Automated image analysis of keratin 7 staining can predict disease outcome in primary sclerosing cholangitis

Sjöblom, Nelli

2023-04

ORIGINAL ARTICLE

Hepatology Research

Automated image analysis of keratin 7 staining can predict disease outcome in primary sclerosing cholangitis

Nelli Sjöblom[1] ⓘ | Sonja Boyd[1] | Anniina Manninen[2] | Sami Blom[2] |
Anna Knuuttila[2] | Martti Färkkilä[3] | Johanna Arola[1]

[1]Department of Pathology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

[2]Aiforia Technologies Oyj, Helsinki, Finland

[3]Department of Gastroenterology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
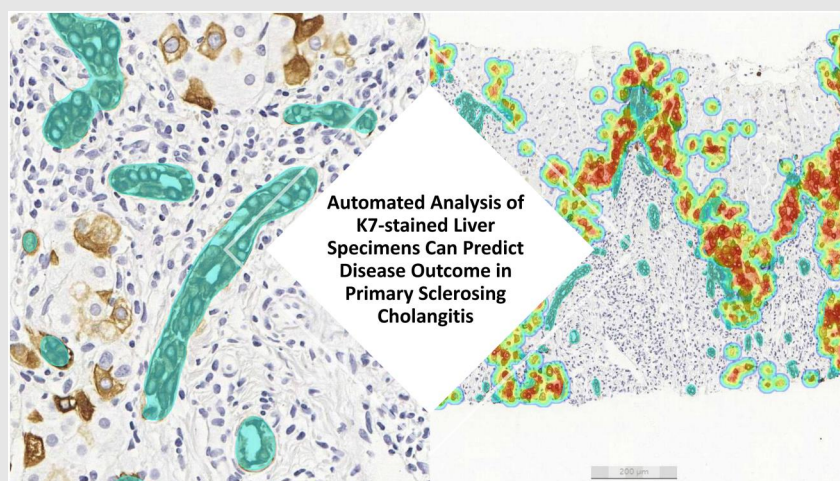
Correspondence
Nelli Sjöblom, Department of Pathology, University of Helsinki and Helsinki University Hospital, Haartman-institute, Haartmaninkatu 3, 00290 Helsinki, Finland.
Email: nelli.sjoblom@hus.fi and nelli.sjoblom@helsinki.fi

Graphical Abstract

Automated Analysis of K7-stained Liver Specimens Can Predict Disease Outcome in Primary Sclerosing Cholangitis

Primary sclerosing cholangitis (PSC) is a chronic cholestatic liver disease that can lead to liver cirrhosis or cholangiocarcinoma. Here, we developed a K7-AI model 2.0 to analyze K7-stained liver specimens of patients with PSC. We found that the K7-AI model 2.0 can serve as a prognostic tool for clinical endpoints, since it was able to provide significant information on disease outcomes based on different histological features.

Abstract

Background and Aims: Primary sclerosing cholangitis (PSC) is a chronic cholestatic liver disease that obstructs the bile ducts and causes liver cirrhosis and chol-angiocarcinoma. Efficient surrogate markers are required to measure disease

progression. The cytokeratin 7 (K7) load in a liver specimen is an independent prognostic indicator that can be measured from digitalized slides using artificial intelligence (AI)-based models.

Methods: A K7-AI model 2.0 was built to measure the hepatocellular K7 load area of the parenchyma, portal tracts, and biliary epithelium. K7-stained PSC liver biopsy specimens ($n = 295$) were analyzed. A compound endpoint (liver transplantation, liver-related death, and cholangiocarcinoma) was applied in Kaplan–Meier survival analysis to measure AUC values and positive likelihood ratios for each histological variable detected by the model.

Results: The K7-AI model 2.0 was a better prognostic tool than plasma alkaline phosphatase, the fibrosis stage evaluated by Nakanuma classification, or K7 score evaluated by a pathologist based on the AUC values of measured variables. A combination of parameters, such as portal tract volume and area of K7-positive hepatocytes analyzed by the model, produced an AUC of 0.81 for predicting the compound endpoint. Portal tract volume measured by the model correlated with the histological fibrosis stage.

Conclusions: The K7 staining of histological liver specimens in PSC provides significant information on disease outcomes through objective and reproducible data, including variables that cannot be measured by a human pathologist. The K7-AI model 2.0 could serve as a prognostic tool for clinical endpoints and as a surrogate marker in drug trials.

KEYWORDS
artificial intelligence, ductular reaction, liver histology, primary sclerosing cholangitis, surrogate marker

INTRODUCTION

Primary sclerosing cholangitis (PSC) is a chronic cholestatic liver disease that predisposes patients to liver cirrhosis and cholangiocarcinoma (CCA).[1] Liver biopsy is no longer regarded necessary for the diagnosis of typical cases of PSC,[2] as the diagnostics rely on imaging modalities, mainly on endoscopic retrograde cholangiography (ERC) and/or magnetic resonance cholangiopancreatography (MRI-MRCP). Despite its well-known limitations, liver histology has been used as an endpoint in several recent clinical drug trials for PSC, such as sintuzumab,[3] nor-UDCA (NCT03872921), and cilofexor (NCT03890120).[4,5]

Currently, there is a lack of internationally validated scoring systems that strictly rely on the prognostic histological features of PSC. Liver fibrosis, assessed histologically by Ludwig stage, has been demonstrated to be independently associated with death and/or liver transplantation in several studies.[6–9] More recently, independent prognostic values of the Nakanuma, Ishak, and Ludwig staging systems were documented in PSC, with the Nakanuma system showing the strongest predictive value in a multicenter cohort study.[10] In particular, the disease stage and tissue fibrosis evaluated by the

Nakanuma system have shown strong correlations with clinical endpoints.[2,3] This result was later validated in our recent work, but with a larger patient cohort and major clinical endpoints (CCA, liver transplantation, and liver-related death), where we reported the prognostic value of a novel histological scoring system for PSC liver specimens, the PSC histoscore.[11]

