# The role of imaging in determining prognosis for primary sclerosing cholangitis: A systematic review

Dan Segal, Paul Marotta[1], Mahmoud Mosli[2], Guangyong Zou[3], Brian G. Feagan[1,3], Bandar Al-Judaibi[4]

Department of Medicine, McMaster University, Hamilton, Ontario, Canada, [1]Department of Medicine, London Health Sciences Centre, Western University, London, Ontario, Canada, [2]Department of Medicine, King Abdulaziz University, Jeddah, Saudi Arabia, [3]Robarts Clinical Trials, Robarts Research Institute, London, Ontario, Canada, [4]Department of Medicine, University of Rochester, Rochester, New York, USA

## Abstract

**Background/Aims:** Primary sclerosing cholangitis (PSC) is a chronic, progressive, fibrotic bile duct disease. Resultant complications include infection, progressive liver disease and cancer. While diagnosis relies extensively on imaging, the role of imaging in determining prognosis is unclear. The aim of this study was to systematically review existing imaging indices and features that predict PSC progression.

**Materials and Methods:** We performed a systematic review of imaging features that predict PSC progression. PubMed, EMBASE, MEDLINE, Clinicaltrials.gov and the Cochrane Library were searched from inception to November 2018 for relevant studies. Pertinent data were extracted and assessed. Study quality was evaluated using the Newcastle-Ottawa scale (NOS).

**Results:** The search returned 2504 results. Nine studies were included in the final review. Four studies evaluated the prognostic value of imaging features and five evaluated prognostic algorithms. The mean NOS score was $4.44 \pm 0.98$ on a scale of 0 to 9. Imaging features that were of prognostic value were degree of intrahepatic duct narrowing, the presence of a dominant biliary duct stricture and percentage of narrowed intraheptic ducts. Three imaging indices (one endoscopic retrograde cholangiopancreatography (ERCP)-based and two magnetic resonance-based) had been derived. The ERCP index was validated in a second cohort and subsequently updated to improve its predictive ability. The magnetic resonance cholangiopancreatography (MRCP) index was validated in two studies and was found to be predicative of transplant-free survival. A modified MRCP index (MRCP-risk score) was evaluated in a prospective multicenter study and was found to be predicative of PSC-related disease progression.

**Conclusion:** In conclusion, ERCP and MRCP-based indices have short-term prognostic value in PSC. However, more studies are required to validate their predictability of disease-related progression, such as liver decompensation, ascending cholangitis, cholangiocarcinoma and liver transplantation.

**Keywords:** Endoscopic retrograde cholangiopancreatography, magnetic resonance imaging, score, survival, transplant-free survival

**Address for correspondence:** Dr. Bandar Al-Judaibi, Department of Medicine, Rochester, New York, United States of America.
E-mail: Bandaraljudaibi@gmail.com

| Access this article online | |
|---|---|
| **Quick Response Code:** | **Website:** www.saudijgastro.com |
| | **DOI:** 10.4103/sjg.SJG_478_18 |

# INTRODUCTION

Primary sclerosing cholangitis (PSC) is a disease of chronic bile duct destruction[1] that predisposes patients to multiple complications, such as cholangitis, end-stage liver disease and cholangiocarcinoma.[2] While elevated cholestatic biochemical markers are suggestive of this disease, they are neither diagnostic nor do they provide an accurate assessment of disease activity or prognosis. Liver biopsy is the gold standard for diagnosis but is generally avoided unless small duct disease is being considered. Accordingly, diagnosis relies on imaging studies including historically, endoscopic retrograde cholangiopancreatography (ERCP) and currently more commonly magnetic resonance cholangiopancreatography (MRCP).[3]

Hepatic transplantation remains the mainstay for the management of progressive disease, as there is no medical therapy that alters the disease's natural history. Currently median-transplant free survival ranges between 12 and 18 years.[4] Robust prognostic models could aid in understanding an individual's specific disease course, predict time to transplant and aid health care systems in resource planning. However, no existing model has been widely accepted or endorsed.[2]

It is unclear what items should be included in a prognostic model for PSC. Most models incorporate biochemical indices such as alkaline phosphatase and bilirubin that can vary widely within a short time frame.[5] Although histopathologic disease severity on liver biopsy is associated with survival, it is a suboptimal tool due to the patchy distribution of the disease and the potential for procedural-related morbidity.[6] Alternatively, radiological studies can be noninvasive, are less susceptible to short-term variability, and are widely available. Imaging studies are uniquely positioned to be an effective tool for prognostication.

