Research Article

JOURNAL OF HEPATOLOGY

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

Elisabeth M.G. de Vries[1], Joanne Verheij[2], Stefan G. Hubscher[3], Mariska M.G. Leeflang[4], Kirsten Boonstra[1], Ulrich Beuers[1], Cyriel Y. Ponsioen[1,*]

[1]*Department of Gastroenterology and Hepatology, Academic Medical Center, Amsterdam, The Netherlands;* [2]*Department of Pathology, Academic Medical Center, Amsterdam, The Netherlands;* [3]*School of Cancer Sciences, University of Birmingham and Department of Cellular Pathology, Queen Elizabeth Hospital, Birmingham, United Kingdom;* [4]*Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Center, Amsterdam, The Netherlands*

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

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

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

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



Keywords: Primary sclerosing cholangitis; Histology; Histologic scoring system; Prognosis; Surrogate endpoint.

*Received 13 February 2015; received in revised form 13 May 2015; accepted 8 June 2015; available online 18 June 2015*

* Corresponding author. Address: Dept. of Gastroenterology and Hepatology, Academic Medical Center, Meibergdreef 9, 1105 AZ, Amsterdam, The Netherlands. Tel.: +31 205666012; fax: +31 206917033.

*E-mail address:* c.y.ponsioen@amc.uva.nl (C.Y. Ponsioen).

*Abbreviations:* PSC, primary sclerosing cholangitis; LTx, liver transplantation; CRS, cirrhosis related symptoms; CI, confidence interval; MRCP, magnetic resonance cholangiopancreatography; ERCP, endoscopic retrograde cholangiopancreatography; AIH, autoimmune hepatitis; CBP, copper binding protein; PBC, primary biliary cirrhosis; IQR, interquartile range; AST, aspartate aminotransferase; ALT, alanine aminotransferase; ALP, alkaline phosphatase; γGT, gamma-glutamyl transferase; HE, haematoxylin and eosin; CA, cholangitis activity; HA, hepatitis activity; HR, hazard ratio; xULN, times upper limit of normal; IQR, interquartile range; r, correlation coefficient; UDCA, ursodeoxycholic acid.

## Introduction

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

Histologic changes seen in PSC include the characteristic periductal concentric fibrosis leading to bile duct obliteration, infiltration of inflammatory cells in portal tracts, loss of bile ducts, bile ductular reaction and focal accumulation of copper binding protein (CBP) [1,14,15]. Using liver histology, disease severity and progression can be assessed in terms of grade and stage [16]. Grade is usually used to describe the degree of

necroinflammatory activity as measurement of the severity of the underlying disease process, while stage generally reflects the degree of fibrosis and cirrhosis as measurement of disease progression [16].

If a PSC appropriate histologic scoring system would have prognostic significance in terms of predicting the occurrence of solid clinical endpoints, liver histology may be an important candidate for the evaluation of treatment efficacy in therapeutic trials. Evaluation of treatment efficacy by solid clinical endpoints such as death or liver transplantation (LTx), is hindered by the chronic disease course and the low prevalence of PSC [18]. Therefore, surrogate endpoints, including clinical biomarkers, biochemical biomarkers, Mayo risk score, and liver histology have often been used, but were never validated [7–13]. Currently, the lack of properly validated surrogate endpoints for clinical trials is one of the major challenges in PSC research. An important asset of liver histology is its face-validity, meaning that liver biopsy directly measures the degree of disease severity in the affected organ. However, it is currently unknown if face-validity is maintained in PSC livers where the patchy distribution may give rise to confounding sampling variability [19].

At present, there is no specific PSC histologic scoring system with clinical significance to evaluate both disease grade and stage. Commonly, the Ludwig and Ishak systems have been used to grade and stage histologic disease severity in PSC [17,20]. A drawback of the Ludwig staging system is that it was designed primarily to assess disease progression of primary biliary cirrhosis (PBC) [17]. Furthermore, the Ludwig system does not separately score disease grade, and instead incorporates features such as portal and periportal inflammation, which are probably better regarded as manifestation of disease grade rather than stage.

Recently, Nakanuma *et al.*, have proposed a new grading and staging system for PBC, which takes into account particular features that are shared with PSC, such as the presence of CBP and loss of bile ducts [21].

