Surrogate markers of bile duct disease progression in primary sclerosing cholangitis – A prospective study with repeated ERCP examinations

Authors
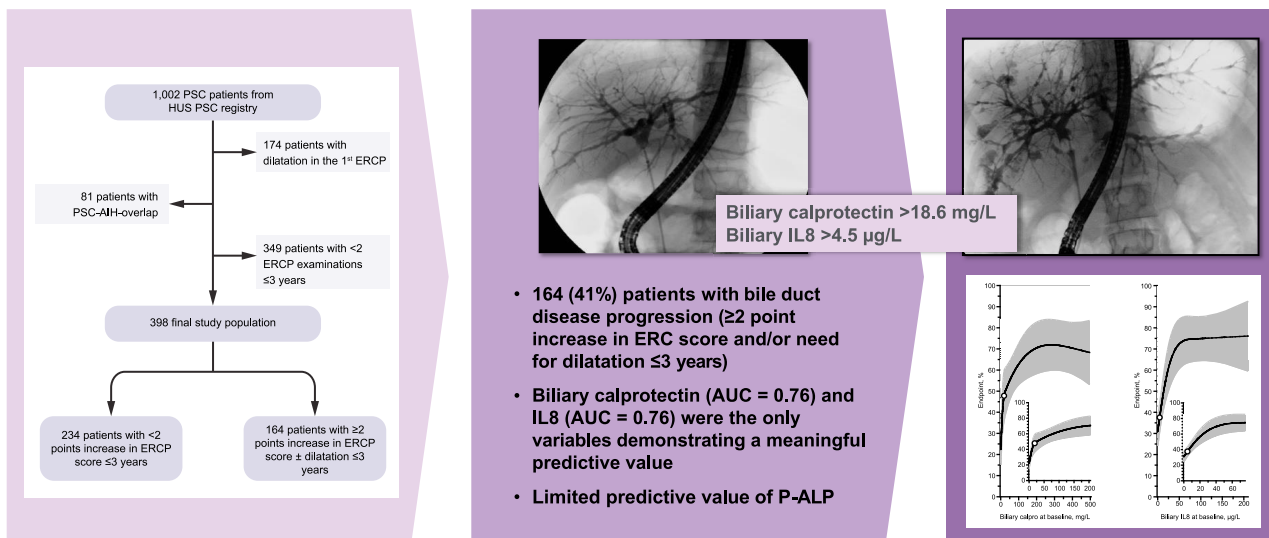
Martti Färkkilä, Fredrik Åberg, Henrik Alfthan, …, Lauri Puustinen, Hannu Kautiainen, Andrea Tenca

Correspondence

martti.farkkila@hus.fi (M. Färkkilä).

Graphical abstract

The most sensitive markers for bile duct disease progression were calprotectin and interleukin 8 in bile.

Biliary calprotectin >18.6 mg/L
Biliary IL8 >4.5 µg/L

- 164 (41%) patients with bile duct disease progression (≥2 point increase in ERC score and/or need for dilatation ≤3 years)
- Biliary calprotectin (AUC = 0.76) and IL8 (AUC = 0.76) were the only variables demonstrating a meaningful predictive value
- Limited predictive value of P-ALP

Highlights:

- At the 3-year follow-up, 41% of patients demonstrated bile duct disease progression.

- The most sensitive markers for progression were biliary calprotectin and biliary interleukin-8.

- The value of conventionally used serum liver function tests was limited in assessing bile duct disease progression.

- This study could improve the management of PSC by identifying reliable surrogate markers for short-term disease progression.

Impact and implications:

Validated prognostic tools for estimating short-term bile duct disease progression in primary sclerosing cholangitis are lacking. In this prospective study, based on sequential endoscopic retrograde cholangiopancreatography examinations, biliary calprotectin and IL8 levels turned out to be more sensitive for predicting bile duct progression than traditional liver function tests, such as alkaline phosphatase, in the short term. These findings could lead to more personalized patient surveillance and improve clinical practice by providing a more accurate method for monitoring disease progression and treatment responses. Additionally, these markers have potential as surrogate endpoints in clinical drug trials. The limitation is that measurement of biliary IL8 and calprotectin requires endoscopic retrograde cholangiopancreatography with bile sampling.

Surrogate markers of bile duct disease progression in primary sclerosing cholangitis – A prospective study with repeated ERCP examinations

Martti Färkkilä[1,4,*], Fredrik Åberg[2], Henrik Alfthan[3], Kalle Jokelainen[4], Lauri Puustinen[4], Hannu Kautiainen[5], Andrea Tenca[4]

Check for updates

Background & Aims: Validated prognostic tools for estimating short-term bile duct disease progression in primary sclerosing cholangitis (PSC) are lacking. We evaluated the predictive value of serum and biliary biochemistry for the progression of bile duct disease in PSC using repeated endoscopic retrograde cholangiopancreatography (ERCP) examinations to identify surrogate markers for more personalized surveillance.

Methods: We conducted a prospective analysis including patients with PSC who underwent ERCP for confirmation of diagnosis, monitoring of disease progression, or dysplasia surveillance. ERCP findings were scored, and dilatation was performed if a dominant stricture was diagnosed or if a cytology brush could not be passed. Bile samples were aspirated for biliary IL8 and calprotectin. We analysed optimal cut-off values and AUCs for 20 laboratory markers and evaluated their association with the time to an ERCP score increase of ≥2 points or first dilatation, whichever came first. Of the 1,002 patients, 653 had ≥2 ERCP examinations and ≥3 years of follow-up. After excluding patients with PSC-overlap syndrome or initial dilatation, 398 patients were included.

Results: Of the patients included, 62% had mild or moderate and 38% had advanced bile duct disease. During follow-up, 41% of patients demonstrated progression of disease. Biliary calprotectin (AUC 0.76; 95% CI 0.69 to 0.82) and IL8 (AUC 0.76; 95% CI 0.69 to 0.84) were the only variables that demonstrated predictive value for disease progression and/or need for dilatation.

Conclusions: Biliary calprotectin and IL8 are promising surrogate markers for identifying patients with PSC at risk of progression and determining the timing for subsequent imaging. Conventional liver function tests may not be sensitive or specific enough to monitor PSC progression, particularly in the short term.

