groups from youngest to oldest: 89.4%, 87.0%, 82.4%, 77.7%, and 64.1% (*P*<.001). To gain additional insight into the effect of age on transplant-free survival of PBC patients, they were assessed relative to a general population (matched according to age, sex, and birth year) within each age group. On life table analysis, the PBC population within each age group had significantly lower transplant-free survival than the matched general population (**Figure 2A**). Interestingly, the transplant-free survival HR relative to a general population significantly decreased with advancing age (*P*<.0001) (**Figure 2B**). PBC patients who were ≤35 years old had the highest HR (HR 14.59, 95% CI 9.66-22.02, *P*<.0001) and patients >65 years of age had the lowest (HR 1.39, 95% CI 1.23-1.57, *P*<.0001). The distribution of clinical events from the 5-year transplant-free survival (n=67) was also significantly variable with age, because younger patients more often received a liver transplant and older patients experienced increased mortality that was primarily liver-unrelated (*P*<.0001) (**Supplementary Figure 3A**).

Transplant-free survival among male and female patients

On crude analysis of overall transplant-free survival, male patients had a significantly lower 10-year transplant-free survival rate than female patients (67.7% vs 80.1%, *P*<.0001) (**Supplementary Figure 4A**). However, after adjusting for age at UDCA initiation, year of diagnosis, bilirubin, ALP, platelet count, and center, the increased risk for liver transplantation or death in male patients was no longer significant (HR 1.19, 95% CI 0.99-1.44, *P*=.07) (**Table 4, Supplementary Figure 4B**). There was also no significant difference in the distribution of clinical events from the 5-year transplant-free survival between male and female patients (**Supplementary Figure 3B**).

Table 4. Multivariable Cox regression for factors affecting transplant-free survival

Baseline variable	Entire cohort (n=4349)		
	HR	95% CI	P-value
Male sex	1.19	0.99-1.44	.07
Age at UDCA initiation (per 10-year increase)	1.55	1.47-1.66	<.0001
Year of diagnosis (per 10-year increase)	0.83	0.75-0.92	.001
Baseline log bilirubin (×ULN)	7.34	6.03-8.93	<.0001
Baseline log ALP (×ULN)	1.88	1.47-2.40	<.0001
Platelet count (per 50×10^9/L increase)	0.88	0.83-0.93	<.0001

HR, hazard ratio; CI, confidence interval; ULN, upper limit of normal; ALP, alkaline phosphatase.

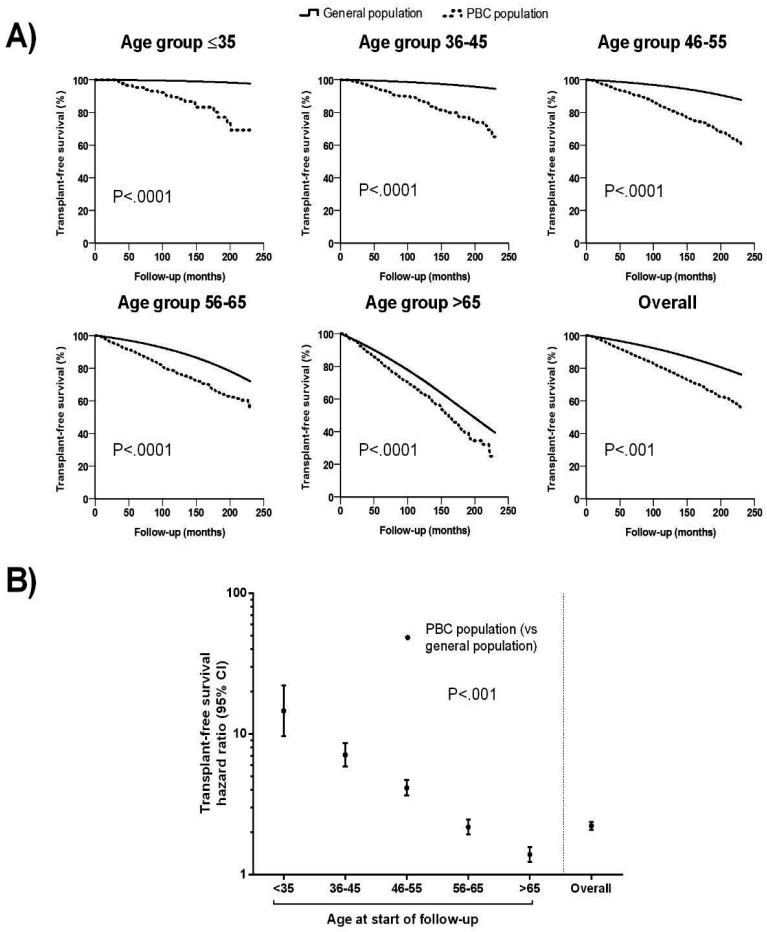


Figure 2. Transplant-free survival according to age at the start of ursodeoxycholic acid (UDCA) of PBC patients compared with age-, sex-, and birth year-matched general population. (A) Life table analysis of transplant-free survival in different age groups relative to matched general population. (B) Transplant-free survival hazard ratios (95% CI) obtained from Cox regression analyses of PBC patients relative to matched general population and in different age groups. Age was a significant determinant of the transplant-free survival hazard ratio relative to matched general population (P<.001). CI, confidence interval; PBC, primary biliary cholangitis.

DISCUSSION

The results of our study confirm that a younger age at presentation confers an impaired biochemical response to UDCA compared with older patients, even after adjusting for sex and disease severity. Despite manifesting less severe biochemical and histologic disease, younger patients exhibit more pronounced biochemical hepatitic activity, as evident by significantly greater serum transaminases levels.[19] Moreover, younger age is also associated with markedly lower transplant-free survival relative to a matched general population. Conversely, patient sex does not appear to be an independent determinant of biochemical response or transplant-free survival, but rather, male patients present with more advanced disease, a known cause of diminished treatment response and prognosis in PBC.[2,20]

Similar to other diseases of autoimmune origin, the pathogenesis of PBC appears to be driven by fundamental differences in susceptibility across male and female patients, as well as different age groups. The inherent challenges posed by the epidemiology of PBC have led to an elusive understanding of whether male patients or younger patients have a more aggressive disease phenotype. Owing to the size of the cohort and statistical validation through the use of center-specific stratification and multiple imputation, our study is strongly positioned to explore outcomes in small subgroups while minimizing bias.

In line with our study's findings, Carbone *et al.* found that when response was stratified by sex, it appeared that older female patients have significantly better response than their younger counterparts, whereas male patients have weak age-associated response rates.[7] However, we found that the effect of age on response to UDCA is similar in both sexes. There are several potential reasons for lower rates of biochemical response in younger patients. Although the relationship between age and medication compliance is complex, one possibility is that younger patients have reduced compliance, as demonstrated in other chronic disease literature.[21,22] Alternatively, disparities in response may be related to underlying disease pathology. Patients with ductopenia have been previously demonstrated to have diminished response to UDCA[23], and descriptions of a severe ductopenic variant of PBC all involved patients younger than 50 years of age.[24] Younger patients were more likely to present with severe disease, as determined by ALP levels above 4×ULN and thus it is possible that they also have a predominantly ductopenic phenotype that is particularly resistant to UDCA treatment. In addition, patients in our cohort younger than the age of 45 appeared to have higher AST and ALT, which may suggest more exuberant histologic inflammation. Interestingly, Carbone *et al.* found that younger patients were more likely to fail therapy based on transaminase criteria[7], which collectively implies a more hepatitic phenotype. Alternatively, it may reflect a more advanced disease because of the association of AST elevations with

cirrhosis. Indeed, AST was an independent predictor of response in our cohort. Furthermore, we demonstrate that the effect of age on response rates does not vary according to their status at baseline (criteria for response evaluated at baseline).

The life expectancy of asymptomatic patients diagnosed at 55 years or older has been shown to be comparable with a matched population.[25] Similarly, we found that when our cohort of PBC patients was matched to a general population, the risk for liver transplantation and death incrementally decreased with increasing age. Taking into account that younger patients are less likely to respond to UDCA, this suggests that younger patients could have lower transplant-free survival than their older counterparts as a consequence of diminished treatment response.

Earlier studies have demonstrated that male patients present with more advanced disease, reflected by their higher rates of jaundice, variceal bleeding, and thrombocytopenia at presentation.[5,7] Asymptomatic male patients also present at an older age than female patients, with a mean difference of approximately 5 years.[5,7] The UK-PBC cohort also showed that male sex was an independent predictor of biochemical response. In contrast, our study demonstrated that sex was not independently associated with biochemical response or transplant-free survival. In a previous study of a Dutch population, sex was also not an independent predictor of response to UDCA.[26] The lack of association between male sex and clinical outcomes in PBC suggests that sex is not an inherent determinant of treatment response or prognosis, but rather that males are at greater risk of presenting with more advanced disease, with a greater degree of hepatic synthetic dysfunction and portal hypertension. A possible factor to explain this finding is that the diagnosis of PBC is not sufficiently considered in male patients presenting with features of liver disease. However, this is highly speculative, and it may well be that male patients develop less frequent or less severe symptoms and therefore remain undiagnosed until later in the course of the disease. Lastly, although this is the largest study of the impact of male sex on transplant-free survival, it is possible that we were insufficiently powered to detect a small effect size. This would suggest that despite adequate biochemical response, additional factors are leading to decreased transplant-free survival in male patients. This highlights the need for further research evaluating sex-specific factors in the outcome of PBC patients, both from an epidemiologic standpoint as well as clinical trials.

In conclusion, patient age irrespective of sex has significant impact on biochemical response and transplant-free survival. Our data suggest that younger patients should be monitored carefully, with early consideration for additional therapies, because they appear to be at greatest risk of biochemical non-response to UDCA, liver transplantation, and death. The

presence of more overt biochemical hepatitic activity suggests a more aggressive and inflammatory phenotype in younger compared with older patients. Conversely, males appear to be diagnosed at a more advanced disease stage, putatively accounting for the differences in biochemical response rates compared with female patients. It is thus important to prevent diagnostic delays by maintaining a high index of suspicion for PBC in male patients and aggressively managing any potential concomitant causes of progressive fibrosis. Further studies are required to unravel the mechanisms underlying the diminished treatment response to UDCA and transplant-free survival in young patients.

REFERENCES

1. Lindor KD, Gershwin ME, Poupon R et al. Primary biliary cirrhosis. Hepatology 2009;50:291-308.
2. Poupon RE, Lindor KD, Cauch-Dudek K et al. Combined analysis of randomized controlled trials of ursodeoxycholic acid in primary biliary cirrhosis. Gastroenterology 1997;113:884-90.
3. Boonstra K, Beuers U, Ponsioen CY. Epidemiology of primary sclerosing cholangitis and primary biliary cirrhosis: a systematic review. J Hepatol 2012;56:1181-8.
4. 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;50:1884-92.
5. Rubel LR, Rabin L, Seeff LB et al. Does primary biliary cirrhosis in men differ from primary biliary cirrhosis in women? Hepatology 1984;4:671-7.
6. Lucey MR, Neuberger JM, Williams R. Primary biliary cirrhosis in men. Gut 1986;27:1373-6.
7. 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;144:560-9.
8. Trivedi PJ, Lammers WJ, van Buuren HR et al. Stratification of hepatocellular carcinoma risk in primary biliary cirrhosis: a multicentre international study. Gut 2016;65:321-9.
9. 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.
10. European Association for the Study of the Liver. EASL Clinical Practice Guidelines: The diagnosis and management of patients with primary biliary cholangitis. J Hepatol 2017;67:145-72.
11. 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;149:1804-12.
12. Scheuer P. Primary biliary cirrhosis. Proc R Soc Med 1967;60:1257-60.
13. Ludwig J, Dickson ER, McDonald GS. Staging of chronic nonsuppurative destructive cholangitis (syndrome of primary biliary cirrhosis). Virchows Arch A Pathol Anat Histol 1978;379:103-12.
14. ter Borg PC, 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;101:2044-50.
15. 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;59:327-35.
16. Levy C, Zein CO, Gomez J et al. Prevalence and predictors of esophageal varices in patients with primary biliary cirrhosis. Clin Gastroenterol Hepatol 2007;5:803-8.
17. Rubin DB. Multiple imputation for nonresponse in surveys. John Wiley & Sons, Inc.: Toronto, Canada, 1987.
18. Rubin DB. Multiple Imputation after 18+ Years. J Am Stat Assoc 1996;91:473-89.

19. Hiramatsu K, Aoyama H, Zen Y et al. Proposal of a new staging and grading system of the liver for primary biliary cirrhosis. Histopathology 2006;49:466-78.

20. 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;136:1281-7.

21. Cohen MJ, Shaykevich S, Cawthon C et al. Predictors of medication adherence postdischarge: the impact of patient age, insurance status, and prior adherence. J Hosp Med 2012;7:470-5.

22. Rolnick SJ, Pawloski PA, Hedblom BD et al. Patient characteristics associated with medication adherence. Clin Med Res 2013;11:54-65.

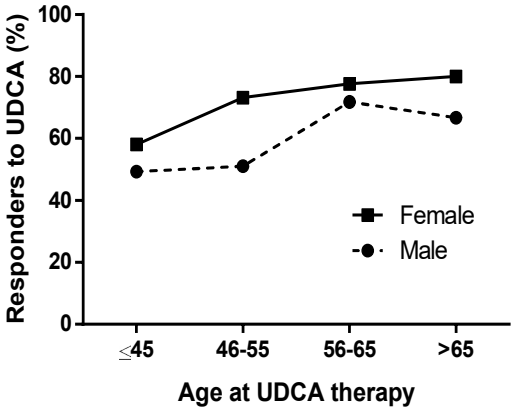
23. Kumagi T, Guindi M, Fischer SE et al. Baseline ductopenia and treatment response predict long-term histological progression in primary biliary cirrhosis. Am J Gastroenterol 2010;105:2186-94.

24. Vleggaar FP, van Buuren HR, Zondervan PE et al. Jaundice in non-cirrhotic primary biliary cirrhosis: the premature ductopenic variant. Gut 2001;49:276-81.

25. Kubota J, Ikeda F, Terada R et al. Mortality rate of patients with asymptomatic primary biliary cirrhosis diagnosed at age 55 years or older is similar to that of the general population. J Gastroenterol 2009;44:1000-6.

26. 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;74:240-6.

SUPPLEMENTARY DATA



Supplementary Figure 1. The impact of age on response to UDCA according to GLOBE score stratified by sex.

Supplementary Table 1. Multivariable logistic regression of the effect of age on biochemical response to UDCA according to various published criteria

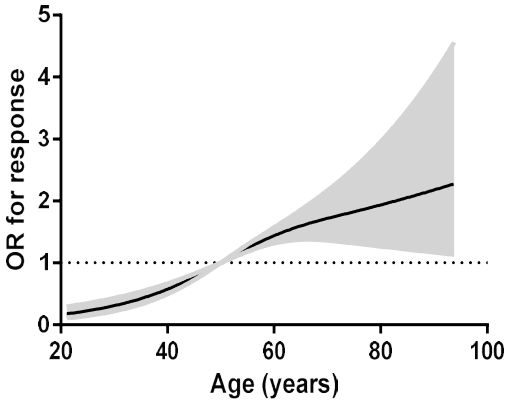
Criteria	OR[a]	95% CI	P-value
Barcelona	1.12	1.05-1.18	<.0001
Paris-I	1.15	1.06-1.23	.001
Rotterdam	0.94	0.88-1.02	.14
Toronto	1.25	1.16-1.34	<.0001
Paris-II	1.15	1.08-1.23	<.0001

UDCA, ursodeoxycholic acid; OR, Odds ratio; CI, confidence interval.
[a] OR corresponds to a 10-year increase in age.

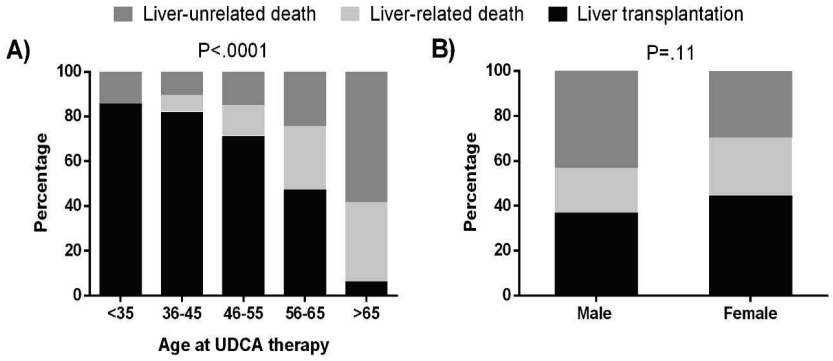
Supplementary Table 2. Multivariable logistic regression of the effect of sex on biochemical response to UDCA according to various published criteria

Criteria	OR[a]	95% CI	P-value
Barcelona	1.05	0.84-1.31	.66
Paris-I	0.88	0.66-1.16	.35
Rotterdam	0.80	0.61-1.05	.10
Toronto	0.80	0.60-1.05	.10
Paris-II	1.06	0.82-1.37	.67

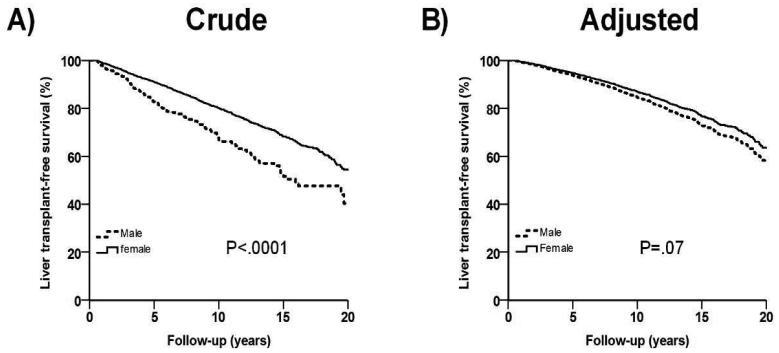
UDCA, ursodeoxycholic acid; OR, Odds ratio; CI, confidence interval.
[a] OR corresponds to male sex.



Supplementary Figure 2. The association between age and the odds for response to UDCA according to the GLOBE score modelled by a restricted cubic spline function. The age reference is 50 years. The test for curvature was curvature was significant (P=.004).



Supplementary Figure 3. The distribution of clinical events at 5 years stratified by age at the start of ursodeoxycholic acid and sex. The distribution of liver-unrelated death, liver-related death, and liver transplantations at 5 years (n=67) according to (A) age at the start of ursodeoxycholic acid and (B) sex were compared by Chi-square tests.



Supplementary Figure 4. Crude and adjusted transplant-free survival curves of males and females. Survival curves of (A) unadjusted (crude) transplant-free survival and (B) adjusted transplant-free survival between males and females. Cox regression analysis (n=4349) was adjusted for center, age at the start of ursodeoxycholic acid, year of diagnosis, serum alkaline phosphatase levels, serum bilirubin levels, and platelet count.

CHAPTER 5

5

Fibrosis stage is an independent predictor of outcome in primary biliary cholangitis despite biochemical treatment response

Murillo Perez CF, Hirschfield GM, Corpechot C, Floreani A, Mayo MJ, van der Meer A, Ponsioen CY, Lammers WJ, Parés A, Invernizzi P, Carbone M, Battezzati PM, Nevens F, Kowdley KV, Thorburn D, Mason AL, Trivedi PJ, Lindor KD, Bruns T, Dalekos GN, Gatselis NK, Verhelst X, Janssen HLA, Hansen BE, Gulamhusein A

Alimentary Pharmacology & Therapeutics, 2019

ABSTRACT

Background: Fibrosis stage predicts prognosis in patients with chronic liver disease independent of disease aetiology, although its precise role in risk stratification in patients with primary biliary cholangitis (PBC) remains undefined.

Aims: To assess the utility of baseline fibrosis stage in predicting long-term outcomes in the context of biochemical risk stratification.

Methods: In a large and globally representative cohort of patients with PBC, liver biopsies performed from 1980 to 2014 were evaluated. The predictive ability of histologic fibrosis stage in addition to treatment response at 1 year (Toronto/Paris-II criteria), as well as non-invasive markers of fibrosis (AST/ALT ratio [AAR], AST to platelet ratio index [APRI], FIB-4), for transplant-free survival was assessed with Cox proportional-hazards models.

Results: There were 1828 patients with baseline liver biopsy. Advanced histologic fibrosis (stage 3/4) was an independent predictor of survival in addition to non-invasive measures of fibrosis with the hazard ratios ranging from 1.59-2.73 ($P<0.001$). Patients with advanced histologic fibrosis stage had worse survival despite biochemical treatment response, with a 10-year survival of 76.0-86.6% compared to 94.5-95.1% depending on the treatment response criteria used. Poor correlations were observed between non-invasive measures of fibrosis and histologic fibrosis stage.

Conclusion: Assessment of fibrosis stage grants prognostic value beyond biochemical treatment response at 1 year. This highlights the need to incorporate fibrosis stage in individual risk stratification in patients with PBC, partly to identify those that may derive benefit from novel therapies.

INTRODUCTION

Primary biliary cholangitis (PBC) is a chronic autoimmune biliary disease characterized by immune-mediated destruction of intrahepatic bile ducts, cholestasis, portal inflammation and over time, a progression to end-stage biliary cirrhosis. While the histologic features can serve as diagnostic parameters, for the majority of patients, in keeping with consensus treatment guidelines, diagnosis is made based on the co-incidence of cholestatic serum liver tests and specific anti-mitochondrial antibodies. This has meant the utility of routine baseline biopsy for patients living with PBC has been questioned; yet liver biopsy findings, beyond diagnosis, may be of value in determining the risk of end-stage liver disease complications, by the understanding offered from directly evaluating stage of liver fibrosis at presentation. Yet relatively little formal evaluation exists exploring how the information gained from a baseline liver biopsy in PBC as regards to histologic stage, can contribute to long-term risk stratification for individual patients. While on its own, advanced histologic stage is an independent predictive factor for transplant-free survival conferring a 1.5-fold increased risk for liver transplantation or death,[1,2] less is known of its utility in the context of biochemical risk tools; and by extension, the value baseline liver fibrosis stage may offer for refining timely adoption of therapies designed to optimize outcomes of patients.[1,3–6]

In PBC, most laboratory tools designed to allow clinical risk stratification rely on biochemical markers, including treatment response to first-line therapy with ursodeoxycholic acid (UDCA). Bilirubin and alkaline phosphatase (ALP) are surrogate markers that, either in binary, or as part of continuous models, predict outcome including transplant-free survival in patients with PBC.[7,8] Several easily measurable non-invasive markers of fibrosis (AST/ALT ratio [AAR], AST to platelet ratio index [APRI], and FIB-4) have also been described,[5,9,10] and while they were initially developed and validated in patients with chronic viral hepatitis, their utility in fibrosis assessment and prediction of prognosis in PBC has also been suggested.[5]

Using the large dataset afforded by the Global PBC Study Group, for which long-term outcome is reported, as well as comprehensive biochemical stratifiers, we sought to robustly explore the utility for patients offered by understanding at baseline the stage of liver fibrosis. To do so, we characterized histologic fibrosis stage in a large international cohort of PBC patients; determined whether histologic fibrosis staging improves risk stratification in addition to, and independently of, biochemical non-invasive measures of fibrosis and biochemical treatment response; evaluated the correlation between histology and biochemical non-invasive markers of fibrosis; and established optimal PBC-specific thresholds of non-invasive markers of fibrosis.

METHODS

Population and study design

Data from the GLOBAL PBC Study Group database, an international cohort of patients with PBC from Europe and North America was used for this study. We included UDCA-treated patients diagnosed with PBC according to established criteria in whom a liver biopsy was performed at study entry.[11] Biopsies conducted in the 24 months prior to study entry and up to 12 months after study entry were considered as baseline in order to maximize the number of eligible patients while minimizing selection bias that may already exist in those who are biopsied. These criteria were employed given established supportive data demonstrating that short-term treatment with UDCA is not associated with regression of histologic stage.[12,13] A sub-analysis limited to patients with biopsies 6 months prior to UDCA initiation but not after, was also performed. Patients with a short follow-up (<6 months), UDCA discontinuation, unknown dates of clinical events, autoimmune hepatitis overlap syndrome, or another concomitant liver disease were excluded from the study. Patients with a biopsy outside the specified time frames were also excluded. This study was conducted in accordance with the 1975 Declaration of Helsinki. The protocol was approved by the institutional research board at all participating centers as per local regulations.

Data collection

Baseline (study entry) was defined as the date of UDCA initiation. At study entry, the following data were collected: sex, age at diagnosis, anti-mitochondrial (AMA) antibody serological status, biochemical disease stage according to Rotterdam criteria[14], and histologic stage, for which the most commonly used methods for staging were Ludwig and Scheuer's criteria.[15,16] Given the variability in staging systems used across centers, for the purposes of this study patients were categorized as having early fibrosis stage if reported as having stage 1 or 2 disease and advanced fibrosis stage if reported as having stage 3 or 4 disease or evidence for cirrhosis. The following laboratory parameters were collected every 6-12 months: total bilirubin, ALP, albumin, aspartate aminotransferase (AST), alanine aminotransferase (ALT), and platelet count.

Statistical Analyses

The primary endpoint was a composite of liver transplantation and death. The influence of the various measures of fibrosis on transplant-free survival were assessed with multivariable Cox proportional hazards' regressions (hazard ratio [HR] with 95% confidence interval [CI]) while adjusting for age at the start of UDCA, sex, center, and year of diagnosis. In order to determine whether histologic fibrosis stage confers additional prognostic information to the non-invasive

markers of fibrosis, histologic fibrosis stage≥3 was inputted into the model in a forward step approach. The impact of histologic fibrosis stage was assessed as stage 1/2 vs stage 3/4 due to potential sampling and interpretative errors associated with liver biopsy. Furthermore, given that the non-invasive tests are used as measures of fibrosis, we aimed to limit the discussion to fibrosis stage. Each non-invasive marker was evaluated as a continuous and dichotomous variable according to established thresholds that have been associated with advanced fibrosis/cirrhosis (AAR=1, APRI=2, and FIB-4=3.25),[17] with the GLOBE score analysed based on age-specific thresholds.[7] Additionally, thresholds established based on the Youden index in our cohort were also employed. APRI and FIB-4 were not normally distributed and thus were log transformed (natural logarithm [LN]) for all analyses. For elaborative purposes, the components of each non-invasive marker were evaluated independently to determine which component grants the highest predictive value for transplant-free survival, as determined by the concordance statistics (C-statistics).

In order to assess the prognostic impact of baseline histologic fibrosis stage in the context of biochemical response after 1 year of UDCA therapy, patients were stratified based on histologic fibrosis stage (early vs advanced fibrosis) and treatment response defined by Toronto (ALP≤1.67× upper limit of normal [ULN])[18] and Paris-II criteria (ALP≤1.5×ULN, AST≤1.5×ULN, and normal bilirubin)[19] to estimate their survival with a Kaplan-Meier curve. These criteria were employed because they are the most widely used to determine response to UDCA, including in clinical trials.[18,20] Advanced fibrosis was defined as stage≥3 and according to the thresholds of the non-invasive markers of fibrosis (AAR/APRI/ FIB4) established with the Youden index. A Cox regression analysis was employed to estimate the HR associated with these categories. Subsequently, we aimed to determine whether differences exist in ability to predict transplant-free survival based on methods of fibrosis assessment (histologic fibrosis stage, AAR, APRI, or FIB-4) while considering treatment response 1 year.

Non-invasive markers of fibrosis, namely AAR, APRI, and FIB-4, were calculated at baseline. The diagnostic performance of these markers for the diagnosis of histologic fibrosis stage ≥2, stage≥3, and stage=4 was assessed with a receiver operating curve. Additionally, the potential correlation between Rotterdam biochemical disease stage and the GLOBE score with histologic fibrosis stage was also assessed. A grid search of the area under the receiver operating curve (AUROC) was also used to establish disease-specific thresholds for APRI and FIB-4 with a 95% sensitivity and 95% specificity for the diagnosis of cirrhosis. The negative predictive value (NPV) and positive predictive value (PPV) of the conventional thresholds and newly determined thresholds for cirrhosis were calculated.[17] Conventional thresholds of FIB-4 to exclude advanced fibrosis, namely FIB-4≤1.45, and to rule in advanced fibrosis, FIB-4>3.25,

were used.[17] Additionally, the threshold with the highest sensitivity and specificity for the diagnosis of advanced fibrosis (stage≥3) was determined with the Youden index (J) for each non-invasive marker of fibrosis.

Multiple imputation was completed by the Markov Chain Monte Carlo (MCMC) method for missing data with SAS version 9.4 (SAS Institute Inc., Cary, NC, USA). Ten imputed data sets for missing biochemical values were generated to reduce sampling variability of laboratory results at baseline and yearly thereafter up to 15 years (SAS Proc MI, MCMC method).[21] The imputation was performed based on the assumption that data were missing at random, in which variables predicting outcomes and outcomes themselves were included in the imputation model. Rubin's rules were used to estimate the parameter and standard error.[22,23]

A *P*-value of < 0.05 was considered statistically significant. All analyses were two-sided and were performed using IBM SPSS Statistics for Windows, version 25.0 (IBM Corp., Armonk, NY, USA) and SAS version 9.4 (SAS Institute Inc., Cary, NC, USA).