This systematic review aims to evaluate the evidence for using imaging features and models for determining PSC prognosis.

# MATERIALS AND METHODS

## Search strategy
MEDLINE (Ovid, PUBMED), EMBASE (Ovid), Clinicaltrials.gov and the Cochrane Library (CENTRAL) were searched from inception to November 2018 for imaging features or indices that predict PSC disease progression. The databases were searched for "sclerosing cholangitis" AND multiple imaging terms including "PSC radiology index", "magnetic resonance" (MR), "MRI", "MRCP", OR "cholangiopancreatography." This search strategy was modified to match each database.

## Study eligibility criteria and study selection
Randomized controlled trials, case-controlled trials and cohort studies that evaluated the value of PSC imaging indices for assessment of prognosis were included. Only studies that included transplant or death as the definitive outcome measure for disease progression were selected. References were manually scanned for additional studies that were missed by electronic search. Citations and their abstracts were screened according to predetermined exclusion criteria to select studies for full text review. Two reviewers (D.S. and B.A.) assessed the studies for inclusion and resolved disagreements by consensus.

## Risk of bias assessment
Given that all the studies employed observational designs, they were assessed for quality using the Newcastle-Ottawa scale (NOS) studies, a method endorsed by the Cochrane collaboration for assessment of cohort studies.[7,8] Quality scores generated by this scale range from 0 to 9 with higher scores reflecting higher study quality. Two reviewers (D.S. and B.A.) independently assessed the studies for quality and resolved disagreements by consensus.

# RESULTS

The literature search identified 2497 records. Seven additional studies were accrued from the manual review yielding 2504 records in total. After duplicates were removed 1366 references remained. After screening the titles and abstracts according to pre-specified criteria, nine studies were selected for full text review. Of these, nine studies were finally selected; studies four studies described individual prognostic imaging items and five studies described three multiplex imaging indices [Figure 1]. The studies were assessed for quality using the NOS [Table 1]. The mean NOS score was 4.44 ± 0.98. The following summarizes key features of the studies.

## Craig *et al*.[9]
This retrospective cohort study from the Mayo Clinic identified 129 patients with PSC who had undergone a conventional cholangiogram between 1970 and 1984.[9] A radiologist blinded to clinical information, reviewed and categorized cholangiograms by bile duct features of both the intra- and extrahepatic trees. Bile duct strictures were graded by the length of the stricture, percent narrowing of the duct and extent of the stricturing (defined as localized or diffuse based on the percent of ducts involved). Bile duct dilation was also evaluated. By the study censure date of 1986, 28% of the cohort had died, while 11% had
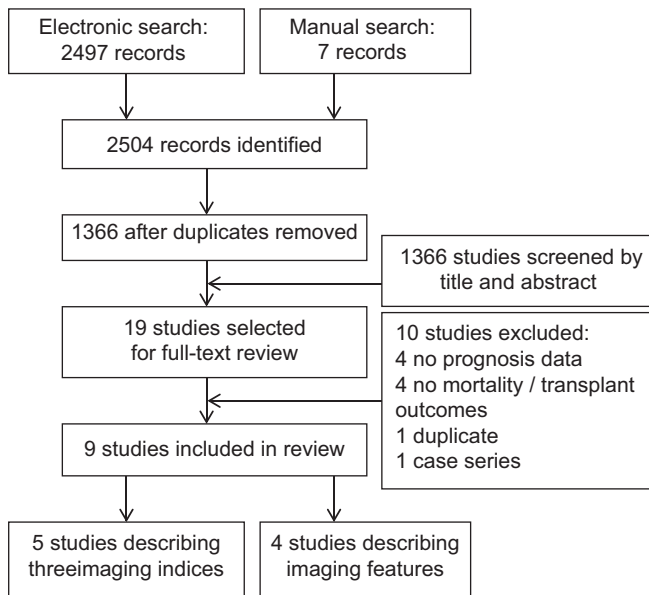
**Figure 1:** Search results