Ductular reaction (DR) and bile duct regeneration frequently encountered histological findings in PSC.[12–14] DR, characterized by periportal proliferation and hyperplasia of bile ducts, has been described as a regenerative process of liver tissue in a variety of chronic cholestatic liver diseases and autoimmune hepatitis.[15,16] DR in injured liver tissue can be visualized using cytokeratin 7 (K7).[12,17] DR has been shown to be a histological prognostic marker in primary biliary cholangitis.[13]

K7-positivity in liver specimens has been established as a histological marker of chronic cholestasis in a variety of chronic liver diseases because of its tendency to highlight cholestatic hepatocytes in a distinct manner.[12,17–21] In addition, the evaluation of liver specimens of patients with primary biliary cholangitis has shown potential value as a marker of disease progression.[22–24] In our previous study, we showed that the quantity of K7-positive hepatocytes

in a cohort of PSC liver biopsy specimens correlated with a common clinical marker, the plasma alkaline phosphatase (P-ALP) levels used to assess PSC prognosis.[25]

In a recently published paper by Saffioti et al., the collagen proportionate area, measured by an automated image analysis method, produced a significant prognostic value in a cohort of patients with PSC.[26] These results imply that a more sophisticated analysis of liver specimens for the evaluation of its role as a prognostic marker is warranted.

Image analysis by deep learning neural networks is a promising development, and many applicable solutions in different platforms have been created to assist in complex digital image analysis tasks, routinely performed by human pathologists.[27-30] The ability to produce quantitative, objective, and rapidly available data is the key component of this type of technology of higher quality and better prognostic value to aid in diagnostic processes.[31] By applying these previously described methods within Aiforia's cloud-embedded platform, we have recently trained and validated an artificial intelligence (AI) model for automated image analysis to quantify and locate cytokeratin 7 (K7)-positive hepatocytes in K7-stained histological liver specimens. These over 200 liver specimens were of patients diagnosed with PSC.[25]

Here, we aimed to study the correlation between tissue K7 load and DR quantity with major clinical endpoints in a cohort of patients with PSC. We aimed to build a modified and improved version of the recently developed and validated AI model for automated detection of K7-positive hepatocytes, a K7-AI model.[25] The object was that the modified K7-AI model 2.0 would be able to recognize the DR, the area of the portal tracts, as well as various parameters of hepatocellular K7 load, which cannot be quantified with the human eye.

METHODS

Patient cohort and endoscopic retrograde cholangiography

The Helsinki University Hospital (HUH) and the Department of Gastroenterology hold a registry of over 1000 patients with PSC. We selected a cohort of 295 patients diagnosed with PSC between 1988 and 2018, to ensure adequate follow-up time. Laboratory parameters within 3 months from the date of liver biopsy (P-ALP, serum bilirubin [S-Bil], plasma gamma-glutamyl transferase [P-GGT], plasma AST [P-AST], plasma albumin [P-ALB], and plasma creatinine), ERC scores, and pathology reports (histology and cytology) were retrieved from the registry. Patients with overlapping syndromes were excluded from this study.

PSC diagnosis was based on MRI-MRCP and ERC. The indications for ERC at our center were: (1) documentation of PSC diagnosis owing to elevated or fluctuating S-ALP levels in conjunction with inflammatory bowel disease (IBD); (2) MRCP findings or liver histology suggestive of PSC; or (3) surveillance of disease progression and/or biliary dysplasia (EASL). The Helsinki score (modified Amsterdam score, mERC score) was used to evaluate ERC images.[32] If only mild intrahepatic changes are seen (intrahepatic score ≤4), a liver biopsy is performed to a verify the PSC diagnosis.

Written informed consent was obtained from all the patients. The study protocol conformed to the ethical guidelines of the 1975 Declaration of Helsinki, as reflected in a priori approval by the institution's human research committee (license number 278/13/03/01/09, Ethical Statement of the Internal Medicine § 305, HUH).

Liver biopsy specimens and immunohistochemistry

In our center, a histological core needle liver biopsy is sampled from most patients with PSC to exclude concomitant autoimmune hepatitis or furthermore to confirm PSC diagnosis in cases where only mild intrahepatic disease is detected in ERC. Liver biopsy specimens used in this study were collected from the archives of the Biobank of Helsinki.

K7 staining of the PSC liver biopsy specimens was performed routinely to evaluate chronic cholestasis and bile duct loss. In addition, Herovici staining (which highlights tissue collagen) was performed to evaluate the stage of fibrosis. K7 and Herovici staining were performed as previously described.[25]

Liver biopsy specimens scanning and scoring

K7-stained sections were scanned using a 3DHistec Pannoramic SCAN II 150 scanner to obtain digital whole slide images (WSIs). The original pixel size was 0.221 μm and the object magnification was ×20.

All slides were first digitally evaluated by a liver pathologist (S.B.) for chronic cholestasis according to the adapted Nakanuma classification.[33] For scoring of the K7-stained slides, please see Appendix A. In addition, the Herovici-stained and scanned liver biopsy specimens of the same cohort were re-evaluated for fibrosis stage as described previously.[25] Biopsy specimens with at least six portal tracts were included.

AI model and software analysis of parameters from the liver biopsy specimens

The WSIs were uploaded to Aiforia Cloud, which is a commercial cloud-based platform where it is possible to view WSIs and train AI models based on convolutional neural networks (CNNs). The AI model design was based on the K7-AI model validated in a recent publication.[25] The threshold of a K7-positive hepatocyte was as described previously; however, the existing K7-AI model was updated to Version 2.0, which included the following modifications:

1. An additional CNN was inserted in the AI model design to detect bile duct epithelium histology, and this CNN was trained to detect

biliary epithelium, DR, and bile ducts from K7-stained liver biopsy specimens.

2. Annotations were added to the liver parenchyma and portal areas in a training set of 70 histological images of PSC liver specimens as input data to the neural network via supervised learning, to improve the performance of the CNNs of the AI model. The number of included annotations and the final design of the K7-AI model 2.0 is shown in Appendices B and C, respectively. The ground truth for annotations per CNN is shown in Appendix D.