RESULTS

Study population characteristics

A total of 1828 patients with a liver biopsy with assessment of histologic fibrosis stage at baseline were included, whose baseline characteristics are reported in **Table 1**. The majority of biopsies were conducted within 6 months of study entry (n=1496, 81.8%). Within the cohort, 66.9% (n=1223) had early histologic fibrosis stage (stage 1 or 2) and 33.1% (n=605) had advanced histologic fibrosis stage (stage 3 or 4). The median follow-up time was 8.4 years (4.9-12.4). The primary endpoint was met in 310 patients (194 deaths and 116 liver transplantations). The biopsies in this study were performed from 1980 to 2014, and reflective of practice, with a bias towards 1993 to 2004. Due to the unknown indications for the biopsies performed, the characteristics of patients included in this study were compared to those without a biopsy to examine a potential selection bias. Patients with a biopsy were younger and had significantly higher ALP, bilirubin, AST, and ALT, but similar albumin (**Supplementary Table 1**). While the 10-year transplant-free survival was higher for those biopsied (84.3%) compared to those not biopsied (73.8%) (**Supplementary Figure 1**), an absence of biopsy was not a significant predictor of outcome in multivariable analysis when considering age, sex, year of diagnosis, UDCA treatment, and center (HR 1.03, 95% CI 0.76-1.40, P=0.85). Response rates between the biopsied group and the non-biopsied group were similar (**Supplementary Table 1**), and even when considering response status, having a biopsy was not predictive of outcome.

Table 1. Baseline characteristics of patients with a biopsy at baseline

Baseline characteristics	Total (n=1828)
Age at study entry, mean (SD)	52.4 (11.3)
Female, n (%)	1649 (90.2)
AMA-positive, n (%)	1586/1761 (90.1)
Year of diagnosis, median (P25-75)	1998 (1993-2004)
Year of biopsy, median (P25-75)	1997 (1993-2003)
Histologic stage, n (%)	
1	649 (35.5)
2	574 (31.4)
3	289 (15.8)
4	316 (17.3)
Biochemical disease stage†, n (%)	
Early	728 (71.8)
Moderately advanced	234 (23.1)
Advanced	52 (5.1)
Laboratory parameters‡, median (P25-75)	
Total bilirubin, ×ULN	0.59 (0.43-0.91)
ALP, ×ULN	2.12 (1.34-3.83)
Albumin, ×LLN	1.16 (1.06-1.26)
AST, ×ULN	1.52 (1.06-2.32)
ALT, ×ULN	1.72 (1.09-2.70)
Platelet count, $\times 10^9$/L	253 (204-300)

SD, standard deviation; UDCA, ursodeoxycholic acid; AMA, antimitochondrial antibody; ULN, upper limit of normal; ALP, alkaline phosphatase; LLN, lower limit of normal; AST, alanine amminotransferase; ALT, alanine aminotransferase.

†Biochemical disease stage according to Rotterdam criteria[14] was available for 1014 (55.5%) patients.
‡Laboratory parameters were not available for all patients: Total bilirubin (n=1305, 71.4%), ALP (n=1370, 74.9%), albumin (n=1057, 57.8%), AST (n=1336, 73.1%), ALT (n=1330, 72.8%), platelet count (n=1098, 60.1%).

Histologic fibrosis stage is an independent predictor of transplant-free survival

We sought to confirm that assessment of fibrosis stage histologically and/or by non-invasive testing predicts outcome. Analyses of histology were limited to stage 1/2 vs stage 3/4, but incremental changes in disease stage were associated with prognosis (**Figure 1A**). The transplant-free survival at 10 years of patients with early histologic fibrosis stage was significantly improved to those with advanced stage, 91.8% vs 70.2% (P<0.001; **Figure 1**). On multivariable analysis, fibrosis stage histologically and as assessed by non-invasive markers were significantly associated with outcome with advanced histologic fibrosis stage and increased biochemical non-invasive scores being associated with an increased risk for liver transplantation or death (**Table 2**). Each of the non-invasive measures independently predicted outcome but of the ones assessed, the Rotterdam criteria was the strongest predictor of outcome (C statistic 0.78, 95% CI 0.74-0.81), followed by FIB-4 (C statistic 0.75, 95% CI 0.72-0.78), APRI (C statistic 0.71, 95% CI 0.68-0.75), and finally AAR (C statistic 0.69, 95% CI 0.65-0.72). The laboratory parameter that was most predictive of survival for Rotterdam criteria,

APRI, FIB-4, and AAR were bilirubin, platelet count, platelet count, and AST respectively. For reference, the C statistic for histologic stage 1/2 vs 3/4 was 0.72 (95% CI 0.69-0.75).

In order to assess whether histologic fibrosis stage offers prognostic value beyond that gained from non-invasive markers of fibrosis, histologic fibrosis stage≥3 was included in a forward step approach. Histologic fibrosis stage was included in each model and was an independent significant predictor of transplant-free survival, irrespective of the non-invasive markers of fibrosis used as continuous or categorical variables (**Table 2, Supplementary Table 2**). The prognostic value of non-invasive markers of fibrosis (FIB-4 and APRI) in addition to response to UDCA was also validated in the cohort of patients without a liver biopsy (**Supplementary Table 3**).

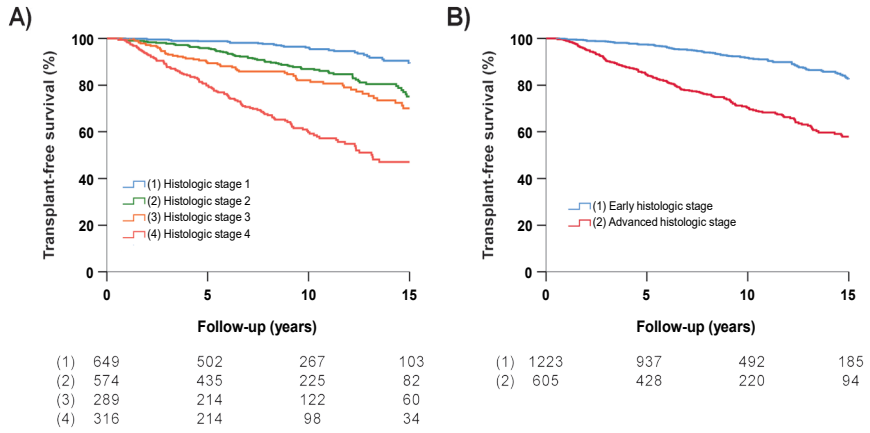


Figure 1. Transplant-free survival estimates according to baseline histology stratified by A) histologic stage and B) early and advanced histologic fibrosis stage.

Table 2. Multivariable Cox regression models that depict the independent association of baseline histologic fibrosis stage and non-invasive markers of fibrosis with transplant-free survival

	Multivariable†		Multivariable (histologic fibrosis stage) ‡		
	HR (95% CI)	**P value**		**HR (95% CI)**	**P value**
Histologic fibrosis stage					
Stage 1/2	1.00		-	-	-
Stage 3/4	2.85 (2.22-3.67)	<0.001	-	-	-
AAR	1.39 (1.25-1.54)	<0.001	AAR	1.34 (1.20-1.50)	<0.001
			Stage 3/4	2.73 (2.12-3.51)	<0.001
APRI (LN)	2.39 (2.00-2.86)	<0.001	APRI (LN)	2.11 (1.75-2.53)	<0.001
			Stage 3/4	2.07 (1.59-2.70)	<0.001
FIB-4 (LN)	3.19 (2.58-3.95)	<0.001	FIB-4 (LN)	2.76 (2.20-3.47)	<0.001
			Stage 3/4	1.96 (1.50-2.56)	<0.001
GLOBE score	2.85 (2.48-3.27)	<0.001	GLOBE score	2.68 (2.31-3.11)	<0.001
			Stage 3/4	1.59 (1.21-2.10)	0.001
Rotterdam criteria			**Rotterdam criteria**		
Mild	1.00		Mild	1.00	
Moderate	3.66 (2.73-4.89)	<0.001	Moderate	3.17 (2.36-4.26)	<0.001
Advanced	10.98 (7.09-17.00)	<0.001	Advanced	9.40 (6.00-14.72)	<0.001
			Stage 3/4	2.06 (1.57-2.69)	<0.001

HR, hazard ratio; CI, confidence interval; AAR, AST to ALT ratio; APRI, AST to platelet ratio index.

†Adjusted for age at the start of follow-up, sex, year of diagnosis, and center.
‡Adjusted for age at the start of follow-up, sex, year of diagnosis, and center. Histologic fibrosis stage (3/4 vs 1/2) was input in a forward selection approach.

Table 3. Multivariable Cox regression models that depict the combined association of response to treatment after 1 year and baseline histologic fibrosis stage with transplant-free survival

Group	HR (95% CI) †	P value
Toronto criteria‡		
Response + stage 1/2	1.00	
Non-response + stage 1/2	2.94 (1.97-4.40)	<0.001
Response + stage 3/4	3.08 (2.09-4.53)	<0.001
Non-response + stage 3/4	6.29 (4.37-9.04)	<0.001
Paris-II criteria§		
Response + stage 1/2	1.00	
Non-response + stage 1/2	2.98 (1.95-4.55)	<0.001
Response + stage 3/4	2.17 (1.29-3.64)	0.003
Nonresponse + stage 3/4	7.59 (5.14-11.22)	<0.001

HR, hazard ratio; CI, confidence interval.

†Multivariable model also adjusted for age the start of follow-up, sex, year of diagnosis, and center.
‡Response defined as ALP≤1.67×ULN.
§Response defined as ALP≤1.5×ULN, AST≤1.5×ULN, and normal bilirubin.

Histologic fibrosis stage predicts prognosis despite biochemical treatment response

The prognostic impact rendered by baseline fibrosis stage was assessed in addition to biochemical treatment response to UDCA according to Toronto and Paris-II criteria at 1year. Patients with advanced histologic fibrosis stage had a worse transplant-free survival despite treatment response, with a 10-year survival of 76.0% as compared to 94.5% according to Toronto criteria and 86.6% as compared to 95.1% for Paris-II criteria (**Figure 2**). Even with response defined as ALP normalization, patients with advanced histologic stage and normal ALP have worse transplant-free survival compared to those with early histologic stage and normal ALP (P<0.001), yet similar to those with early histologic stage and abnormal ALP (P=0.12), (**Supplementary Figure 2**). Furthermore, outcomes of patients with advanced histologic stage and normal ALP were also worse than those with early fibrosis and treatment response defined by the Toronto and Paris-II criteria (P<0.001).

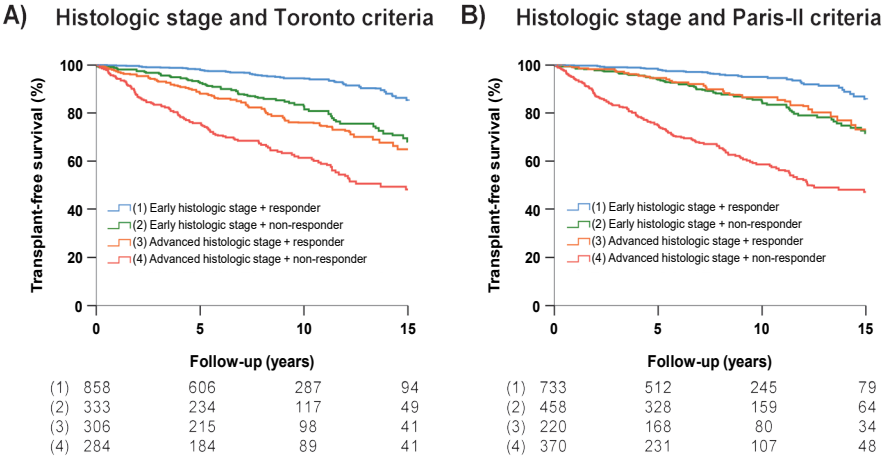


Figure 2. Transplant-free survival estimates after 1 year of treatment with ursodeoxycholic acid stratified by baseline histologic fibrosis stage and biochemical response according to A) Toronto criteria and B) Paris-II criteria. In patients with complete response after 1 year, those with early histologic stage have worse survival as compared to those with advanced histologic stage.

In multivariable analysis, a similar trend was noted (**Table 3**). All pairwise comparisons were statistically significant (P<0.01), except the risk in patients with incomplete treatment response and early histologic fibrosis stage compared to those with adequate treatment response and advanced histologic fibrosis stage (Toronto: P=0.82; Paris-II: P=0.18).

The thresholds for AAR, APRI, and FIB-4 with the greatest sensitivity and specificity in the prediction of advanced fibrosis according to the Youden index (J) were 0.9 (J=0.17), 0.8

(J=0.27), and 1.8 (J=0.29), respectively. A similar trend in outcomes was observed with Toronto criteria when advanced fibrosis was defined with these thresholds (**Supplementary Table 4**). The association of baseline histologic fibrosis stage with transplant-free survival was maintained after adjusting for ALP and AAR/APRI/FIB-4 at 1 year (**Supplementary Table 5**).

The method of assessment of fibrosis at baseline that was most predictive of outcome in addition to treatment response at 1 year was assessed. This revealed that assessment of fibrosis stage using FIB-4 in addition to biochemical response after 1 year of UDCA therapy was most predictive of transplant-free survival (Toronto: C statistic 0.78, 95% CI 0.75-0.80; Paris-II: C statistic 0.78, 95% CI 0.75-0.81). However, it was not significantly different from the other means of assessing fibrosis (histologic fibrosis stage, APRI, and AAR), suggesting all are similarly predictive of prognosis.

The distribution of events (liver transplantation, liver-related death, or liver-unrelated death) according to biochemical response and histologic stage were significantly different between these groups (**Supplementary Figure 3**). The proportion of events that were liver-related (liver transplantation or liver-related death) was higher in those with an advanced histologic compared to those with an early histologic stage, irrespective of response status. The predominance of liver-related events in patients with advanced histologic stage, even responders, further supports the importance of assessing histologic fibrosis stage in the context of biochemical response and importantly, further emphasizes that this population remains at risk for clinically relevant liver-related endpoints.

Biochemical non-invasive markers correlate poorly with histologic fibrosis stage in PBC

Correlation between biochemical markers of fibrosis and histologic fibrosis stage was assessed by the AUROC for stage≥2, stage≥3, and stage=4 for AAR, APRI, FIB-4, Rotterdam disease stage, and GLOBE score, which are shown in **Table 4**. None of the biochemical non-invasive measures of fibrosis evaluated demonstrated strong correlation with histologic fibrosis stage, although among those assessed, FIB-4, APRI, and GLOBE score showed the greatest correlation. The correlations between histologic fibrosis stage and non-invasive measures of fibrosis were lower than previously reported in the literature. When the cohort was restricted to only patients with biopsies conducted in the 6 months prior to, but not after UDCA initiation, similar findings were observed for the correlation between histologic fibrosis stage and non-invasive markers, as well as the survival associated with patients with advanced histologic stage and complete response to UDCA (**Supplementary Table 6, Supplementary Figure 4**).

Table 3. Multivariable Cox regression models that depict the combined association of response to treatment after 1 year and baseline histologic fibrosis stage with transplant-free survival

Group	HR (95% CI) †	P value
Toronto criteria‡		
Response + stage 1/2	1.00	
Non-response + stage 1/2	2.94 (1.97-4.40)	<0.001
Response + stage 3/4	3.08 (2.09-4.53)	<0.001
Non-response + stage 3/4	6.29 (4.37-9.04)	<0.001
Paris-II criteria§		
Response + stage 1/2	1.00	
Non-response + stage 1/2	2.98 (1.95-4.55)	<0.001
Response + stage 3/4	2.17 (1.29-3.64)	0.003
Nonresponse + stage 3/4	7.59 (5.14-11.22)	<0.001

HR, hazard ratio; CI, confidence interval.

†Multivariable model also adjusted for age the start of follow-up, sex, year of diagnosis, and center.
‡Response defined as ALP≤1.67×ULN.
§Response defined as ALP≤1.5×ULN, AST≤1.5×ULN, and normal bilirubin.

Table 4. The area under the curve for the diagnosis of histologic fibrosis stage ≥ 3 and stage = 4 of various biochemical non-invasive markers in our cohort

		AUROC (95% CI)	
Non-invasive marker	**Stage ≥ 2 (n=1828)**	**Stage ≥ 3 (n=1828)**	**Stage = 4 (n=1828)**
AAR	0.54 (0.51-0.57)	0.60 (0.57-0.64)	0.63 (0.59-0.67)
APRI	0.64 (0.61-0.67)	0.68 (0.65-0.71)	0.69 (0.65-0.72)
FIB-4	0.64 (0.61-0.67)	0.69 (0.67-0.72)	0.73 (0.69-0.76)
Rotterdam criteria	0.61 (0.58-0.64)	0.63 (0.61-0.66)	0.65 (0.61-0.68)
GLOBE score	0.66 (0.63-0.69)	0.70 (0.67-0.73)	0.70 (0.66-0.74)

AUROC, area under receiver operating curve; CI, confidence interval; AAR, AST to ALT ratio; APRI, AST to platelet ratio index.

Optimal thresholds for AAR, APRI, and FIB-4 for prediction of cirrhosis in PBC patients

The thresholds with 95% sensitivity in excluding cirrhosis (stage 4) for APRI and FIB-4 established in our cohort were 0.26 and 0.70, respectively, which were lower than the conventional thresholds (**Supplementary Table 7**). The conventional thresholds to 'rule out' advanced fibrosis/cirrhosis demonstrated low sensitivity but NPV was approximately 90% (**Supplementary Table 7**). Meanwhile, the thresholds with 95% specificity for APRI and FIB-4 were 2.00 and 4.03, respectively. As for APRI, the threshold developed in our cohort was equal to the conventional threshold (2.0); whereas the FIB-4 threshold developed in our cohort was higher and demonstrated improved specificity and PPV compared to the conventional threshold of 3.25. Although these thresholds are highly sensitive and specific, cirrhosis could only be excluded or included in a small proportion of patients from the total cohort (7-10%).

DISCUSSION

In PBC there remain challenges for patients and clinicians as regards the best approaches to understanding the disease course, the risk of adverse events, and identification of those patients who will get the maximum benefit from disease modifying therapies.[24] We sought to understand how baseline liver fibrosis stage could contribute to our understanding of PBC risk. We demonstrate that histologic fibrosis stage categorized as early (stage 1/2) vs. advanced (stage 3/4) fibrosis is an independent predictor of transplant-free survival while adjusting for the non-invasive markers, and importantly, that its association with outcome persists after 1 year, despite treatment response, that is, it adds prognostic information to a group with biochemical response to first-line treatment, potentially not presently readily identified. We further show that histologic fibrosis stage correlates poorly with biochemical non-invasive measures of fibrosis, namely FIB-4, APRI, and AAR; we finally establish disease-specific thresholds for cirrhosis for APRI and FIB-4.

Despite poor correlation, both histologic fibrosis stage and non-invasive markers of fibrosis were independently predictive of patient outcome suggesting that each measure captures in different ways relevant context about disease biology, be it fibrosis, cholestasis (bilirubin), inflammation in addition to fibrosis (AST), and/or portal hypertension (platelet count). Nevertheless, while there are now other modalities such as elastography that have clinical utility beyond biopsy, it is telling that our data robustly demonstrates how the assessment of fibrosis stage at baseline, by histology as well as by biochemical non-invasive measures of fibrosis, predicts prognosis independent of UDCA treatment response after 1 year. In choosing the time point for second-line therapies this raises the question of when, and for whom, to add second-line therapy. Our data identified patients with advanced histologic fibrosis stage at baseline as having ongoing poor outcome *despite* biochemical treatment response and interestingly similar outcomes to those patients with inadequate biochemical response with early fibrosis – both groups demonstrating 9-19% lower 10-year transplant-free survival as compared to patients with early fibrosis and complete response. Similar findings were seen in a smaller series of 342 patients when response was assessed according to Toronto criteria after 2 years, yet the 10-year survival difference was 4% and 13%, with the greater difference in those with adequate response and advanced fibrosis stage. [6] Importantly, the need for second-line therapies is currently largely based on an assessment of biochemical response after 1 year of therapy with UDCA and does not incorporate an understanding of fibrosis stage in risk assessment. Our findings support the use of fibrosis stage histologically (only if clinically indicated) or by non-invasive means, in addition to the current biochemical criteria to identify a less readily recognized cohort at risk of impactful outcomes. Implementation can help identify

additional patients who may benefit from second-line therapies, namely biochemical responders that have an advanced fibrosis stage.

Our findings emphasize that while biochemical markers of cholestasis are important in risk stratification, determining fibrosis stage histologically, or arguably now via non-invasive measures that are increasingly available including tools such as vibration controlled transient elastography (VCTE), must also be considered in the risk stratification of patients with PBC.[25,26] VCTE has shown strong correlation with advanced fibrosis stage in contemporary cohorts of patients PBC, though this data were not available in our cohort and were not readily available in many centers at the time of baseline liver biopsy. Additionally, although the access to elastography has increased significantly with time, it is still not a universally applicable tool for all patients. Although not part of our study, clear ongoing efforts to evaluate elastography in the management of patients with PBC are of clinical interest. Further, while liver biopsy at baseline is clinically logical, the timing of elastography as a surrogate is challenged by the combined effects of fibrosis and cholestasis on liver stiffness. While identification of patients with advanced stage disease is clearly important, there requires further study to understand whether new licenced therapies such as obeticholic acid or alternative off-label agents including bezafibrate, offer the same benefit to this subgroup with potentially normal (or near-normal) ALP values.[20,27]

Similar to prior smaller series,[28] our large cohort study confirmed suboptimal correlation between histologic fibrosis stage and non-invasive markers. These correlations are weaker than previously described, which may be explained by our large sample size, differences in fibrosis staging systems used between studies, or differences in patient cohorts included, with some prior reports including patients with PBC-AIH overlap syndrome, prolonged UDCA use (6.9 years), and those receiving second-line adjunctive therapies in addition to UDCA.[29] These data are in contrast to stronger correlations reported between APRI and FIB-4 and histologic fibrosis stage in other aetiologies of liver disease such as viral hepatitis; a recent systematic review reported an AUROC to detect cirrhosis of 0.84 for APRI and 0.87 for FIB-4 in patients with chronic hepatitis C.[30] We established disease-specific thresholds for APRI and FIB-4 for excluding and including cirrhosis. It is not clear how much utility these will prove to have; only a small percentage (7-10%) of the total population met these criteria, emphasizing the suboptimal correlation with histologic fibrosis stage.

Several potential explanations for discordance between liver biopsies and biochemical non-invasive measures of fibrosis have been reported in the literature and include factors such as variability in necro-inflammatory burden across different aetiologies of liver disease, for example in HDV/HBV viral co-infection as compared to monoinfection,[31] sampling error,[32] inter-

observer variability between pathologists,[33] and timing of non-invasive measure assessment relative to the time of biopsy.[34] Although sampling bias and inter-observer variability could not be completely excluded, the large sample size of our cohort can minimize its potential influence on our findings. Furthermore, staging was assessed as early and advanced to account for interpretive or inter-observer variability, and some of the potentially associated confounding factors associated with it, such as center and diagnosis year, which were adjusted for in multivariable analyses. Additionally, correlation between histologic stage and biochemical non-invasive measures of fibrosis is similar when the cohort is limited to biopsies 6 months prior to UDCA initiation.

Our study represents a large cohort of well-characterized patients with PBC, for whom histologic staging was available with contemporaneous biochemical variables. Prior studies included smaller cohorts with heterogeneous patient populations that may have introduced greater bias into evaluations. Despite the size of our study aiding investigation of this rare and slowly progressive disease, we wish to have changed aspects were we able to. For example, the indications for biopsy were not available and nor were additional histologic variables other than fibrosis stage including inflammatory burden, ductopenia and sinusoidal fibrosis, which may also be associated with prognosis in PBC. Although biochemical data at baseline were not available for all patients due to the real-life nature of the cohort, valid multiple imputation methods were implemented to minimize bias. In addition, biopsy data at baseline were only available for a subset of patients within the Global PBC Study Group database, although an effort was made to consider a potential bias by comparing the patients included in the study to those who did not have a biopsy. While those who were biopsied did tend to have higher liver tests, they did not have overlap syndrome with autoimmune hepatitis by diagnosis, nor did they meet Paris criteria for this, and their survival was comparable to those without a biopsy.

In conclusion, in evaluating the utility of baseline liver biopsy in patients with PBC, we demonstrate that an assessment of fibrosis stage, be it histologically *or* via non-invasive measures of fibrosis, adds meaningful information to individualized patient risk stratification beyond biochemical treatment response after 1 year. Whilst liver biopsy is not indicated for diagnosis, our data highlights the importance of incorporating available technologies and risk stratifiers for liver fibrosis, alongside biochemical markers of cholestatic therapy response, to best identify patients at risk of poor outcomes and in need of second-line therapies.

REFERENCES

1. 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.
2. Corpechot C, Abenavoli L, Rabahi N, et al. Biochemical response to ursodeoxycholic acid and long-term prognosis in primary biliary cirrhosis. Hepatology 2008;48:871-877.
3. 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;78:236-246.
4. 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;20:707-713.
5. 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. J Hepatol 2014;60:1249-1258.
6. Lammert C, Juran BD, Schlicht E, et al. Biochemical Response to Ursodeoxycholic Acid Predicts Survival in a North American Cohort of Primary Biliary Cirrhosis Patients. J Gastroenterol 2014;49:1414-1420.
7. 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.
8. 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.
9. Alempijevic T, Krstic M, Jesic R, et al. Biochemical markers for non-invasive assessment of disease stage in patients with primary biliary cirrhosis. World J Gastroenterol 2009;15:591-594.
10. Su WC, Chan CC, Hung HH, et al. Predictive value of aspartate aminotransferase to alanine aminotransferase ratio for hepatic fibrosis and clinical adverse outcomes in patients with primary biliary cirrhosis. J Clin Gastroenterol 2009;43:876-883.
11. Lindor KD, Bowlus CL, Boyer J, et al. Primary Biliary Cholangitis: 2018 Practice Guidance from the American Association for the Study of Liver Diseases. Hepatology 2019;69:394-419.
12. Poupon RE, Lindor KD, Parés A, et al. Combined analysis of the effect of treatment with ursodeoxycholic acid on histologic progression in primary biliary cirrhosis. J Hepatol 2003;39:12-16.
13. Corpechot C, Carrat F, Bonnand AM, et al. The effect of ursodeoxycholic acid therapy on liver fibrosis progression in primary biliary cirrhosis. Hepatology 2000;32:1196-1199.
14. 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.
15. Ludwig J, Dickson ER, McDonald GSA. Staging of chronic nonsuppurative destructive cholangitis (Syndrome of primary biliary cirrhosis). Virchows Arch A Pathol Anat Histol 1978;379:103-112.
16. Scheuer P. Primary biliary cirrhosis. Proc R Soc Med 1967;60:1257-1260.

17. European Association for Study of Liver, Asociacion Latinoamericana para el Estudio del Higado. EASL-ALEH Clinical Practice Guidelines: Non-invasive tests for evaluation of liver disease severity and prognosis. J Hepatol 2015;63:237-264.

18. Nevens F, Andreone P, Mazzella G, et al. A placebo-controlled trial of obeticholic acid in primary biliary cholangitis. N Engl J Med 2016;375:631-643.

19. Corpechot C, Chazouillères O, Poupon R. Early primary biliary cirrhosis: Biochemical response to treatment and prediction of long-term outcome. J Hepatol 2011;55:1361-1367.

20. Corpechot C, Chazouillères O, Rousseau A, et al. A placebo-controlled trial of bezafibrate in primary biliary cholangitis. N Engl J Med 2018;378:2171-2181.

21. 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.

22. Rubin DB. Multiple imputation after 18+ years. J Am Statitical Assoc 1996;91:473-489.

23. Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York: Wiley, 1987.

24. Gerussi A, Carbone M, Invernizzi P. Editorial: liver transplantation for primary biliary cholangitis-the need for timely and more effective treatments. Aliment Pharmacol Ther 2019;49:472-473.

25. Mayo MJ, Parkes J, Adams-Huet B, et al. Prediction of Clinical Outcomes in Primary Biliary Cirrhosis by Serum Enhanced Liver Fibrosis (ELF) Assay. Hepatology 2008;48:1549-1557.

26. Poupon R. Non-Invasive Assessment of Liver Fibrosis Progression and Prognosis in Primary Biliary Cholangitis. Dig Dis 2015;33:115-117.

27. Corpechot C, Chazouillères O, Lemoinne S, et al. Letter: reduction in projected mortality or need for liver transplantation associated with bezafibrate add-on in primary biliary cholangitis with incomplete UDCA response. Aliment Pharmacol Ther 2019;49:236-238.

28. Floreani A, Cazzagon N, Martines D, et al. Performance and utility of transient elastography and noninvasive markers of liver fibrosis in primary biliary cirrhosis. Dig Liver Dis 2011;43:887-892.

29. 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;56:198-208.

30. Chou R, Wasson N. Blood tests to diagnose fibrosis or cirrhosis in patients with chronic hepatitis C virus infection: A systematic review. Ann Intern Med 2013;158:807-820.

31. Takyar V, Surana P, Kleiner DE, et al. Non-invasive Markers for Staging Fibrosis in Chronic Delta Hepatitis. Aliment Pharmacol Ther 2017;45:127-138.

32. Shin WG, Park SH, Jang MK, et al. Aspartate aminotransferase to platelet ratio index (APRI) can predict liver fibrosis in chronic hepatitis B. Dig Liver Dis 2008;40:267-274.