**Table 1: Newcastle-Ottawa score**

| Study | Selection (/4) | Comparability (/2) | Outcome (/3) | Total (/9) |
|---|---|---|---|---|
| Craig[9] | 2 | 0 | 2 | 4 |
| Olsson[10] | 3 | 0 | 2 | 5 |
| Tischendorf[11] | 2 | 0 | 3 | 5 |
| Rudolph[12] | 3 | 0 | 3 | 6 |
| Ponsoien[13] | 2 | 0 | 2 | 4 |
| Ponsoien[14] | 2 | 0 | 3 | 5 |
| Lemoinne[18] | 2 | 0 | 1 | 3 |
| Cazzagon[19] | 2 | 0 | 2 | 4 |
| Muir[20] | 2 | 0 | 2 | 4 |

undergone liver transplantation, resulting in an overall 5-year actuarial estimate of 57% for event-free survival.

Survival curves with various cholangiographic findings were compared using a log-rank test. Two imaging items were determined to have prognostic importance. Patients with high-grade intrahepatic duct narrowing (>75% loss of lumen) had a 19% decrease in transplant-free survival at three years ($P = 0.05$). Those with diffuse intrahepatic strictures (>25% of ducts involved) had a 16% decrease in transplant-free survival at 3 years ($P = 0.012$).

### Olsson and Asztely[10]

A retrospective study conducted in Gothenburg, Sweden, reviewed 94 patients with PSC who had undergone conventional cholangiography, including many who had previously remained undiagnosed.[10] The cholangiograms were scored for bile duct stricture length, narrowing, extent, as well as duct dilation using the same criteria as Craig *et al*.[9] Median follow-up was 54 months and the 5-year event-free survival rate was 72%. Multivariable regression confirmed that narrowing of intrahepatic ducts was an independent prognostic feature of death or liver transplantation [relative risk (RR) 2.48 (1.31–4.72); $P = 0.005$]. In contrast to the Mayo clinic study, no other radiological items were independently associated with a poor outcome.

### Tischendorf *et al*.[11]

This retrospective single center study from Germany evaluated 273 patients with PSC.[11] Patient data including ERCP images, laboratory results and biopsy findings were collected and follow-up was censored at time of death or liver transplant. Median event-free survival for

the cohort was just over 9 years. Multivariate analysis of ERCP items identified the presence of combined intrahepatic and extrahepatic duct changes of any type (hazard ratio = 2.46 [1.65–3.66]; $P < 0.001$) and a dominant biliary stricture (HR = 2.29 [1.60 − 3.27]; $P < 0.001$) to be independently associated with a poor outcome. The investigators derived a multiplex prognostic score that incorporated age, low albumin, bilirubin, hepatosplenomegaly and the two ERCP items.

### Rudolph *et al*.[12]

This retrospective single center study from Heidelberg, Germany, evaluated the influence of dominant stenosis on survival, in a cohort of 171 PSC patients who were enrolled in prospective trial evaluating the efficacy of ursodeoxycholic acid.[12] A dominant stenosis was defined as a stenosis with a diameter <1.5 mm in the common bile duct or <1 mm in a hepatic duct within 2 cm of the bifurcation. After 18 years of follow-up, those with a dominant stenosis based on ERCP had a 25% chance of transplant-free survival compared to 75% for those without a dominant stenosis ($P = 0.011$). Even when patients with a dominant stenosis who went on to develop cholangiocarcinoma were removed from the analysis, the survival difference between the two groups persisted.