By applying the K7-AI model 2.0, we were able to analyze: (1) the areas of liver parenchyma, including only hepatocytes (μm^2); (2) K7-positive hepatocytes (μm^2) per biopsy specimen; (3) number of K7-positive hepatocytes and (4) their mean distance from the closest portal area (μm); (5) area of portal tracts per liver specimen; and (6) area of biliary epithelium, including the original bile ducts and DR.

Clinical endpoints

Because we had a limited number of clinical endpoints ($n = 34$), we created a compound endpoint that included liver transplantation ($n = 21$), CCA ($n = 12$), and/or PSC-related deaths ($n = 1$). The risk for CCA in PSC is increased by disease duration with an annual risk of 0.5%–1.5%.[34] Thus, CCA was included as one of the important PSC-related clinical endpoints. In our center, the indications for liver transplantation were repeatedly documented biliary dysplasia and/or high-grade dysplasia ($n = 5$), severe symptoms ($n = 4$), or end-stage liver disease ($n = 12$). Patients who did not reach the endpoint ($n = 261$) were followed up to October 30, 2020. Cirrhosis was clinically determined by imaging or by diagnosing esophageal varices. Patients with an endpoint of less than a year from the biopsy date were not included in the cohort.

In our cohort, most of the patients with PSC are diagnosed at an early stage, mostly due to structured surveillance protocols of patients with IBD, where IBD-associated PSC cases are found as soon as elevated liver examinations are discovered. Moreover, in our biopsy cohort, there were several patients, who had no or a limited amount of fibrosis as an indicator of disease progression in their samples.

Statistical analysis

The analyses were performed by applying the R statistical programming software.[35–38] Combinations of variables used in the analyses were (1) laboratory examinations and compound endpoint, and (2) variables determined by the K7-AI model and compound endpoint by the human K7 score determined by the pathologist and compound endpoint.

For the optimal cut point, the parameters given for the cutpointr[39] function used were the maximum-metric as the method and Youden as the metric. The AUC, sensitivity, and specificity were obtained from the results of the cut-off point function. Positive and negative likelihood ratios were calculated based on the sensitivity and specificity values produced by the cut-off point function. For the Kaplan–Meier analysis, the patients were divided into two groups based on the obtained cut-off point value.

First, we calculated the optimal cut-off points for the K7-AI model 2.0 variables, laboratory examinations, and human K7-score values versus compound endpoint. Based on the cut-off points, we divided the values into two groups for every K7-AI model variable, laboratory examination, and their endpoint combinations. For the human K7 score, the cut points were either 1 or 2, and for the fibrosis score (0–3) the cut point was 1.

The specimens that had values under the cut-off point formed one group, and the specimens that had values larger than the cut-off point formed another group. These groups were used in the Kaplan–Meier analysis to determine whether there was a difference in survival between these two groups. The Kaplan–Meier method was applied to estimate the cumulative survival for the compound endpoint and ERC/dilatation endpoint. Statistical significance was set at $p < 0.005$.

RESULTS

K7-AI model 2.0 detects ductular reaction and K7 load

By using K7-stained liver specimens from a large cohort of patients with PSC (please see Table 1 for demographics), we were able to build an AI model that detects and measures several parameters from histological images. The first version of the K7-AI model could identify liver tissue, portal tracts, and liver parenchyma, in addition to K7-positive hepatocytes and their distance from the closest portal tract. The additional histological parameters detected by the K7-AI model 2.0 are the size of portal tracts and the volume of biliary epithelium, including the DR. Figure 1 demonstrates the results of the further developed AI model with the inference masks and how the AI model works in each nested CNN.

In addition to the visual results and discriminative inference masks, we were able to retrieve numeric data from the AI model on the histological parameters that were detected from the histological images. The quantities of each histological variable found by the AI model are listed in the study material in Appendix E.

K7-AI model 2.0 is superior in predicting disease outcome compared with the laboratory parameters or with the human score

Histological parameters detected by the K7-AI model 2.0, such as the K7 load, correlated with clinical outcome—the compound endpoint—and the predictive value was higher than that of traditional scoring by pathologists. The AUC for human K7 scores was 0.72 for the

TABLE 1 Clinical characteristics and laboratory values of patients based on compound endpoint

Variable	Compound endpoint reached	
	Yes, $n = 34$	No, $n = 261$
Female, n (%)	9 (26)	110 (42)
Age at end of follow-up, mean (SD)	47 (12)	44 (14)
Age at PSC diagnosis, mean (SD)	34 (15)	35 (13)
P-Crea, U/L, median (IQR)	63 (56, 72)	70 (61, 79)
P-ALB, g/L, median (IQR)	36 (32, 39)	39 (36, 41)
P-GGT, U/L, median (IQR)	235 (100, 336)	119 (42, 278)
P-ALP, U/L, median (IQR)	247 (175, 367)	131 (97, 214)
P-AST, U/L, median (IQR)	73 (47, 103)	37 (28, 56)
P-Bil, total, μmol/L, median (IQR)	77 (48, 161)	51 (27, 91)
ERC score at baseline, median (IQR)	7 (6, 10)	4 (2, 6)
ERC score at the end of follow-up[a], median (IQR)	9.5 (6, 12)	4 (2, 6)
Nakanuma stage of fibrosis 0 in liver specimen	3	83
Nakanuma stage of fibrosis 1 in liver specimen	4	87
Nakanuma stage of fibrosis 2 in liver specimen	21	83
Nakanuma stage of fibrosis 3 in liver specimen	6	8
Chronic cholestasis score 0 in liver specimen	5	153
Chronic cholestasis score 1 in liver specimen	6	47
Chronic cholestasis score 2 in liver specimen	6	33
Chronic cholestasis score 3 in liver specimen	17	28

Note: The laboratory examinations were obtained from the PSC registry +/− 3 months based on the biopsy date.