33. Adams LA, George J, Bugianesi E, et al. Complex non-invasive fibrosis models are more accurate than simple models in non-alcoholic fatty liver disease. J Gastroenterol Hepatol 2011;26:1536-1543.

34. Shah AG, Lydecker A, Murray K, et al. Comparison of noninvasive markers of fibrosis in patients with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol

2009;7:1104-1112.

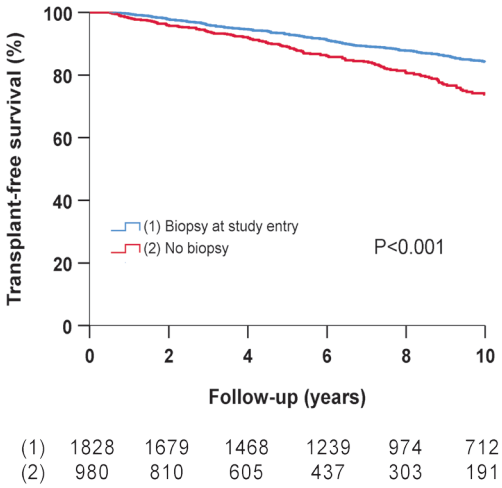
SUPPLEMENTARY DATA

Supplementary Table 1. Baseline characteristics of patients with a biopsy that were included in the study and those excluded due to a lack of biopsy

	No biopsy N=980	Biopsy at entry N=1828	P value
Age at study entry, mean (SD)	59.1 (13.0)	52.4 (11.3)	<0.001
Female, n (%)	881 (89.9)	1649 (90.2)	0.79
UDCA, n (%)	765/926 (82.6)	1828/1828 (100.0)	<0.001
AMA-positive, n (%)	905/966 (93.7)	1586/1761 (90.1)	0.001
Year of diagnosis, median (P25-75)	2003 (1997-2007)	1998 (1993-2004)	<0.001
Biochemical disease stage†			0.43
Mild	458 (72.1)	728 (71.8)	
Moderately advanced	136 (21.4)	234 (23.1)	
Advanced	41 (6.5)	52 (5.1)	
Laboratory parameters ‡, median (P25-75)			
Total bilirubin, ×ULN	0.40 (0.55-0.86)	0.59 (0.43-0.91)	0.01
ALP, ×ULN	1.67 (1.08-2.71)	2.12 (1.34-3.83)	<0.001
Albumin, ×LLN	1.17 (1.06-1.26)	1.16 (1.06-1.26)	0.23
AST, ×ULN	1.23 (0.87-1.85)	1.52 (1.06-2.32)	<0.001
ALT, ×ULN	1.28 (0.77-1.92)	1.72 (1.09-2.70)	<0.001
Platelet count, $\times 10^9$/L	240 (184-296)	253 (204-300)	0.01
Biochemical response at 1 year			
Toronto	490/723 (67.8)	1165/1781 (65.4)	0.22
Paris-II	387/723 (53.5)	953/1781 (53.5)	0.71

SD, standard deviation; UDCA, ursodeoxycholic acid; AMA, antimitochondrial antibody; ULN, upper limit of normal; ALP, alkaline phosphatase; LLN, lower limit of normal, AST, aspartate aminotransferase; ALT, alanine aminotransferase.

†Biochemical disease stage according to Rotterdam criteria.[14]
‡Laboratory parameters were not available for all patients in the no biopsy group: Total bilirubin (n=745), ALP (n=730), albumin (n=645), AST (n=701), ALT (n=613), platelet count (n=518).



Supplementary Figure 1. 10-year transplant-free survival estimates of patients with a biopsy at baseline and those without a biopsy.

Supplementary Table 2. Multivariable Cox regression models that depict the independent association of baseline histologic fibrosis stage and non-invasive markers of fibrosis with transplant-free survival

	Multivariable†		Multivariable (histologic fibrosis stage) ‡		
	HR (95% CI)	**P value**		**HR (95% CI)**	**P value**
Stage 3/4	2.85 (2.22-3.67)	<0.001	-	-	-
AAR>1	2.14 (1.65-2.76)	<0.001	AAR>1	1.94 (1.48-2.53)	<0.001
			Stage 3/4	2.66 (2.07-3.42)	<0.001
AAR>0.9§	1.86 (1.43-2.42)	<0.001	AAR>0.9	1.70 (1.30-2.21)	<0.001
			Stage 3/4	2.70 (2.10-3.47)	<0.001
APRI>2	4.31 (3.01-6.15)	<0.001	APRI>2	3.41 (2.38-4.89)	<0.001
			Stage 3/4	2.44 (1.89-3.16)	<0.001
APRI>0.8§	2.78 (2.13-3.63)	<0.001	APRI>0.8	2.35 (1.79-3.09)	<0.001
			Stage 3/4	2.35 (1.82-3.04)	<0.001
FIB-4>3.25	3.99 (2.87-5.55)	<0.001	FIB-4>3.25	3.16 (2.22-4.52)	<0.001
			Stage 3/4	2.27 (1.75-2.95)	<0.001
FIB-4>1.8§	2.56 (1.93-3.39)	<0.001	FIB-4>1.8	2.17(1.63-2.88)	<0.001
			Stage 3/4	2.47 (1.91-3.19)	<0.001
GLOBE score > age-specific threshold	4.60 (3.26-6.49)	<0.001	GLOBE score > threshold	3.93 (2.76-5.58)	<0.001
			Stage 3/4	2.08 (1.58-2.74)	<0.001

HR, hazard ratio; CI, confidence interval; AAR, AST to ALT ratio; APRI, AST to platelet ratio index.

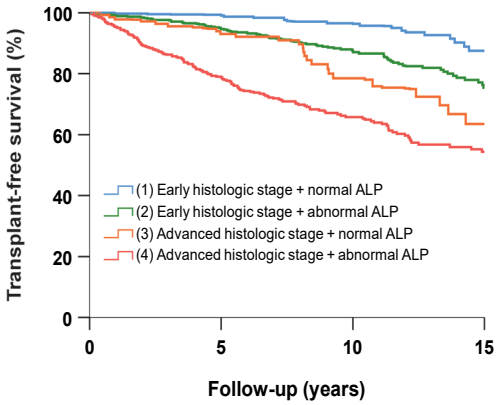
†Adjusted for age at the start of follow-up, sex, year of diagnosis, and center.
‡Adjusted for age at the start of follow-up, sex, year of diagnosis, and center. Baseline histologic fibrosis stage (3/4 vs 1/2) was input in a forward selection approach.
§Threshold derived from our cohort and are based on the highest sensitivity and specificity for the diagnosis of advanced fibrosis (stage 3/4).

Supplementary Table 3. The association of non-invasive biomarkers at baseline and response to UDCA at 1 year with transplant-free survival in patients without a biopsy

	HR (95% CI) †	P value
Toronto‡		
Non-response	1.99 (1.28-3.11)	0.002
APRI >0.8 at baseline	2.26 (1.49-3.45)	<0.001
Non-response	2.08 (1.33-3.26)	0.001
FIB-4>1.8 at baseline	2.22 (1.29-3.83)	0.005
Non-response	1.96 (1.26-3.04)	0.003
APRI (LN) at baseline	1.77 (1.39-2.25)	<0.001
Non-response	1.99 (1.27-3.24)	0.003
FIB-4 (LN) at baseline	2.37 (1.74-3.24)	<0.001
Paris-II§		
Non-response	2.28 (1.38-3.76)	0.001
APRI >0.8 at baseline	2.03 (1.32-3.14)	0.001
Non-response	2.44 (1.50-3.97)	<0.001
FIB-4>1.8 at baseline	1.06 (1.21-3.49)	0.008
Non-response	2.17 (1.32-3.58)	0.003
APRI (LN) at baseline	1.64 (1.28-2.11)	<0.001
Non-response	2.12 (1.30-3.45)	0.003
FIB-4 (LN) at baseline	2.20 (1.61-3.00)	<0.001

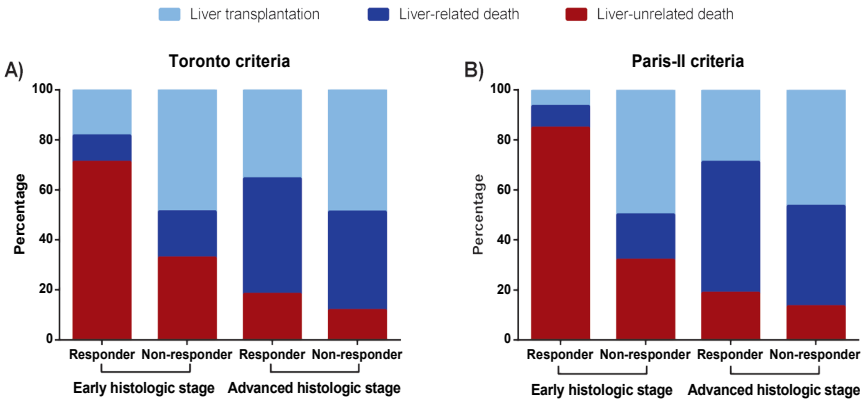
HR, hazard ratio; CI, confidence interval; APRI, AST to platelet ratio index; LN, natural logarithm.

†Multivariable model also adjusted for age the start of follow-up, sex, year of diagnosis, and center.
‡Response defined as ALP≤1.67×ULN.
§Response defined as ALP≤1.5×ULN, AST≤1.5×ULN, and normal bilirubin.



(1)	471	329	158	54
(2)	720	511	246	89
(3)	139	106	43	15
(4)	451	293	144	67

Supplementary Figure 2. Transplant-free survival estimates after 1 year of treatment with ursodeoxycholic acid stratified by baseline histologic fibrosis stage and alkaline phosphatase normalization. All pairwise comparisons were significantly different (log-rank: P<0.001), except that of early histologic stage + abnormal ALP vs. advanced histologic stage + normal ALP (P=0.12).



Supplementary Figure 3. Distribution of events associated with 10-year survival according to biochemical response and histologic fibrosis stage. Percentage of events attributed to liver transplantation, liver-related deaths, and liver-related deaths; Chi square: P<0.001 for Toronto and Paris-II criteria.

Supplementary Table 4. Multivariable Cox regression models that depict the combined association of response to treatment after 1 year and baseline fibrosis stage assessed by non-invasive markers with transplant-free survival

Group	HR (95% CI) †	P value
Toronto criteria‡		
Response + AAR ≤0.9	1.00	
Non-response + AAR ≤0.9	2.49 (1.67-3.72)	<0.001
Response + AAR >0.9	1.70 (1.15-2.52)	0.008
Non-response + AAR >0.9	4.79 (3.32-6.90)	<0.001
Response + APRI ≤0.8	1.00	
Non-response + APRI ≤0.8	2.43 (1.58-3.73)	<0.001
Response + APRI >0.8	2.49 (1.67-3.71)	<0.001
Non-response + APRI >0.8	5.34 (3.78-7.52)	<0.001
Response + FIB-4 ≤1.8	1.00	
Non-response + FIB-4 ≤1.8	2.78 (1.86-4.16)	<0.001
Response + FIB-4 >1.8	2.48 (1.61-3.82)	<0.001
Non-response + FIB-4 >1.8	5.56 (3.82-8.10)	<0.001
Paris-II criteria§		
Response + AAR ≤0.9	1.00	
Non-response + AAR ≤0.9	2.78 (1.81-4.26)	<0.001
Response + AAR >0.9	1.29 (0.77-2.14)	0.337
Non-response + AAR >0.9	5.70 (3.83-8.50)	<0.001
Response + APRI ≤0.8	1.00	
Non-response + APRI ≤0.8	2.29 (1.46-3.60)	<0.001
Response + APRI >0.8	1.33 (0.70-2.51)	0.377
Non-response + APRI >0.8	5.72 (3.96-8.27)	<0.001
Response + FIB-4 ≤1.8	1.00	
Non-response + FIB-4 ≤ 1.8	2.64 (1.70-4.09)	<0.001
Response + FIB-4 >1.8	1.51 (0.87-2.62)	0.146
Non-response + FIB-4 >1.8	6.14 (4.10-9.18)	<0.001

HR, hazard ratio; CI, confidence interval, AAR, AST to ALT ratio; APRI, AST to platelet ratio index.

†Multivariable model also adjusted for age the start of follow-up, sex, year of diagnosis, and center.
‡Response defined as ALP≤1.67×ULN.
§Response defined as ALP≤1.5×ULN, AST≤1.5×ULN, and normal bilirubin.

Supplementary Table 5. Multivariable Cox regression models that depict the independent association of baseline histologic fibrosis stage, and non-invasive fibrosis markers and alkaline phosphatase at 1 year with transplant-free survival

AAR	**HR (95% CI) †**	**P value**
AAR at 1 year (LN)	1.18 (1.06-1.31)	0.002
ALP>1.67 at 1 year	2.39 (1.83-3.14)	<0.001
Baseline histologic fibrosis stage 3/4	2.48 (1.91-3.21)	<0.001
APRI	**HR (95% CI) †**	**P value**
APRI at 1 year (LN)	2.69 (2.27-3.19)	<0.001
ALP>1.67 at 1 year	1.47 (1.10-1.97)	0.010
Baseline histologic fibrosis stage 3/4	1.73 (1.32-2.27)	<0.001
FIB-4	**HR (95% CI) †**	**P value**
FIB-4 at 1 year (LN)	3.15 (2.58-3.84)	<0.001
ALP>1.67 at 1 year	1.88 (1.42-2.50)	<0.001
Baseline histologic fibrosis stage 3/4	1.69 (1.28-2.23)	<0.001

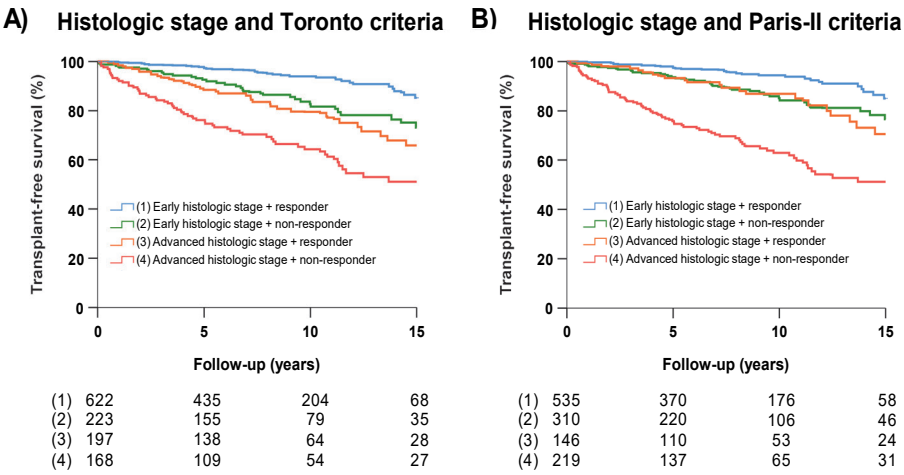
AAR, AST/ALT ratio; HR, hazard ratio; CI, confidence interval; ALP, alkaline phosphatase APRI, AST to platelet ratio index.

†Multivariable model also adjusted for age the start of follow-up, sex, year of diagnosis, and center (n=1729).

Supplementary Table 6. The area under the curve for the diagnosis of histologic fibrosis stage ≥ 3 and stage = 4 of various biochemical non-invasive markers in the whole cohort and sub-group with biopsies 6 months prior to UDCA initiation

	AUROC (95% CI)	
	Overall cohort (n=1828)	**6-month biopsies (n=1240)**
Non-invasive marker	**Stage ≥ 3**	
AAR	0.60 (0.57-0.64)	0.61 (0.57-0.71)
APRI	0.68 (0.65-0.71)	0.66 (0.62-0.70)
FIB-4	0.69 (0.67-0.72)	0.68 (0.64-0.71)
Rotterdam criteria	0.63 (0.61-0.66)	0.62 (0.58-0.66)
GLOBE score	0.70 (0.67-0.73)	0.69 (0.65-0.72)
Non-invasive marker	**Stage = 4**	
AAR	0.63 (0.59-0.67)	0.63 (0.57-0.69)
APRI	0.69 (0.65-0.72)	0.68 (0.63-0.73)
FIB-4	0.73 (0.69-0.76)	0.72 (0.67-0.76)
Rotterdam criteria	0.65 (0.61-0.68)	0.64 (0.58-0.69)
GLOBE score	0.70 (0.66-0.74)	0.68 (0.63-0.73)

AUROC, area under receiver operating curve; CI, confidence interval; AAR, AST to ALT ratio; APRI, AST to platelet ratio index.



Supplementary Figure 4. Transplant-free survival estimates after 1 year of treatment with ursodeoxycholic acid stratified by baseline histologic fibrosis stage and biochemical response according to A) Toronto criteria and B) Paris-II criteria. Patients included had biopsies up to 6 months before, but not after UDCA initiation. All pairwise comparisons were significantly different (P<0.01), except that of early histologic stage + non-responder vs. advanced histologic stage + responder (log-rank: Toronto, P=0.34; Paris-II, P=0.74).

Supplementary Table 7. The performance of high and low cut-offs of non-invasive markers of fibrosis for the exclusion and inclusion of cirrhosis

	Low cut-offs ('rule out' cirrhosis)				
	Frequency	**Sensitivity**	**NPV**	**Specificity**	**PPV**
Conventional					
APRI≤1.0	1336 (73.1%)	48.9	87.9	77.7	31.4
FIB-4≤1.45†	917 (50.1%)	75.0	91.4	55.4	26.0
Global PBC					
APRI≤0.26	172 (9.4%)	95.3	91.3	10.4	18.2
FIB-4≤0.70	186 (10.2%)	95.2	91.8	11.3	18.3
	High cut-offs ('rule in' cirrhosis)				
	Frequency	**Sensitivity**	**NPV**	**Specificity**	**PPV**
Conventional					
APRI>2.0	145 (7.9%)	21.2	85.2	94.9	46.4
FIB-4>3.25†	226 (12.3%)	33.1	86.8	92.0	46.4
Global PBC					
APRI>2.0	145 (7.9%)	21.2	85.2	94.9	46.4
FIB-4>4.03	150 (8.2%)	23.1	85.5	94.9	48.7

NPV, negative predictive value; PPV, positive predictive value; APRI, AST to platelet ratio index.

†FIB-4 threshold relates to advanced fibrosis.

CHAPTER 6

6

A comparison of prognostic scores (Mayo, UK-PBC and GLOBE) in primary biliary cholangitis

Goet JC, Murillo Perez CF, Harms MH, Floreani A, Cazzagon N, Bruns T, Prechter F, Dalekos GN, Verhelst X, Gatselis NK, Lindor KD, Lammers WJ, Gulamhusein AF, Reig A, Carbone M, Nevens F, Hirschfield GM, van der Meer AJ, van Buuren HR, Hansen BE*, Parès A*

**shared last authorship*

Submitted

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.

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] 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).

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*(((altast12*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.

Statistical analyses were performed with IBM SPS Statistics version 22.0 (IBM Corp. Released 2013, IBM Corp, Armon, NY). SAS version 9.4 (SAS Institute Inc., Cary, NC) was used to generate 10 imputed datasets of laboratory results as described in a previous study.[25–28] 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.**

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.

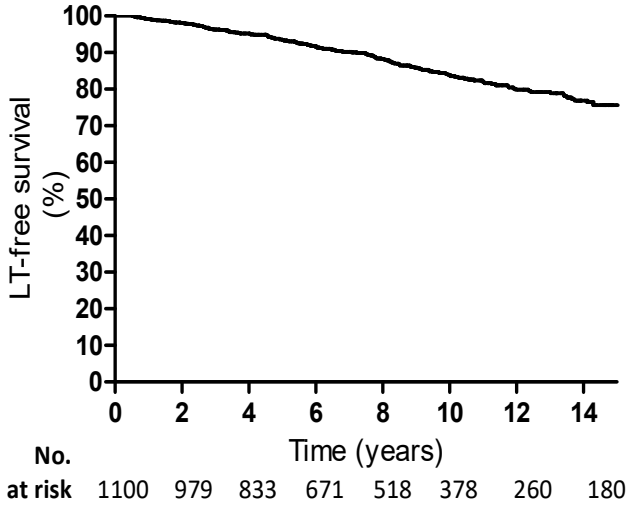


Figure 1. Kaplan-Meier estimate of transplant-free survival in this cohort

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**.

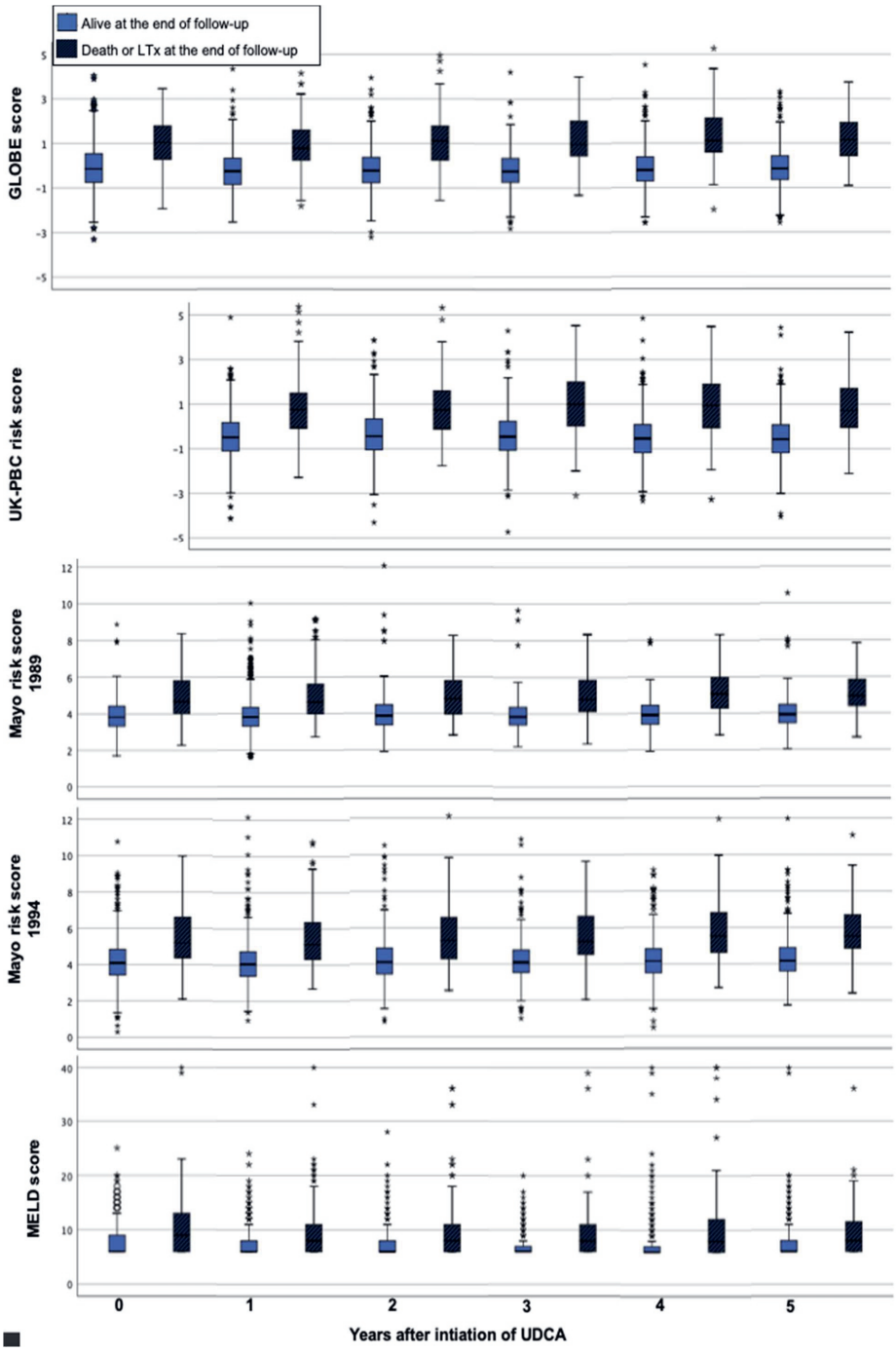


Figure 2. Boxplots of the various risk prediction scores from initiation of UDCA therapy to 5 years according to whether they experienced a clinical outcome at the end of follow-up.

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 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.

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.

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.

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.

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.

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 constrast, the addition of all scores to the MELD score yielded an increase in C-statistic, ranging from 0.07 to 0.12.

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.

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**).

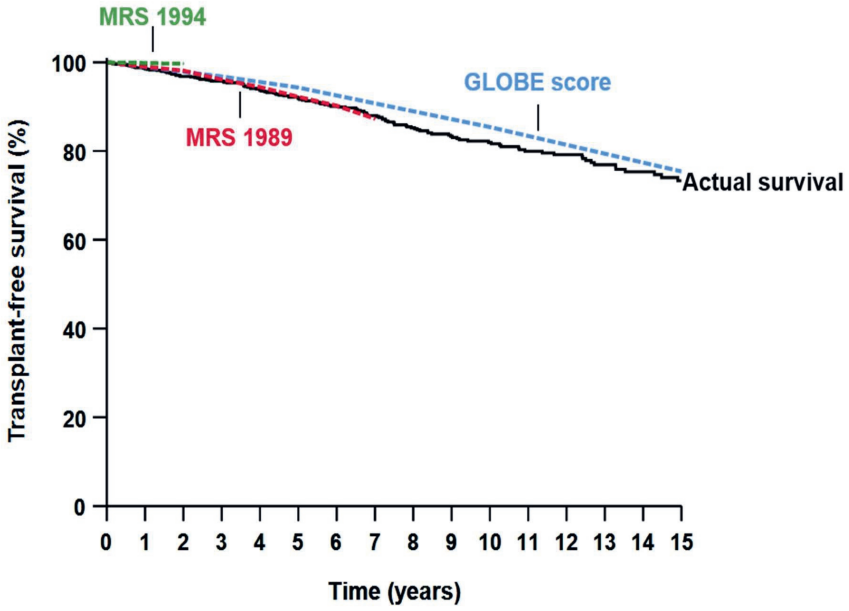


Figure 3. Predicted versus observed liver transplant-free survival for the GLOBE score and Mayo Risk Score (MRS 1989 and 1994). Figure shows prediction accuracy (calibration) of the GLOBE score and MRS up to 15 years of follow-up after 1 year of ursodeoxycholic acid (UDCA) therapy (N=905). Solid line = 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, 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. Second, due to the retrospective nature of the current study a proportion of data was missing. To overcome this problem multiple imputation techniques were used.[26] 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.

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;67:1920-1930.

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.

CHAPTER 7

7

Simplified care-pathway selection for non-specialist practice: the GLOBAL Primary Biliary Cholangitis Study Group ABA risk assessment tool

Murillo Perez CF, Gulamhusein A, Carbone M, Trivedi PJ, van der Meer AJ, Corpechot C, Battezzati PM, Lammers WJ, Cazzagon N, Floreani A, Parés A, Nevens F, Lleo A, Mayo MJ, Kowdley KV, Ponsioen CY, Dalekos GN, Gatselis NK, Thorburn D, Mason AL, Janssen HLA, Verhelst X, Bruns T, Lindor KD, Chazouillères O, Invernizzi P, Hansen BE, Hirschfield GM

European Journal of Gastroenterology & Hepatology, 2020

ABSTRACT

Background: Opportunity to redefine the care journeys for those living with primary biliary cholangitis (PBC) includes facilitating access to enhanced (PBC-dedicated) programmes by non-specialist risk 'flagging' of patients.

Objective: To develop a non-expert PBC stratification tool to help care pathway choices (standard vs. enhanced) choices in PBC.

Methods: We included ursodeoxycholic acid (UDCA)-treated patients with PBC from the Global PBC Study Group. The performance of baseline and one-year clinical markers with transplant-free survival was assessed to develop the 'ABA' tool using Age (A), Bilirubin (B), and Alkaline phosphatase (A). Added value of fibrosis estimation was assessed.

Results: 'ABA' classification mapped three risk groups (n=2226): low (Age>50 years, bilirubin ≤1×ULN, alkaline phosphatase [ALP] ≤3×ULN), high (Age≤50 years, bilirubin >1×ULN, ALP >3×ULN), and intermediate (other). Transplant-free survival at 10 years in the low, intermediate, and high-risk groups were 89%, 77%, and 59% at baseline and 86%, 76%, and 40% at 1 year, respectively. We propose that high-risk patients at baseline be directly triaged to enhanced (PBC-dedicated) care and the remaining be reassessed at 1 year. Modelling showed after 1 year 46% patients were proposed to enhanced care and 54% to standard care. The 'ABA' mapped pathways facilitated identification of patients at risk based on a young age, as compared to traditional liver biochemical stratification. In patients proposed to standard care, estimated fibrosis stage had ongoing prognostic value.

Conclusion: Non-specialist use of the 'ABA' risk tool could prioritise care journey choices for patients with PBC.

INTRODUCTION

Most patients with primary biliary cholangitis (PBC) are looked after by general internists, gastroenterologists or primary care providers who individually have experience of very few numbers of patients. As such, clinicians may not be exposed to the entire range of disease heterogeneity, and complexity particularly arising with new therapy and growing choices for disease modification and symptom control. Equally, multiple efforts to develop increasingly sophisticated risk stratification tools are confusing for the generalists faced with multiple competing algorithms.