Introduction

Primary sclerosing cholangitis (PSC) is a chronic inflammatory disease of the biliary epithelium that leads to strictures of the intra- and extrahepatic bile ducts and eventually to secondary biliary cirrhosis.[1] Chronic inflammation is associated with increased proliferation of biliary epithelial cells, development of bile duct strictures, and a markedly increased risk of biliary dysplasia and cholangiocarcinoma, with relative risks ranging from 161- to 973-fold.[2,3] The ESGE/EASL recommends[4] magnetic resonance cholangiography (MRCP) over endoscopic retrograde cholangiopancreatography (ERCP) as the primary diagnostic modality for PSC. However, MRCP has a low sensitivity for identifying early changes in intrahepatic PSC.[5,6] Earlier studies on the role of imaging in predicting PSC prognosis have been based on late endpoints such as liver transplantation and death,[7] rather than disease progression in the short term, for more personalized clinical management and in clinical trials. MRCP has been studied to evaluate predictive features of bile duct disease progression based on liver

dysmorphology and signs of portal hypertension, which are all late signs.[8] The use of the MR risk score in combination with transient elastography and survival without transplantation or decompensation as the primary endpoint has been shown to identify risk of developing adverse outcomes.[9] The recently published DISTRICT score evaluating intrahepatic and extrahepatic ductal changes in three-dimensional MRCP images is also based on late outcomes, transplantation, or liver-related death.[10]

Several previous studies have evaluated serum biomarkers,[11] e.g., decreases in plasma (P-) alkaline phosphatase (ALP),[12] interleukin-8 (IL8),[13] and ELF (enhanced liver fibrosis) test,[14] in addition to combined indices, e.g., the UK-PSC score,[15] Amsterdam-Oxford PSC model (AOM),[16] and Pres-To,[17] using late endpoints such as hepatic decompensation, cirrhosis, transplantation-free survival or death. The recent EASL clinical practice guidelines on PSC recommended risk assessment at the time of diagnosis and sequentially based on phenotypic factors and non-invasive tests, including standard

biochemical tests; MRI with MRCP; and liver elastography or serum fibrosis tests.[18] Liver biopsy, commonly used as a primary endpoint in clinical trials for drug development,[11,19] is not optimal due to the patchy nature of histological findings in PSC. At present, there are no validated prognostic tools for estimating short-term indicators of bile duct disease progression and the development of new strictures in individual patients with PSC.

We aimed to identify surrogate markers for short-term PSC progression for more personalized patient surveillance. We evaluated the ability of clinical, serum and biliary biochemistry parameters and bile duct inflammation to predict the progression of bile duct disease based on changes in scores on sequential ERCP examinations at the 3-year follow-up.

Patients and methods

Clinical data and ERCP examinations

Patients referred for ERCP examination with suspicion or documented PSC were included in this prospective registry study. The diagnosis of PSC was made according to the EASL clinical practice guidelines.[18] The indications for ERCP were to 1) confirm the diagnosis, 2) follow-up disease progression, and 3) for the surveillance of dysplasia. ERCP was performed by four experienced gastroenterologists (MF, KJ, AT, LP) using the balloon occlusion technique to ensure adequate and standardized filling of intra- and extrahepatic bile ducts in successive ERCP procedures. Images were evaluated using the Helsinki score (modified Amsterdam score);[20] see Fig. 1. Dominant stricture (DS) was defined as a stenosis with a diameter ≤1.5 mm in the common duct or ≤1.0 mm in the hepatic duct within 2 cm of the bifurcation.[21] Dilatation was performed when DS was diagnosed or when the cytology brush could not be passed through the stenosis. Disease progression was defined as the time to increase to an ERCP score ≥2 and/or

to the time to dilatation from the first ERCP, whichever came first in the 3-year follow-up.

In total, 1,002 patients were included, 653 of whom had ≥2 ERCP examinations performed within a follow-up of ≤3 years after the first ERCP. After excluding patients with PSC-autoimmune hepatitis (AIH) overlap syndrome and those with dilatation at their first ERCP, 398 patients were identified (Fig. 2). The numbers of ERCP procedures and dilatations performed are presented in Fig. 3.

Brush cytology

Brush cytology (BC) was collected systematically regardless of possible DS to grade inflammation and detect biliary dysplasia using a brush with a guide wire (RX Cytology Brush, Boston Scientific, MA, USA). The brush tip was cut, and the brush and the fluid from the brush catheter shaft were flushed with 50% ethanol into a vial containing 50% ethanol. The sample was analysed as described previously.[2] Neutrophilic inflammation in BC was evaluated semiquantitatively (0 = neutrophils/epithelial cells <0.05, 1 = neutrophils/epithelial cells 0.05-0.4, 2 = neutrophils/epithelial cells >0.4) and likewise for lymphocytes.

Plasma and bile acid samples

Plasma and serum clinical chemistry parameters were obtained from the Helsinki University Central Laboratory (HUSLab). Bile samples were aspirated using a balloon occlusion catheter and stored at -80 °C.

Biliary calprotectin was quantitated with an immunometric assay from CALPRO AB, Lysaker, Norway (CalproLab™ ELISA (ALP)). Before analysis, each sample was diluted threefold in sample dilution buffer (1:10, 1:300 and 1:9,000). The standard curve covers the calprotectin concentration range of 7.8–500 µg/L. The detection limit of the assay was 5 µg/L. The intra-assay coefficient of variation (C.V.) was <6%, and the

| Amsterdam score[28]	Helsinki score[20]	Intrahepatic
0	0	No visible abnormalities
I	1	Ductular irregularities
	2	Multiple caliber change; minimal dilatation
II	3	Multiple strictures; saccular dilatations, decreased arborization
III	4	Only central branches filled despite adequate filling pressure; severe pruning
		Extrahepatic
0	0	No visible abnormalities
I	1	Slight irregularities of duct contour, no stricture
II	2	Segmental strictures
III	3	Strictures of almost entire length of duct
IV	4	Extremely irregular margins: diverticulum like outpouchings
Total	16	

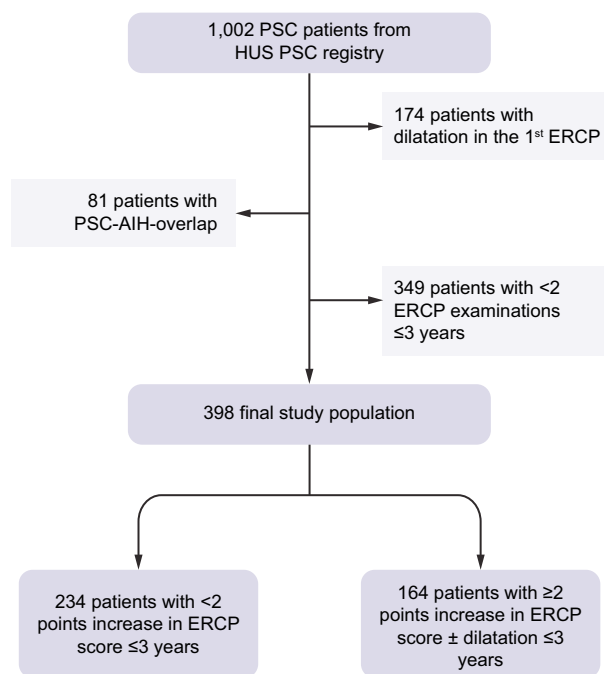Helsinki score: 3+2+1+2=8
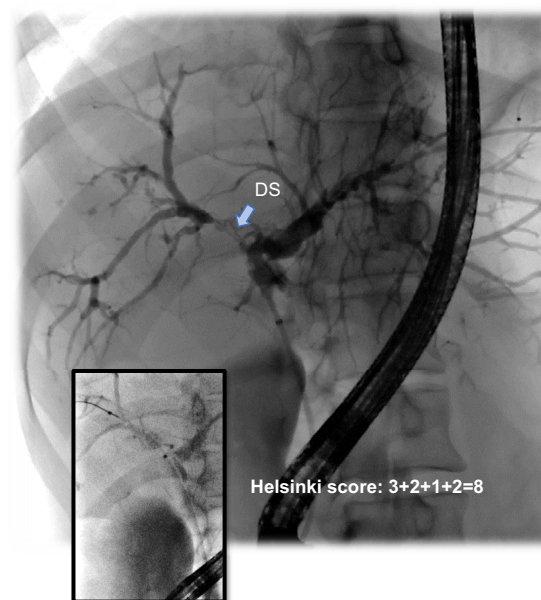
Fig. 1. Scoring endoscopic retrograde cholangiopancreatography findings: Comparison of Amsterdam and Helsinki scoring systems.