Abbreviations: CD, Crohn's disease; ERC, endoscopic retrograde cholangiography; IBD, inflammatory bowel disease; ID, indeterminate disease; P-ALB, plasma albumin; P-ALP, plasma alkaline phosphatase; P-AST, plasma AST; P-Bil, plasma bilirubin; P-Crea, plasma creatinine; P-GGT, plasma gamma-glutamyl transferase; PSC, primary sclerosing cholangitis; UC, ulcerative colitis.

[a]Follow-up ended on October 30th, 2020, or on the date of the reached endpoint..

compound endpoint, whereas all AUC values for parameters assessed by the K7-AI model were ≥0.8 for (1) area of K7-positive hepatocytes, (2) number of K7-positive hepatocytes, (3) number of K7-positive hepatocytes per area of liver specimen, and (4) number of K7-positive hepatocytes per area of liver specimen × mean distance of K7-positive hepatocytes closest to the portal tract, (5) area of portal tracts and biliary epithelium within the parenchyma, and (6) area of bile duct epithelium with DR had AUC values of 0.78, and 0.79, respectively. Only the mean distance of K7-positive hepatocytes from the closest portal tract had an AUC value of 0.63 for the compound endpoint (see Table 2 and Appendix F). See Table 2 for AUC values for the compound endpoint of the human K7 score, of each clinical variable (P-ALP, S-Bil, P-GGT), and of each histological parameter determined by the K7-AI model.

Appendix G shows the Kaplan–Meier survival curves and *p*-values for each laboratory parameter (P-ALP, P-GGT, S-Bil, P-AST, and P-ALB), human K7, and fibrosis scores by applying the compound endpoint. P-ALP, P-AST, and P-ALB were statistically significant predictors of disease outcome.

Combining the P-ALP values with the results of the AI model, the predictive value did not improve; in fact, it decreased. The combination of P-ALP values and K7 load assessed by the K7-AI model 2.0 produced an AUC value of 0.80 for the compound endpoint, whereas the AI model produced an AUC of 0.81. Please see Appendix H for the Kaplan–Meier survival curves for the compound variable.

Combination of histological features measured by the K7-AI model 2.0

Using a combination of several parameters measured by the AI model, we were not able to show significant improvement in the AUC values for the compound endpoint. The highest AUC (0.81) was produced by applying several different compound variables. With a combination of (1) number of K7-positive hepatocytes per area of liver specimen, (2) area of K7-positive hepatocytes within a liver specimen, and (3) volume of biliary epithelium, including DR a positive likelihood ratio of 5.7 was reached, when specificity was 90%.
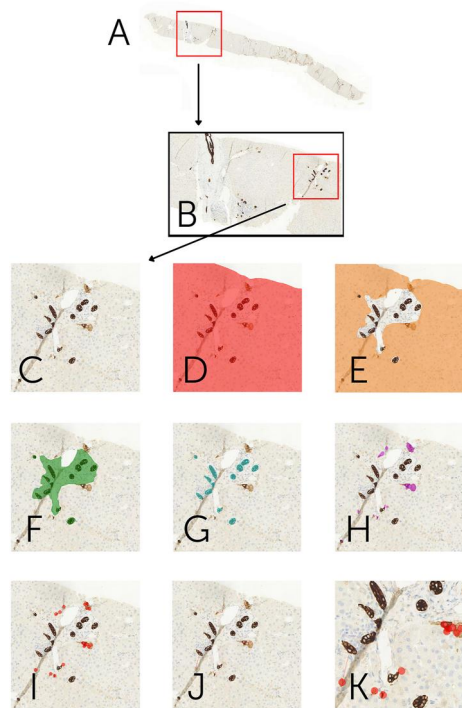
FIGURE 1 (a–c) A K7-stained core needle liver biopsy specimen and a normal portal tract with a higher magnification. The portal tract contains vascular structures and bile ducts visualized by cytokeratin 7 (intense brown). In the periportal zone, faint cytoplasmic positivity can be detected within the cytoplasm of cholestatic hepatocytes. (d) Red area indicates what is detected as liver tissue by the K7-AI model 2.0. (e) Orange area indicates what is detected as liver parenchyma. (f) Green area indicates what is detected as portal tracts and biliary epithelium within the parenchyma. (g) Turquoise indicates what is detected as biliary epithelium. (h,i) Purple shows K7-positive cholestatic hepatocytes, and red circles are marking them as single objects that can be calculated by the K7-AI model 2.0. (j,k) The K7-AI model 2.0 can calculate the distance of each K7-positive hepatocyte from the closest portal tract

See Figure 2 for the Kaplan–Meier survival analyses for all histological variables evaluated using the K7-AI model 2.0.

Liver fibrosis assessed by the K7-AI model: A novel feature in the K7-AI model

The K7-AI model was able to estimate indirectly the liver fibrosis from K7 staining, as we trained K7-AI model 2.0 to identify the portal tract volume. Table 2 shows the Spearman coefficient to demonstrate the correlation between fibrosis stage determined from Herovici-stained liver specimens by Nakanuma classification systems and the histological variables determined by K7-AI model 2.0 (areas of portal tracts and of bile duct epithelium). This correlation was statistically significant ($p < 0.001$). Boxplots of the correlation analysis are shown in Figure 3a,b. The Nakanuma fibrosis stage had an AUC of 0.61 for the compound endpoint.

DISCUSSION

Chronic cholestasis highlighted by K7 staining in liver biopsy specimens has proven to be an independent prognostic marker of disease outcomes in PSC,[10,11,17] and it correlates well with clinical markers of chronic cholestasis.[25] Thus, it is recommended to apply K7 staining of PSC liver specimens as part of the examinations included in the routine examination protocol of these patients.

The quantity of K7-positive periportal hepatocytes, namely the K7 load, can be measured by the recently developed K7-AI model.[25] Here, we trained this AI model further to produce a K7-AI model 2.0 to detect portal tracts and biliary epithelium, including DR, from PSC liver specimens. When applying this model, we found that several single parameters, such as the area of the biliary epithelium and quantity of DR, which are practically unmeasurable by the human eye, can produce prognostic information for patients with PSC. A specific target of our study was to investigate the prognostic value of individual histological features, as well as their combination, detected by the K7-AI model 2.0, in PSC.