Care pathways are used in many chronic diseases but have not yet been widely adopted for patients with PBC, with an opportunity to highlight early in the course of disease, those patients who would benefit from early 'PBC-specialist/PBC-dedicated' review. Integrated care pathways yield opportunity to harness different models of health care delivery, for instance through digital health technologies.

Patient age at presentation is an omission to current dichotomous approaches to risk stratification in PBC; this is a factor globally identified as relevant to patient outcome.[1–7] In a Japanese study, an age at onset of <55 years confers a mortality risk >7-fold to that of an age- and sex-matched control population, which was not significantly different in older age presentations.[8] These findings echoed the UK-PBC study group results that demonstrated how age at diagnosis was inversely correlated with likelihood of response to UDCA.[9] Our own dataset suggests that older age at time of UDCA initiation had increased odds of a biochemical response to UDCA therapy, based on the GLOBE score, compared to younger patients.[10] Reciprocally, the risk of liver transplantation (LT) or death (compared to a general matched population) decreased significantly with advancing age.

In the present analysis, we sought to provide readily identifiable and easy to remember triggers that could be applied at diagnosis, and after 1 year of UDCA therapy. Our goal was to facilitate patient triage into proposed standard (i.e. primary and secondary providers) vs enhanced care pathways (i.e. care programmes with more focused PBC-experienced clinical teams). We demonstrate that i) a very simple risk algorithm that captures age, and easily recallable dichotomous classification of serum liver tests, readily highlights patients for enhanced care; ii) such an approach ensures age at presentation is not omitted in determining a care journey for patients; and iii) there remains independent information that baseline evaluation of liver fibrosis has to offer, highlighting the importance of improving the use of tools such as elastography and serum fibrosis markers in liver disease management generally.

METHODS

Population and study design

We utilized the GLOBAL PBC Study Group database, an international cohort of patients with PBC from Europe and North America, to model a simplified baseline and 1-year screening tool in patients with PBC. We included UDCA-treated adult patients diagnosed with PBC according to established criteria and with available bilirubin and ALP values at baseline and 1 year. Patients with follow-up less than 1 year, UDCA discontinuation, unknown dates of clinical events, autoimmune hepatitis (Paris criteria[11]) or another concomitant liver disease were excluded. We elected to use three components with simple, recallable whole integer thresholds, of which two are biochemical parameters and the third being age at presentation, due to the inherent risk this associates with in PBC.[9,10]

Our study was conducted in accordance with the 1975 Declaration of Helsinki. The protocol was approved by the institutional research board at all participating centers as per local regulations.

Statistical Analyses

Data collection has been described in previous publications.[12] The primary endpoint of the study was a composite of LT and all-cause mortality. Patients without an event at the end of follow-up or who were lost to follow-up were censored at their last visit. The predictive performance of liver biochemistry (total bilirubin, ALP, albumin, AST, ALT, and platelet count) for transplant-free survival in the Global PBC cohort at baseline were determined as continuous variables, followed by an effort to establish the most predictive threshold of each biochemical parameter. Meanwhile, the normal threshold was employed for bilirubin and albumin. The discriminatory ability of biochemical parameters and their thresholds was assessed by multivariable Cox regression analyses (hazard ratio [HR] and 95% confidence interval [CI]) and concordance statistics (Harrell's C-statistics).

We assessed multiple models that included age and all possible combinations of biochemical parameters as dichotomous variables at baseline and 1 year. The selected threshold for age was 50 being that patients who are younger than 50 at presentation have the lowest response rates to UDCA.[9,10] Meanwhile, the thresholds for biochemical variables were those selected by C-statistics, and with reference to existing thresholds established. Given elevated biochemical parameters are associated with increased risk, one point was given for each of the following criteria met at time of assessment: age ≤50 years, biochemical parameter 1 >threshold, biochemical parameter 2> threshold. Patients were subsequently categorized into three risk

groups: low (score=0), intermediate (score=1 or 2), and high (score=3). The discriminatory ability of each combination was assessed with C-statistics for the selection of the model.

The transplant-free survival associated with each risk group at baseline and 1 year was estimated with a Kaplan-Meier curve and compared with the log-rank test. Changes in risk group from baseline to 1 year were evaluated and a complementary care pathway was proposed. The proportion of patients that would be allocated to standard and enhanced care at baseline, 1 year, and yearly thereafter up to 5 years of follow-up were described. Risk assessments from years 2-5 were carried out with the imputed database.

To highlight the added value of incorporating age, our risk stratification tool was compared to two of the most applied criteria: the 'Toronto' criteria (ALP>1.67×ULN or abnormal bilirubin) and the established Paris-II criteria (ALP>1.5×ULN or AST>1.5×ULN or abnormal bilirubin). Further, to consider the role of fibrosis stage, we selected patients proposed to standard care after 1 year and evaluated their transplant-free survival according to baseline histologic stage and the presence of cirrhosis, defined by baseline histologic stage IV or non-invasive markers of cirrhosis measured at 1 year (APRI>2 or FIB-4>4.03) using the imputed database.[12]

Multiple imputation was conducted with by the Markov Chain Monte Carlo method for missing data with SAS version 9.4 (SAS Institute Inc., Cary, NC, USA). Ten imputed data sets for missing biochemical values were generated to reduce sampling variability of laboratory results.[13] Imputation was performed based on the assumption that data were missing at random, in which variables predicting outcomes and outcomes themselves were included in the model. Rubin's rules were used to estimate the parameter and standard error.[14,15]

A P-value less than 0.05 was considered statistically significant. Multivariable Cox regression analyses were adjusted for sex, year of diagnosis, and center. All analyses were two-sided and performed using IBM SPSS Statistics for Windows, version 25.0 (IBM Corp., Armonk, NY, USA) and SAS version 9.4 (SAS Institute Inc., Cary, NC, USA).

RESULTS

Development of the simplified ABA risk assessment tool

The model was developed utilizing all data from the Global PBC Study Group. The patient population in this study was representative of a typical PBC population (**Table 1**).

At presentation, all evaluated biochemical variables were significant predictors of transplant-free survival, of which total bilirubin demonstrated the highest discriminatory ability, while ALT demonstrated the lowest (**Table S1**). The thresholds with the highest predictive ability for ALP, transaminases (AST and ALT), and platelet count were 3×ULN, 2×ULN, and 150×10^9/L,

respectively, although there were no significant differences compared to other thresholds (**Table S2**).

Implementation of these thresholds at baseline for categorization into three risk groups according to various combinations of biochemical factors, along with age, resulted in similar proportions allocated to each group and respective transplant-free survival estimates at 10 years (**Table S3**). There were also no major differences between the models when the risk algorithms were applied after 1 year of UDCA (**Table S4**). Considering the equivalency in discriminatory ability of trialed models and to provide the most uniform and simple utility to a non-specialist audience, the model of choice was that which included age, bilirubin, and ALP.

Table 1. Baseline characteristics

Baseline characteristics	Total (n=2226)
Age at the start of UDCA, mean (SD)	54.9 (11.8)
Female, n (%)	2030 (91.2)
AMA-positive, n (%)	1987/2174 (91.4)
Year of diagnosis, median (IQR)	1998 (1990-2004)
Biochemical disease stage, n (%)	n=1817
Early	1292 (71.1)
Moderately advanced	423 (23.3)
Advanced	102 (5.6)
Laboratory parameters, median (IQR)	
Total bilirubin, ×ULN (n=2226)	0.62 (0.44-0.95)
ALP, ×ULN (n=2226)	2.20 (1.34-3.87)
Albumin, ×LLN (n=1817)	1.16 (1.06-1.26)
AST, ×ULN (n=2166)	1.53 (1.00-2.30)
ALT, ×ULN (n=2106)	1.70 (1.05-2.64)
Platelet count, $\times 10^9$/L (n=1700)	251 (199-303)

UDCA, ursodeoxycholic acid; anti-mitochondrial antibody; ALP, alkaline phosphatase; AST, aspartate aminotransferase; ALT, alanine aminotransferase.

Application of the simplified ABA risk assessment tool at baseline and 1 year stratifies patients based on outcome

Patients were given a score according to the three-tiered algorithm of Age, Bilirubin, and ALP (ABA), in which a point was obtained for each of the following criteria: age ≤50 years, bilirubin >1×ULN, ALP >3×ULN. Three risk groups were defined: low ABA risk (score=0), intermediate (score=1 or 2), and high ABA risk (score=3). Transplant-free survival at 10 years in the low, intermediate, and high-risk groups were 89%, 77%, and 59% at baseline and 86%, 76%, and

40% at 1 year, respectively (**Figure 1**). With a threshold of 2×ULN for ALP, the transplant-free survival of each risk group was similar to that with a threshold of 3×ULN at baseline and 1 year (**Table S5**). In multivariable analyses, the intermediate and high risk groups were significantly associated with an increased risk for LT or death at baseline and 1 year compared to the low-risk group (Baseline: intermediate risk [HR 1.61, 95% CI 1.28-2.03, P<0.001; high risk [HR 3.11, 95% CI 2.26-4.30]; 1 year: intermediate risk [HR 1.63, 95% CI 1.33-2.00, P<0.001; high risk [HR 4.15, 95% CI 2.87-6.01]).

In the low-risk group at 1 year, there were 462 (38.0%) patients with ALP≤1×ULN, 413 (34.0%) from 1.0-1.5×ULN, 189 (15.5%) from 1.5-2.0×ULN, and 152 (12.5%) from 2.0-3.0×ULN. Low risk ABA is distinct from response to UDCA, such that 387/1141 (33.9%) of these patients are incomplete responders according to Paris-II, 261/1216 patients (21.5%) have an ALP >1.67×ULN, and 303/1090 patients (27.8%) are 'complete' responders with normal ALP, AST, and ALT.

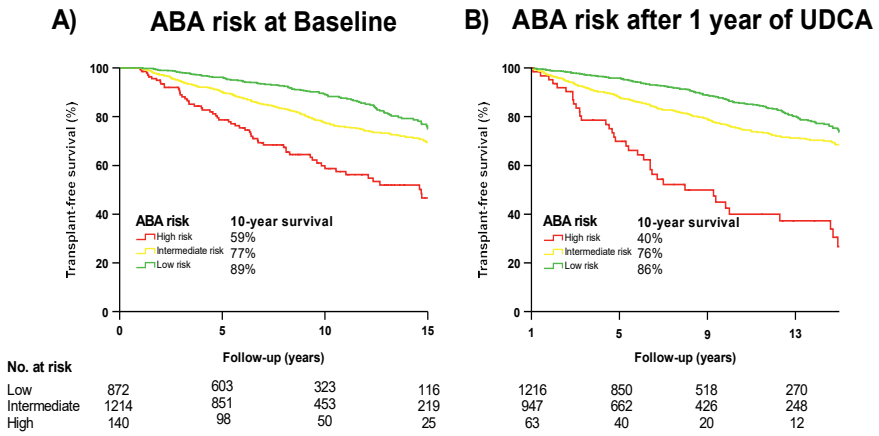


Figure 1. Transplant-free survival estimates at baseline and after 1 year of ursodeoxycholic acid according to the Age, Bilirubin, Alkaline phosphatase (ABA) risk tool. Allocation into the three risk groups was as follows: low ABA risk (age >50 years, bilirubin ≤1×ULN, ALP ≤3×ULN), high ABA risk (age ≤50 years, bilirubin >1×ULN, ALP >3×ULN), intermediate ABA risk (all other combinations).

Utility of the simplified ABA risk assessment tool in guiding care pathways for patients with PBC

There were 140 patients identified as high-risk at baseline, of whom 86 had a change in risk at 1 year, the majority to an intermediate risk (**Figure 2**). Of 1214 patients with intermediate

risk at baseline, 372 (31%) patients transitioned predominantly to the low-risk group. Meanwhile, most patients with low risk at baseline remained so at 1 year, as only 25 patients' risk changed to intermediate. We propose that patients with a high risk at baseline be directly introduced to enhanced care (i.e. care programmes with focused PBC-experienced clinical teams) while those with an intermediate risk be reassessed at 1 year (**Figure 3**). Furthermore, patients of an intermediate and high risk at 1 year or thereafter should be proposed to enhanced care.

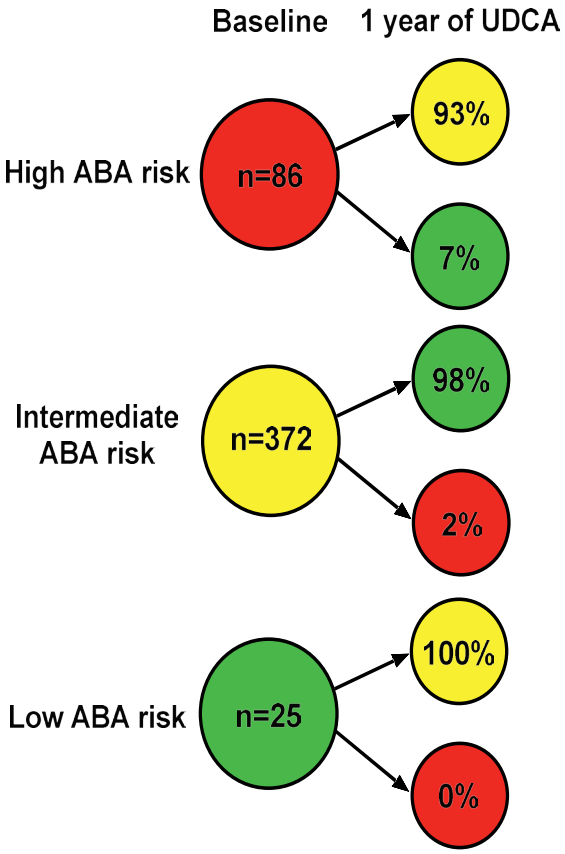


Figure 2. Age, Bilirubin, Alkaline phosphatase **(**ABA) risk group changes from baseline to 1 year of ursodeoxycholic acid treatment.

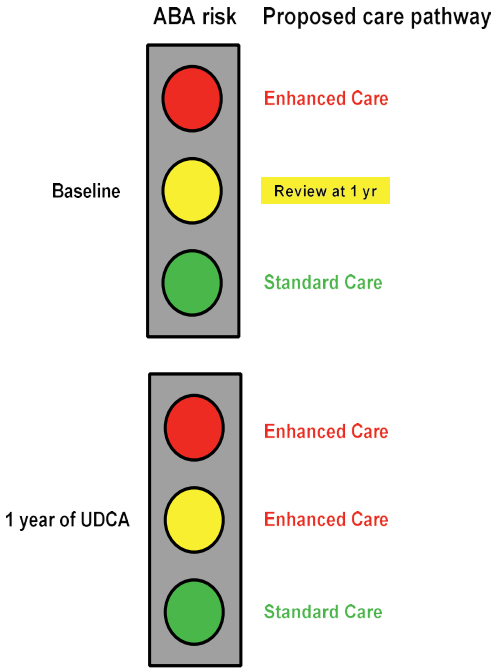


Figure 3. Proposed care pathway for standard and enhanced care based on Age, Bilirubin, Alkaline phosphatase (ABA) risk stratification at baseline and 1 year. Red signifies high risk, amber signifies intermediate risk, and green signifies low risk.

The simplified ABA risk assessment tool identifies nearly half of patients to enhanced care after 1 year

Based on the identification of 140 (6%) high-risk patients at baseline, these were directly proposed to enhanced care. Reassessment of risk in patients who remained in standard care (n=2086) led to 876 (42%) patients proposed to enhanced care at 1 year (**Figure 4**). Overall, 1016 (46%) patients were proposed to enhanced care and 1210 (54%) to standard care after 1 year. The proportion of patients proposed to enhanced care with the ABA risk assessment tool were compared to the proportion of incomplete responders to UDCA at 1 year according to various published criteria (**Table S6**), which ranged from 19-51%. Continual application of the risk algorithm to guide care pathways at yearly intervals after 1 year led to enhanced care proposal in 6.5%, 5.4%, 4.8%, and 4.7% at each consecutive year.

In order to demonstrate the application of this tool, the risk assessment was carried out with the ABA tool and subsequent care pathway determined for a specific patient. For example, a patient who at baseline is 55 years old, with normal bilirubin, and ALP of 1.5×ULN would

receive a score of 0 and be considered low risk. Thus, this patient would initiate UDCA and be reassessed at 1 year. After 1 year, with an age of 56 years, normal bilirubin, and ALP at 1.2×ULN, this patient remains in the low-risk group, thus would continue in standard care (primary care).

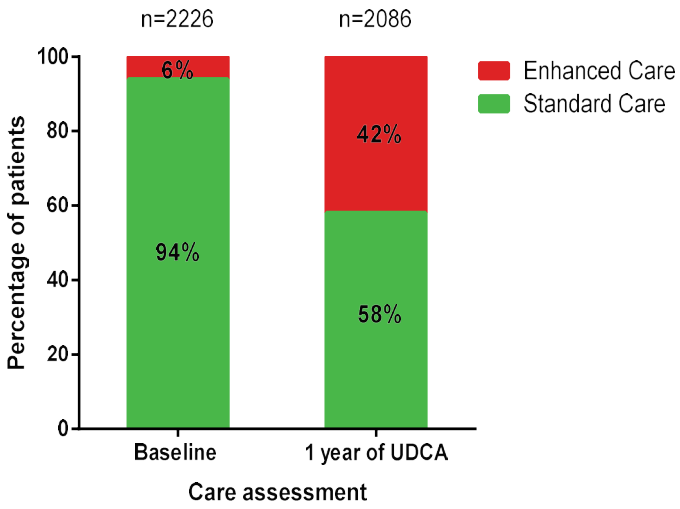


Figure 4. Proportion of patients proposed to standard and enhanced care at baseline and 1 year. At baseline, high ABA risk patients were triaged to enhanced care, while intermediate and low risk patients to standard care. After reassessment of these patients at 1 year, intermediate and high ABA risk patients were triaged to enhanced care.

The simplified ABA risk assessment tool highlights patients under 50 years of age at diagnosis for specialty care

Risk identification by our three-tiered risk algorithm was contrasted with standard criteria (ALP>1.67×ULN and/or abnormal bilirubin) at 1 year while considering age (**Table S7**). All patients in the high-risk group according to our risk assessment tool met standard criteria of risk (ALP>1.67×ULN and/or abnormal bilirubin). Meanwhile, of those classified to the intermediate-risk group and who would be proposed to enhanced care, 62% met standard criteria for risk and 48% did not. Patients not identified as being at risk by standard criteria but so with our risk assessment tool are those younger than age 50, with a median age of 43. Similarly, in patients of intermediate or high risk at 1 year and with available Paris-II response status (n=984), 287 (29%) patients were not considered at risk based on Paris-II complete response and had a median age of 43 (IQR 38-47). In contrast, the median age of those who had an incomplete response to Paris-II was higher at 49 years (IQR 43-59).

DISCUSSION

With evolving PBC care and delivery, including the opportunity to harness different models of care such as virtual health, it is opportune to consider how care pathways in PBC could be changed. This is aligned to increasing complexity in therapeutic choices for patients with PBC, alongside ongoing interest in developing new therapies, as well as organizational recognition that rare disease care can be enhanced structurally. At the same time, non-expert PBC clinicians report (through personal communications) confusion due to a myriad of biochemical response criteria and risk assessment tools. Finally, there is recognition that primary care should remain involved in care of patients, and should be in a position to understand, and advocate as needed, whether their patient with PBC has access to appropriate expertise. Thus, we modelled and report a simplified risk assessment tool, predominantly to demonstrate how care pathways in PBC could look in the future, agnostic to individual therapy choices. In so doing, we prioritized features of disease with recognized impact on risk as well as clinical utility for non-expert gastroenterologists and primary care providers, principally age, bilirubin, and alkaline phosphatase. Notedly, ALP and bilirubin are considered surrogate endpoints in PBC and have formed the cornerstone of educational programmes for clinicians and patients.[16]

Our analyses show how the simplified ABA risk assessment tool can be used at baseline and 1 year to guide care pathway choices for non-expert gastroenterologists and primary care providers without added time burden for clinicians. Its application would propose an estimated 46% of patients to enhanced care after 1 year. These risk groups should not be confused with the definitions of suboptimal response to UDCA, as patients in the low-risk group may need additional therapies after 1 year given the broad range of ALP levels in this group. The ABA tool identifies additional patients at risk based on a young age as compared to liver biochemistry alone, such as standard criteria[17] (ALP/bilirubin) or Paris-II criteria[18]. Our tool also highlights an ongoing need to assess liver fibrosis to best determine the need for enhanced care.

Age is an important risk factor in PBC, but apart from the GLOBE score, it is seldom captured in risk stratification models, nor sufficiently highlighted in clinician education. Thus, we pre-specified age as one of the components and demonstrated that its inclusion can highlight patients under the age of 50 who are not otherwise readily identified with traditional criteria that solely rely on biochemical parameters. Of those with an intermediate or high risk at 1 year based on our risk assessment tool, 29-36% do not meet criteria for being at risk, depending on the response criteria used. This suggests that some patients under the age of 50 with biochemical response may incorrectly be classified as low risk.

As non-invasive testing for liver fibrosis has evolved, there remains a need to best understand how to use such tools in PBC. In patients proposed to standard care after 1 year, there would be added prognostic value in assessing liver fibrosis stage. Less than 10% of patients had a suggestion of cirrhosis either by histologic stage or non-invasive markers that was associated with significantly diminished transplant-free survival. This is in line with our previous findings that patients with advanced fibrosis have diminished transplant-free survival, despite biochemical response.[12] In both instances, the importance of liver fibrosis evaluation when determining risk is highlighted, which should therefore be considered when determining care pathways in PBC. Although some of these patients may have decompensated liver disease and therefore readily identified, patients with compensated cirrhosis are more difficult to identify. Our goal was to determine what percentage of patients would require prompt specialty care but were not captured with the ABA risk tool. There is a need for non-invasive methods to determine advanced fibrosis and cirrhosis outside of specialist liver care settings and prior to referral, particularly to identify patients with well-compensated cirrhosis in need of enhanced care. Tools such as elastography and serum fibrosis markers need development in representative cohorts to determine best use.

There are limitations and potential biases to this study. Our tool does not provide information about the need for therapy and is focused solely on care pathways. We recognize that symptoms are important for patients, and that this type of tool, may not adequately capture this. Further, it is clear that care pathways need better definition, as health care structures seek definitions of ‘expertise’ Lastly, fibrosis stage was not included in the score as histology was not readily available in the entire population, however, there remains a need to determine best use of tools that assess liver fibrosis to also guide care pathways, for which the best candidate is transient elastography.[19]

In conclusion, the simplified GLOBAL PBC Study Group ABA risk assessment tool is intended and readily applicable to non-expert settings. It promptly flags a care pathway for a patient with PBC and the need for referral, applying simple and easily recallable components of risk, namely age, bilirubin, and alkaline phosphatase. Implementation of very simple clinical tools in PBC care has potential to enhance outcomes for patients.

REFERENCES

1. 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.
2. 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;136:1281-1287.
3. 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;105:2186-2194.
4. Corpechot C, Chazouillères O, Poupon R. Early primary biliary cirrhosis: Biochemical response to treatment and prediction of long-term outcome. J Hepatol 2011;55:1361-1367.
5. 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.
6. Corpechot C, Abenavoli L, Rabahi N, et al. Biochemical response to ursodeoxycholic acid and long-term prognosis in primary biliary cirrhosis. Hepatology 2008;48:871-877.
7. 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.
8. Kubota J, Ikeda F, Terada R, et al. Mortality rate of patients with asymptomatic primary biliary cirrhosis diagnosed at age 55 years or older is similar to that of the general population. J Gastroenterol 2009;44:1000-1006.
9. 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;144:560-569.e7; quiz e13-4.
10. 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. Clin Gastroenterol Hepatol 2019;17:2076-2084.e2.
11. Chazouillères O, Wendum D, Serfaty L, et al. Primary biliary cirrhosis-autoimmune hepatitis overlap syndrome: clinical features and response to therapy. Hepatology 1998;28:296-301.
12. 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.
13. 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.
14. Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York: Wiley, 1987.
15. Rubin DB. Multiple imputation after 18+ years. J Am Statitical Assoc 1996;91:473-489.
16. 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.

17. Nevens F, Andreone P, Mazzella G, et al. A placebo-controlled trial of obeticholic acid in primary biliary cholangitis. N Engl J Med 2016;375:631-643.

18. Corpechot C, Chazouillères O, Rousseau A, et al. A placebo-controlled trial of bezafibrate in primary biliary cholangitis. N Engl J Med 2018;378:2171-2181.

19. 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;56:198-208.

SUPPLEMENTARY DATA

Table S1. Performance of baseline liver biochemistry at the start of UDCA for liver transplantation and death

Predictive factor[a]	n	C-statistic (95% CI)
Bilirubin, ×ULN	2809	0.76 (0.74-0.78)
Albumin, ×LLN	2329	0.71 (0.69-0.74)
Platelet count, $\times 10^9$/L	2269	0.70 (0.67-0.73)
AST, ×ULN	2764	0.68 (0.65-0.70)
ALP, ×ULN	2827	0.66 (0.64-0.69)
ALT, ×ULN	2693	0.64 (0.61-0.66)

CI, confidence interval; ULN, upper limit of normal; LLN, lower limit of normal; AST, aspartate aminotransferase, ALP, alkaline phosphatase; UDCA, ursodeoxycholic acid; ALT, alanine aminotransferase.

[a]All analyses correspond to multivariable Cox regression analysis adjusted for sex, diagnosis year, and center. Variables that were not normally distributed were transformed with the natural logarithm (bilirubin, AST, ALP, ALT).

Table S2. Assessment of the performance of thresholds of various biochemical markers for liver transplantation and death

Biochemical marker	C-statistic (95% CI)[a]
ALP threshold, xULN	N=2827
1.0	0.68 (0.66-0.71)
1.5	0.69 (0.67-0.71)
1.67	0.70 (0.67-0.72)
2.0	0.70 (0.67-0.72)
3.0*	0.70 (0.68-0.72)
AST threshold, xULN	N=2764
1.0	0.70 (0.68-0.72)
2.0*	0.72 (0.70-0.74)
3.0	0.70 (0.68-0.73)
ALT threshold, xULN	N=2693
1.0	0.69 (0.67-0.71)
2.0*	0.69 (0.67-0.72)
3.0	0.69 (0.67-0.72)
Platelet count, $\times 10^9$/L	N=2269
150*	0.73 (0.70-0.76)
200	0.72 (0.69-0.75)
250	0.70 (0.68-0.73)

[a]Multivariable Cox regression analyses adjusted for age, sex, center, diagnosis year.

*Signifies the threshold with highest C-statistic.

Table S3. Risk assessment according to various combinations of parameters at baseline

Combination[a]	n	C-statistic (95% CI)[b]	Low, n	10-yr surv. (%)	Intermediate, n	10-yr surv. (%)	High, n	10-yr surv. (%)
Age, bilirubin, ALP	**2226**	**0.67 (0.63-0.70)**	**872**	**89.1**	**1214**	**77.4**	**140**	**58.9**
Age, bilirubin, albumin	1584	0.70 (0.66-0.73)	726	88.9	822	72.6	36	45.6
Age, bilirubin, platelet count	1426	0.68 (0.64-0.72)	697	90.7	715	78.6	14	43.5
Age, bilirubin, AST	2009	0.69 (0.65-0.72)	847	89.0	1022	77.8	140	64.9
Age, albumin, platelet count	1118	0.68 (0.64-0.72)	578	87.9	535	75.9	5	40.0
Age, ALP, platelet count	1402	0.67 (0.62-0.71)	590	90.2	802	80.6	10	64.0
Age, albumin, AST	1465	0.67 (0.64-0.71)	608	86.5	822	74.4	35	56.8
Age, ALP, AST	2204	0.66 (0.63-0.69)	858	86.6	1117	79.9	229	76.8
Age, platelet count, AST	1386	0.66 (0.61-0.70)	612	88.8	763	81.1	11	60.6
Age, ALP, albumin	1516	0.66 (0.63-0.69)	600	84.8	877	75.7	39	64.3

CI, confidence interval; AST, aspartate aminotransferase; ALP, alkaline phosphatase.

[a]The following thresholds were implemented for each variable: Age (50 years), bilirubin (1×ULN), albumin (1×ULN), ALP (3×ULN), platelet count (150×10^9/L), AST (2×ULN). With parameters whose elevation signifies and increased risk, allocation into each group was as follows: low (age > 50 years, biochemical parameter 1 ≤ threshold, biochemical parameter 2 ≤ threshold), high (age ≤ 50 years, biochemical parameter 1 > threshold, biochemical parameter 2 > threshold), and intermediate (other combinations).

[b]Multivariable Cox regression analyses were adjusted for sex, center, and diagnosis year.

Table S4. Risk assessment according to various combinations of parameters at 1 year

Combination[a]	n	C-statistic (95% CI)[b]	Low, n	10-yr surv. (%)	Intermediate, n	10-yr surv. (%)	High, n	10-yr surv. (%)
Age, bilirubin, ALP	**2226**	**0.67 (0.64-0.70)**	**1216**	**86.4**	**947**	**76.2**	**63**	**39.9**
Age, bilirubin, albumin	1584	0.70 (0.67-0.73)	833	88.0	733	71.4	18	9.3
Age, bilirubin, platelet count	1426	0.69 (0.65-0.73)	755	91.4	659	76.8	12	42.3
Age, bilirubin, AST	2009	0.69 (0.66-0.72)	1124	86.9	832	77.2	53	42.9
Age, albumin, platelet count	1118	0.69 (0.65-0.73)	615	87.1	497	75.9	5	50.0
Age, ALP, platelet count	1402	0.67 (0.63-0.71)	751	89.0	642	79.5	9	62.2
Age, albumin, AST	1465	0.68 (0.65-0.71)	807	84.5	646	72.4	12	38.2
Age, ALP, AST	2204	0.66 (0.63-0.69)	1271	84.3	849	80.5	84	65.9
Age, platelet count, AST	1386	0.67 (0.63-0.72)	769	88.1	608	80.1	9	51.9
Age, ALP, albumin	1516	0.67 (0.64-0.70)	821	83.9	676	73.2	19	64.8

CI, confidence interval; AST, aspartate aminotransferase; ALP, alkaline phosphatase.