### Amsterdam model[13,14]

This ERCP-based model was derived from a retrospective cohort of 174 patients with PSC, in the Netherlands.[13] Median time from diagnosis to death or transplant was 18 years. Index cholangiograms were scored according to the Amsterdam cholangiographic classification of PSC [Table 2]. This system, described by Majoie *et al*,[14] was a modification of an earlier classification.[15] Analysis was performed to create a prognostic model that incorporated the ERCP classification system and relevant historical details. The model relies on two variables SUMIHDEHD" and AGEERCP.

SUMIHDEHD" is an assigned value derived from the Amsterdam classification of the intra- and extrahepatic ducts adjusted to combine stages that exhibited no

**Table 2: Amsterdam cholangiographic classification of PSC by ERCP**

| Type of duct involvement/Classification | Cholangiographic abnormalities |
|---|---|
| Intrahepatic | |
| 0 | No visible abnormalities |
| I | Multiple strictures; normal caliber of bile ducts or minimal dilation |
| II | Multiple strictures, saccular dilation, decreased arborization |
| III | Only central branches filled despite adequate filling pressure; severe pruning |
| Extrahepatic | |
| 0 | No visible abnormalities |
| I | Slight irregularities of duct contour; no stenosis |
| II | Segmental stricture |
| III | Stricture of almost entire length of duct |
| IV | Extremely irregular margin; diverticulum-like outpouching |

survival difference [Table 3]. AGEERCP is the age of the patient at index ERCP. The equation is described below:

$$\text{Prognostic index} = 1.13 \times X_3 + 1.98 \times X_4 + 0.024 \times Y$$

$X_3$ and $X_4$ are determined by the SUMIHDEHD score [Table 4], while Y is the AGEERCP in years. The prognostic index ranging from 0 to 3.5 can then be used to estimate median survival from a graph.[13] The C statistic for the model was 0.706. A C statistic is equal to the area under a receiver operator characteristic curve and is a measure of goodness of fit. A C statistic of 0.5 suggests a model's ability to predict an outcome equal to random chance; while a C statistic of 1 suggests that a model can perfectly predict an outcome.

This prototypic model was subsequently validated in a second cohort of 111 patients from Norway.[16] The model was able to predict patient survival in the Norwegian cohort similarly in the validation data set as it did during derivation with a C statistic of 0.703. The prognostic index equation was updated to reflect data from both cohorts:

**Table 3: SUMIHDEHD" score derived from cross-referencing the assigned IHD and EHD scores as per the Amsterdam classification; IHD intrahepatic duct, EHD extrahepatic duct**

| | IHD | | | |
|---|---|---|---|---|
| | 0 | I | II | III |
| EHD | | | | |
| 0 | – | 2 | 3 | 3 |
| I | 1 | 2 | 3 | 3 |
| II | 2 | 3 | 3 | 4 |
| III | 3 | 3 | 4 | 5 |
| IV | 3 | 3 | 4 | 5 |

$$\text{Prognostic index} = 0.89 \times X_3 + 1.59 \times X_4 + 0.028 \times Y$$

Ponsoien *et al.*[14] published a nomogram to simplify how to derive survival data for an individual patient [Figure 2].

### Ruiz model[17-19]

Ruiz *et al.* created a prognostic model using three-dimensional MRCP imaging using data from a retrospective review of 64 patients who had each undergone at least two MR studies.[17] MRCPs were scored according to criteria developed in part from Craig *et al.*'s classification system.[9] Liver-related items including dysmorphy (defined as lobar atrophy, lobular surface changes, or an abnormal caudate to right lobe volume ratio) and portal hypertension were evaluated. An overall rating of disease severity derived by counting the number of features that were present was assigned to each MRI. MR studies were compared sequentially by their overall rating to identify an individual's radiological course. Over time, the disease was radiologically stable in 42% of patients, worsened in 58%, and improvement was noted in none.