Using the remodeled and robust K7-AI model 2.0, several specific histological features were measured. When the compound endpoint (liver-related death, CCA, and liver transplantation) was applied as the primary endpoint, several of these features had a higher prognostic impact than any of the clinical markers, such as P-ALP. In addition, the variables measured by the K7-AI model 2.0 produced a higher likelihood of reaching the clinical endpoints compared with the human K7-score, namely the K7 load evaluated by a pathologist. Most of the histological features evaluated by K7-AI model 2.0 cannot be measured by the traditional eyeballing method applied in the clinical workflow. Thus, this automated image analysis tool provides a unique method for evaluating the prognosis of patients with PSC, based on liver biopsies.

The histological fibrosis stage has been shown to have significant prognostic value in PSC.[10,33] We found that the correlation between histological fibrosis stage (Nakanuma score) and area of portal tracts and/or biliary epithelium measured by the K7-AI model 2.0 was statistically significant ($p < 0.001$). Even if K7 staining is by no means a traditionally applied staining method for evaluating tissue fibrosis, the area of portal tracts measured by the AI model correlates with the tissue fibrosis evaluated by a pathologist. Thus, by applying the K7-AI model 2.0, we were able to measure indirectly the disease outcome (the compound endpoint) from the K7-stained liver specimens, because the histological fibrosis stage has been shown to have significant prognostic value in PSC.[10,11,33] Similar results have been published by measuring collagen proportionate area using an automated method.[26] Furthermore, the AUC values for the compound endpoint were higher when the portal tract (AUC 0.78) or biliary epithelium (AUC 0.79) volumes were measured by the K7-AI model 2.0, than when the fibrosis stage was measured by a pathologist (AUC 0.62). This indicates that the prognostic value of these features measured by K7-AI model 2.0 is higher than that of the traditional fibrosis stage evaluated by a pathologist.

TABLE 2 AUC, sensitivity, specificity, positive and negative likelihood ratios, p-values, and cutpoints for each clinical and histological variable for compound endpoint

Clinical variables	AUC	Sensitivity	Specificity	Positive likelihood ratio	Negative likelihood ratio	p < 0.005	Cutpoint
P-ALP	0.72	0.74	0.68	2.31	0.39	yes	180
P-GGT	0.63	0.71	0.55	1.58	0.53	yes	135
S-Bil	0.67	0.88	0.42	1.52	0.28	yes	41
P-AST	0.74	0.85	0.57	2.01	0.26	yes	42
P-ALB	0.69	0.68	0.64	1.87	0.51	yes	37
P-Crea	0.62	0.79	0.43	1.4	0.48	no (p = 0.009)	73
Histological variables determined by the K7-AI model							
Area of K7-positive hepatocytes 5	0.81	0.76	0.73	2.85	0.32	yes	37547
Number of K7-positive hepatocytes per area of liver specimen 1	0.81	0.79	0.69	2.53	0.54	yes	1.63e-05
(Number of K7-positive hepatocytes per area of liver specimen) × mean distance of K7-positive hepatocytes from closest portal tract 2	0.8	0.79	0.68	2.46	0.3	yes	0.0007
Number of K7-positive hepatocytes	0.81	0.85	0.69	2.75	0.21	yes	195
Mean distance of K7-positive hepatocytes from closest portal tract	0.63	0.65	0.66	2.33	0.32	yes	53.6
Area of portal tracts and biliary epithelium within parenchyma 3	0.78	0.79	0.66	2.33	0.31	yes	922509
Area of bile duct epithelium including DR 4	0.79	0.68	0.79	3.21	0.41	yes	124214
Compound variables determined by the K7-AI model							
Compound variable of 1 and 4 and 5	0.81	0.59	0.9	5.69	0.46	yes	13.65
Compound variable of 1 and 2 and 5	0.81	0.79	0.69	2.58	0.3	yes	3
Compound variable of 1 and 3 and 5	0.81	0.85	0.64	2.39	0.23	yes	2.3
Compound variable of 1 and 2	0.81	0.76	0.72	2.68	0.32	yes	2.88
Compound variable of 1 and 5	0.81	0.76	0.7	2.56	0.35	yes	1.48
Compound variable of 1 and 2 and 3	0.81	0.85	0.64	2.38	0.23	yes	2.86
Compound variable of 1 and 2 and 4	0.81	0.79	0.69	2.55	0.30	yes	3.39
Compound variable (K7-AI model + clinical Variables)							
P-ALP + area of K7-positive hepatocytes	0.8	0.85	0.67	2.59	0.22	yes	13.6
Histological variables determined by pathologist							
Human K7 score (0–1 = 0, 2–3 = 1)	0.72	0.68	0.66	2.91	0.42	yes	2
Human K7 score (0 = 0, 1–3 = 1)	0.72	0.85	0.59	2.07	0.25	yes	1
Fibrosis (Nakanuma stage, 0–3)	0.61	0.91	0.32	1.33	0.28	yes	1

TABLE 2 (Continued)

Clinical variables	AUC	Sensitivity	Specificity	Positive likelihood ratio	Negative likelihood ratio	p < 0.005	Cutpoint
Spearman correlation coefficients	Area of portal tracts evaluated by the AI	Area of bile duct epithelium evaluated by the AI					
Nakanuma fibrosis (0–3) by human pathologist	0.73, p < 0.0001	0.68, p < 0.0001					