[a]The following thresholds were implemented for each variable: Age (50 years), bilirubin (1×ULN), albumin (1×ULN), ALP (3×ULN), platelet count (150×10^9/L), AST (2×ULN). With parameters whose elevation signifies and increased risk, allocation into each group was as follows: low (age > 50 years, biochemical parameter 1 ≤ threshold, biochemical parameter 2 ≤ threshold), high (age ≤ 50 years, biochemical parameter 1 > threshold, biochemical parameter 2 > threshold), and intermediate (other combinations).

[b]Multivariable Cox regression analyses were adjusted for sex, center, and diagnosis year.

Table S5. Risk assessment according to Age (50 years), Bilirubin (1×ULN) and ALP (2×ULN) at baseline and 1 year

Baseline	**Low N=622**	**Intermediate N=1436**	**High N=168**
10-year survival	90.3	79.1	60.9
1 year	**Low N=1064**	**Intermediate N=1078**	**High N=84**
10-year survival	87.5	77.0	43.4

Table S6. Proportion of patients that are incomplete responders after 1 year of UDCA according to various published criteria

Response criteria	**Incomplete responders n (%)**
Barcelona	886/2226 (39.8)
Rotterdam	306/1646 (18.6)
Paris-I	588/2079 (28.3)
Paris-II	1084/2125 (51.0)
Toronto[a]	805/2226 (36.2)

[a]Assessed at 1 year

Table S7. Age of patients according to Age, Bilirubin, Alkaline phosphatase (ABA) risk assessment tool and standard criteria at 1 year of ursodeoxycholic acid.

Risk stratification at 1 year	**n**	**Overall age[a]**	**Criteria met, n (%)[b]**	**Age in those meeting criteria[a]**	**Age in those who do not meet criteria[a]**
Low ABA risk	1216	61 (56-68)	261 (21.5)	60 (55-67)	61 (56-69)
Intermediate ABA risk	947	47 (41-55)	588 (62.1)	51 (45-60)	43 (38-47)
High ABA risk	63	42 (38-47)	63 (100.0)	42 (38-47)	-

[a]Data presented as median (IQR).
[b]ALP>1.67×ULN and/or bilirubin>1×ULN at 1 year.

Fibrosis stage needs to be considered ultimately to optimize care pathway choices

The added role of baseline fibrosis assessment for risk determination and guiding care pathways was assessed in the sub-group of patients proposed to standard care after 1 year (n=1210), given that the outcome of patients with PBC is a balance determined by cholestatic injury and liver fibrosis (existing at initial presentation and subsequently developing). First, we evaluated histologic stage in patients with an available baseline biopsy (n=547), of whom 46 patients had histologic stage IV that was associated with a significantly lower transplant-free survival compared to stages I-III (P≤0.01) (**Figure S1**). Since histology was not available in all patients, non-invasive markers for cirrhosis (FIB-4>4.03 or APRI>2) were also utilized to define cirrhosis, which resulted in 119 (9.8%) patients meeting criteria for cirrhosis either by histologic stage or non-invasive markers. The transplant-free survival at 10 years of these patients was 52%, in contrast to 91% in those who did not meet either criteria (P<0.01).

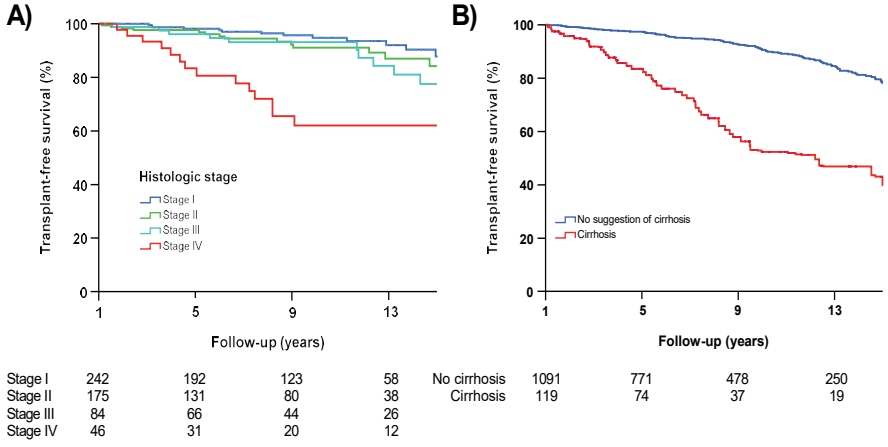
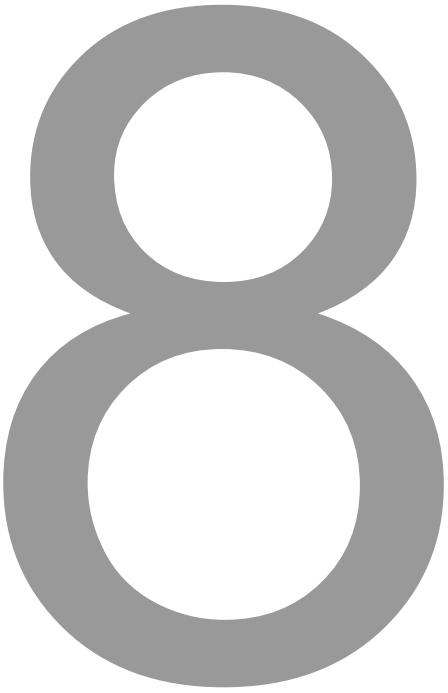


Figure S1. Transplant-free survival in patients proposed to standard care after 1 year according to A) histologic stage in patients with a baseline biopsy (n=547) and B) cirrhosis defined by histologic stage IV or surrogates of cirrhosis (APRI>2 or FIB-4>4.03).

CHAPTER 8



Goals of treatment for improved survival in primary biliary cholangitis: Treatment targets should be bilirubin within the normal range and alkaline phosphatase normalization

Murillo Perez CF, Harms MH, Lindor KD, van Buuren HR, Hirschfield GM, Corpechot C, van der Meer AJ, Feld JJ, Gulamhusein A, Lammers WJ, Ponsioen CY, Carbone M, Mason AL, Mayo MJ, Invernizzi P, Battezzati PM, Floreani A, Lleo A, Nevens F, Kowdley KV, Bruns T, Dalekos GN, Gatselis NK, Thorburn D, Trivedi PJ, Verhelst X, Parés A, Janssen HLA, Hansen BE

American Journal of Gastroenterology, 2020

ABSTRACT

Background: In primary biliary cholangitis (PBC), bilirubin and alkaline phosphatase (ALP) are widely established as independent predictors of prognosis. Current treatment goals do not aim for normalization of surrogate markers because their association with survival has not been defined.

Methods: The patient cohort from the GLOBAL PBC Study Group was used, comprising of long-term follow-up data from European and North American centers. Ursodeoxycholic acid-treated and untreated patients with bilirubin levels ≤1× upper limit of normal (ULN) at baseline or 1 year were included. The association of normal ALP with transplant-free survival was assessed in a subgroup with ALP≤1.67×ULN at 1 year. Optimal thresholds of bilirubin and ALP to predict liver transplantation (LT) or death were evaluated.

Results: There were 2,281 patients included in the time zero cohort and 2,555 patients in the 1-year cohort. The bilirubin threshold with the highest ability to predict LT or death at 1 year was 0.6×ULN (hazard ratio 2.12, 95% CI 1.69-2.66, *P*<.001). The 10-year survival rates of patients with bilirubin ≤0.6×ULN and >0.6×ULN were 91.3% and 79.2%, respectively (*P*<.001). The risk for LT or death was stable below the bilirubin levels of 0.6×ULN, yet increased beyond this threshold. Ursodeoxycholic acid-induced reduction in bilirubin below this threshold was associated with an 11% improvement in 10-year survival. Furthermore, ALP normalization was optimal, with 10-year survival rates of 93.2% in patients with ALP≤1×ULN and 86.1% in those with ALP 1.0-1.67×ULN.

Conclusion: Attaining bilirubin levels ≤0.6×ULN or normal ALP are associated with the lowest risk for LT or death in patients with PBC. This has important implications for treatment targets.

INTRODUCTION

Primary biliary cholangitis (PBC) is an autoimmune cholestatic liver disease that is characterized by chronic nonsuppurative inflammation of the small intrahepatic bile ducts.[1] The disease usually has a slow progressive course, which may eventually lead to cirrhosis and ultimately liver failure or premature death in the absence of liver transplantation (LT). However, the prolonged number of years it may take for patients to develop such clinical outcomes poses a significant obstacle in randomized controlled trials that aim to evaluate the clinical benefit of therapeutic interventions. Owing to these feasibility concerns, various surrogate markers have been evaluated for their prognostic value on clinical outcomes.[2] Such surrogate markers can allow risk stratification of patients without the need for extended follow-up and can be implemented by health care providers or in clinical trials to promptly assess the need and benefit of a therapeutic agent.

It is established that bilirubin is an independent predictor of prognosis in both ursodeoxycholic acid (UDCA)-treated and untreated patients with PBC.[2–4] The normalization of bilirubin prompted by UDCA has been associated with improved transplant-free survival.[4] Furthermore, bilirubin and alkaline phosphatase (ALP) have been established as surrogate endpoints that are "reasonably likely to predict clinical benefit," and the widely accepted thresholds are 1× the upper limit of normal (ULN) and 1.67×ULN, respectively. Normal bilirubin is also a component of multiple response criteria, such as the Rotterdam, Paris-I, and Paris-II criteria.[5–7] Abnormal bilirubin levels are observed during later stages of PBC and are indicative of impaired liver function.[8] Over the past decades, however, there has been an increase in the proportion of patients who present with normal bilirubin levels, and this group now represents most patients with PBC.[9] Because bilirubin is usually not elevated above the ULN until later stages of the disease, it is considered to be an inadequate marker for risk stratification in early stage PBC. The prognostic value of bilirubin and ALP below the ULN has not been previously assessed. We sought to evaluate whether bilirubin or ALP levels within the normal range (≤1×ULN) are associated with survival in patients with PBC to optimize treatment goals and the number of patients who may benefit from second-line therapy.

PATIENTS AND METHODS

Population and study design

We used the GLOBAL PBC Study Group database to study the predictive value of normal bilirubin for survival in PBC. The Global PBC Study Group is an international collaboration of 17 centers across Europe and North America that includes long-term follow-up data of patients with PBC. During the time frame analyzed, no patients were on obeticholic acid. We included UDCA-treated and untreated patients diagnosed with PBC according to the internationally accepted guidelines and whose bilirubin levels were normal (≤1×ULN as defined by each local center) at time zero or 1 year. [8,10,11] Those with short follow-up (<6 months for time zero cohort; <1 year for 1-year cohort), UDCA discontinuation, unknown clinical event dates, autoimmune hepatitis overlap, or other concomitant liver diseases were excluded from the study. Patients were allocated to 2 independent cohorts based on the time point(s) at which their bilirubin levels were normal (time zero and 1 year) in which inclusion is not mutually exclusive. This study was conducted in accordance with the 1975 Declaration of Helsinki. The protocol was approved by the institutional research board at all participating centers as per local regulations.

Data collection

Time zero (study entry) is defined as the date UDCA was initiated in treated patients and the date of the first visit in untreated patients. At study entry, the following data were available: sex, age at diagnosis, anti-mitochondrial antibody serological status, liver histology, biochemical disease stage (Rotterdam criteria[5]), and UDCA therapy. Three biochemical disease stages, mild, moderately advanced, and advanced, were defined by normal bilirubin and albumin, abnormal bilirubin or albumin, and abnormal bilirubin and albumin, respectively. As per standard of care, the following laboratory parameters were collected every 6-12 months: total bilirubin, alkaline phosphatase (ALP), albumin, aspartate aminotransferase, alanine aminotransferase, and platelet count.[8,11] Histological data obtained from liver biopsies were staged according to the criteria of Ludwig et al. and Scheuer.[12,13]

Statistical Analysis

The primary endpoint was a composite of LT and all-cause mortality. Survival was defined as an absence in LT and all-cause mortality. Patients without an event at the end of follow-up or who were lost to follow-up were censored at their last visit. The survival rates across quartiles corresponding to each cohort were estimated with a Kaplan-Meier curve and compared with a log-rank test. Multivariable Cox proportional hazards' regression (hazard ratio [HR] with 95% CI) analyses were performed to adjust for potential confounding variables.

To test the hypothesis of a threshold and to determine the optimal threshold for bilirubin within the normal range 2 approaches were followed: i) bilirubin at baseline and 1 year were dichotomized according to various thresholds ranging from 0.3 to 0.9 ×ULN in 0.01 increments. Multivariable Cox proportional hazards' regression analyses were used to estimate the risk for LT or death associated with each threshold. The C-statistic was calculated to evaluate the performance of each threshold in predicting survival and the threshold with the best performance was determined by the highest C-statistic. ii) To assess bilirubin on a continuous spectrum and test the hypothesis that the risk for LT or death increases at the predetermined bilirubin threshold, bilirubin was analyzed as a restricted cubic spline function with 4 knots. Patients with abnormal bilirubin were included to illustrate their risk for a poor prognosis relative to those with normal bilirubin. The restricted spline function was repeated with crude bilirubin levels (mg/dL).

Sensitivity analyses of the predetermined bilirubin threshold by multivariable Cox regression were performed in subgroups: bilirubin ULN (75^{th} percentile: 1.2mg/dl [21.0µmol/L]), age at study entry (≤45, 46-55, 56-65, and >65 years), sex, UDCA treatment, histologic stage (I/II and III/IV), and ALP (≤1.67×ULN and >1.67×ULN). Of note, sensitivity analyses according to histologic stage were conducted in a subset of patients with available histologic stage at baseline. Furthermore, sensitivity analyses were performed for bilirubin at 2-5 years after the start of follow-up.

Kaplan-Meier analyses were conducted to illustrate the survival rates associated with bilirubin at baseline and 1 year (normal bilirubin [≤/> the threshold] and abnormal bilirubin [reference purposes]). The distribution of the clinical events (LT, liver-related death, or liver-unrelated death) at 10 years within each bilirubin group was also evaluated. Of interest was the impact of bilirubin change in UDCA-treated patients with normal bilirubin at baseline, which was assessed after 1 year.

The pattern of bilirubin (mean and 95% CI) over the first 5 years was evaluated in patients with normal bilirubin at time zero and stratified on whether they experienced a late clinical event (LT or death from 5 to 10 years) or no clinical event in the first 10 years of follow-up.

The association of ALP with survival was assessed in a subgroup of UDCA-treated patients with ALP≤1.67×ULN at 1 year by dichotomization according to thresholds ranging from 0.5 to 1.6×ULN. ALP was assessed at 1 year in order to maximize the number of patients and emphasize the development of thresholds as treatment targets. The risk for LT and death for each threshold was estimated with Cox regression analyses. Within the normal range of ALP, the survival associated with quartiles was also emphasized.

Multiple imputation was conducted with by the Markov chain Monte Carlo method for missing data with SAS version 9.4 (SAS Institute Inc., Cary, NC, USA). Ten imputed data sets for missing biochemical values were generated to reduce sampling variability of laboratory results.[14] Imputation was performed based on the assumption that data were missing at random, in which variables predicting outcomes and outcomes themselves were included in the imputation model. Rubin's rules were used to estimate the parameter and standard error.[15,16]

All multivariable analyses were adjusted for age at study entry, sex, year of diagnosis, UDCA therapy, ALP, and geographical region. Biochemical markers that were not normally distributed were log transformed. A *P* value less than 0.05 was considered statistically significant. All analyses were two-sided and were performed using IBM SPSS Statistics for Windows, version 25.0 (IBM Corp., Armonk, NY, USA) and Statistical Analysis System (SAS) version 9.4 (SAS Institute Inc., Cary, NC, USA).

RESULTS

Study population characteristics

At baseline, there were 2,545 adult PBC patients with normal bilirubin, of whom 264 met exclusion criteria (54 for autoimmune hepatitis overlap, 76 for UDCA discontinuation, and 134 for short follow-up). Similarly, from 2,797 adult patients with PBC with normal bilirubin at 1 year, 242 were excluded (64 for autoimmune hepatitis overlap, 47 for UDCA discontinuation, and 131 for short follow-up). Overall, a total of 3,059 patients with normal bilirubin at baseline or one year were included and 2 cohorts were constructed: time zero cohort (n=2,281) and 1-year cohort (n=2,555). An overlap of 1,777 patients exists between these cohorts. There were 297 and 344 primary endpoints according to each respective cohort. Patient characteristics per cohort are presented in **Table 1**.

Table 1. Characteristics of PBC patients in each normal bilirubin cohort

Parameter	Time zero cohort [a, b] (n=2281)	1-year cohort (n=2555)
Follow-up time, y, median (IQR)	7.9 (4.3-12.7)	7.3 (3.7-11.5)
Age at study entry, mean ± SD	55.3 ± 12.0	54.6 ± 11.8
Female, no. (%)	2086 (91.5)	2354 (92.1)
AMA-positive, no. (%)	2036/2222 (91.6)	2273/2485 (91.5)
Year of diagnosis, median (range)[c]	1998 (1961-2014)	1997 (1961-2013)
UDCA-treated, no. (%)	1979/2223 (89.0)	2345/2523 (92.9)
Laboratory parameters, median (IQR)[d]		
Total bilirubin, ×ULN	0.53 (0.40-0.70)	0.50 (0.38-0.67)
ALP, ×ULN	1.99 (1.27-3.32)	1.26 (0.88-1.96)
Albumin, ×LLN	1.17 (1.09-1.26)	1.17 (1.09-1.26)
AST, ×ULN	1.30 (0.93-1.93)	0.87 (0.65-1.20)
ALT, ×ULN	1.51 (0.98-2.35)	0.83 (0.58-1.33)
Platelet count, $\times 10^9$/L	255 (207-308)	250 (202-304)
Bilirubin ULN (mg/dl), median (IQR)[e]	1.1 (1.0-1.2)	1.17 (1.0-1.2)

PBC, primary biliary cholangitis; IQR, interquartile range; SD, standard deviation; AMA, antimitochondrial antibody; UDCA, ursodeoxycholic acid; ULN, upper limit of normal; ALP, alkaline phosphatase; AST, aspartate aminotransferase; ALT, alanine aminotransferase.

[a]Histological disease stage at study entry available for 898 patients (39.4%)- stage I/II: 686 patients (76.4%), stage III/IV: 212 patients (23.6%).
[b]Biochemical disease stage at study entry available for 1809 patients (79.3%) – mild: 1670 (92.3%), moderate: 139 (7.7%), advanced: 0 patients.
[c]Year of diagnosis was available for 2279 in the time zero cohort and 2554 patients in the 1-year cohort.
[d]Laboratory parameters other than bilirubin were not available for all patients:
Time zero cohort: ALP (n=2119), albumin (n=1809), AST (n=2023), ALT (n=1937), platelet count (n=1627).
1-year cohort: ALP (n=2413), albumin (n=1678), AST (n=2138), ALT (n=2142), platelet count (n=1201).
[e]The upper limit of normal for bilirubin was variable per center.

Normal bilirubin quartiles are associated with survival

In patients with normal bilirubin at time zero, the cumulative 10-year survival rate decreased with higher bilirubin quartiles and was 93.3%, 89.9%, 87.7%, 81.3% from quartiles 1-4 (Q1-Q4), respectively (**Figure 1**). In pairwise comparisons, Q4 was significantly different from Q1-Q3 (all *P*<.005). In addition, Q1 was significantly different from Q3 (*P*=.04). Similar results were obtained in the Kaplan-Meier analysis of the 1-year cohort, in which the 10-year survival rates with increasing bilirubin quartiles were 92.0%, 92.3%, 86.1%, and 78.2%. Q3 and Q4 were significantly different from one another and from the remaining quartiles (all *P*<.01). In multivariable Cox regression analyses of the time zero cohort, the risk for LT or death increased with higher bilirubin quartiles: Q1 (reference), Q2 (HR 1.12, 95% CI 0.73-1.72, *P*=.61), Q3 (HR 1.34, 95% CI 0.89-2.01, *P*=.16), Q4 (HR 1.83, 95% CI 1.24-2.71, *P*=.003). A similar trend was observed in the 1-year cohort: Q1 (reference), Q2 (HR 0.97, 95% CI 0.65-1.45, *P*=.88), Q3 (HR 1.46, 95% CI 1.02-2.10, *P*=.04), Q4 (HR 2.20, 95% CI 1.56-3.10, *P*<.001).

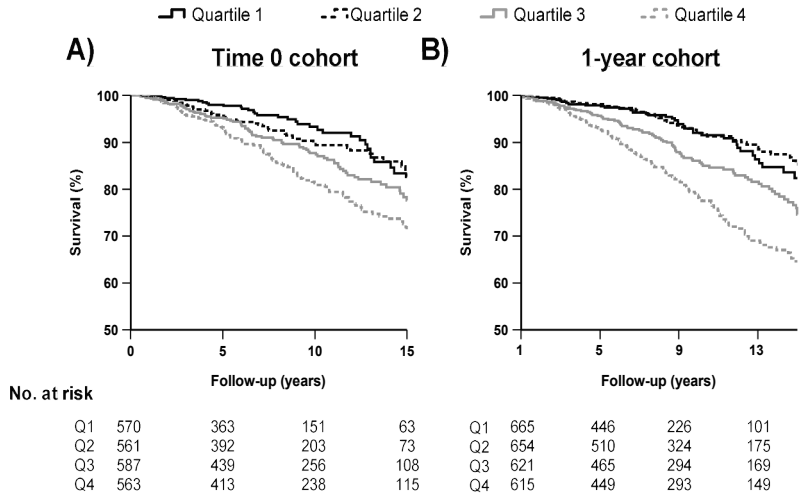


Figure 1. Survival estimates of bilirubin quartiles in patients with normal bilirubin. Kaplan-Meier estimates of bilirubin quartiles from (A) time zero and (B) 1 year.

Bilirubin threshold within the normal range to predict survival

On exploration of the optimal threshold of bilirubin within the normal range at 1 year, all bilirubin thresholds (0.3-0.9×ULN) were significant predictors of survival in that patients with bilirubin above each threshold had an increased risk for LT or death (**Supplementary Table 1**). The bilirubin threshold at 1 year with the highest ability to predict LT or death was 0.6×ULN (C-

statistic 0.7429, 95% CI 0.7144-0.7713). The 10-year survival of patients with normal bilirubin ≤0.6×ULN, normal bilirubin >0.6×ULN, and abnormal bilirubin at 1 year were 91.3%, 79.2%, and 37.3%, respectively (*P*<.001) (**Figure 2A**). At baseline, the 10-year survival rates were 91.7%, 85.6%, and 49.5% (*P*<.001). Discordant survival rates were also observed at 15 years (**Supplementary Table 2**). We evaluated the distribution of clinical events from the 10-year survival rates associated with each bilirubin group. Clinical events in patients with bilirubin from 0.6 to 1.0×ULN were characterized by a significantly increased proportion of LT and liver-related deaths, alongside a decreased proportion of liver-unrelated deaths as compared to patients with bilirubin ≤0.6×ULN (*P*<.001) (**Supplementary Figure 1**). Of 1,934 UDCA-treated patients with normal baseline bilirubin, 1-year bilirubin was available for 1,644 (85%). A UDCA-induced reduction in bilirubin (≤0.6×ULN) at 1 year was associated with a 10.5% improvement in 10-year survival (93.7% vs 83.2%) and 17.1% in 15-year survival (86.5% vs 69.4%) as compared to bilirubin that remained above the threshold (**Figure 2B**).

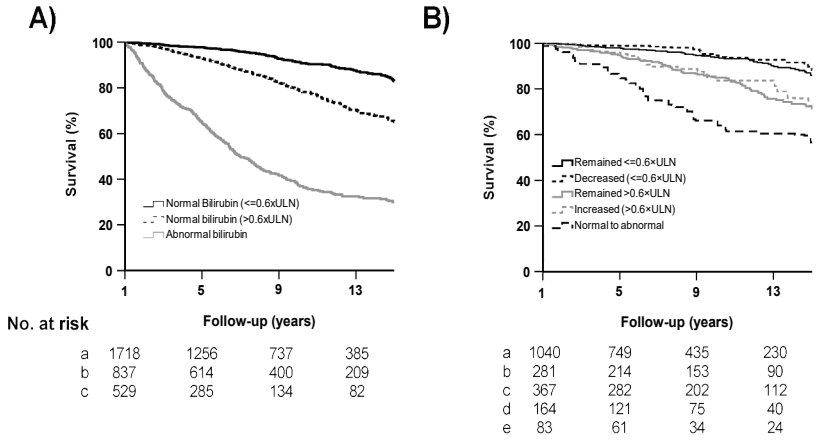


Figure 2. Survival estimates in patients with normal bilirubin (stratified by 0.6×ULN threshold) and abnormal bilirubin. (A) Kaplan-Meier estimates of survival rates in patients with normal bilirubin (stratified by 0.6×ULN threshold) and abnormal bilirubin at 1 year. (B) Additional analysis of the survival rates in ursodeoxycholic acid-treated patients with normal bilirubin levels at baseline and stratified according to the change in bilirubin from baseline to 1 year. ULN, upper limit of normal.

The threshold was evaluated in various subgroups of patients who had normal bilirubin at 1 year, all of which confirmed that patients with bilirubin >0.6×ULN have an increased risk for LT or death (**Figure 3**). Surpassing the bilirubin threshold of 0.6×ULN was associated with an increased risk for liver transplantation or death according to histologic stage, but demonstrated a trend for significance (Stage I/II: *P*=0.05; Stage III/IV:*P*=0.07) that may be because of the

limited availability of baseline histologic stages in our cohort. The association remained when patients in whom the ULN was defined as ≥1.2mg/dl (21.0μmol/L) were excluded from the analyses (HR 2.10, 95% CI 1.54-2.85, *P*<.001). In addition, the threshold was a significant factor in a subgroup (n=495) with known UDCA dosage ≥13mg/kg (HR 1.85, 95% CI 1.02-3.34, *P*=.04). There were no significant interactions between the bilirubin threshold and the variables explored in subgroup analyses (**Figure 3**).

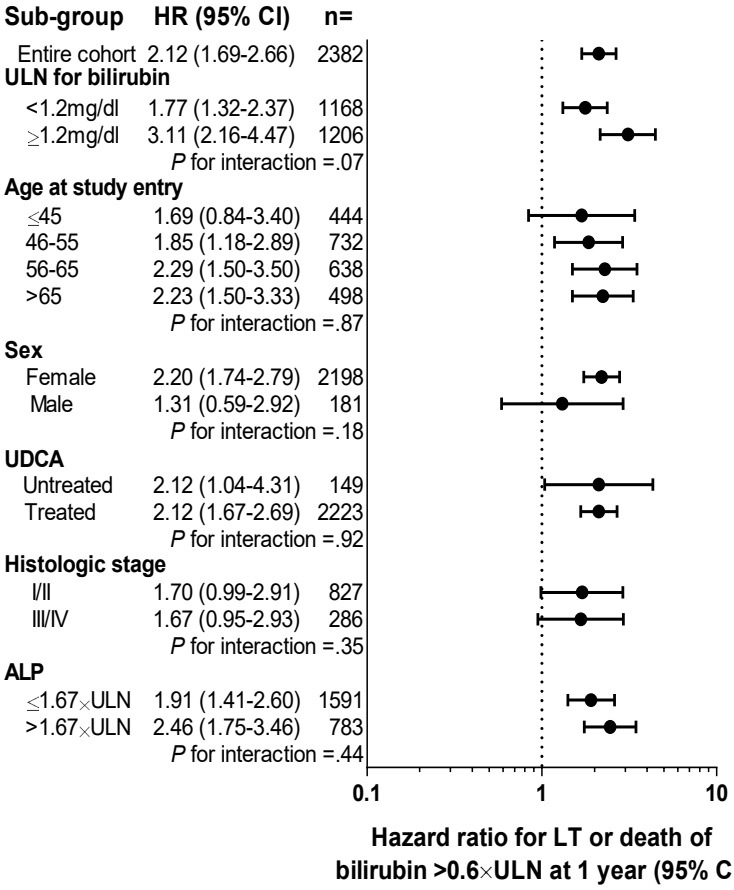


Figure 3. Subgroup analyses based on the bilirubin threshold of 0.6×ULN in patients with normal bilirubin at 1 year. Hazard ratio for liver transplantation or death (95% CI) obtained from multivariable Cox proportional hazards regression analyses in patients with normal bilirubin in various subgroups. The hazard ratios correspond to bilirubin levels >0.6×ULN (versus bilirubin ≤0.6×ULN). The *P* value for interaction corresponds to an interaction between the bilirubin threshold and associated variable.