Multivariable regression analysis was performed to identify the items associated with radiological progression. The predictor variables were incorporated into scores to determine risk of disease progression on reimaging:

Score (MRI without gadolinium) =1 × intrahepatic bile duct dilation + 2 × dysmorphy + 1 × portal hypertension

Score (MRI with gadolinium) =1 × dysmorphy + 1 × parenchymal enhancement heterogeneity

IHBD dilation can be scored as 0 (IHBD <4 mm), 1 (IHBD = 4 mm), or 2 (IHBD >4 mm), while the other variables are either present,[1] or absent.[2]

The first model (without gadolinium) predicted radiological progression for those with a score of three or higher. The sensitivity was 87%; the specificity was 63%. The second model (with gadolinium) predicted progression if the score was two. The sensitivity was 91%; the specificity was 72%.

A second study assessed the operating properties of these scores for prediction of transplant-free survival.[18] A total

**Table 4: SUMIHDEHD" scores correspond to specific $X_3$ and $X_4$; SUMIHDEHD" Scores of 1 and 5 were not present in the cohort**

| | $X_3$ | $X_4$ |
|---|---|---|
| SUMIHDEHD" | | |
| 2 | 0 | 0 |
| 3 | 1 | 0 |
| 4 | 0 | 1 |

**Figure 2: A nomogram to determine survival based on the Amsterdam model. The points for the age at index ERCP and points corresponding to the assigned EHD and IHD score are combined. The total points then vertically align with expected survival at different times**

| AGE | 20 | 30 | 40 | 50 | 60 | 70 | | | |
|---|---|---|---|---|---|---|---|---|---|
| Points | 0 | 2 | 5 | 8 | 11 | 14 | | | |
| SUMEHDIHD" | 2 | 3 | 4 | | | | | | |
| Points | 0 | 8 | 15 | | | | | | |
| Total Points | 0 | 5 | 10 | 15 | 20 | 5 | 30 | 35 | 40 |
| 1-yr survival (%) | 98 | 96 | 94 | 90 | 85 | 76 | 63 | 47 | 29 |
| 5-yr survival (%) | 94 | 90 | 84 | 76 | 63 | 47 | 28 | 13 | 3.3 |
| 10-yr survival (%) | 89 | 82 | 72 | 59 | 42 | 24 | 9.2 | 2.0 | 0.2 |

of 67 PSC patients were retrospectively reviewed and censured at time of death or liver transplant. Multivariate analysis revealed that the score for MRI without gadolinium predicted survival, while the score for MRI with gadolinium did not. The C statistic was 0.818, with a sensitivity of 90% and specificity of 56% when the cutoff is set at two.[18]

A third large retrospective multicenter study confirmed the prognostic value of MRI (with or without gadolinium) scores in 238 PSC patients.[19] In this study, two cohorts of PSC patients were evaluated, a derivation cohort ($n = 119$) from Paris and an external validation cohort ($n = 119$) from Birmingham, Padova and Montreal. Decompensated cirrhosis and liver transplant-free survival were the primary outcomes. During the median follow-up of 4.4 and 3.8 years, 20 and 25 patients underwent liver transplant, 9 and 5 patients died, and 18 and 24 patients developed cirrhotic decompensation in the derivation and validation cohorts, respectively. According to univariate analysis, items associated with event-free survival were: total bilirubin, AST, ALT, GGT, albumin, MR score without gadolinium and MR score with gadolinium. Predictive performances of MR scores without and with gadolinium assessed by c-statistic were 0.89 IC95% [0.84–0.95] and 0.75 IC95% [0.64–0.87].

## Muir *et al.*[20]

In this study, the association between biliary severity on MRCP and disease progression in PSC patients was evaluated prospectively in phase 2, placebo-controlled trial of simtuzumab. MRC was performed at baseline in 234 PSC patients.[20] Consensus reading of MRCPs by two radiologists was performed to characterize the Ruiz model.[17] The association between the Ruiz model's features and PSC-related clinical events (decompensation, ascending cholangitis, cholangiocarcinoma and liver transplant) was determined using Cox regression and MRCP risk score (MRCP-RS) derived based on factors with independent prognostic value. At baseline, 40% of patients had bridging fibrosis and 11% had cirrhosis. The median follow-up was 23 months and 47 (20%) patients developed PSC-related clinical events.[20]

Based on multivariate analysis, PSC-related events were associated with baseline hepatic dysmorphy (HR = 3.11 [1.22–7.92]), signs of portal hypertension (HR = 2.31 [1.28–4.17]), and perihepatic lymph nodes (HR = 2.14 [1.20-3.81]). Based on the model coefficients, an MRCP-RS assigning 1-point for each of the three variables was derived (range, 0–3), which accurately predicted clinical events (c-statistic 0.71; 95% CI 0.63–0.79). During follow-up, the risk of clinical events increased according to baseline MRCP - RS: 0 (6%), 1 (14%), 2 (30%), and 3 (56%); $P < 0.001$.