Abbreviations: P-ALB, plasma albumin; P-ALP, plasma alkaline phosphatase; P-AST, plasma AST; P-Crea, plasma creatinine; P-GGT, plasma gamma-glutamyl transferase; S-Bil, serum bilirubin.
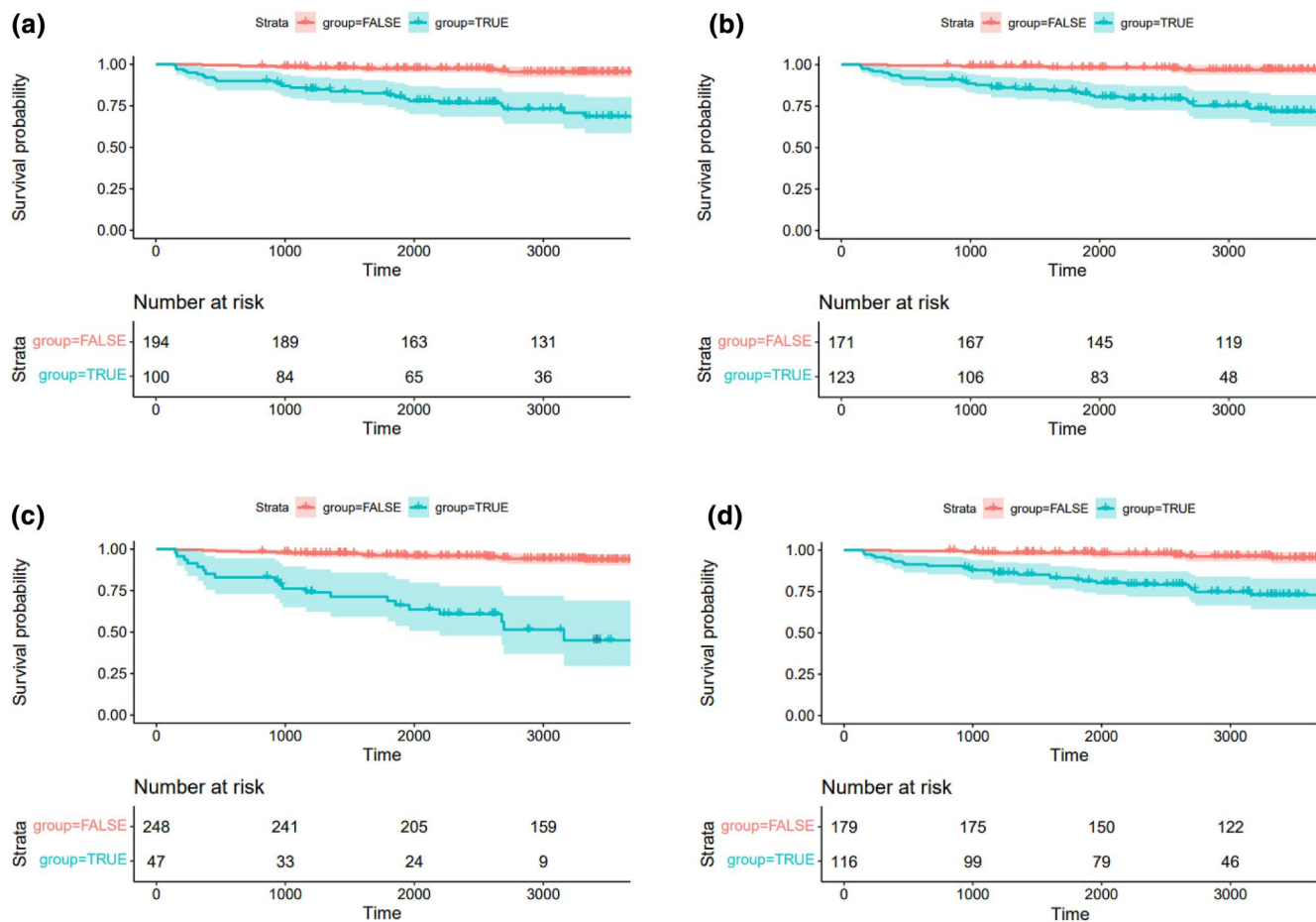