Fig. 2. Derivation of the study cohort. ERCP, endoscopic retrograde cholangiopancreatography; PSC, primary sclerosing cholangitis.

interassay C.V. was <8% in the concentration range 17–395 µg/L.

Biliary IL8 was quantitated with an immunometric assay from R&D Systems® (Quantikine® ELISA, Human IL8/CXCL8).

Before analysis, each sample was diluted threefold in RD5P kit buffer (1:10, 1:100 and 1:1,000).

The standard curve covers the IL8 concentration range of 31–2,000 ng/L. The detection limit of the assay was 3.5 ng/L. The intra-assay C.V. was <8%, and the interassay C.V. was <10% in the concentration range 105–1,090 µg/L. For both assays, the diluted sample with the closest fit to the middle of the standard curve was used to calculate the quantitative result.

Statistics

Summary statistics are presented as mean (SD), median (IQR), or count as a percentage. The unadjusted hypothesis of linearity was tested using the Cochran–Armitage test, analysis of variance or Cuzick test depending on the distribution of the outcome. Unadjusted differences between the groups were evaluated using the t test, Mann–Whitney test, or chi-square test. When adjusting for disease duration between the groups, the non-parametric Koch's test was used. Measurement ratios between groups were evaluated using linear regression, logistic regression, or median regression (least-absolute-value) models. The possible non-linear relationships between the endpoint and the ERCP score, and laboratory values were assessed by using 4-knot restricted cubic spline logistic regression models. The lengths of the distribution of knots were located at the 5th, 35th, 65th and 95th percentiles. Models for endpoints included sex and disease duration as covariates. Receiver-operating characteristic curves were constructed to determine the cut-off point with a bias-corrected accelerated bootstrap confidence interval (10,000 repetitions). AUC values ranging from 0.7 to 0.8 indicated
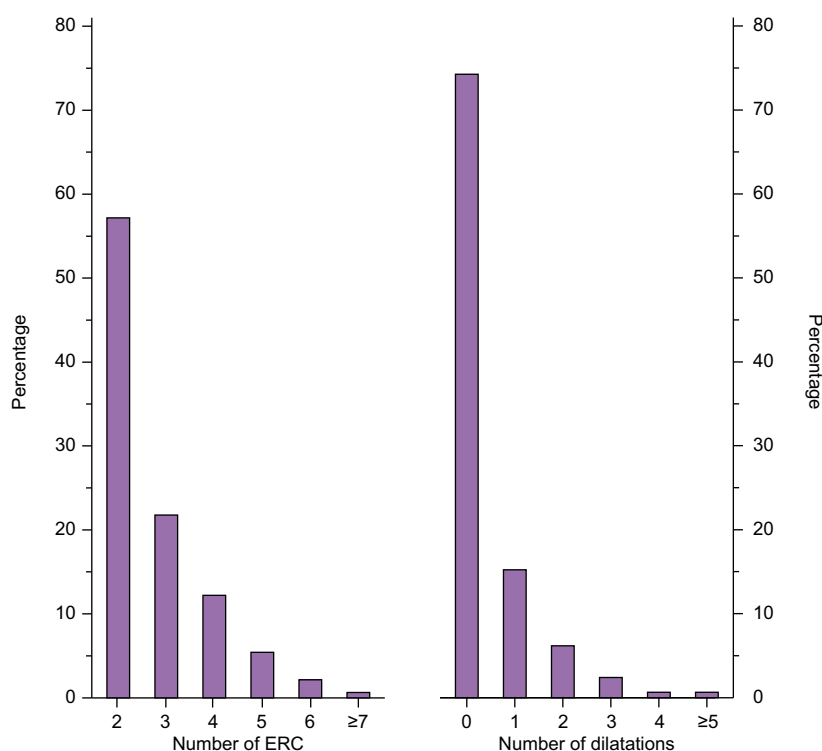
Fig. 3. Distribution of percentages of ERCP examinations and dilatations performed during the follow-up period of 3 years. ERCP, endoscopic retrograde cholangiopancreatography.

reasonable discrimination, and values exceeding 0.8 indicated good discrimination. We defined the best cut-off value as the value at which the Liu method maximized the product of sensitivity and specificity. Diagnostic accuracy statistics and 95% CI values were calculated. Differences between the AUCs were evaluated using the DeLong algorithm. The STATA 18.0 and StataCorp LP (College Station, TX, USA) statistical packages were used for the analyses. The Mayo risk score,[22] AOM[16] and FIB4 (fibrosis-4 index)[23] were calculated as previously described.

Ethics

All patients included in the PSC registry provided written informed consent. The study was performed following the principles of the GCP and in accordance with the ethical guidelines of the Declaration of Helsinki (6th revision, 2008). The study protocol was approved by the Helsinki University Hospital Ethical Committee IV, HUS/1566/2020.

Results

Demographic and laboratory data

Clinical characteristics and laboratory results according to bile duct disease severity at baseline are presented in Table S1. Of the whole study cohort (N = 398), 115 patients had mild (Helsinki ERCP score 2-3), 130 had moderate (score 4-6), and 153 had advanced bile duct disease (score ≥7). The majority of patients (73-81%) used ursodeoxycholic acid, with no differences between groups. Plasma aminotransferases, gamma-glutamyltransferase and bilirubin differed markedly across patients with mild, moderate, or advanced disease, but no difference was detected in P-ALP levels. However, the markers of biliary inflammation: biliary (Bi)-calprotectin and Bi-IL8 levels, and neutrophil and lymphocyte counts were significantly linear across patients with mild to advanced bile duct disease. In total, 164 (41%) patients demonstrated progression based on an increase in the ERCP score of ≥2 points and/or the need for dilatation of at least one stricture. Table 1 shows the baseline characteristics and laboratory results of non-progressors (n = 234) and progressors (n = 164). Patients who reached the endpoint presented significantly more frequently with both intra- and extrahepatic disease and higher baseline ERCP scores (p <0.001). P-ALP, the Mayo score, FIB4 or AOM did not differ between the groups. The median ratios of progression to non-progression were 11.92 (4.73 to 30.93) for Bi-calprotectin and 7.10 (3.44 to 12.86) for Bi-IL8.