## DISCUSSION

In this review, we identified nine studies that used individual radiographic items or multiplex indices to predict transplant-free survival in patients with PSC. Craig *et al.* found that the degree of intrahepatic duct narrowing and proportion of ducts involved on ERCP were significant prognostic factors.[9] Olsson and Asztely's results corroborated the finding that intraheptic duct narrowing was associated with transplant-free survival; however, they did not confirm the association with the presence of diffuse intraheptic strictures.[10] Alternatively, both German ERCP studies only identified the presence of a dominant stenosis as an independent prognostic feature.[11,12]

Three multiplex prognostic instruments based on radiological studies have been described.[13,17,20] Ponsoien *et al.* created an ERCP-based model that was validated with modification in a second cohort of patients. Calculation of the index score requires multiple complex calculations; thus, a nomogram was generated to facilitate scoring.[13,14] It is important to highlight that the Amsterdam index is ERCP-based and has not been validated for MRCP–generated items. MRCP has supplanted ERCP as the diagnostic imaging study of choice, and thus it is critical to determine whether the ERCP generated items identified as components of the index remain valid if generated by MRCP. In this regard, the index generated by Ruiz *et al.* using MRCP-defined items provides some insight. Their study evaluated items previously evaluated in ERCP-based studies including severity of strictures, stricture length, hypertension and degree of duct involvement, and evaluated extraductal items not visible on ERCP including lymph nodes, portal hypertension, and parenchymal dysmorphy. Intrahepatic duct dilatation was the only item associated with radiological progression on MRCP that could potentially be defined on ERCP.[17] Several studies have shown that the MRCP-based model is predictive of transplant-free survival.[13,17,20] However, whether or not MRI-based gadolinium has similar predictability

of transplant-free survival compared to MRI without gadolinium, requires further investigation. In addition, a modified MRCP-risk score (MRCP-RS) has been developed in a prospective phase-2 study.[20] The MRCP-RS accurately predicts PSC-related disease progression in the clinical setting over 96 weeks. However, further studies are required to validate the MRCP-RS.

This paper is the first systematic review of radiographic studies of prognosis in PSC. Strengths of this study include the systematic methodology employed in collecting the data and our focus on studies that used transplant-free survival as an outcome. A notable study limitation is that this review relies on individual studies that are retrospective with suboptimal methods for minimizing bias with respect to data retrieval. Furthermore, all of the studies were conducted from tertiary care centers and as a result might suffer from referral bias. Due to the variability in study design, data collected, and length of follow-up, we were unable to compare study findings directly or pool results.

This review is limited to studies that described radiological items for prognosis. Strengths of a radiographic-based instrument include the frequency of imaging performed in PSC, the lack of day-to-day variability, and potential simplicity of a tool that relies on a single modality. Nevertheless, multimodality models do exist and were excluded from the review. It is noteworthy to mention that in the context of other similar autoimmune diseases, such as inflammatory bowel disease, multiplex models function better, likely because they take into account both clinical and biochemical assessments. On the other hand, multiplex scoring systems are sometimes difficult to interpret.

The most widely cited prognostic model is the Mayo Risk Score[5] comprised of age, bilirubin, albumin, AST and history of variceal bleeding. While this instrument has been widely used, it has some weaknesses. The derivation and validation cohorts had a median follow-up of only 4 years. Also, patients were only censured at death and not at the time of liver transplant. Instead, patients who underwent liver transplant had an assumed survival of 1-year posttransplant.[21] In clinical practice, transplant-free survival is considered a more relevant end point. There may also be issues related to variability of specific serum markers like AST, on a day-to-day basis.