FIGURE 2 (a–d) Kaplan–Meier survival curves for the compound variables, and the portal tract volume evaluated by the K7-AI model for the compound endpoint. (a) A compound variable versus compound endpoint. The compound variable includes (1) the area of K7-positive hepatocytes, and (2) (Number of K7-positive hepatocytes per area of liver biopsy specimen) × mean distance of K7-positive hepatocytes from the closest portal tract. Group 0, specimens where the values were below the optimal cutpoint, and in group 1 were specimens that had values over the optimal cutpoint value. (b) A compound variable versus compound endpoint. The compound variable includes (1) the area of K7-positive hepatocytes, (2) (Number of K7-positive hepatocytes per area of liver biopsy specimen) × mean distance of K7-positive hepatocytes from the closest portal tract, and (3) the area of portal tracts within each sample. Group 0, specimens where the values were below the optimal cutpoint, and in group 1 were specimens that had values over the optimal cutpoint value. (c) A compound variable versus compound endpoint. The compound variable includes (1) the area of K7-positive hepatocytes, (2) (Number of K7-positive hepatocytes per area of liver biopsy specimen) × mean distance of K7-positive hepatocytes from the closest portal tract, and (3) the area of biliary epithelium within each sample. Group 0, specimens where the values were below the optimal cutpoint, and in group 1 were specimens that had values over the optimal cutpoint value. (d) The area of portal tracts within a liver biopsy specimen versus compound endpoint. Group 0, specimens where the area values were below the optimal cutpoint, and in group 1 were specimens that had area values over the optimal cutpoint value
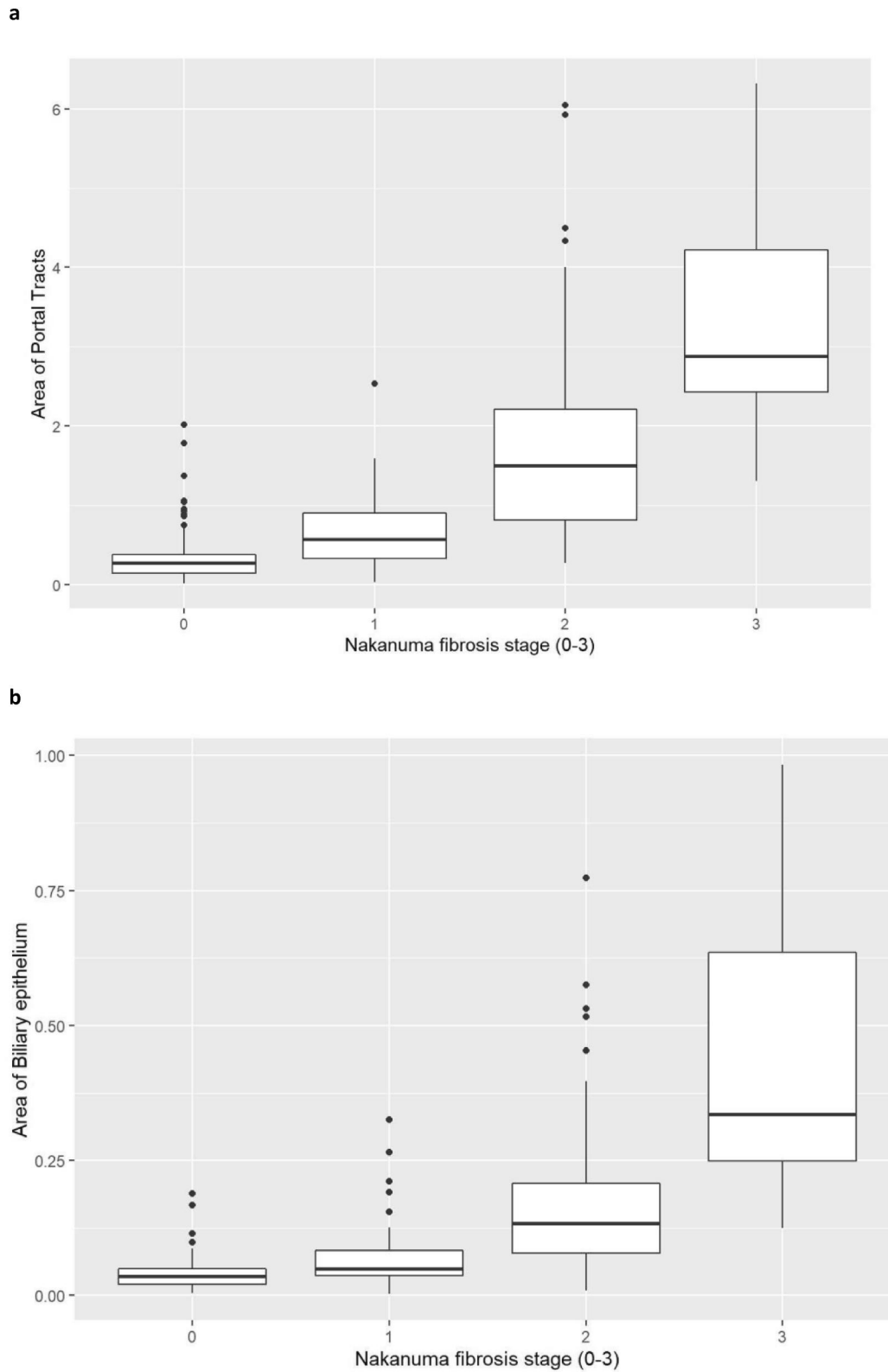
a

b

F I G U R E 3   Boxplots for Nakanuma fibrosis in correlation with K7-AI model's analysis of (1) area of portal tracts, and (2) area of biliary epithelium. (a) Nakanuma fibrosis stage (0–3) in correlation (Spearman) with the area of portal tracts. (b) Nakanuma fibrosis stage (0–3) in correlation (Spearman) with the area of biliary epithelium

The origin of the DR in liver pathology and the determinant factor that initiates the regenerative process remain unclear.[15,16] Recent findings have shown that progenitor stem-like cells play a crucial role in biliary epithelium restoration and hepatocyte regeneration.[40,41] As progenitor cells express unique regenerative qualities and seem to be associated with liver tissue fibrosis, they have become an interesting target for the development of novel therapeutic options to tackle biliary injury.[42,43] With K7-AI model 2.0, DR can be quantified precisely and reproducibly.

When assessing the prognostic value of compound variables, including several histological parameters measured by the K7-AI model 2.0, the AUC value for the compound endpoint was as high as 0.81 (better than Nakanuma fibrosis evaluated by pathologist, AUC 0.62), suggesting their potential to serve as surveillance markers to predict disease outcome. These compound variables measure different aspects of liver disease pathogenesis in cholestatic liver disease: (1) number of K7-positive hepatocytes per area of liver specimen, (2) number of K7-positive hepatocytes per area of liver specimen multiplied by the mean distance of K7-positive hepatocytes from the closest portal tract, (3) volume of biliary epithelium including DR, (4) volume of portal tracts, or (5) area of K7-positive hepatocytes (K7-load). Several different combinations were tested between these parameters, and all produced AUC values of 0.81. The K7 load demonstrates the quantity of tissue cholestasis, whereas the mean distance of K7-positive hepatocytes from the closest portal tract indirectly measures the same biological event, but from a spatial perspective. DR is a phenomenon characterized by liver regeneration, and the volume of the portal tracts is an indicator of tissue fibrosis, because the portal tracts tend to spread when liver fibrosis increases. Liver fibrosis, on the other hand, is a major histological prognostic marker in PSC, as shown in several previous publications.[10,11]

Combining P-ALP values for the analysis, the prognostic value decreased, as did the AUC values for the compound endpoint (AUC 0.80). This is an interesting finding because P-ALP is commonly regarded as a clinical surrogate marker of disease progression and a clinical indicator of chronic cholestasis.[2,44,45] Other commonly applied clinical markers of cholestasis are P-GGT and S-Bil. None of these markers showed a significant prognostic value (AUC) in this study compared with the K7-AI model 2.0. P-ALP did, however, show the highest positive likelihood ratio (2.31) of the laboratory examinations to reach the compound endpoint. The K7 load scored by the pathologist showed similar prognostic values than the P-ALP values (AUC 0.72).