The risk for LT or death increases at bilirubin levels of 0.6×ULN

We assessed bilirubin on a continuous spectrum with a restricted spline function to evaluate whether the predetermined threshold is the point at which the HR for LT or death increases. The reference in each cohort was the predetermined threshold of 0.6×ULN. In both cohorts, the risk for LT or death remained stable below 0.6×ULN (**Figure 4**). However, beyond this threshold, a linear relationship was observed between bilirubin and the risk for LT or death that continued past the normal range. The test for curvature that establishes a significant deviation from a linear relationship was significantly different for the time zero (*P*=.03) and 1-year cohorts (*P*=.05). As sensitivity analyses, the restricted spline function was repeated using crude bilirubin levels (mg/dL) and with normal bilirubin levels from 2 to 5 years. (**Supplementary Figures 2 and 3**).

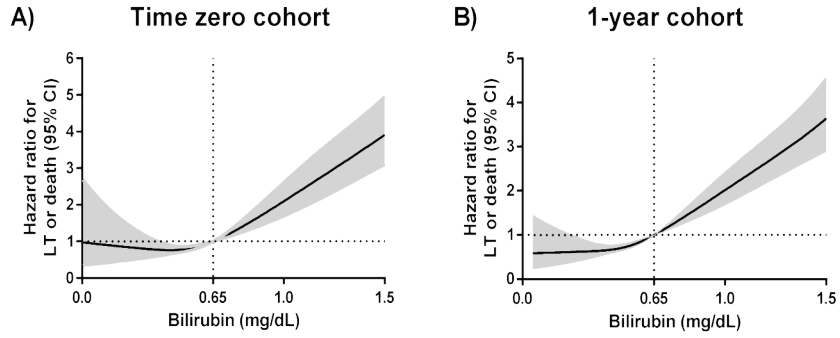


Figure 4. The association between bilirubin levels (×ULN) and risk for liver transplantation or death. Hazard ratios and 95% CI were estimated by a restricted cubic spline function in (A) the time zero cohort and (B) the 1-year cohort. The bilirubin reference in each cohort is 0.6×ULN. CI, confidence interval; LT, liver transplantation; and ULN, upper limit of normal.

Patients who remain below 0.6×ULN over time have good long-term prognosis

To assess how the trajectory of bilirubin over time may be related with the development of LT or death, bilirubin was evaluated over 5 years in patients with normal bilirubin at time zero. Patients were stratified according to whether they developed a late clinical event from 5 to 10 years (n=103) or did not develop a clinical event in the first 10 years of follow-up (n=848). Patients who had no clinical event within 10 years presented with a mean bilirubin level of 0.55×ULN (95% CI 0.54-0.56) that remained stable (below 0.6×ULN) over 5 years (**Supplementary Figure 4**). By contrast, patients who reached a clinical endpoint presented

with slightly higher mean bilirubin levels (0.61×ULN, 95% CI 0.57-0.65, *P*=.01) and exhibited a gradual increase within the normal range that precluded the occurrence LT or death.

ALP normalization is associated with improved survival

In a subgroup of patients with ALP ≤1.67×ULN from the normal bilirubin cohort at 1 year (n=1,523), the optimal ALP threshold was 1.0×ULN (C-statistic 0.7552, 95% CI 0.7151-0.7953). The HR for LT or death was 1.44 in those with ALP>1×ULN (95% CI 1.04-2.00, *P*=.03). Patients with ALP ≤1×ULN had the highest survival rate (93.2% at 10 years and 84.1% at 15 years), which was significantly different from those with ALP between 1.0 and 1.67×ULN (86.1% at 10 years and 76.4% at 15 years) and ALP>1.67×ULN (85.4% at 10 years and 73.8% at 15 years), *P*<.005 (**Supplementary Figure 5**). Survival between the group with ALP 1.0-1.67×ULN and that with ALP>1.67×ULN was similar (*P*=.64). Quartiles within the normal range of ALP (n=773) were not associated with survival.

ALP and bilirubin levels below 0.6×ULN

Implementing both ALP and bilirubin thresholds established, the prognosis of patients with bilirubin >0.6×ULN was dependent on ALP normalization (**Supplementary Figure 6**). Given normal ALP levels, their survival rates were similar to those with bilirubin ≤0.6×ULN, however, if ALP was 1.0-1.67×ULN, their survival was diminished to 74.2% at 10 years and 63.4% at 15 years (*P*<.001 compared to remaining groups).

DISCUSSION

This study reports that bilirubin levels within the normal range are associated with the risk for LT or death in patients with PBC. We demonstrated that bilirubin levels ≤0.6×ULN at baseline and 1 year were associated with a decreased risk for LT or death as compared to patients with bilirubin above this threshold. Although the risk for LT or death was stable when bilirubin levels were below 0.6×ULN, beyond this threshold, a positive linear relationship was observed between bilirubin and the risk for a clinical event. In patients with already normal bilirubin at baseline, a reduction below this threshold after 1 year of UDCA therapy was associated with an 11% improvement in 10-year survival. These results were confirmed in several subgroups of patients. These findings suggest that the interpretation of not being at risk if bilirubin is within the normal range needs to be revised. In addition, ALP normalization was also associated with prolonged survival. This might have implications in the number of patients eligible for inclusion in clinical trials for novel second-line therapies since ALP>1.67×ULN/abnormal bilirubin are common eligibility requirements. These results are in line with the change in criteria for disease remission in autoimmune hepatitis. Previously, one of the requirements for disease remission was transaminase levels below twice the ULN. However, the definition for disease remission now includes normal transaminases[17,18] because failure to normalize these liver enzymes is associated with an increased risk for relapse after treatment withdrawal[19,20], histological worsening or progression to cirrhosis[19,21], and poor clinical outcomes.[21–23]

Although the ULN of bilirubin is reported as the most predictive for survival in patients with PBC[2], we found that the risk for LT or death is already increased when bilirubin is above 0.6×ULN. The optimum bilirubin cutoff associated with survival has been previously suggested to be lower than the ULN.[24] The current ULN of bilirubin represents the 97.5 percentile cutoff in the general population, yet this may not be the best approach to determine an optimal threshold because levels below this threshold are not reflective of an absence of increased risk.[25] This may partly be explained by the high percentage (3%-10%) of individuals with Gilbert's syndrome in the general population.[26] In addition, the current ULN of bilirubin may be suboptimal for risk stratification in PBC due to the female predominance of the disease, whereas sex differences in bilirubin are present in the general population.[27] An American study based on the Third National Health and Nutrition Examination Survey assessed the serum bilirubin levels in 16,865 adults from the general population and reported that mean serum bilirubin levels are significantly lower in women (0.52 mg/dL ± 0.003) than in men (0.72 mg/dL ± 0.004).[25] Consequently, the 97.5 percentile cutoff was 0.5 mg/dL higher in men. Other studies have reported similar sex differences in bilirubin levels in the general population.[28,29] Thus, the overall ULN of bilirubin may be skewed higher because of the inclusion of both men and

women. These considerations suggest that the ULN for bilirubin may need to be stratified by sex, as has been previously implemented for aspartate aminotransferase.[30,31]

We found that the predictive value of the bilirubin threshold of 0.6×ULN was irrespective of age, UDCA treatment, and ALP levels. Importantly, it remained significantly predictive at various independent time points. Furthermore, in UDCA-treated patients with normal bilirubin levels above 0.6×ULN at treatment initiation, a reduction below 0.6×ULN was associated with prolonged survival. This suggests that besides the predictive value of bilirubin within the normal range, a treatment-induced reduction of bilirubin within the current normal range is beneficial for long-term prognosis, which could have important implications for current patient care, but also for the design and interpretation of future clinical trials of potential second-line therapies in PBC. Although recent clinical trials have often included normalization of bilirubin as a primary endpoint, it might be preferable to aim for lower bilirubin levels.[32,33]

The pattern of bilirubin over time may also be relevant, as there was an overall increase of 0.47×ULN in mean bilirubin after 5 years in patients who reached a clinical endpoint after an extended follow-up. Although rapid increases in bilirubin have been shown to preclude death in untreated patients, these results suggest that there is an association between the trajectory of bilirubin and clinical outcomes, even within the normal range.[3] The finding that mean bilirubin levels of patients who did not experience a clinical event remained below 0.6×ULN over time further supports an incentive to aim for bilirubin levels below our proposed threshold of 0.6×ULN and emphasizes the importance of continuous bilirubin evaluation even in early stage disease.

A robust analysis of the predictive value of bilirubin within the normal range would not be possible without the large number of patients and extended follow-ups available from the Global PBC Study Group cohort. Furthermore, bilirubin was assessed at multiple independent time points to confirm that bilirubin levels obtained during a random follow-up assessment could also be used for risk stratification. Nonetheless, some study limitations should be noted. Whereas total serum bilirubin levels in healthy patients are primarily composed of unconjugated bilirubin, it is predominantly of conjugated form in patients with PBC.[34] Therefore, evaluating the influence of solely the conjugated form could be of interest. However, the real-world nature of our cohort only allowed us to evaluate total bilirubin, since independent measurements of the conjugated and unconjugated forms is not part of routine standard of care in most laboratories. The methodological limitations to determining accurate conjugated bilirubin measurements need consideration because direct bilirubin also measures delta bilirubin. In addition, some patients may also be affected by Gilbert's syndrome and can

potentially lead to an underestimation of the association of bilirubin above 0.6×UN with survival, albeit this is not expected to play a major role because of the large sample size.

Although bilirubin was analyzed based on the ULN defined by local centers, which ranged from 0.6 to 1.7mg/dL, sensitivity analyses were performed to address this. The analyses with crude bilirubin levels (mg/dL) and the one excluding patients with an ULN above 1.2mg/dL (21.0 μmol/L) confirmed our initial findings and exclude the possibility that patients with bilirubin levels above 0.6×ULN have worse survival because of the utilization of high ULNs.

In this multicenter international follow-up study of patients with PBC, bilirubin levels below the current ULN were shown to be predictive of survival and 0.6×ULN was established as the threshold from which point on the risk for LT or death increases. In addition, treatment-induced reduction of normal bilirubin below 0.6×ULN was associated with prolonged survival. By contrast, ALP normalization was the established threshold for improved survival. Our proposed thresholds of 0.6×ULN for bilirubin and normalization for ALP represent a refinement of previous criteria with an aim to optimize survival and identify patients at risk for poor outcome. Moreover, their implementation can broaden the patient population included in intervention studies who may benefit from therapeutic agents.

REFERENCES

1. Kaplan MM, Gershwin ME. Primary biliary cirrhosis. N Engl J Med 2005;353:1261-1273.
2. 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.
3. Shapiro JM, Smith H, Schaffner F. Serum bilirubin: a prognostic factor in primary biliary cirrhosis. Gut 1979;20:137-140.
4. 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.
5. 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.
6. Corpechot C, Abenavoli L, Rabahi N, et al. Biochemical response to ursodeoxycholic acid and long-term prognosis in primary biliary cirrhosis. Hepatology 2008;48:871-877.
7. Corpechot C, Chazouillères O, Poupon R. Early primary biliary cirrhosis: Biochemical response to treatment and prediction of long-term outcome. J Hepatol 2011;55:1361-1367.
8. 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.
9. 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;67:1920-1930.
10. Reshetnyak VI. Primary biliary cirrhosis : Clinical and laboratory criteria for its diagnosis. World J Gastroenterol 2015;21:7683-7708.
11. Lindor KD, Bowlus CL, Boyer J, et al. Primary Biliary Cholangitis: 2018 Practice Guidance from the American Association for the Study of Liver Diseases. Hepatology 2019;69:394-419.
12. Ludwig J, Dickson ER, McDonald GSA. Staging of chronic nonsuppurative destructive cholangitis (Syndrome of primary biliary cirrhosis). Virchows Arch A Pathol Anat Histol 1978;379:103-112.
13. Scheuer P. Primary biliary cirrhosis. Proc R Soc Med 1967;60:1257-1260.
14. 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.
15. Rubin DB. Multiple imputation after 18+ years. J Am Statitical Assoc 1996;91:473-489.
16. Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York: Wiley, 1987.
17. Manns MP, Czaja AJ, Gorham JD, et al. Diagnosis and management of autoimmune hepatitis. Hepatology 2010;51:2193-2213.
18. European Association for the Study of the Liver. EASL clinical practice guidelines: Autoimmune hepatitis. J Hepatol 2015;63:971-1004.
19. Verma S, Gunuwan B, Mendler M, et al. Factors predicting relapse and poor outcome

in type I autoimmune hepatitis: role of cirrhosis development, patterns of transaminases during remission and plasma cell activity in the liver biopsy. Am J Gastroenterol 2004;99:1510-1516.

20. Montano-Loza AJ, Carpenter HA, Czaja AJ. Improving the end point of corticosteroid therapy in type 1 autoimmune hepatitis to reduce the frequency of relapse. Am J Gastroenterol 2007;102:1005-1012.

21. Muratori L, Muratori P, Lanzoni G, et al. Application of the 2010 American Association for the study of liver diseases criteria of remission to a cohort of Italian patients with autoimmune hepatitis. Hepatology 2010;52:1857-1857.

22. Miyake Y, Iwasaki Y, Terada R, et al. Persistent normalization of serum alanine aminotransferase levels improves the prognosis of type 1 autoimmune hepatitis. J Hepatol 2005;43:951-957.

23. Hoeroldt B, McFarlane E, Dube A, et al. Long-term outcomes of patients with autoimmune hepatitis managed at a nontransplant center. Gastroenterology 2011;140:1980-1989.

24. 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. J Hepatol 2014;60:1249-1258.

25. Zucker SD, Horn PS, Sherman KE. Serum bilirubin levels in the U.S. population: Gender effect and inverse correlation with colorectal cancer. Hepatology 2004;40:827-835.

26. Bosma PJ, Roy Chowdhury J, Bakker C, et al. The genetic basis of the reduced expression of bilirubin UDP-glucuronosyltransferase 1 in gilbert's syndrome. N Engl J Med 1995;333:1171-1175.

27. Podda M, Selmi C, Lleo A, et al. The limitations and hidden gems of the epidemiology of primary biliary cirrhosis. J Autoimmun 2013;46:81-87.

28. White GL, Nelson JA, Pedersen DM, et al. Fasting and gender (and altitude?) influence reference intervals for serum bilirubin in healthy adults. Clin Chem 1981;27:1140-1142.

29. Rosenthal P, Pincus M, Fink D. Sex- and age-related differences in bilirubin concentrations in serum. Clin Chem 1984;30:1380-1382.

30. Terrault NA, Bzowej NH, Chang K, et al. AASLD guidelines for treatment of chronic hepatitis B. Hepatology 2016;63:261-283.

31. Prati D, Taioli E, Zanella A, et al. Updated definitions of healthy ranges for serum alanine aminotransferase levels. Ann Intern Med 2002;137:1-10.

32. Nevens F, Andreone P, Mazzella G, et al. A placebo-controlled trial of obeticholic acid in primary biliary cholangitis. N Engl J Med 2016;375:631-643.

33. Corpechot C, Chazouillères O, Rousseau A, et al. A placebo-controlled trial of bezafibrate in primary biliary cholangitis. N Engl J Med 2018;378:2171-2181.

34. Levitt DG, Levitt MD. Quantitative assessment of the multiple processes responsible for bilirubin homeostasis in health and disease. Clin Exp Gastroenterol 2014;7:307-328.

SUPPLEMENTARY DATA

Supplementary Table 1. Multivariable Cox regression analyses of various normal bilirubin thresholds at 1 year for the prediction of liver transplantation and death

Bilirubin at 1 year (n=2382)				
Threshold (×ULN)	**C-statistic (95%CI)**	**HR (95% CI)**	***P* value**	**No. of patients ≤/> threshold**
0.30	0.7223 (0.6930-0.7515)	1.63 (1.04-2.58)	0.04	302/2080
0.40	0.7240 (0.6948-0.7531)	1.51 (1.12-2.06)	0.008	712/1670
0.50	0.7366 (0.7081-0.7651)	1.85 (1.46-2.36)	<0.001	1243/1139
0.55	0.7357 (0.7073-0.7642)	1.90 (1.51-2.40)	<0.001	1416/966
0.59	0.7400 (0.7114-0.7686)	2.02 (1.61-2.54)	<0.001	1573/809
0.60	0.7429 (0.7144-0.7713)	2.12 (1.69-2.66)	<0.001	1619/763
0.61	0.7423 (0.7137-0.7710)	2.09 (1.67-2.62)	<0.001	1630/752
0.62	0.7385 (0.7095-0.7676)	2.00 (1.60-2.50)	<0.001	1676/706
0.63	0.7385 (0.7095-0.7675)	2.02 (1.61-2.52)	<0.001	1687/695
0.65	0.7351 (0.7061-0.7642)	1.89 (1.51-2.37)	<0.001	1751/631
0.66	0.7354 (0.7063-0.7645)	1.90 (1.52-2.38)	<0.001	1755/627
0.67	0.7361 (0.7070-0.7652)	1.92 (1.53-2.40)	<0.001	1821/561
0.68	0.7354 (0.7064-0.7644)	1.89 (1.51-2.37)	<0.001	1824/558
0.69	0.7341 (0.7051-0.7631)	1.88 (1.50-2.36)	<0.001	1854/528
0.70	0.7344 (0.7055-0.7633)	1.91 (1.52-2.40)	<0.001	1889/493
0.75	0.7346 (0.7052-0.7640)	1.96 (1.54-2.49)	<0.001	1999/383
0.80	0.7336 (0.7045-0.7626)	2.14 (1.67-2.75)	<0.001	2085/297
0.85	0.7291 (0.6997-0.7584)	1.89 (1.41-2.52)	<0.001	2175/207
0.90	0.7253 (0.6959-0.7546)	1.86 (1.34-2.59)	<0.001	2242/140

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

Supplementary Table 2. Survival estimates at 15 years according to bilirubin quartiles and bilirubin threshold of 0.6×ULN

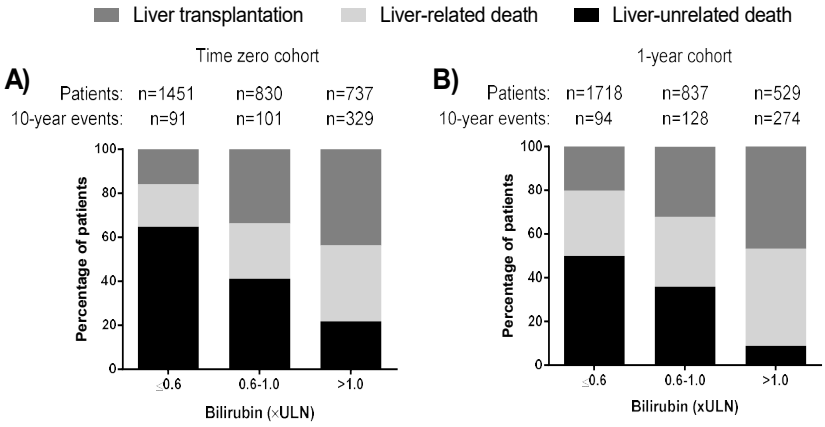
	15-year survival rates (%)	
Bilirubin quartiles	**Time zero cohort**	**1-year cohort**
Q1	83.4	82.3
Q2	82.6	85.4
Q3	77.6	74.6
Q4	71.3	64.6
Bilirubin threshold	**Time zero cohort**	**1-year cohort**
<0.6×ULN	82.0	82.9
0.6-1.0×ULN	73.0	65.1
>1×ULN	33.4	30.0

ULN, upper limit of normal.

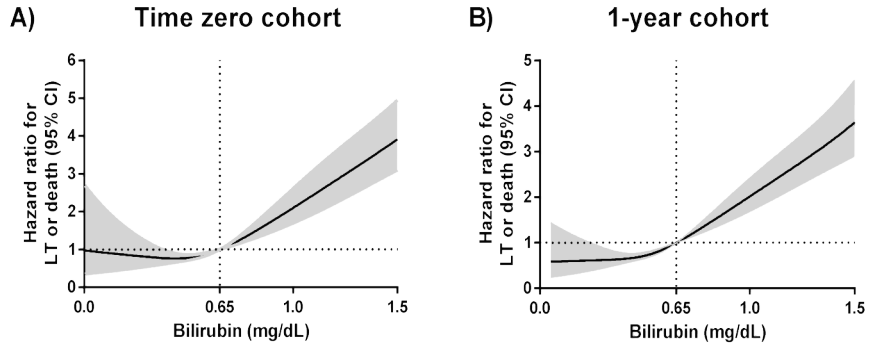
Supplementary Table 3. Multivariable Cox regression analyses of various normal bilirubin thresholds at time zero for the prediction of liver transplantation and death

Threshold (×ULN)	C-statistic (95% CI)	HR (95% CI)	*P* value	No. of patients ≤/> threshold
0.30	0.7335 (0.7021-0.7649)	1.48 (0.90-2.43)	.13	228/1824
0.40	0.7363 (0.7053-0.7674)	1.43 (1.00-2.05)	.05	522/1530
0.50	0.7404 (0.7098-0.7709)	1.56 (1.19-2.05)	.001	979/1073
0.55	0.7380 (0.7069-0.7692)	1.44 (1.11-1.86)	.006	1135/917
0.60	0.7403 (0.7093-0.7712)	1.47 (1.15-1.89)	.002	1323/729
0.65	0.7467 (0.7162-0.7772)	1.71 (1.33-2.19)	<.001	1435/617
0.66	0.7471 (0.7166-0.7776)	1.70 (1.33-2.17)	<.001	1439/613
0.67	0.7452 (0.7146-0.7757)	1.59 (1.24-2.03)	<.001	1492/560
0.68	0.7419 (0.7112-0.7726)	1.52 (1.18-1.95)	.001	1503/549
0.69	0.7421 (0.7115-0.7728)	1.52 (1.18-1.96)	.001	1527/525
0.70	0.7425 (0.7120-0.7730)	1.54 (1.20-1.98)	.001	1551/501
0.75	0.7414 (0.7112-0.7717)	1.60 (1.22-2.08)	.001	1681/371
0.80	0.7402 (0.7093-0.7711)	1.57 (1.18-2.09)	.002	1769/283
0.85	0.7398 (0.7090-0.7706)	1.66 (1.22-2.27)	.001	1860/192
0.90	0.7431 (0.7125-0.7737)	2.17 (1.53-3.09)	<.001	1929/123

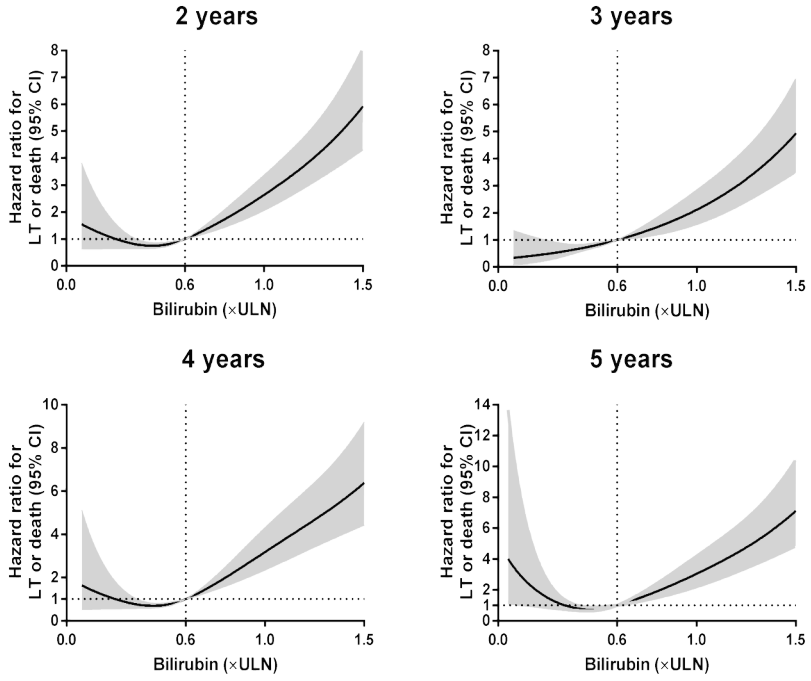
ULN, upper limit of normal; HR, hazard ratio; CI, confidence interval.



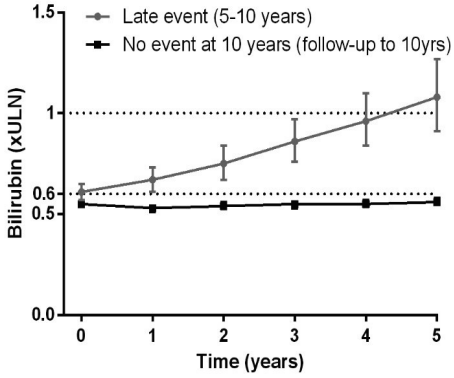
Supplementary Figure 1. Distribution of clinical events from the 10-year survival rates associated with each bilirubin group. Distribution of liver transplantation, liver-related death, and liver-unrelated death at (A) time zero and (B) 1 year. There was a significantly different distribution in the type of event according to bilirubin group at baseline and 1 year: $P < .001$ (Pearson Chi-square).



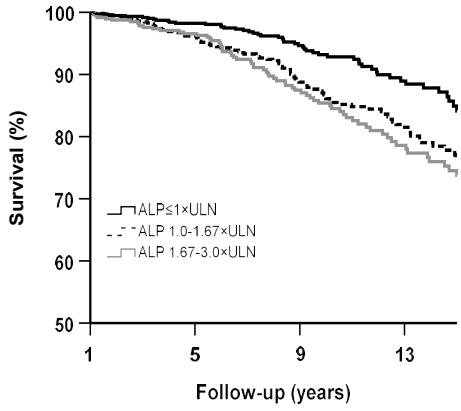
Supplementary Figure 2. The association between bilirubin levels (mg/dL) and risk for liver transplantation or death. Hazard ratios and 95% CI were estimated by a restricted cubic spline function in (A) the time zero cohort and (B) the 1-year cohort. The bilirubin reference in each cohort is 0.65 mg/dL (11.1μmol/L).



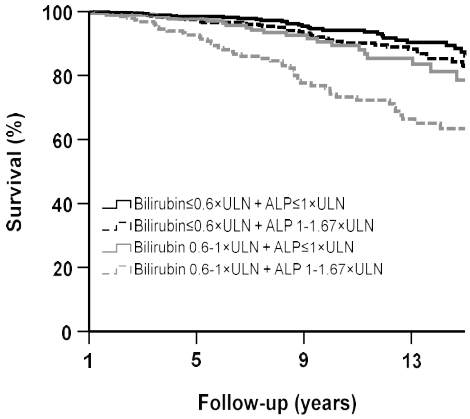
Supplementary Figure 3. The association between bilirubin level and risk for liver transplantation or death from 2-5 years. Hazard ratios and 95% CI were estimated by a restricted cubic spline function at 2-5 years. The bilirubin reference in each cohort is 0.6×ULN.



Supplementary Figure 4. Mean bilirubin levels over 5 years in patients with normal bilirubin at study entry and stratified by outcome. Trajectory of the mean bilirubin levels (×ULN) and 95% CI over the first 5 years depending on whether they experienced a late clinical event between 5 and 10 years (n=132) or no event within the first 10 years of follow-up (n=979). Clinical event is defined as liver transplantation or death. All patients without a clinical event had a follow-up of at least 10 years.



Supplementary Figure 5. Survival estimates stratified by ALP levels in patients with normal bilirubin at 1 year. Survival rates were significantly different between ALP≤1×ULN and 1.0-1.67×ULN (P = .001) and 1.67-3.0×ULN (P < .001), yet there was no significant difference between ALP 1.0-1.67 and ALP 1.67-3.0×ULN (P = .64). Survival was compared with the log-rank test.



a	579	407	243	123
b	552	416	238	126
c	194	148	97	48
d	198	143	97	51

Supplementary Figure 6. Survival estimates according to the established thresholds for bilirubin and alkaline phosphatase. Patients with normal bilirubin (≤1×ULN) and ALP≤1.67×ULN at 1 years were included.

CHAPTER 9

9

General discussion and conclusions

Background

Primary biliary cholangitis (PBC) is a chronic cholestatic liver disease that predominantly affects 1 in 1000 women over the age of 40. Due its relative rarity and slow progression to clinical events, studying this disease in a robust manner would require inclusion of a large number of patients with long-term follow-up, which can be difficult from single-center studies. Thus, a culmination of patient data from multiple centers and in an international manner can be greatly informative and aid in the generalizability of findings. Such efforts have been achieved with national and international groups such as the UK-PBC Consortium, CaNAL, and as evaluated in this thesis, the Global PBC Study Group. Importantly, the predictive ability of main liver biomarkers, alkaline phosphatase (ALP) and bilirubin, as surrogate end points for clinical outcomes in PBC patients was studied in the Global PBC Study group database.[1] These are now widely used in clinical practice and clinical trials, both to determine the need for second-line therapies and to evaluate the efficacy of treatment. Ultimately, the goal is to optimize patients' prognosis through guidance of patient management, which can be accomplished with care pathways, identification of risk factors, determination of the need of second-line therapies, and establishment of biochemical treatment targets.