Fig. 4 shows the proportion of patients with bile duct disease progression based on the baseline ERCP score. The most marked difference between progressors and non-progressors was observed in mild or moderate disease. We combined these two groups to further study the factors associated with progression.

The clinical characteristics and laboratory results of patients with and without disease reaching the endpoint at the end of follow-up are shown in Table 2. The most marked changes were observed in Bi-calprotectin and Bi-IL8 levels.

Optimal cut-off values predicting endpoint

We then analysed the optimal cut-off values, AUCs, positive likelihood ratios and accuracies of 20 different laboratory variables and their relationships with the endpoint in the whole study population and separately in groups with an ERCP score <7 or ≥7 at baseline; see Table S2.

Of the variables included in the analysis, only Bi-calprotectin (AUC 0.76 [95% 0.69 to 0.84]), and Bi-IL8 (AUC 0.76 [95% 0.69 to 0.84]) were associated with disease progression. The odds ratio was 7.24 for Bi-calprotectin (3.89 to 13.49) and 10.24 for Bi-IL8 (5.12 to 20.48). The accuracies were 74% (68 to 80) and 77% (71 to 83). P-ALP, gamma-glutamyltransferase or -aminotransferase levels were not significantly associated with progression, with AUCs varying from 0.55-0.62. For mild disease only (score 2-3), the cut-off values for Bi-calprotectin were ≥6.5 mg/L and AUC 0.80 (0.68 to 0.91), and for Bi-IL8 ≥0.40 µg/L and AUC 0.77 (0.65 to 0.88), respectively.

We then calculated the relationships of Bi-calprotectin, Bi-IL8 and P-ALP with the probability of reaching the endpoint (Fig. 5) and found that there was a close relationship with baseline Bi-calprotectin up to a level of 200 mg/L and Bi-IL8 up to a level of 75 µg/L but not with P-ALP. In fact, the relationship with P-ALP was like a U-shaped curve, showing an inverse relationship up to the inflection point of 140 U/L (upper limit of normal [ULN] <105 U/L); thereafter, there was a linear relationship with the probability of reaching the endpoint. At higher levels, both Bi-calprotectin (n = 35, 8.7%) and Bi-IL8 (n = 24, 2.4%) lost their relationship with the endpoint.

Discussion

The disease course of PSC is variable and unpredictable.[24] At present, there are scarce data on markers for predicting disease activity and the development of new strictures in clinical practice. To our knowledge, this is the first study to evaluate a wide variety of clinical and laboratory parameters for predicting bile duct disease progression in the short term in a large cohort of patients with PSC. In this study, including sequential ERCP examinations, the most accurate markers were Bi-calprotectin and Bi-IL8 in patients with mild or moderate disease (ERCP score <7). In contrast, conventionally used liver function tests, and demographic data[18] have very limited or no value in predicting disease progression. In a study by Trivedi et al.,[25] the authors reported large interindividual and intraindividual variations in ALP activity, and ALP was not associated with disease progression assessed by the development of cirrhosis or an increase in fibrosis over a 2-year period. The authors concluded that ALP imparted inconsistent prognostic utility longitudinally, which limits its use as a surrogate endpoint. In a study evaluating the prognostic value of ALP at the time of diagnosis of PSC and 1 year later,[12] the authors found that the hazard of reaching an endpoint approached a plateau at an ALP >2.0x ULN when using a composite endpoint of transplantation and PSC-related death. However, only a few patients with an ALP >2.0x ULN were included in their cohort. In the present study, we found a U-shaped association between P-ALP and disease progression and P-ALP >140 U/L was associated with reaching the endpoint.

Earlier studies have demonstrated the prognostic role of cholangiographic abnormalities in ERCP using either death or liver transplantation[26] or time to death or liver transplantation and the first appearance of jaundice as endpoints.[27] A Dutch study[28] demonstrated that cholangiographic scoring was inversely associated with survival. However, none of the earlier

Table 1. The clinical characteristics and laboratory results of progressors and non-progressors at baseline.