The aim of this study was to determine the prognostic value of three different scoring systems, designed primarily to assess disease grade and/or stage in chronic hepatitis (Ishak) or PBC (Ludwig and Nakanuma), for grading and/or staging of PSC [17,20–22].

## Patients and methods

### *Study design*

This cohort study is part of the 'Epi PSC PBC project', a large population-based cohort study of PSC and PBC in the Netherlands. All PSC patients alive on January 1st 2000, and living in a geographically defined area of six adjacent provinces comprising 50% of the Dutch population were included in this study, between January 1st 2008 and December 31st 2011. The case-finding and case-ascertainment methods have been described previously [18]. The protocol was approved by the central Committee for Research Ethics in Utrecht and all 44 local ethics committees of the participating hospitals in the Netherlands (trialregister.nl number, NTR2813).

PSC diagnosis was based on: 1) elevated alkaline phosphatase and gamma-glutamyl transferase (γGT), not explained otherwise; 2) presence of characteristic bile duct changes with multifocal strictures and segmental dilatations on ERC or magnetic resonance cholangiography (MRC); and/or 3) liver histology and; 4) no evidence for secondary sclerosing cholangitis. When criteria 1, 3, and 4 were fulfilled in the absence of cholangiographic abnormalities on MRC or ERC, cases were diagnosed as small duct PSC [6]. AIH overlap syndrome (PSC-AIH) is ill defined. A diagnosis of PSC-AIH was made in patients with a characteristic cholangiogram who, in addition, met the simplified AIH criteria [23].

PSC patients from the Epi PSC PBC cohort, who underwent diagnostic liver biopsy or liver biopsy to assess disease severity at time of diagnosis between 1978 and 2011, were included. Patients with PSC-AIH overlap syndrome were excluded. Biochemical values of aspartate aminotransferase (AST), alanine aminotransferase (ALT), alkaline phosphatase (ALP), γGT and total bilirubin at time of liver biopsy (range two months before until one month after liver biopsy) were retrieved from hospital databases and Mayo risk score was calculated. Clinical data reflecting liver CRS at follow-up were collected from patient files. CRS included gastro-esophageal varices and variceal bleeding, both assessed by gastrointestinal endoscopy, ascites, and splenomegaly assessed clinically and/or by imaging, and hepatic encephalopathy.

### *Tissue preparation and histologic evaluation*

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

### *Grading*

Grading was scored according to the Ishak system, evaluating degree of interface hepatitis (score 0–4), confluent necrosis (score 0–6), lobular inflammation (score 0–4) and portal inflammation (score 0–4) [20]. Furthermore slides were scored according to the Nakanuma system, encompassing degree of cholangitis activity (CA; score 0–3) and hepatitis activity (HA; score 0–3) [21] (Supplementary Table 1).

### *Staging*

Staging was performed according the Ishak, as well as the Nakanuma and Ludwig systems [17,20,21]. With Ishak staging system degree of fibrosis is evaluated (0–6). The Nakanuma staging system is based on semi-quantitative scoring of three histological features; fibrosis (score 0–3), bile duct loss (score 0–3) and deposition of orcein positive granules (score 0–3). The final Nakanuma stage is obtained from the total score of these three features. Stage I (no or minimal progression) is a score of 0, stage II (mild progression) a score of 1–3, stage III (moderate progression) a score of 4–6 and stage IV (advanced progression) a score of 7–9 [24]. Ludwig staging system consists of four stages; stage I, cholangitis or portal hepatitis; stage II, periportal fibrosis or hepatitis; stage III, septal fibrosis, bridging necrosis or both; and stage IV, biliary cirrhosis (Supplementary Table 2).

### *Endpoints*

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

### *Statistical analysis*

Patient characteristics were expressed as mean ± standard deviation and median and interquartile range (IQR) where ever appropriate.

Association of histologic grade and stage with transplant-free survival, time to LTx and development of CRS was estimated using Kaplan–Meier survival curve and Wilcoxon log-rank test. Due to relatively small sample size, and the resulting small amount of patients per grade and stage, survival analyses were performed in grouped subcategories. In this reclassification the order of severity of grade/stage was maintained; those subgroups with very few or no patients were grouped together with the grade/stage of most similar severity. Ishak grading components interface hepatitis, lobular inflammation and portal inflammation were reclassified in score 0, 1, ⩾2 (original score: 0–4), component confluent