Recently, a multiplex prognostic score known as the Amsterdam-Oxford model was described[22] that was derived from a broad cohort largely presenting to community hospitals and was subsequently validated in a tertiary care center. The predictive model based on Cox regression modeling includes PSC subtype (large duct or small duct), age at diagnosis, albumin, platelets, AST, ALP and bilirubin. Following 1 year of follow-up, the C statistic in the derivation cohort was 0.68 and 0.67 in the validation cohort. The process for deriving a probability for transplant-free survival requires multiple complex calculations.

A number of new prognostic tests have been proposed. Corpechot *et al.*[23] assessed the utility of transient elastography (TE) in PSC patients. TE was well correlated to fibrosis stage and cutoff values were established. The rate of change in a patient's TE result was correlated with survival though the underlying results were not published. Elevated biliary calprotectin levels have been associated with the presence of a dominant biliary stenosis and need for biliary intervention. It has not borne out as an independent risk factor for disease activity.[24]

Serum markers that reflect fibrosis and inflammation have become an area of research interest. The enhanced liver fibrosis (ELF) score was evaluated in PSC.[25] The ELF consists of three serum proteins that are expressed during liver collagen deposition (hyaluronic acid, tissue inhibitor of metalloproteinases-1, propeptide of type III procollagen). The ELF score was able to predict transplant-free survival in a robust manner with a hazard ratio of 1.9 (95% CI 1.4–2.5). The ELF test however is not widely available. The same group identified the cytokine IL-8, a macrophage derived chemokine, as a novel prognostic marker. It was able to predict survival, but was not as prognostically robust as the ELF score or the Mayo risk score.[26]

In summary, at this time there is no single marker or prognostic algorithm that can be recommended for PSC. While the Mayo risk score is the most well-known model, the MRCP-based score and the ELF score may have potential value.

### Conflicts of interest
There are no conflicts of interest.

### REFERENCES

1. Hirschfield GM, Karlsen TH, Lindor KD, Adams DH. Primary

sclerosing cholangitis. Lancet 2013;382:1587-99.

2.  Chapman R, Fevery J, Kalloo A, Nagorney DM, Boberg KM, Shneider B, *et al.* Diagnosis and management of primary sclerosing cholangitis. Hepatology 2010;51:660-78.

3.  Lazaridis KN, LaRusso NF. Primary Sclerosing Cholangitis. N Engl J Med 2016;375:1161-70.

4.  Silveira MG, Lindor KD. Primary sclerosing cholangitis. Can J Gastroenterol 2008;22:689-98.

5.  Kim WR, Therneau TM, Wiesner RH, Poterucha JJ, Benson JT, Malinchoc M, *et al.* A revised natural history model for primary sclerosing cholangitis. Mayo Clin Proc 2000;75:688-94.

6.  de Vries EM, de Krijger M, Farkkila M, Arola J, Schirmacher P, Gotthardt D, *et al.* Validation of the prognostic value of histologic scoring systems in primary sclerosing cholangitis: An international cohort study. Hepatology 2017;65:907-19.

7.  Reeves BC DJ, Higgins JP, Wells GA. Chapter 13: Including non-randomized studies. In: Higgins JP, Green S, editor. Cochrane Handbook for Systematic Reviews of Interventions version 5.1.0 [handbook.cochrane.org]; 2011.

8.  Wells GA SB, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa scale (NOS) for assessing the quality of nonrandomized studies in meta-analyses. 2019. [www.ohri. ca/ programs/clinical_epidemiology/oxford.asp].

9.  Craig DA, MacCarty RL, Wiesner RH, Grambsch PM, LaRusso NF. Primary sclerosing cholangitis: Value of cholangiography in determining the prognosis. AJR Am J Roentgenol 1991;157:959-64.

10. Olsson RG, Asztely MS. Prognostic value of cholangiography in primary sclerosing cholangitis. Eur J Gastroenterol Hepatol 1995;7:251-4.