Variables	ΔERCP score ≥2 or dilatation in 3 years		p value	Ratio between progressors/non-progressors (95% CI)*
	Non-progressors, n = 234	Progressors, n = 164		
Male	151 (65)	105 (64)	0.92	0.99 (0.85 t o1.15)
Age at PSC diagnosis, mean (SD), years	37 (13)	38 (15)	0.44	1.03 (0.95–1.10)
Intra- and extrahepatic PSC	90 (38)	103 (63)	<0.001	1.63 (1.34–2.00)
IBD diagnosed before PSC	129 (55)	86 (52)	0.60	0.95 (0.79–1.15)
Colectomy before diagnosis of PSC	25 (11)	23 (14)	0.31	1.31 (0.77–2.23)
UDCA use during follow-up	175 (75)	136 (83)	0.053	1.11 (1.00–1.23)
ERCP score	4 (2–8)	6 (4–9)	<0.001	1.50 (1.12–1.88)
Mayo score	0.87 (0.39–1.34)	0.91 (0.46–1.58)	0.13	1.05 (0.86–1.29)
FIB4	0.98 (0.68–1.57)	1.01 (0.63–1.73)	0.99	1.03 (0.89–1.19)
AOM	1.54 (1.22–1.91)	1.46 (1.20–2.01)	0.64	0.95 (0.88–1.08)
B-Hb, g/L	140 (128–149)	136 (124–145)	0.061 [n.s.]	0.97 (0.94–1.01)
B-platelets, 10⁹/L	256 (210–313)	277 (208–336)	0.11 [n.s.]	1.08 (0.89–1.14)
P-AST, U/L	35 (27–52)	41 (28–68)	0.025 [n.s.]	1.17 (1.08–1.41)
P-ALT, U/L	43 (24–74)	48 (25–104)	0.037 [n.s.]	1.11 (1.01–1.45)
P-ALP, U/L	137 (99–212)	150 (98–312)	0.10 [n.s.]	1.09 (0.91–1.31)
P-GGT, U/L	98 (33–247)	110 (52–318)	0.048 [n.s.]	1.12 (0.82–1.72)
P-Bil, μmol/L	11.0 (8.0–16.0)	11.0 (7.0–18.0)	0.75 [n.s.]	1.00 (0.99–1.20)
P-Alb, g/L	38.2 (35.5–41.0)	37.3 (34.5–40.8)	0.045 [n.s.]	0.97 (0.95–1.02)
P-TT, %	102 (87–117)	102 (88–119)	0.86 [n.s.]	1.00 (0.96–1.07)
P-IgG, g/L	12.2 (10.5–14.4)	12.5 (10.7–15.1)	0.25 [n.s.]	1.02 (0.96–1.09)
S-IgG4, g/L	0.62 (0.28–1.14)	0.69 (0.33–1.21)	0.24 [n.s.]	1.11 (0.90–1.45)
S-IL8, pg/L	38.2 (17.7–105.6)	56.4 (23.4–193.4)	0.054 [n.s.]	1.47 (0.87–2.79)
S-CEA, μg/L	1.30 (1.00–2.20)	1.40 (1.00–2.05)	0.36 [n.s.]	1.08 (0.93–1.25)
S-CA19-9, IU/L	5.0 (3.0–9.0)	6.0 (3.0–13.5)	0.052 [n.s.]	1.20 (0.83–1.50)
S-cANCIF positivity	15 (8)	9 (6)	0.45 [n.s.]	0.73 (0.33–1.63)
S-pANCIF positivity	93 (50)	79 (52)	0.72 [n.s.]	1.04 (0.84–1.28)
Biliary calprotectin, mg/L	2.6 (0.3–16.0)	30.9 (4.0–155.2)	<0.001 [<0.001]	11.92 (4.73–30.93)
Biliary IL8, μg/L	1.1 (0.1–5.7)	7.6 (1.0–32.8)	<0.001 [<0.001]	7.10 (3.44–12.86)
BC-neutrophils ≥1	117 (50)	125 (76)	<0.001 [<0.001]	1.52 (1.30–1.77)
BC-lymphocytes ≥1	135 (58)	122 (74)	<0.001 [0.014]	1.28 (1.11–1.48)
BC-plasma cells+	8 (3)	13 (8)	0.050 [n.s.]	2.30 (0.98–5.42)
BC- IEL+	113 (48)	118 (72)	<0.001 [0.018]	1.48 (1.26–1.754)

Alb, albumin; ALP, alkaline phosphatase; ALT, alanine aminotransferase; AOM, Amsterdam-Oxford PSC model; AST, aspartate aminotransferase; B-, blood; BC, brush cytology; Bil, bilirubin; cANCIF, anti-neutrophil cytoplasmic antigen, immunofluorescence antibodies; ERCP, endoscopic retrograde cholangiopancreatography; FIB4, fibrosis-4 index; GGT, gamma-glutamyltransferase; IBD, inflammatory bowel disease; IEL, intraepithelial lymphocytes; IL8, interleukin-8; P-, plasma; pANCIF, anti-neutrophil perinuclear antigen, immunofluorescence antibodies; S-, serum; TT%, thrombotest; UDCA, ursodeoxycholic acid.
Data are presented as n (%) or median (IQR) unless otherwise stated.
In square brackets, p values were adjusted for multiplicity using Hochberg's step-up procedure for p value adjustment in the laboratory values.
Unadjusted differences between the groups were evaluated using the t test, Mann–Whitney test, or chi-square test. When adjusting for disease duration between the groups, the non-parametric Koch's test was used.
*Ratio between the progressors and non-progressors regarding measurements (mean, median or percentage).

studies analysed the factors associated with changes in cholangiographic scores. The scheduled ERCP procedure with endoscopic balloon dilatation in patients with PSC and dominant strictures demonstrated that transplantation-free survival was greater in patients receiving scheduled ERCP than in those receiving ERCP only on demand, even in asymptomatic patients.[29] The Hannover score is so far the only prognostic index accounting for ERCP findings using the time of diagnosis to death or transplantation as the endpoint.[30]

Most published studies analysing the role of different biomarkers, or their combinations have also used very late endpoints, such as cirrhosis, decompensation, transplantation-free survival, cholangiocarcinoma, or liver-related deaths. In the present study, prognostic scores such as the Mayo score or AOM did not identify patients with progressive or nonprogressive bile duct disease. Several new promising biomarkers have been identified for predicting PSC outcome.[31] Recently, anti-glycoprotein 2 IgA and anti-neutrophil cytoplasmic antibodies to serine proteinase 3 have been shown to predict more severe disease, poorer survival, and cholangiocarcinoma.[32] In this study, anti-neutrophil cytoplasmic antibody positivity (45 to 55%), was not associated with disease progression.

PSC is a chronic inflammatory cholangiopathy characterized by an activated phenotype of the biliary epithelium with expansion of the peribiliary gland system, leading to the development of peribiliary inflammation, concentric periductal fibrosis and strictures.[19,33] In our previous study, we demonstrated that the grade of inflammation in liver biopsies correlated with increasing ERCP load.[34] IL8 is a chemokine produced by cell subsets, such as macrophages and epithelial cells, that possess Toll-like receptors.[35] IL8 induces chemotaxis in primarily neutrophils and leads to the activation and release of a wide variety of substances, including calprotectin.[13] Lipopolysaccharide treatment of normal human cholangiocytes leads to significantly increased mRNA expression of IL8,[36] suggesting a possible link to the leaky gut hypothesis.[13] Patients with elevated biliary calprotectin concentrations above a cut-off of 11.6 μg/ml had significantly shorter transplantation-free survival than those with lower concentrations.[37]

Analysis of Bi-IL8 levels at different stages of PSC demonstrated that the median IL8 concentration in ductal bile was markedly elevated compared to that in controls, suggesting that an ongoing inflammatory stimulus drives IL8 production.
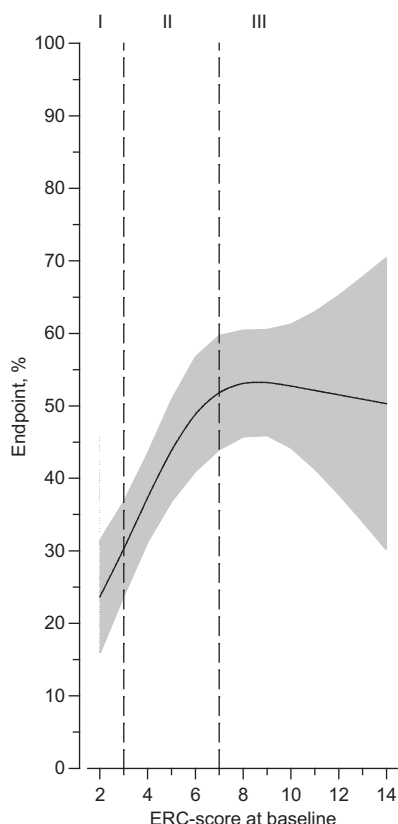
Fig. 4. Percentage of patients reaching the endpoint (increase in ERCP score ≥2 or dilatation) based on the baseline ERCP score. The curve was derived from a 4-knot-restricted cubic splines logistic regression. The models were adjusted for sex and disease duration. The grey area represents 95% CIs. ERCP, endoscopic retrograde cholangiopancreatography.