Temporal and Spatial Trends in PBC

There have been multiple studies evaluating temporal trends for PBC, particularly with regards to its incidence and prevalence. A systematic review demonstrated that the prevalence for PBC has increased over time.[2] The pattern for incidence over time differs between studies, as some reported increased incidence and others stability.[3,4] Yet, seldom have patient characteristics and clinical outcomes of patients been described temporally. A single-center study from Padova, Italy suggested that the mean age at diagnosis increased from 48 years in patients diagnosed in 1973 to 64 years in those diagnosed in 2007.[5] Similarly, in a Japanese population, the median age at diagnosis increased from 59 year in 1999 to 63 years in 2004.[6] A Canadian population-based study of patients diagnosed from 1996 through 2002 did not report a significant difference in survival according to year of diagnosis.[7] **Chapter 2** evaluated the changes in patient characteristics and clinical outcomes of patients with PBC from a globally representative cohort over a 44-year period according to decades. The mean age at diagnosis increased by 2-3 years per decade from 47 years in the 1970s to 57 years beyond 2010. Furthermore, at presentation, patients were of a reduced disease severity in recent decades, as indicated by biochemical disease stage, histology, and ALP and bilirubin levels. The incidence of decompensation decreased over time and improvements in 10-year transplant-free survival were noted. These findings demonstrate how the patient population has evidently changed over time. Independent of whether this is due to an aging population, a

changing environmental trigger, increased disease awareness, increased routine testing of liver function, the introduction of ursodeoxycholic acid (UDCA), or a combination thereof, it highlights that PBC has not been a static disease. Given that the introduction of UDCA in the 1990s likely played a role in the improvement in survival, novel second-line therapies for those who do not respond to UDCA are still needed.[8] There is the possibility for continued changes in the patient population as a result of greater improvements in patient management and the introduction of novel second-line therapies such as obeticholic acid (OCA), which is currently being studied in relation to its effect on clinical outcomes (COBALT).

Like calendar time trends, those according to geographical region have primarily focused on incidence and prevalence. Different regions have varying incidence and prevalence of PBC, with some of the highest reported in Northeast England, Iceland, Calgary, and Minnesota.[7,9–11] In **Chapter 3**, the patient characteristics and clinical outcomes of patients from the European centers in the Global PBC database (Western Europe) were assessed according to geographical region. There were differences in decompensation and transplant-free survival according to region, albeit no differences in patient age and sex were noted. Increased hepatic decompensation rates and decreased transplant-free survival were observed for the Northwestern region evaluated, pertaining to the United Kingdom, even when adjusting for various markers of disease severity. Even though the contributing factors for the varying rates in clinical outcomes are not clear at this time, it raises the possibility that an environmental factor is at play. Understanding and acknowledging differences in patient outcomes within Europe and beyond can highlight at-risk populations based on region.

Risk Stratification

Some of the common factors that drive risk stratification in PBC are demographic factors and liver biochemistry.[12–18] Demographic factors such as age and sex have been found to be associated with prognosis. Male sex has been identified as a risk factor for incomplete response to UDCA, irrespective of age, portal hypertension, and biochemical indices of disease severity[19], and male patients appear to have diminished transplant-free survival compared to females.[7,20] On the other hand, increasing age is associated with increased response to UDCA, notably in female patients.[19] **Chapter 4** assessed the impact of age and sex on response to UDCA and clinical outcomes. The prognostic value of age on response to UDCA and transplant-free survival was validated as an increased age was associated with increased likelihood of response, while the risk for liver transplant or death decreased with advanced age when compared to a general population matched for age, sex, and birth year. In contrast, male sex was not associated with response nor survival in multivariable analysis. These results emphasize that younger patients ought to be monitored closely as they appear

to be at greater risk for incomplete response to UDCA, liver transplantation, and death. Indeed, the influence of age on response has been underlined in the UDCA Response Score, which is applied at baseline to predict the probability of response at 1 year to consider the need for second-line therapies in a timely manner.[21] Age is also included in the GLOBE score to predict transplant-free survival.[12] On the contrary, the worse prognosis of males can be explained by the fact that they are diagnosed at a more advanced disease stage, indicated by higher bilirubin and the greater proportion with moderately advanced and advanced biochemical stage. This is potentially due to delays in diagnosis given that they represent the minority of the affected population. Therefore, it is important to maintain high vigilance for a timely diagnosis in male patients with PBC.

Fibrosis stage is a well-established risk factor in PBC as an advanced histologic stage is associated with an increased risk for liver transplantation or death.[5,14] However, risk stratification assessed by published response criteria largely rely on biochemical markers and do not consider fibrosis stage. In **Chapter 5,** the utility of baseline fibrosis stage in determining transplant-free survival in the context of response to UDCA was evaluated. The correlations between non-invasive biochemical markers and histologic stage were poor. Advanced fibrosis stage at baseline, assessed histologically and via non-invasive biochemical markers, granted additional prognostic value to response to UDCA at 1 year. Despite biochemical response at 1 year, those with an advanced histologic stage have diminished transplant-free survival compared to those with an early fibrosis stage. This study highlights the need to incorporate fibrosis stage in patient risk stratification to identify the need for second-line therapies in addition to biochemical markers. Implementation can select a subgroup of patients who otherwise may not have been candidates for second-line therapies based on biochemistry alone. Yet, whether novel second-line therapies may benefit these patients remains to be studied. Importantly, there is a need to establish reliable non-invasive measures of fibrosis. Liver biopsies can aid in staging of disease, but they are not essential for a diagnosis of PBC. A promising approach to assess liver fibrosis is measuring liver stiffness by transient elastography, which demonstrates strong correlations with fibrosis stage.[22] The timing of elastography as a surrogate, be it at baseline or 1 year, needs further consideration due to the combined effect of fibrosis and cholestasis on liver stiffness.

Multiple risk scores that can predict the risk for clinical outcomes have been proposed over the years. Earlier risk scores include the Mayo Risk Score 1994 (MRS) and Model for End-stage Liver Disease (MELD) score, which were developed to predict short-term survival at 2 years and 3 months, respectively.[23,24] The more recent GLOBE and UK-PBC scores predict transplant-free survival and liver transplantation and liver-related death, respectively.[12,13] **Chapter 6** compared the performance of risk scores developed to predict outcome in PBC

(MRS, UK-PBC, and GLOBE score) in a UDCA-treated population. The GLOBE score was found to have consistently better performance in discriminatory ability for the prediction of liver transplantation and death in comparison to the remaining risk scores at various time points. However, this was not statistically different in comparison to the other scores. Importantly, the population was of a predominantly early biochemical disease stage, a population that represents the majority of patients in recent decades (**Chapter 1**). Therefore, implementation of risk scores in PBC should be based on clinical context.

Recently developed risk scores include similar risk factors in the model to predict survival, which are mainly liver biochemistry (ALP, bilirubin, albumin, platelet count, transaminases) and age. An improvement of these risk scores may be possible by considering additional variables that are associated with survival. Novel prognostic factors that are being increasingly evaluated are autoantibodies. Antinuclear antibodies (ANAs) are associated with disease severity and prognosis, as patients who are positive for these autoantibodies tend to have more severe biochemical and histological disease.[25] Anti-gp210 is associated with marked cholestasis and impaired liver function, as well as more severe interface hepatitis and lobular inflammation.[26–29] A meta-analysis that evaluated the prognostic value of anti-gp210 reported an association between anti-gp210 positivity and progression to liver failure and mortality.[30] Another prognostic factor that may be beneficial is neutrophil-to-lymphocyte ratio (NLR), for which a high NLR above 2.46 has been associated with reduced transplant-free survival in a Korean population.[31] Further research is required to establish the association between these prognostic factors and survival, but emphasize that efforts ought to be made to include other variables beyond demographics and liver biochemistry to obtain a more accurate and complete prediction model for clinical outcomes in patients with PBC.

Management of Patients

There are various risk stratification tools available for patients with PBC, both continuous risk scores and binary response criteria.[12–18] These have greatest value and applicability in specialized centers for PBC, as these may be confusing for non-specialist clinicians due to their complexity and abundance of options. It is important that there be the opportunity to highlight early in the course of disease those who should be offered enhanced care (programmes with focused PBC-experienced physicians), which can be achieved with care pathways. There is lack of a simple and rapid risk assessment tool to guide care pathways at diagnosis and assist non-expert gastroenterologists and primary care providers with the decision for referral. As demonstrated from previous research and in **Chapter 4**, increasing age is associated with an increased likelihood of response to UDCA and a smaller deviation from the survival of a general population. Age has not been consistently included in all risk

stratification tools, albeit present in the GLOBE score and UDCA Response Score.[12,21] In **Chapter 7**, a simple risk algorithm that captures age and serum liver tests was developed to guide care pathways. The ABA risk tool was proposed, which include age, bilirubin, and ALP, to three risk groups: low (Age>50 years, bilirubin ≤1×ULN, alkaline phosphatase [ALP] ≤3×ULN), high (Age≤50 years, bilirubin >1×ULN, ALP >3×ULN), and intermediate (other combinations). High risk patients were proposed to enhanced care and the remaining reassessed at 1 year, for which those in the intermediate and high risk were proposed to enhanced care. As an extension of the results from **Chapter 5**, fibrosis stage was assessed in those proposed to standard care (primary care) after 1 year, which was found to provide additional prognostic value. Overall, it was demonstrated that implementing a simple clinical risk tool in PBC in non-expert settings can promptly provide patients with augmented care if deemed necessary. While consideration of fibrosis stage is still important and to date transient elastography has been strongly correlated with fibrosis stage, it is not universally available across primary care settings. This could therefore hinder its use outside a liver care setting, particularly since the sole reliance on non-invasive biochemical markers such as APRI or FIB-4 is not optimal due to their poor correlation with fibrosis stage in PBC (**Chapter 5**). The takeaway message is that while ABA may not be determined as the optimal tool, it can be used as a basis to develop a simple, yet efficient care pathway in PBC.

Bilirubin is an established risk factor for PBC in treated and untreated patients.[1,32,33] It was first found to be associated with prognosis in untreated patients, in which a sharp increase in bilirubin preceded death.[32] Inclusion of bilirubin in various prognostic models and response criteria highlights its importance. Yet, bilirubin is thought to be of limited value early in the course of disease due to the fact that elevations are not observed until later stages of the disease.[34] A previous study by the Global PBC Study group assessed ALP and bilirubin as surrogate end points for clinical outcomes.[35] While this study determined normal bilirubin to be the optimal threshold, thresholds below the upper limit of normal were not evaluated. The majority of patients who are included in clinical trials have normal bilirubin, as demonstrated in the Phase 3 Study of Obeticholic Acid in Patients With Primary Biliary Cirrhosis (POISE) trial, for which 6-10% of patients had abnormal bilirubin.[36] Further, even within the normal range, patients experienced reductions in bilirubin. Whether bilirubin within the normal range or reductions in bilirubin were associated with clinical outcomes in PBC was unclear. In **Chapter 8**, the association between normal bilirubin and reductions within the normal range were examined, along with the predictive value of ALP below a threshold of 1.67×ULN. In this study, the treatment targets for ALP and bilirubin, which are main liver biomarkers in clinical trials, were redefined. A threshold of 0.6×ULN for bilirubin and ALP normalization were found to be optimal. Below 0.6×ULN of bilirubin, the lowest risk for liver transplantation or death is

observed, after which there is a linear increase in risk. A UDCA-induced reduction in bilirubin below this threshold at 1 year was also associated with improvements in transplant-free survival. As a result, implementing these thresholds can broaden the patient criteria for whom second-line therapies may be considered, as well as redefine treatment targets to optimize patient survival. While this study only focused on the main biomarkers of PBC, it is plausible that a similar association may be observed for additional liver biochemistry, which remains to be studied.

CONCLUSIONS

This thesis contributes a range of knowledge to the existing literature in PBC. Over the years, the patient characteristics and outcomes of patients with PBC has changed, characterized by an increase in age at diagnosis, milder disease stage, and improved clinical outcomes. It is plausible that geographical region may have an impact on clinical outcomes of PBC, particularly in the UK region within Western Europe. Age is an independent predictor for response to UDCA and transplant-free survival, while male patients tend to present with more advanced disease stage at baseline. Fibrosis stage is an important risk factor to consider in addition to biochemical response when risk stratifying patients, as those with biochemical response but advanced fibrosis have worse transplant-free survival. When comparing the various risk scores in PBC, the MRS, UK-PBC and GLOBE scores are comparable in the prediction of liver transplantation and death. In non-expert settings, the proposal of patients early in the course of disease to a care pathway and determination of the need for referral can be communicated with a simple risk assessment tool that requires fibrosis stage to be incorporated in future iterations. Bilirubin within the normal range and ALP normalization are associated with improved transplant-free survival. This suggests more patients ought to be treated to optimize survival and warrants a refinement of treatment biochemical targets.

REFERENCES

1. 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.
2. Boonstra K, Beuers U, Ponsioen CY. Epidemiology of primary sclerosing cholangitis and primary biliary cirrhosis: a systematic review. J Hepatol 2012;56:1181-1188.
3. 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:e31-e38.
4. McNally RJQ, James PW, Ducker S, et al. No Rise in Incidence but Geographical Heterogeneity in the Occurrence of Primary Biliary Cirrhosis in North East England. Am J Epidemiol 2014;179:492-498.
5. 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.
6. Sakauchi F, Oura A, Ohnishi H, et al. Comparison of the clinical features of Japanese patients with primary biliary cirrhosis in 1999 and 2004: utilization of clinical data when patients applied to receive public financial aid. J Epidemiol 2007;17:210-214.
7. 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;50:1884-1892.
8. 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;71:357-365.
9. James OF, Bhopal R, Howel D, et al. Primary biliary cirrhosis once rare, now common in the United Kingdom? Hepatology 1999;30:390-394.
10. Kim WR, Lindor KD, Locke GR 3rd, et al. Epidemiology and natural history of primary biliary cirrhosis in a US community. Gastroenterology 2000;119:1631-1636.
11. Baldursdottir TR, Bergmann OM, Jonasson JG, et al. The epidemiology and natural history of primary biliary cirrhosis: a nationwide population-based study. Eur J Gastroenterol Hepatol 2012;24:824-830.
12. 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.
13. 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.
14. Corpechot C, Abenavoli L, Rabahi N, et al. Biochemical response to ursodeoxycholic acid and long-term prognosis in primary biliary cirrhosis. Hepatology 2008;48:871-877.
15. Corpechot C, Chazouillères O, Poupon R. Early primary biliary cirrhosis: Biochemical response to treatment and prediction of long-term outcome. J Hepatol 2011;55:1361-1367.
16. 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.
17. 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;136:1281-1287.

18. 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;105:2186-2194.

19. 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;144:560-569.e7; quiz e13-4.

20. 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.

21. Carbone M, Nardi A, Flack S, et al. Pretreatment prediction of response to ursodeoxycholic acid in primary biliary cholangitis: development and validation of the UDCA Response Score. Lancet Gastroenterol Hepatol 2018;3:626-634.

22. Gómez-Dominguez E, Mendoza J, García-Buey L, et al. Transient elastography to assess hepatic fibrosis in primary biliary cirrhosis. Aliment Pharmacol Ther 2008;27:441-447.

23. 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.

24. 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.

25. Rigopoulou EI, Davies ET, Pares A, et al. Prevalence and clinical significance of isotype specific antinuclear antibodies in primary biliary cirrhosis. Gut 2005;54.

26. Gatselis NK, Zachou K, Norman GL, et al. Clinical significance of the fluctuation of primary biliary cirrhosis-related autoantibodies during the course of the disease. Autoimmunity 2013;46:471-479.

27. Hu CJ, Song G, Huang W, et al. Identification of new autoantigens for primary biliary cirrhosis using human proteome microarrays. Mol Cell Proteomics 2012;11:669-680.

28. Muratori P, Muratori L, Ferrari R, et al. Characterization and clinical impact of antinuclear antibodies in primary biliary cirrhosis. Am J Gastroenterol 2003;98:431-437.

29. 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;45:118-127.

30. Huang C, Han W, Wang C, et al. Early Prognostic Utility of Gp210 Antibody-Positive Rate in Primary Biliary Cholangitis: A Meta-Analysis. Dis Markers 2019;2019.

31. Yoo JJ, Cho EJ, Lee B, et al. Prognostic Value of Biochemical Response Models for Primary Biliary Cholangitis and the Additional Role of the Neutrophil-to-Lymphocyte Ratio. Gut Liver 2018;12:714-721.

32. Shapiro JM, Smith H, Schaffner F. Serum bilirubin: a prognostic factor in primary biliary cirrhosis. Gut 1979;20:137-140.

33. 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.

34. Lindor KD, Gershwin ME, Poupon R, et al. Primary biliary cirrhosis. Hepatology 2009;50:291-308.

35. van Buuren HR, Lammers WJ, Harms MH, et al. Surrogate Endpoints for Optimal Therapeutic Response to UDCA in Primary Biliary Cholangitis. Dig Dis 2015;33:118-124.

36. Nevens F, Andreone P, Mazzella G, et al. A placebo-controlled trial of obeticholic acid in primary biliary cholangitis. N Engl J Med 2016;375:631-643.

CHAPTER 10

10

Discussie en conclusies

Achtergrond

Primaire biliaire cholangitis (PBC) is een chronische, cholestatische leverziekte waarbij voornamelijk vrouwen zijn aangedaan; 1 op de 1000 vrouwen boven de 40 jaar heeft deze ziekte. Robuust onderzoek naar deze ziekte is lastig, omdat het relatief zeldzaam is en het beloop tot het eindstadium van de leverziekte of het overlijden aan de ziekte een langzaam progressief beloop kent. Om kwalitatief goed onderzoek te verrichten zijn grote patiëntengroepen nodig met een jarenlange follow-up duur en dit is nauwelijks te realiseren vanuit één onderzoekscentrum. Daarom is een samenwerking tussen multiple centra wereldwijd noodzakelijk en dit heeft als extra voordeel dat de resultaten beter generaliseerbaar zijn. Naast de Global PBC Study Group, waarvan meerdere studies opgenomen zijn in dit proefschrift, hebben meerdere samenwerkingsverbanden, zoals de UK-PBC Consortium en CaNAL, zichzelf reeds bewezen. De Global PBC Study Group richtte zich in eerste instantie op het valideren van biomarkers, zoals alkalische fosfatase (ALP) en bilirubine, als surrogaat voor klinische eindpunten in PBC patiënten.[1] Deze worden nu wereldwijd gebruikt in de klinische praktijk en in klinische trials waarbij de noodzaak voor tweedelijnsbehandeling wordt onderzocht, alsmede de effectiviteit van potentiële medicatie. Het uiteindelijke doel van internationale samenwerkingsverbanden is om de prognose van PBC patiënten te optimaliseren met behulp van heldere behandelstrategieën, door het opstellen van zorgpaden, het identificeren van risicofactoren, het vroegtijdig inzetten van tweedelijnsbehandelingen en het vaststellen van biochemische behandeldoelen.

Trends in tijd en plaats

Er zijn meerdere studies die trends over de tijd hebben beschreven, met name met betrekking tot incidentie en prevalentie. Zo toonde een systemische review aan dat de prevalentie van PBC is toegenomen over de tijd.[2] De incidentiecijfers over de tijd verschillen tussen studies, waarbij sommige studies stabiele cijfers beschrijven en anderen een toename.[3,4] Verschillen in karakteristieken en klinische uitkomsten van PBC patiënten over de tijd zijn echter sporadisch beschreven. Een studie uit Padova in Italië suggereert dat de gemiddelde leeftijd waarop de ziekte PBC wordt vastgesteld, is toegenomen van 48 jaar in 1973 tot 64 jaar in 2007.[5] Een Japanse studie laat een vergelijkbaar resultaat zijn, waarbij de gemiddelde leeftijd bij diagnose toenam van 59 jaar in 1999 tot 63 jaar in 2004.[6] Een Canadese population-based studie van patiënten gediagnosticeerd tussen 1996 en 2002 liet geen verschillen in overleving zien tussen de verschillende diagnosejaren.[7] **Hoofdstuk 2** toont de resultaten van een internationale studie die de trends in patiëntkarakteristieken en klinische uitkomsten van PBC patiënten beschrijft over een periode van 44 jaar. De gemiddelde leeftijd bij diagnose nam toe van 47 jaar in 1970 tot 57 jaar in 2010; een toename van 2-3 jaar per decade. Bovendien bleek

dat patiënten over de tijd in een eerder biochemisch of histologisch ziektestadium worden gediagnosticeerd. Ook de incidentie van gedecompenseerde leverziekte daalde en de 10-jaar transplantatievrije overleving verbeterde over de tijd. De resultaten laten duidelijk zien hoe de patiëntenpopulatie is veranderd en dat PBC geen statische ziekte is, onafhankelijk of dit wordt veroorzaakt door een ouder wordende populatie, verandering in omgevingsfactoren, meer bewustzijn voor de ziekte, de introductie van ursodeoxycholzuur (UDCA), of een combinatie van deze factoren. De introductie van UDCA in de jaren '90 heeft bijgedragen aan de verbetering van de overleving van het grootste deel van de PBC patiënten en daarom zijn tweedelijnsbehandelingen noodzakelijk voor degenen die niet adequaat op dit medicijn reageren.[8] Waarschijnlijk zal de patiëntenpopulatie ook in de toekomst blijven veranderen als een gevolg van verbeteringen in de behandeling van patiënten en de introductie van een tweedelijnsbehandeling, zoals obeticholzuur (OCA), waarvan het effect op klinische uitkomsten op dit moment wordt onderzocht in de COBALT studie.

Ook studies die zich richtten op het beschrijven van trends in een bepaalde geografisch gebied, focusten zich met name op incidentie- en prevalentiecijfers. Het blijkt dat er een duidelijke variatie is tussen regio's, waarbij de hoogste aantallen werden geobserveerd in Noordoost-Engeland, IJsland, Calgary en Minnesota.[7,9-11] In **hoofdstuk 3** worden de resultaten beschreven van een studie die verschillen in patiëntkarakteristieken en klinische uitkomsten heeft onderzocht tussen PBC patiënten uit verschillende West-Europese regio's. Ondanks dat er geen leefijds- en geslachtsverschillen waren tussen de diverse gebieden, werden er wel verschillen gevonden in de mate van decompensatie en transplantatievrije overleving. In het Noordwesten van het Verenigde Koninkrijk werd een hoger percentage gedecompenseerde leverziekte en een afgenomen transplantatievrije overleving gevonden, onafhankelijk van de verschillen in ernst van de ziekte tussen de regio's. Desondanks is het niet met zekerheid te zeggen welke factoren werkelijk bijdragen aan de regionale verschillen in klinische uitkomsten; mogelijk spelen omgevingsfactoren een belangrijke rol. Begrip van deze factoren kan bijdragen aan het bepalen van risicogroepen op basis van regio.

Risicostratificatie

Demografie en biochemie zijn belangrijke factoren waarmee het risico van PBC patiënten kan worden gestratificeerd.[12-18] Bekende demografische factoren, zoals leeftijd en geslacht, zijn geassocieerd met prognose. Zo is het mannelijke geslacht een risicofactor voor een incomplete respons op UDCA, onafhankelijk van de leeftijd, de aanwezigheid van portale hypertensie, en de ernst van de ziekte op basis van biochemische variabelen.[19] Eveneens hebben mannelijke PBC patiënten een lagere transplantatievrije overleving dan vrouwen.[7,20] Aan de andere kant is een hogere leeftijd geassocieerd met een verhoogde respons op UDCA

en dit geldt met name voor vrouwelijke patiënten.[19] In **hoofdstuk 4** komt de impact van leeftijd en geslacht op de UDCA respons en klinische uitkomsten aan bod. De prognostische waarde van leeftijd op UDCA respons en klinische uitkomsten werd gevalideerd; een hogere leeftijd was geassocieerd met een betere respons op UDCA, terwijl het overlijdensrisico juist lager was in vergelijking met de overleving van een algemene populatie (met een vergelijkbare verdeling van leeftijd, geslacht en geboortejaar). In tegenstelling hieraan werd er in een multivariabele analyse geen verschil gevonden tussen mannelijke en vrouwelijke PBC patiënten betreffende UDCA respons of transplantatievrije overleving. Deze resultaten onderstrepen dat jonge PBC patiënten strikter moeten worden gecontroleerd, omdat ze gemiddeld genomen minder goed reageren op behandeling met UDCA en een hoger risico hebben op leverfalen en dus de noodzaak voor een levertransplantatie of overlijden. Deze bevindingen zijn in lijn met de UDCA Response Score, die bij het starten met de UDCA behandeling voorspelt hoe groot de kans is op een respons na 1 jaar behandeling, om zo tijdig te kunnen ingrijpen met tweedelijnsbehandeling.[21] De variabele 'leeftijd' is ook onderdeel van de GLOBE score; een score die de transplantatievrije overleving in PBC patiënten voorspelt.[21] De matige prognose van mannelijke PBC patiënten wordt met name verklaard door het feit dat zij zich bij diagnose in een verder gevorderd ziektestadium bevinden (op basis van biochemische variabelen) dan vrouwen. Vermoedelijk komt dit doordat het proces van diagnosestelling langer duurt, omdat PBC nu éénmaal veel minder voorkomt bij mannen dan bij vrouwen. Het is dus belangrijk om waakzaam te blijven om zodoende op tijd de diagnose PBC bij een mannelijke patiënt te kunnen stellen.

Het stadium van fibrose is een belangrijke risicofactor in PBC, waarbij patiënten met een verder gevorderd ziektestadium op basis van histologie een lagere transplantatievrije overleving hebben.[5,14] Echter, risicostratificatie op basis van respons criteria is voornamelijk gebaseerd op biochemie en niet op histologie. In **hoofdstuk 5** werd geëvalueerd welke waarde het fibrosestadium heeft in het voorspellen van de transplantatievrije overleving van PBC patiënten in de context van respons op behandeling met UDCA. Er bleek een slechte correlatie te zijn tussen niet-invasieve biochemische waarden en het histologische stadium. Toegenomen fibrosestadium op basis van histologie en biochemie had additionele voorspellende waarde ten aanzien van biochemische respons op UDCA na 1 jaar behandeling. Het bleek namelijk dat responders op UDCA en een gevorderd histologische ziektestadium een minder goede transplantatievrije overleving hadden dan patiënten in een vroeg stadium van fibrose. Deze studie onderstreept het belang van het fibrosestadium in additie op biochemie bij het stratificeren van risico om patiënten te identificeren die in aanmerking komen voor tweedelijnsbehandeling. Zodoende kan een groep worden geïdentificeerd die alleen op basis van biochemie niet in aanmerking zou zijn gekomen voor tweedelijnsbehandeling. Of deze

patiënten baat zullen hebben bij tweedelijnsbehandeling moet nog worden onderzocht. Er is een duidelijke noodzaak voor niet-invasieve markers voor fibrose, omdat histologie niet meer nodig is voor het stellen van de diagnose PBC. Een veelbelovende aanpak om leverfibrose te meten is met behulp van de fibroscan, waarbij er eerder een sterke corelatie met het fibrosestadium is aangetoond.[22] Wanneer de fibroscan moet worden ingezet als surrogaat voor het fibrosestadium, het zij bij het stellen van de diagnose of na één jaar behandeling met UDCA, dient nog te worden onderzocht, mede gezien het feit dat de aanwezigheid van cholestase de uitslagen van de fibroscan negatief beïnvloed.

Er zijn veel risicoscores ontwikkeld over de jaren die klinische uitkomsten voorspellen voor PBC patiënten (MRS, UK-PBC en GLOBE). In **hoofdstuk 6** worden de resultaten weergeven van een studie die het voorspellend vermogen van deze risicoscores vergelijkt in UDCA-behandelde PBC patiënten. De GLOBE score had consequent betere prestaties in discrimininerende vermogen voor de voorspellen van levertransplantatie en overlijden in vergelijking met de resterende risicoscores op diverse punten in de tijd. Echter, dit was niet statistisch anders in vergelijking met de andere scores. Hierbij is het van belang dat het grootste deel van de onderzochte populatie PBC patiënten waren in een vroeg biochemisch ziektestadium; een populatie representatief voor de huidige groep PBC patiënten (**zie hoofdstuk 1**). Daarom moet de implementatie van risicoscores in PBC gebaseerd zijn op klinische context.

Recent ontwikkelde risicoscores maken gebruik van vergelijkbare variabelen in het model voornamelijk bestaande uit leverbiochemie (alkalische fosfatase, bilirubine, albumine, trombocyten en transaminasen) en leeftijd. Om deze scores te optimaliseren kan overwogen worden om niet-biochemische variabelen toe te voegen die geassocieerd zijn met overleving. Tegenwoordig is er veel belangstelling voor autoantilichamen. Antinucleaire antilichamen (ANA's) zijn geassocieerd met de ernst van de ziekte en met prognose, aangezien patiënten positief voor deze autoantilichamen ernstigere leverziekte hebben op basis van biochemie en histologie.[25] Anti-gp210 is geassocieerd met uitgesproken cholestase, verminderde leverfunctie, ernstigere interface hepatitis en lobulaire inflammatie.[26-29] Een meta-analyse die de prognostische waarde van anti-gp210 evalueerde, rapporteerde een associatie tussen de aanwezigheid van anti-gp210 en progressie van leverfalen en mortaliteit.[30] Een andere prognostische factor is de neutrofiel-lymfocyt ratio (NLR), waarbij een NLR boven de 2.46 geassocieerd is met een verminderde transplantatievrije overleving in een Koreaanse populatie.[31] Verder onderzoek is nodig om de associatie te onderzoeken tussen deze prognostische factoren en overleving. Ze maken wel duidelijk dat andere variabelen dan de bekende demografische en biochemische variabelen mogelijk een accuratere voorspelling kunnen geven van klinische uitkomsten in PBC patiënten.