11. Tischendorf JJ, Hecker H, Kruger M, Manns MP, Meier PN. Characterization, outcome, and prognosis in 273 patients with primary sclerosing cholangitis: A single center study. Am J Gastroenterol 2007;102:107-14.

12. Rudolph G, Gotthardt D, Kloters-Plachky P, Kulaksiz H, Rost D, Stiehl A. Influence of dominant bile duct stenoses and biliary infections on outcome in primary sclerosing cholangitis. J Hepatol 2009;51:149-55.

13. Ponsioen CY, Vrouenraets SM, Prawirodirdjo W, Rajaram R, Rauws EA, Mulder CJ, *et al.* Natural history of primary sclerosing cholangitis and prognostic value of cholangiography in a Dutch population. Gut 2002;51:562-6.

14. Majoie CB, Reeders JW, Sanders JB, Huibregtse K, Jansen PL. Primary sclerosing cholangitis: A modified classification of cholangiographic findings. AJR Am J Roentgenol 1991;157:495-7.

15. Chen LY, Goldberg HI. Sclerosing cholangitis: Broad spectrum of radiographic features. Gastrointest Radiol 1984;9:39-47.

16. Ponsioen CY, Reitsma JB, Boberg KM, Aabakken L, Rauws EA, Schrumpf E. Validation of a cholangiographic prognostic model in primary sclerosing cholangitis. Endoscopy 2010;42:742-7.

17. Ruiz A, Lemoinne S, Carrat F, Corpechot C, Chazouilleres O, Arrive L. Radiologic course of primary sclerosing cholangitis: Assessment by three-dimensional magnetic resonance cholangiography and predictive features of progression. Hepatology 2014;59:242-50.

18. Lemoinne S AL, Fankem DK, Housset C, Corpechot C, Chazouilleres O. Ability of a simple radiological score, assessed by three-dimensional magnetic resonance cholangiography (MRC) to predict clinical outcome in patients with primary sclerosing cholangitis. J Hepatol 2015;62:S794.

19. Cazzagon N, Lemoinne S, El Mouhadi S, Trivedi P, Dohan A, Fankem AK, *et al.* Two simple magnetic resonance scores are able to predict survival in patients with primary sclerosing cholangitis. Hepatology 2018;68(Suppl 1):29A-30A.

20. Muir AJ, Taghipour M, Hassanzadeh E, Sahni VA, Sainani N, Ding D, *et al.* A risk prediction score based on magnetic resonance cholangiopancreatography (MRCP) accurately predicts disease progression in patients with primary sclerosing cholangitis (PSC). Hepatology 2017;66(Suppl 1):81A.

21. de Vries EM, Beuers U, Ponsioen CY. Biomarkers for disease progression of primary sclerosing cholangitis. Curr Opin Gastroenterol 2015;31:239-46.

22. de Vries EM, Wang J, Williamson KD, Leeflang MM, Boonstra K, Weersma RK, *et al.* A novel prognostic model for transplant-free survival in primary sclerosing cholangitis. Gut 2018;67:1864-9.

23. Corpechot C, Gaouar F, El Naggar A, Kemgang A, Wendum D, Poupon R, *et al.* Baseline values and changes in liver stiffness measured by transient elastography are associated with severity of fibrosis and outcomes of patients with primary sclerosing cholangitis. Gastroenterology 2014;146:970-9; quiz e15-6.

24. Gauss A, Sauer P, Stiehl A, Rupp C, Krisam J, Leopold Y, *et al.* Evaluation of biliary calprotectin as a biomarker in primary sclerosing cholangitis. Medicine (Baltimore) 2016;95:e3510.

25. Vesterhus M, Hov JR, Holm A, Schrumpf E, Nygard S, Godang K, *et al.* Enhanced liver fibrosis score predicts transplant-free survival in primary sclerosing cholangitis. Hepatology 2015;62:188-97.

26. Vesterhus M, Holm A, Hov JR, Nygard S, Schrumpf E, Melum E, *et al.* Novel serum and bile protein markers predict primary sclerosing cholangitis disease severity and prognosis. J Hepatol 2017;66:1214-22.