Biliary S100A8 has been shown to be the single marker that best distinguishes patients with mild from advanced PSC.[13]

However, samples for Bi-IL8 and Bi-calprotectin, require ERCP, which is associated with known risks.[4] In our recent series, also including mild cases, the overall risk for pancreatitis was 5.7%.[38] According to the recent EASL guidelines,[18] therapeutic endoscopic intervention is recommended for patients with relevant strictures, defined as high-grade strictures in extrahepatic bile ducts and signs or symptoms of cholestasis and/or bacterial cholangitis. To date, the definition of relevant strictures has not been validated in clinical studies. In a study evaluating follow-up strategies in patients with PSC, the hazard ratio (95% CI) for death was 0.64 (0.48–0.86) for ultrasound/MRI and as low as 0.53 (0.37–0.75) including scheduled ERCP.[39] In an unselected cohort of patients with PSC, yearly carbohydrate antigen 19-9 and MRI/MRCP surveillance followed by ERCP, as recommended in the EASL guidelines, was ineffective at detecting cancer early enough to support long-term survival.[40] In patients with PSC wait-listed for transplantation, ERCP with BC was also found to be the most effective tool for correctly ruling out cholangiocarcinoma, with a specificity of 96%.[41]

In our unit, all patients with suspected PSC undergo ERCP to confirm the diagnosis due to the low sensitivity of MRCP for detecting early intrahepatic changes and even advanced extrahepatic lesions.[5] In addition, we use ERCP with BC and bile samples for evaluation of need for endoscopic therapy, individual risk stratification for progression and for exclusion of biliary neoplasia.

Patients with PSC, with elevated Bi-calprotectin and/or Bi-IL8 levels are at increased risk for disease progression. The measurement of these markers can be used to design more individualized surveillance and to schedule future controls. On the other hand, patients with low levels of calprotectin and IL8

Table 2. The ERCP score and laboratory results of progressors and non-progressors at the end of follow-up.

	ΔERCP score ≥2 or dilatation in 3 years		
Variables	No, n = 234	Yes, n = 164	*p* value*
ERCP score	6 (2–8)	10 (8–12)	<0.001
Mayo score	-0.51 (-0.92–0.12)	-0.31 (-0.81–0.40)	0.24
FIB4	1.06 (0.75–1.80)	1.04 (0.69–1.79)	0.69
AOM	1.54 (1.19–1.94)	1.58 (1.20–2.06)	0.33
B-Hb, g/L	142 (129–152)	137 (125–151)	0.25
B-platelets, 10⁹/L	256 (210– 313)	277 (208– 336)	0.66
P-AST, U/L	30 (24–43)	35 (26–57)	0.081
P-ALT, U/L	32 (21–56)	43 (22–73)	0.13
P-ALP, U/L	101 (79–145)	125 (83–210)	0.025
P-GT, U/L	48 (21–131)	77 (32–215)	0.003
P-Bil, μmol/L	12 (8–19)	11 (8–19)	0.32
P-Alb, g/L	39.0 (36.0–41.0)	37.5 (34.8–40.2)	0.069
P-TT, %	102 (85–117)	101 (85–122)	0.92
P-IgG, g/L	11.9 (10.3–13.8)	12.0 (10.3–14.5)	0.81
S-IgG4, g/L	0.54 (0.25–1.05)	0.68 (0.34–1.24)	0.24
S-IL8, pg/L	39.7 (21.5–102.0)	53.1 (23.4–193.4)	0.19
S-CEA, μg/L	1.10 (1.00–2.00)	1.30 (1.00–2.00)	0.087
S-CA19-9, IU/L	6.0 (2.0–12.0)	6.0 (3.0–13.5)	0.63
Biliary calprotectin, mg/L	1.3 (0.2–11.8)	18.7 (1.3–107.5)	<0.001
Biliary IL8, μg/L	0.94 (0.08–5.74)	6.42 (0.73–21.98)	<0.001
Dysplasia, n (%)	11 (5)	33 (20)	<0.001

Data are presented as median (IQR) unless otherwise stated.
**p* values between groups adjusted for baseline value.
Median score and serum levels (bars represent interquartile ratio, IQR or count as percentage). Unadjusted differences between the groups were evaluated using the t test, Mann–Whitney test, or chi-square test. When adjusting for disease duration between the groups, the non-parametric Koch's test was used.
For abbreviations, see Table 1 legend.
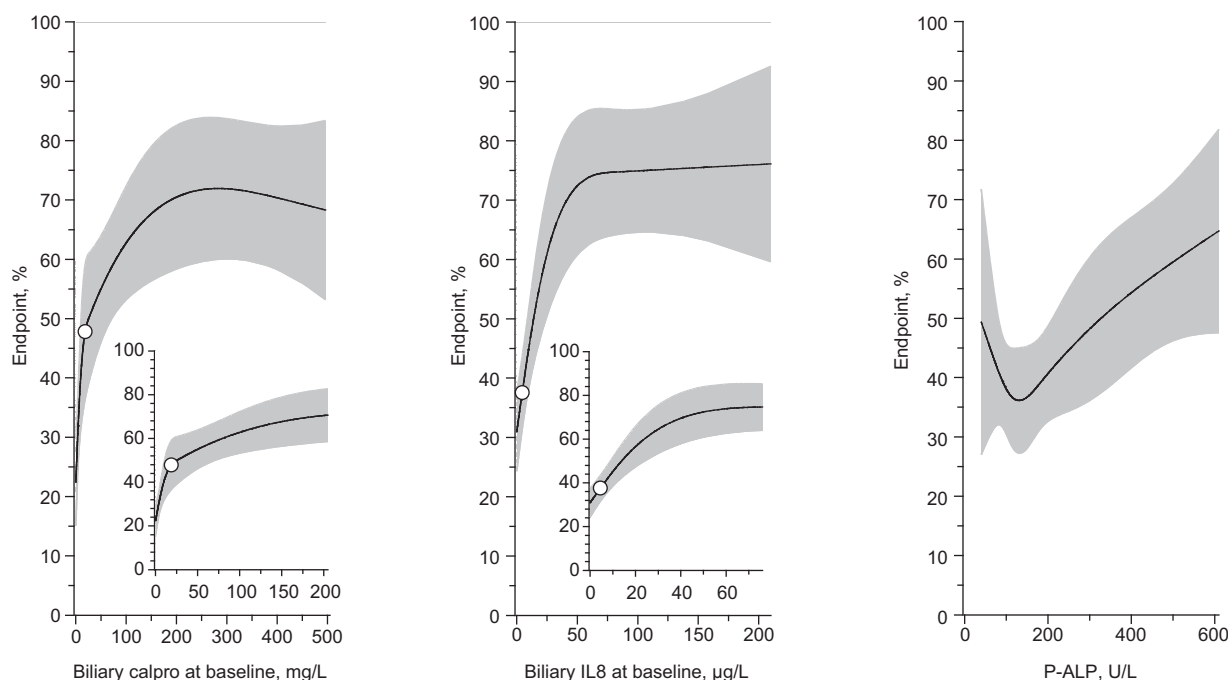
Fig. 5. Relationships of biliary calprotectin, biliary IL8 and P-ALP with bile duct disease progression. The curves were derived from a 4-knot-restricted cubic splines logistic regression. The models were adjusted for sex and disease duration. The white circle shows the optimal cut-off points for biliary calprotectin and biliary IL8 at baseline. IL8, interleukin-8; P-ALP, plasma alkaline phosphatase.