Behandeling van PBC patiënten

Er zijn verschillende middelen beschikbaar voor risicostratificatie in PBC patiënten, waaronder binaire- en continue responscriteria.[12-18] Deze hebben een grote waarde voor de praktijk en ze worden toegepast in gespecialiseerde PBC centra. Echter, voor artsen die PBC patiënten behandelen in minder gespecialiseerde ziekenhuizen kunnen deze modellen te complex en uitgebreid zijn. Het is van belang om vroegtijdig in het beloop van de ziekte patiënten te selecteren die verwezen dienen te worden voor gespecialiseerde zorg en hiervoor zouden zorgpaden kunnen worden gebruikt. Daarom zou een simpel en snel te berekenen score behulpzaam kunnen zijn, echter zo'n score is op dit moment niet beschikbaar. Zoals reeds beschreven in **hoofdstuk 4** en gedemonstreerd door andere studies, is een hogere leeftijd geassocieerd met een toegenomen kans op UDCA respons en een overleving die dichter bij die van een algemene bevolking ligt dan van jonge PBC patiënten. Op dit moment is leeftijd alleen geïncludeerd in de GLOBE score en de UDCA Response Score.[12,21] In **hoofdstuk 7** wordt een simpel algoritme (ABA score) gepresenteerd inclusief leeftijd en biochemie. Deze score onderscheid 3 risicogroepen: laag risico (leeftijd >50 jaar, bilirubine ≤1xULN, alkalische fosfatase ≤3xULN), hoog risico (leeftijd ≤50 jaar, bilirubine >1xULN en ALP>3xULN) en gemiddeld risico (andere combinaties van leeftijd, bilirubine en alkalische fosfatase). Op basis van de resultaten van **hoofdstuk 5** werd de toegevoegde waarde van fibrosestadium na 1 jaar behandeling met UDCA geanalyseerd in de laagrisico groep en dit bleek het geval te zijn. In het algemeen werd met deze studie aangetoond dat implementatie van deze ABA-score in niet-in-PBC-gespecialiseerde ziekenhuizen patiënten kan selecteren die meer gespecialiseerde zorg nodig hebben. Ondanks dat de aanvullende waarde van het fibrosestadium bewezen is en de fibroscan een goede niet-invasief medium is om fibrose vast te stellen, is deze fibroscan niet altijd beschikbaar. Helaas hebben andere niet-invasieve factoren, zoals de APRI of FIB-4 scores, geen sterke associatie met de aan- of afwezigheid van fibrose in PBC patiënten (**hoofdstuk 5**). Ondanks dat de ABA score hierdoor niet optimaal is, kan deze score de basis vormen voor de ontwikkeling van een simpel en efficiënt zorgpad in PBC.

Bilirubinewaarden zijn geassocieerd met prognose in onbehandelde en behandelde PBC patiënten.[1,32,33] In eerste instantie werd deze associatie beschreven bij onbehandelde patiënten, waarbij een sterke stijging in bilirubinewaardes voorafging aan het overlijden van PBC patiënten.[32] Het is niet voor niets dat bilirubine geïncorporeerd is in multipele predictiemodellen. Bilirubine heeft echter voornamelijk belangrijk prognostische betekenis bij gevorderde leverziekte, omdat de bilirubinewaardes pas laat in het ziekteproces stijgen.[34] Een eerdere studie van de Global PBC Study Group onderzocht reeds de waarde van alkalische fosfatase en bilirubine als surrogaten voor klinische eindpunten.[35] Ondanks dat deze studie

normale bilirubinewaardes als optimaal beschouwde, werd de prognostische betekenis van bilirubinewaarden onder de bovengrens van normaal niet onderzocht. Tegenwoordig heeft het merendeel van de patiënten in klinische studies normale bilirubinewaarden, zoals gedemonstreerd in de fase 3 studie voor obeticholzuur in PBC (POISE trial), waar slechts 6-10% een abnormale bilirubinelevel had.[36] Patiënten in deze studie lieten een daling van het bilirubine zien binnen de grenzen van normaal. Het is echter onduidelijk of deze daling zich ook daadwerkelijk laat vertalen in verbeterde transplantatievrije overleving. In **hoofdstuk 8** werd de associatie tussen normale bilirubinewaardes en een reductie binnen de normaalwaarden onderzocht, alsmede de predictieve waarde van alkalische fosfatase waarden onder de 1.67xULN. In deze studie werden nieuwe behandeldoelen op basis van alkalische fosfatase en bilirubine vastgesteld; de belangrijkste biochemische variabelen die gebruikt worden in klinische studies. Geconcludeerd werd dat de meest optimale afkapwaarden voor bilirubine en alkalische fosfatase 0.6xULN en 1.0xULN respectievelijk waren. Patiënten met bilirubinewaarden onder de 0.6xULN hebben het laagste risico op de noodzaak voor een levertransplantatie of overlijden en het risico neemt lineair toe met stijgende bilirubinewaarden. Een reductie van bilirubinewaarden onder deze afkapwaarde na 1 jaar behandeling met UDCA was ook geassocieerd met een verbeterde transplantatievrije overleving. Op basis van deze resultaten zouden huidige behandeldoelen moeten worden bijgesteld, zodat de overleving van PBC patiënten wordt geoptimaliseerd. Deze studie richtte zich op de belangrijkste leverwaarden, alkalische fosfatase en bilirubine; het is goed mogelijk dat andere leverbiochemie een vergelijkbare trend laat zien. Om dit aan te tonen zijn verdere studies noodzakelijk.

Conclusies

Dit proefschrift voegt relevante kennis toe aan de bestaande literatuur over PBC. Door de tijd heen zijn de karakteristieken van PBC patiënten alsmede de prognose veranderd en dit wordt geïllustreerd door een toename van de leeftijd bij de diagnosestelling, een mildere ziekte en een verbetering van klinische uitkomsten door de jaren heen. Het is aannemelijk dat geografische verschillen een invloed hebben op de klinische uitkomsten in PBC, met name in het Verenigde Koninkrijk ten opzicht van de rest van West-Europa. Leeftijd is een belangrijke onafhankelijke voorspeller voor de respons op UDCA en voor de transplantatievrije overleving, terwijl mannen zich voornamelijk presenteren met een verder gevorderd ziektestadium ten tijde van de diagnosestelling. Wanneer het gaat om risicostratificatie is het fibrosestadium van toegevoegde waarde op biochemische respons op UDCA en dit wordt onderstreept door de bevinding dat PBC patiënten met een respons op UDCA en een toegenomen fibrosestadium een slechtere transplantatievrije overleving hebben. MRS, UK-PBC en GLOBE risicoscores had vergelijkbaar prestaties in discrimininerende vermogen voor de voorspellen van

levertransplantatie en overlijden. In ziekenhuizen zonder PBC-experts kan er gebruik worden gemaakt van een makkelijke score (ABA-score) om onderscheid te maken tussen patiënten die vervolgens in een zorgpad kunnen worden behandeld, of patiënten die naar een expertcentrum moeten worden verwezen. In de toekomst zal moeten worden onderzocht hoe het fibrosestadium in deze score kan worden geïmplementeerd. Bilirubine binnen de normaalwaarden en normalisatie van alkalische fosfatase zijn geassocieerd met een verbeterde transplantatievrije overleving. Deze bevinding suggereert dat er meer PBC patiënten zijn die baat kunnen hebben bij tweedelijnsbehandeling en dat huidige behandeldoelen moeten worden bijgesteld.

A special thank you to Wim Lammers for helping translate this summary

CHAPTER 11

11

11.1 CONTRIBUTING AUTHORS

Pier Maria Battezzati
Department of Health Sciences
Università degli Studi di Milano
Milan, Italy

Tony Bruns
Department of Internal Medicine IV
Jena University Hospital, Friedrich Schiller University
Jena, Germany
Department of Internal Medicine III
University Hospital RWTH Aachen
Aachen, Germany

Henk R. van Buuren
Department of Gastroenterology and Hepatology
Erasmus University Medical Center
Rotterdam, the Netherlands

Marco Carbone
Division of Gastroenterology and Center for Autoimmune Liver Diseases
San Gerardo Hospital
Department of Medicine and Surgery
University of Milano-Bicocca,
Monza, Italy

Nora Cazzagon
Department of Surgery, Oncology and Gastroenterology
University of Padua
Padua, Italy

Olivier Chazouillères
Reference Center for Inflammatory Biliary Diseases and Autoimmune Hepatitis
Saint-Antoine Hospital, APHP
Sorbonne University
Paris, France

Angela C. Cheung
Division of Gastroenterology and Hepatology
Mayo Clinic
Rochester, MN, United States of America

Christophe Corpechot
Reference Center for Inflammatory Biliary Diseases and Autoimmune Hepatitis
Saint-Antoine Hospital, APHP
Sorbonne University
Paris, France

George N. Dalekos
Department of Medicine and Research Laboratory of Internal Medicine
National Expertise Center of Greece in Autoimmune Liver Diseases
General University Hospital of Larissa
Larissa, Greece

Jordan J. Feld
Toronto Centre for Liver Disease
Toronto General Hospital, University Health Network
Toronto, ON, Canada

Annarosa Floreani
Department of Surgery, Oncology and Gastroenterology
University of Padua
Padua, Italy

Nikolaos K. Gatselis
Department of Medicine and Research Laboratory of Internal Medicine
National Expertise Center of Greece in Autoimmune Liver Diseases
General University Hospital of Larissa
Larissa, Greece

Alessio Gerussi
Division of Gastroenterology and Center for Autoimmune Liver Diseases
San Gerardo Hospital
Department of Medicine and Surgery
University of Milano-Bicocca,
Monza, Italy

Jorn C. Goet
Department of Gastroenterology and Hepatology
Erasmus University Medical Center
Rotterdam, the Netherlands

Aliya Gulamhusein
Toronto Centre for Liver Disease
Toronto General Hospital, University Health Network
Toronto, ON, Canada

Bettina E. Hansen
Toronto Centre for Liver Disease
Toronto General Hospital, University Health Network
Institute of Health Policy, Management and Evaluation
University of Toronto
Toronto, ON, Canada

Maren H. Harms
Department of Gastroenterology and Hepatology
Erasmus University Medical Center
Rotterdam, the Netherlands

Gideon M. Hirschfield

Toronto Centre for Liver Disease

Toronto General Hospital, University Health Network

Toronto, ON, Canada

Pietro Invernizzi

Division of Gastroenterology and Center for Autoimmune Liver Diseases

San Gerardo Hospital

Department of Medicine and Surgery

University of Milano-Bicocca

Monza, Italy

Harry L.A. Janssen

Toronto Centre for Liver Disease

Toronto General Hospital, University Health Network

Toronto, ON, Canada

Kris V. Kowdley

Liver Care Network and Organ Care Research

Swedish Medical Center

Seattle, WA, United States of America

Willem J. Lammers

Department of Gastroenterology and Hepatology

Erasmus University Medical Center

Rotterdam, the Netherlands

Nicholas F. LaRusso

Division of Gastroenterology and Hepatology

Mayo Clinic

Rochester, MN, United States of America

Konstantinos N. Lazaridis
Division of Gastroenterology and Hepatology
Mayo Clinic
Rochester, MN, United States of America

Keith D. Lindor
Department of Gastroenterology and Hepatology
Mayo Clinic
Rochester, MN, United States of America
Arizona State University
Phoenix, AZ, United States of America

Ana Lleo
Division of Internal Medicine and Hepatology
Humanitas Clinical Research Center IRCSS
Humanitas University
Rozzano (Milan), Italy

Xiong Ma
Shanghai Jiao Tong University School of Medicine
Shanghai Renji Hospital
Shanghai, China

Andrew L. Mason
Divison of Gastroenterology and Hepatology
University of Alberta
Edmonton, AB, Canada

Marlyn J. Mayo
Digestive and Liver Diseases Clinic
UT Southwestern Medical Center
Dallas, TX, United States of America

Adriaan J.P. van der Meer
Department of Gastroenterology and Hepatology
Erasmus University Medical Center
Rotterdam, the Netherlands

Frederik Nevens
Department of Hepatology
University Hospitals Leuven, KU Leuven
Leuven, Belgium

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

Cyriel Y. Ponsioen
Department of Gastroenterology and Hepatology
Academic Medical Center
Amsterdam, the Netherlands

Raoul Poupon
Reference Center for Inflammatory Biliary Diseases and Autoimmune Hepatitis
Saint-Antoine Hospital, APHP
Sorbonne University
Paris, France

Florian Prechter
Department of Internal Medicine IV
Jena University Hospital
Jena, Germany

Anna Reig

Liver Unit, Hospital Clínic, CIBERehd, IDIBAPS

University of Barcelona

Barcelona, Spain

Atsushi Tanaka

Department of Medicine

Teikyo University School of Medicine

Tokyo, Japan

Douglas Thorburn

The Sheila Sherlock Liver Centre

The Royal Free Hospital, London

United Kingdom

Palak J. Trivedi

National Institute for Health Research Birmingham Biomedical Research Centre and Centre for Liver and Gastrointestinal Research

Institute of Immunology and Immunotherapy

University of Birmingham

Birmingham, United Kingdom

Xavier Verhelst

Department of Gastroenterology and Hepatology

Ghent University Hospital

Ghent, Belgium

Kalliopi Zachou

Department of Medicine and Research Laboratory of Internal Medicine

University of Thessaly

Larissa, Greece

11.2 ACKNOWLEDGEMENTS

I am grateful to be able to share this thesis that I have worked on with great pride. This would not have been possible without the contribution of many people, whom I would like to thank for their involvement.

Dear prof. dr. Metselaar, I would like to thank you for the opportunity to conduct my PhD research at Erasmus MC in Rotterdam. This has been a great and unique experience for me and your support along the way is greatly appreciated.

Dear *Harry* (prof. dr. H.L.A. Janssen), I am thankful that you gave me the opportunity to commence my research in the liver field. This has opened so many opportunities for me. I am always motivated to do well and achieve high with your words of encouragement.

Dear *Bettina* (dr. B.E. Hansen), I am thankful to have such a kind and down to earth supervisor. We have worked closely over the years, which has allowed me to grow and learn, both academically and personally. Your expertise and knowledge in statistics and research overall have made all the difference in my projects and are always appreciated. I l always look forward to our meetings for your insight!

I would like to give thanks to prof. dr. R.A de Man, prof. dr. G.M. Hirschfield, and prof. dr. U.H.W Beuers for being part of the sub-committee and evaluating this thesis. My gratitude also goes to prof. dr. J.P.H. Drenth, prof. dr. K. van Erpecum, and dr. H.R. van Buuren for being members of the committee for my PhD defense.

Dear *Henk* (dr. H.R. van Buuren) and *Ad* (dr. A.J.P. van der Meer), thank you for your support on this PhD and the comments and feedback that you have provided on my projects.

Dear *Aliya* (dr. A. Gulamhusein) and *Gideon* (dr. G.M. Hirschfield), thank you for your clinical input and always providing critical feedback. You have allowed me to always view things from a critical perspective. I admire your dedication to your patients and research.

Dear *Jordan* (dr. J.J. Feld), I value the commitment you have for research and as a hepatologist. Thank you for having confidence in my abilities and appreciating my work, which you have helped me improve with your valuable feedback.

Dear Global PBC Study Group members, I have always been appreciative of the large and diverse group that has made important contributions to the PBC literature, both prior and during the course of my studies. This has allowed me to work with a team from all over the world, which has continued to grow over the years. I would like to thank you all for your valuable comments and input on all my projects.

Geraldine, you always had a solution for any issue that may arise, thank you for supporting me throughout the years.

Marion, as an international student in the Netherlands, everything was different and new to me and I would often have many questions. Thank you for your patience and striving always to get me an answer.

Eli and Melissa, your contribution is indispensable. Thank you for help over the years.

Heidy and Kattleya, you have been a valuable addition to the autoimmune group. You are diligent and I appreciate your help whenever I have needed it.

To my colleagues in Toronto: I am thankful to have such a diverse and large group, all from different backgrounds but with similar goals. I will look back on all the memories fondly. The wall of pictures summarises all our unique qualities!

Seng, you always made sure the group would have activities to do and organized many social events, we miss you and wish you the best. *Hannah, Surain, and Mina*, I remember our first conference was for AASLD in Washington, which allowed us to get to know each other better and become closer as a team. I always enjoy the talks, sometimes long and philosophical. We all have different career paths and look forward to seeing everyone's journey. *Marwa, Christina*, and *Natalie*, the research team grew further when you arrived! *Marwa*, you are always witty and often have an interesting/funny story to tell, thanks for brightening up my days! *Christina*, it was an adventure going to Ottawa by train last minute, but we enjoyed.

To my colleagues in Rotterdam: It has been such an honour to experience to the Dutch culture. *Lisette*, I know that you are always willing to help whenever someone needs it, like we always helped each other when we had troubles with stats or figures. Your dedication and hard work always stood out to me. *Wim*, it has been great being able to work together. I appreciate that my work was able to build upon yours! Thank you for your help with this thesis. Always appreciate the conversation! *Jorn and Maren,* I feel grateful to have been able to collaborate with you and publish good papers, your help is appreciated. *Rozanne*, thank you for your guidance while in the Netherlands!

To my friends and family, both in Canada in Peru, I want to thank you all for always being there and supporting me.

Paulina, you have been a close friend of mine from the time that I can remember. I never would have thought that this friendship would be so important to me when we met in elementary school all those years back. I admire your dedication to achieve what you put your mind to. *Alex*, an unexpected but beautiful friendship grew from a mutual friendship. Thank you for

being there for me when I need you most, always making see the positive. I can't wait to see what's in store for us.

Majana (Maria Alejandra), as my older cousin, I have always looked up to you since I was younger. You experienced all the milestones in life before me and I was always curious as to how you handled it all. You had the experience of writing your thesis before me, and you saying "I can't, I can't" when you did it, was my reminder that even when things seem hard, if we put our mind to it we can do anything. Thank you for being an inspiration and role model.

Diego, my dear brother, always spontaneous and living life to the fullest. I hope that I drive you to aim higher in life and that all your wishes come true!

Mina y Mino, mis abuelos queridos, gracias por siempre apoyarme en todo. Espero hacerles orgullosos de todos mis logros. *Mamita Concho* y *papito Carlos*, se que siempre nos estan cuidando desde el cielo, los extraño.

Mom and dad, *Rocio and Alfredo*, you are the reason for where I am today. Thank you for providing me with the opportunity for an education and always supporting me in the difficult times. I know that you have made many sacrifices for me and I hope to return the unconditional love, care, and support that you have given and still give me to this day.

11.3 BIBLIOGRAPHY

Cheung AC*, Lammers WJ*, **Murillo Perez CF**, van Buuren HR, Gulamhusein A, Trivedi PJ, Lazaridis KN, Ponsioen CY, Floreani A, Hirschfield GM, Corpechot C, Mayo MJ, Invernizzi P, Battezzati PM, Parés A, Nevens F, Thorburn D, Mason AL, Carbone M, Kowdley KV, Bruns T, Dalekos GN, Gatselis NK, Verhelst X, Lindor KD, Lleo A, Poupon R, Janssen HLA, Hansen BE. Effects of age and sex of response to ursodeoxycholic acid and transplant-free survival in patients with primary biliary cholangitis. Clin Gastroenterol Hepatol 2019;17:2076-2084.e2.

Goet JC, **Murillo Perez CF**, Harms MH, Floreani A, Cazzagon N, Bruns T, Prechter F, Dalekos GN, Verhelst X, Gatselis NK, Lindor KD, Lammers WJ, Gulamhusein AF, Reig A, Carbone M, Nevens F, Hirschfield GM, van der Meer AJ, van Buuren HR, Hansen BE, Parès A. A comparison of prognostic scores (Mayo, UK-PBC and GLOBE) in primary biliary cholangitis. *Submitted*.

Murillo Perez CF, Goet JC, Lammers WJ, Gulamhusein A, van Buuren HR, Ponsioen CY, Carbone M, Mason A, Corpechot C, Invernizzi P, Mayo MJ, Battezzati PM, Floreani A, Pares A, Nevens F, Kowdley KV, Bruns T, Dalekos GN, Thorburn D, Hirschfield G, LaRusso NF, Lindor KD, Zachou K, Poupon R, Trivedi PJ, Verhelst X, Janssen HLA, Hansen BE. Milder disease stage in patients with primary biliary cholangitis over a 44-year period: A changing natural history. Hepatology 2018;67:1920-1930.

Murillo Perez CF, Gulamhusein A, Carbone M, Trivedi PJ, van der Meer AJ, Corpechot C, Battezzati PM, Lammers WJ, Cazzagon N, Parés A, Nevens F, Lleo A, Mayo MJ, Kowdley KV, Ponsioen CY, Dalekos GN, Thorburn D, Mason AL, Gatselis NK, Verhelst X, Bruns T, Lindor KD, Chazouillères O, Invernizzi P, Janssen HLA, Hansen BE, Hirschfield GM. The simplified GLOBAL PBC Study Group ABA risk assessment tool: reducing complexity to encourage care pathway choices in patients with primary biliary cholangitis. European Journal of Gastroenterology & Hepatology 2020 – *in press.*

Murillo Perez CF, Harms MH, Lindor KD, van Buuren HR, Hirschfield GM, Corpechot C, van der Meer AJ, Feld JJ, Gulamhusein A, Lammers WJ, Ponsioen CY, Carbone M, Mason AL, Mayo MJ, Invernizzi P, Battezzati PM, Floreani A, Lleo A, Nevens F, Kowdley KV, Bruns T, Dalekos GN, Gatselis NK, Thorburn D, Trivedi PJ, Verhelst X, Parés A, Janssen HLA, Hansen BE. 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.

Murillo Perez CF, Hirschfield GM, Corpechot C, Floreani A, Mayo MJ, van der Meer A, Ponsioen CY, Lammers WJ, Parés A, Invernizzi P, Carbone M, Battezzati PM, Nevens F, Kowdley KV, Thorburn D, Mason AL, Trivedi PJ, Lindor KD, Bruns T, Dalekos GN, Gatselis NK, Verhelst X, Janssen HLA, Hansen BE, Gulamhusein A. Fibrosis stage is an independent predictor of outcome in primary biliary cholangitis despite biochemical treatment response. Aliment Pharmacol Ther 2019;50:1127-1136.

**shared co-first authorship*

11.4 CURRICULUM VITAE

Carla Fiorella Murillo Perez was born on March 26th, 1993 in Lima, Peru. In 2002, she moved to Canada with her family to pursue new opportunities. After high school, she studied Biology (Biomedical Science Stream) at York University in Toronto, Canada. During her last year of study, she completed an undergraduate Honours thesis that involved evaluating the effect of oxidative stress on total tumour mass in *Drosophila melanogaster*. After completing her undergraduate degree, she started working at Toronto Centre for Liver Disease and obtained her master's degree from the Institute of Medical Science, University of Toronto under the supervision of prof. dr. H.L.A Janssen and dr. B.E. Hansen. At this time, she became very involved with research in primary biliary cholangitis (PBC) as part of the Global PBC Study Group. To extend her research interests, she pursued her PhD in Rotterdam, the Netherlands under the supervision of prof. H.J. Metselaar, prof. dr H.L.A. Janssen and dr. B.E. Hansen that focused on prognostic factors for clinical outcomes in PBC. Fiorella lives in Toronto, Canada with her family (Alfredo, Rocio, Diego, Toby).

11.5 PhD PORTFOLIO

Summary of PhD training and teaching

Name PhD student:	Carla Fiorella Murillo Perez
PhD period:	2018-2020
Erasmus MC Department:	Gastroenterology and Hepatology
Promotors:	Prof. dr. H.J. Metselaar
	Prof. dr. H.L.A. Janssen
Copromotor:	Dr. B.E. Hansen

1. Research training	Year	Workload (hours/ECT)
General courses		
Biostatistics III: Advanced Biostatistical Techniques for Observational studies	2019	20 hours
Advanced clinical trials	2020	1.9 EC
Scientific Integrity	2020	0.3 EC
Seminars and Workshops		
Hepatobiliary rounds	2018-2020	14 hours
Post graduate course at The Liver Meeting AASLD	2019	6 hours
How to Pitch and design a poster	2020	3 hours
Oral Presentations		
Bilirubin is predictive of transplant-free survival even within the normal range in primary biliary cholangitis. GI research day at University of Toronto. Toronto, ON, Canada.	2018	6 hours
Alkaline phosphatase normalization is associated with a decreased risk for liver transplantation and death in patients with primary biliary cholangitis. 33rd Annual Sheila Sherlock Liver Research Day. Toronto, ON, Canada.	2018	6 hours
Optimising trial design in late-stage primary biliary cholangitis: Evaluating options for composite clinical endpoint studies. 34th Annual Sheila Sherlock Liver Research Day. Toronto, ON, Canada.	2019	6 hours
The impact of geographical region on outcomes of patients with primary biliary cholangitis from Western Europe. The Digital International Liver Congress 2020 (EASL).	2020	35 hours

Poster presentations

Bilirubin is predictive of transplant-free survival even with the normal range in patients with primary biliary cholangitis. Canadian Liver Meeting 2018 (CASL). Toronto, ON, Canada.	2018	32 hours
Younger age is associated with lower transplant-free survival relative to a general population in patients with primary biliary cholangitis. Canadian Liver Meeting 2018 (CASL). Toronto, ON, Canada.	2018	32 hours
Younger age is associated with lower transplant-free survival relative to a general population in patients with primary biliary cholangitis. The International Liver Congress 2018 (EASL). Paris, France.	2018	32 hours
Alkaline phosphatase normalization is associated with a decreased risk for liver transplantation and death in patients with primary biliary cholangitis. The Liver Meeting 2018 (AASLD). San Francisco, United States of America.	2018	32 hours
Treatment of patients with primary biliary cholangitis with seladelpar for 52 weeks improves predicted transplant-free survival. The International Liver Congress 2019 (EASL). Vienna, Austria.	2019	32 hours
Raising awareness and messaging risk in patients with primary biliary cholangitis: The rapid Global PBC Screening Test. The International Liver Congress 2019 (EASL). Vienna, Austria.	2019	32 hours
Treatment of patients with primary biliary cholangitis with seladelpar for 52 weeks improves predicted transplant-free survival. The Canadian Liver Meeting (CASL) 2019. Montreal, QC, Canada.	2019	6 hours
Raising awareness and messaging risk in patients with primary biliary cholangitis: The rapid Global PBC Screening Test. The Canadian Liver Meeting (CASL) 2019. Montreal, QC, Canada.	2019	6 hours
Optimising trial design in late-stage primary biliary cholangitis: Evaluating options for composite clinical endpoint studies. The Liver Meeting 2019 (AASLD). Boston, MA, United States of America.	2019	32 hours

Optimising trial design in late-stage primary biliary cholangitis: Evaluating options for composite clinical endpoint studies. The Canadian Liver Meeting 2020 (CASL). Montreal, QC, Canada.	2020	6 hours

National and international conferences

International HBV Cure: HBV workshop 2018	2018	7 hours
The Canadian Liver Meeting (CASL), Toronto, ON, Canada	2018	17 hours
The International Liver Congress 2018, 53rd Annual Meeting of the European Association for the Study of the Liver (EASL). Paris, France.	2018	28 hours
The Liver Meeting 2018, 69th Annual Meeting of the American Association for the study of Liver Diseases (AASLD). San Francisco, California, United States of America.	2018	28 hours
The International Liver Congress 2019, 54th Annual Meeting of the European Association for the Study of the Liver (EASL). Vienna, Austria.	2019	28 hours
The Liver Meeting 2019, 70th Annual Meeting of the American Association for the study of Liver Diseases (AASLD). Boston, MA, United States of America.	2019	28 hours
Single Topic Conference. Ottawa, Ontario, Canada.	2019	11 hours
The Canadian Liver Meeting 2020. Montreal, Canada.	2020	20 hours
The Digital International Liver Congress 2020, 55th Annual Meeting of the European Association for the Study of the Liver (EASL).	2020	28 hours
Dr. Morris Sherman Farewell Symposium	2020	7 hours

Awards and Recognitions

Abstract selected for inclusion in Best of cholestatic and autoimmune liver disease at ILC2017	2017
EASL ILC 2018 Young investigator registration bursary	2018
Poster of distinction at Canadian Liver Meeting 2018 (CASL): Bilirubin is predictive of transplant-free survival even with the normal range in patients with primary biliary cholangitis.	2018

Best clinical research presentation at 33rd Annual Sheila Sherlock Liver Research Day: Alkaline phosphatase normalization is associated with a decreased risk for liver transplantation and death in patients with primary biliary cholangitis.	2018	
Poster of distinction at Canadian Liver Meeting 2019 (CASL): Raising awareness and messaging risk in patients with primary biliary cholangitis: The rapid Global PBC Screening Test.	2019	
Full bursary to International Liver Congress 2019- EASL	2020	
Reviewer for scientific journals		
Including American Journal of Gastroenterology, Clinical Gastroenterology and Hepatology, Liver International		11 hours
Other		
Global PBC Study Group: Investigator meetings, logistics, writing protocol, design of CRF, ethics approval	2018-2020	120 hours
Cymabay Therapeutics: Design and protocol of clinical trial studies	2019	81 hours
Canadian PBC Society: Summary of PBC research at ILC 2019	2019	6 hours

2. Teaching	**Year**	**Workload (Hours/ECT)**
Lecturing		
HepCNet journal club presentations	2018-2020	14 hours
Graduate student Journal club participation	2018-2020	31 hours
Supervision of graduate students	2018-2020	
Supervision of master's students	2018-2020	60 hours