are unlikely to progress, and the time interval to the next control or scheduled ERCP can be prolonged. In drug trials, patients with high levels of biliary duct inflammation are the ones who are most likely to progress, and Bi-calprotectin and Bi-IL8 can be used to select participants for trials. Moreover, these markers may be suitable for monitoring drug response.

A strength of the present study is that it included a large prospective population of patients from local hospitals around the country who underwent sequential ERCP examinations to evaluate bile duct disease progression and thus data were available on a great number of serum and biliary markers. The limitation is that we had to exclude patients who underwent only one ERCP for a 3-year follow-up period (n = 349) due to very mild disease. In addition, we excluded patients with PSC-AIH-overlap syndrome (n = 81) due to limited data on the

impact of AIH features on bile duct disease progression.[42] Earlier studies have shown that patients with PSC-AIH, compared to those with classic PSC, have similar transplant-free survival.[43] However, despite biochemical response and histological improvement, the progression of biliary lesions was seen in a small number of patients with PSC-AIH.[42]

In conclusion, the measurement of Bi-calprotectin and Bi-IL8 levels during ERCP helps to identify individuals at increased risk for PSC progression and for whom the next imaging procedure or ERCP should be performed within 3 years. In addition, our results suggest that conventionally used liver function tests, especially P-ALP, for monitoring PSC progression lack both sensitivity and specificity and are probably not suitable endpoints in clinical drug trials, at least as short-term endpoints.

Affiliations

[1]*Helsinki University, Finland;* [2]*Transplantation and Liver Surgery, Abdominal Center, Helsinki University Hospital and University of Helsinki, Helsinki, Finland;* [3]*Department of Clinical Chemistry, University of Helsinki and Helsinki University Hospital, Helsinki, Finland;* [4]*Department of Gastroenterology, Helsinki University Hospital, Abdominal Center, Helsinki, Finland;* [5]*Folkhälsan Research Center, Helsinki, Finland and Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio, Finland*

Abbreviations

AIH, autoimmune hepatitis; ALP, alkaline phosphatase; AOM, Amsterdam-Oxford PSC model; BC, brush cytology; Bi-, biliary; DS, dominant stricture; ERCP, endoscopic retrograde cholangiopancreatography; IL8, interleukin-8; MRC, magnetic resonance cholangiography; P-, plasma; PSC, primary sclerosing cholangitis; UDCA, ursodeoxycholic acid; ULN, upper limit of normal.

Conflict of interest

The authors of this study declare that they do not have any conflict of interest.

Please refer to the accompanying ICMJE disclosure forms for further details.

Authors' contributions

Study concept and design: MF, HK. Acquisition of data: MF, HA, KJ, AT, LP. Analysis and interpretation of data: HK, MF. Drafting of the manuscript: MF, HK. Critical revision of the manuscript for important intellectual content: MF, HK, FÅ, HA, AT, KJ. Statistical analysis: HK, MF. Obtained funding: MF. Administrative, technical, or material support: MF. Study supervision: MF.

Data availability statement

Data not available due to ethical restrictions. Participants in this study did not give informed consent for their data to be shared publicly.

Acknowledgements

Study nurses Virpi Pelkonen, Pirkko Tuukkala, Susanna Saarinen and Liridiana Pena Zamora are acknowledged for invaluable help.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhepr.2024.101161.

References

[1] Ponsioen CY, Assis DN, Boberg KM, et al. Defining primary sclerosing cholangitis: results from an international primary sclerosing cholangitis study group consensus process. Gastroenterology 2021;161:1764–1775.

[2] Bergquist A, Ekbom A, Olsson R, et al. Hepatic and extrahepatic malignancies in primary sclerosing cholangitis. J Hepatol 2002;36:321–327.

[3] Manninen P, Karvonen AL, Laukkarinen J, et al. Colorectal cancer and cholangiocarcinoma in patients with primary sclerosing cholangitis and inflammatory bowel disease. Scand J Gastroenterol 2015;50:423–428.

[4] Aabakken L, Karlsen TH, Albert J, et al. Role of endoscopy in primary sclerosing cholangitis: European society of gastrointestinal endoscopy (ESGE) and European association for the study of the liver (EASL) clinical guideline. Endoscopy 2017;49:588–608.

[5] Tenca A, Mustonen H, Lind K, et al. The role of magnetic resonance imaging and endoscopic retrograde cholangiography in the evaluation of disease activity and severity in primary sclerosing cholangitis. Liver Int 2018;38:2329–2339.

[6] Weber C, Kuhlencordt R, Grotelueschen R, et al. Magnetic resonance cholangiopancreatography in the diagnosis of primary sclerosing cholangitis. Endoscopy 2008;40:739–745.

[7] Segal D, Marotta P, Mosli M, et al. The role of imaging in determining prognosis for primary sclerosing cholangitis: a systematic review. Saudi J Gastroenterol 2019;25:152–158.

[8] Ruiz A, Lemoinne S, Carrat F, et al. Radiologic course of primary sclerosing cholangitis: assessment by three-dimensional magnetic resonance cholangiography and predictive features of progression. Hepatology 2014;59:242–250.

[9] Cazzagon N, Lemoinne S, El Mouhadi S, et al. The complementary value of magnetic resonance imaging and vibration-controlled transient elastography for risk stratification in primary sclerosing cholangitis. Am J Gastroenterol 2019;114:1878–1885.

[10] Grigoriadis A, Imeen Ringe K, Bengtsson J, et al. Development of a prognostic MRCP-score (DiStrict) for individuals with large-duct primary sclerosing cholangitis. JHEP Rep 2022;4:10059.

[11] Ponsioen CY, Chapman RW, Chazouillères O, et al. Surrogate endpoints for clinical trials in primary sclerosing cholangitis: review and results from an International PSC Study Group consensus process. Hepatology 2016;63:1357–1367.

[12] de Vries EMG, Wang J, Leeflang MMG, et al. Alkaline phosphatase at diagnosis of primary sclerosing cholangitis and 1 year later: evaluation of prognostic value. Liver Int 2016;36:1867–1875.

[13] Vesterhus M, Holm A, Hov JR, et al. Novel serum and bile protein markers predict primary sclerosing cholangitis disease severity and prognosis. J Hepatol 2017;66:1214–1222.

[14] de Vries EMG, Färkkilä M, Milkiewicz P, et al. Enhanced liver fibrosis test predicts transplant-free survival in primary sclerosing cholangitis, a multicentre study. Liver Int 2017;37:1554–1561.

[15] Goode EC, Clark AB, Mells GF, et al. Factors associated with outcomes of patients with primary sclerosing cholangitis and development and validation of a risk scoring system. Hepatology 2019;69:2120–2135.

[16] Goet JC, Floreani A, Verhelst X, et al. Validation, clinical utility, and limitations of the Amsterdam-Oxford model for primary sclerosing cholangitis. J Hepatol 2019;71:992–999.

[17] Eaton JE, Vesterhus M, McCauley BM, et al. Primary sclerosing cholangitis risk estimate tool (PREsTo) predicts outcomes of the disease: a derivation and validation study using machine learning. Hepatology 2020;71:214–224.

[18] Chazouilleres O, Beuers U, Bergquist A, et al. EASL clinical practice guidelines on sclerosing cholangitis. J Hepatol 2022;77:761–806.

[19] Chapman MH, Thorburn D, Hirschfield GM, et al. British Society of Gastroenterology and UK-PSC guidelines for the diagnosis and management of primary sclerosing cholangitis. Gut 2019;68:1356–1378.

[20] Boyd S, Tenca A, Jokelainen K, et al. Screening primary sclerosing cholangitis and biliary dysplasia with endoscopic retrograde cholangiography and brush cytology: risk factors for biliary neoplasia. Endoscopy 2016;48:432–439.

[21] Stiehl A, Rudolph G, Klöters-Plachky PC, et al. Development of dominant bile duct stenoses in patients with primary sclerosing cholangitis treated with ursodeoxycholic acid: outcome after endoscopic treatment. J Hepatol 2002;36:151–156.

[22] Kim WR, Therneau TM, Wiesner RH. al. A revised natural history model for primary sclerosing cholangitis. Mayo Clin Proc 2000;75:688–694.

[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] Karlsen TH, Folseraas T, Thorburn D, et al. Primary sclerosing cholangitis - a comprehensive review. J Hepatol 2017;67:1298–1323.

[25] Trivedi PJ, Muir AJ, Levy C, et al. Inter- and intra-individual variation, and limited prognostic utility, of serum alkaline phosphatase in a trial of patients with primary sclerosing cholangitis. Clin Gastroenterol Hepatol 2021;19:1248–1257.

[26] Craig DA, MacCarty RL, Wiesner RH, et al. Primary sclerosing cholangitis: value of cholangiography in determining the prognosis. Am J Roentgenol 1991;157:959–964.

[27] Olsson RG, Asztely MS. Prognostic value of cholangiography in primary sclerosing cholangitis. Eur J Gastroenterol Hepatol 1995;7:251–254.

[28] Ponsioen CY, Vrouenraets SME, Prawirodirdjo W, et al. Natural history of primary sclerosing cholangitis and prognostic value of cholangiography in a Dutch population. Gut 2002;51:562–566.

[29] Rupp C, Hippchen T, Bruckner T, et al. Effect of scheduled endoscopic dilatation of dominant strictures on outcome in patients with primary sclerosing cholangitis. Gut 2019;68:2170–2178.

[30] Tischendorf JJ, Hecker H, Krüger M, et al. Characterization, outcome, and prognosis in 273 patients with primary sclerosing cholangitis: a single center study. Am J Gastroenterol 2007;102:107–114.

[31] Tornai D, Ven PL, Lakatos PL, Papp M. Serological biomarkers for management of primary sclerosing cholangitis. World J Gastroenterol 2022;28:2291–2301.

[32] Wunsch E, Norman GL, Milkiewicz M, et al. Anti-glycoprotein 2 (anti-GP2) IgA and anti-neutrophil cytoplasmic antibodies to serine proteinase 3 (PR3-ANCA): antibodies to predict severe disease, poor survival, and cholangiocarcinoma in primary sclerosing cholangitis. Aliment Pharmacol Ther 2021;53:302–313.

[33] Carpino G, Cardinale V, Renzi A, et al. Activation of biliary tree stem cells within peribiliary glands in primary sclerosing cholangitis. J Hepatol 2015;63:1220–1228.

[34] Sjöblom N, Boyd S, Kautiainen H, et al. Novel histological scoring for predicting disease outcome in primary sclerosing cholangitis. Histopathology 2022;81:192–204.

[35] Kaplanski G, Farnarier C, Kaplanski S, et al. Interleukin-1 induces interleukin-8 secretion from endothelial cells by a juxtacrine mechanism. Blood 1994;84:4242–4248.

[36] Tabibian JH, O'Hara SP, Splinter PL, et al. Cholangiocyte senescence by way of N-ras activation is a characteristic of primary sclerosing cholangitis. Hepatology 2014;59:2263–2275.

[37] Gauss A, Sauer P, Stiehl A, et al. Evaluation of biliary calprotectin as a biomarker in primary sclerosing cholangitis. Medicine (Baltimore) 2016;95(17):e3510.

[38] Koskensalo V, Aronen P, Färkkilä M, et al. Use of thiopurines is not a risk factor for post-ERCP pancreatitis in patients with primary sclerosing cholangitis. Dig Liver Dis 2021;53:1020–1027.

[39] Bergquist A, Weismüller TJ, Levy C, et al. Impact on follow-up strategies in patients with primary sclerosing cholangitis. Liver Int 2023;43:127–138.

[40] Villard C, Friis-Liby I, Rorsman F, et al. Prospective surveillance for cholangiocarcinoma in unselected individuals with primary sclerosing cholangitis. J Hepatol 2023;78:604–613.

[41] Duggan WP, Brosnan C, Christodoulides N, et al. Outruling cholangiocarcinoma in patients with primary sclerosing cholangitis wait-listed for liver transplantation: a report on the Irish national experience. Surgeon 2023;21(2):e83–e88.

[42] McNair AN, Moloney M, Portmann BC, et al. Autoimmune hepatitis overlapping with primary sclerosing cholangitis in five cases. Am J Gastroenterol 1998;93:777–784.

[43] Weismüller T, Trivedi PJ, Bergquist A, et al. Patient age, sex, and inflammatory bowel disease phenotype associate with course of primary sclerosing cholangitis. Gastroenterology 2017;152:1975–1984.