As the sampling error is one of the most important limitations in the current study, the limited number of clinical endpoints ($n = 34$) might have biased the results. In addition, by applying a compound endpoint, patients with CCA or cirrhosis cannot be evaluated separately. Inflammation activity, shown to have a major prognostic value in disease progression,[11] is unmeasurable in K7-stained liver biopsy specimens. However, the Herovici staining applied at our center as part of the routine evaluation of liver samples might bring additional value in forming a unified conception of the disease status and prognosis of each patient. Furthermore, evident ductopenia was only visible in few samples of our cohort, as most of the cohort represents cases with an early-stage disease, and mild or nonexistent histological changes typical of PSC. As the number of examples of ductopenia in the training set was too low to be learned by the AI model as in independent image feature, ductopenia was excluded from the K7-AI model 2.0, even though, it is included in the original Nakanuma classification.[33] In addition to the previously mentioned limitations, we also acknowledge that, in the development and training of the K7-AI model, substantially lighter staining intensity of K7 staining might cause bias in determining true cholestatic hepatocytes.

As a major strength of this study, we emphasized the significance and size of our patient cohort. At our center, the diagnosis of PSC is based on both ERC and MRCP. Owing to our ERC surveillance program, we have unique data on bile duct disease progress combined with extensive follow-up data. Thus, we have unique material to study the biology of PSC and search for markers of bile duct disease progression. In addition, the vast amount of material includes liver biopsy specimens that were applied as a complex set of training data for the robust K7-AI model 2.0. Having a large cohort of PSC liver specimens exemplifying different stages of the disease serves as an excellent baseline to test surrogate markers.

In conclusion, as a single immunohistochemical staining method, K7 staining of PSC liver biopsy specimens provides significant information on disease progression and could serve as a surrogate marker even in drug trials to predict disease outcome. Using K7-AI model 2.0, several histological variables, which are practically unmeasurable by a human pathologist, can be detected from a single liver specimen objectively and rapidly. In addition, the results were reproducible. By combining these variables, we were able to demonstrate that this automated tool can serve as a surrogate tool for clinical endpoints. The stage of fibrosis and disease stage can be indirectly measured from a K7-stained liver biopsy specimen by measuring the area of the portal tracts or the biliary epithelium, also illustrating the DR.

We recommend including K7 staining as part of the routine follow-up protocol for PSC liver biopsies owing to its ability to produce valuable information regarding prognosis. By applying automated K7-AI model 2.0, unbiased prognostic information can be obtained in a matter of seconds.

ACKNOWLEDGMENTS

CONFLICTS OF INTEREST

The authors have declared that there are no competing interests. M. Färkkilä, S. Boyd, and J. Arola certify that they have no affiliations with or involvement in any organization or entity with any financial interest in the subject matter or materials discussed in this manuscript. S. Blom, A. Knuuttila, and A. Manninen are employees of Aiforia Technologies Oyj. A. Knuuttila is a shareholder at Aiforia Technologies Oyj. The contributing author, N. Sjöblom, has received consultation fees from Aiforia Technologies Oyj concerning projects unrelated to the study described in the manuscript. More detailed information can be provided if needed.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ETHICS STATEMENT

Approval of the research protocol by an Institutional Reviewer Board: License number 278/13/03/01/09, Ethical Statement of the Internal Medicine § 305, Helsinki University Hospital.
Informed Consent: Yes.
Registry and the Registration No. of the study/trial: N/A.
Animal Studies: N/A.
Research involving recombinant DNA: N/A

ORCID

Nelli Sjöblom https://orcid.org/0000-0002-3360-6645

REFERENCES

1. Dyson JK, Beuers U, Jones DEJ, Lohse AW, Hudson M. Primary sclerosing cholangitis. Lancet. 2018;391(10139):2547–59. https://doi.org/10.1016/s0140-6736(18)30300-3
2. Ponsioen CY, Lindor KD, Mehta R, Dimick-Santos L. Design and endpoints for clinical trials in primary sclerosing cholangitis. Hepatology. 2018;68(3):1174–88. https://doi.org/10.1002/hep.29882
3. Muir AJ, Levy C, Janssen HLA, Montano-Loza AJ, Shiffman ML, Caldwell S, et al. Simtuzumab for primary sclerosing cholangitis: phase 2 study results with insights on the natural history of the disease. Hepatology. 2019;69:684–98. https://doi.org/10.1002/hep.30237
4. Rupcic Rubin V, Bojanic K, Smolic M, Rubin J, Tabll A, Smolic R. An update on efficacy and safety of emerging hepatic antifibrotic agents. J Clin Transl Hepatol. 2021;9:60–70.
5. Trauner M, Gulamhusein A, Hameed B, Caldwell S, Shiffman ML, Landis C, et al. The nonsteroidal farnesoid X receptor agonist Cilofexor (GS-9674) improves markers of cholestasis and liver injury in patients with primary sclerosing cholangitis. Hepatology. 2019;70(3):788–801. https://doi.org/10.1002/hep.30509
6. Farrant JM, Hayllar KM, Wilkinson ML, Karani J, Portmann BC, Westaby D, et al. Natural history and prognostic variables in primary sclerosing cholangitis. Gastroenterology. 1991;100(6):1710–17. https://doi.org/10.1016/0016-5085(91)90673-9
7. Wiesner RH, Grambsch PM, Dickson ER, Ludwig J, Maccarty RL, Hunter EB, et al. Primary sclerosing cholangitis: natural history, prognostic factors and survival analysis. Hepatology. 1989;10(4):430–6. https://doi.org/10.1002/hep.1840100406

8. Dickson ER, Murtaugh PA, Wiesner RH, Grambsch PM, Fleming TR, Ludwig J, et al. Primary sclerosing cholangitis: refinement and validation of survival models. Gastroenterology. 1992;103(6):1893–901. https://doi.org/10.1016/0016-5085(92)91449-e
9. Broomé U, Olsson R, Lööf L, Bodemar G, Hultcrantz R, Danielsson A, et al. Natural history and prognostic factors in 305 Swedish patients with primary sclerosing cholangitis. Gut. 1996;38(4):610–15. https://doi.org/10.1136/gut.38.4.610
10. de Vries EM, de Krijger M, Färkkilä 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(3):907–19. https://doi.org/10.1002/hep.28963
11. Sjöblom N, Boyd S, Kautiainen H, Arola J, Färkkilä M. Novel histological scoring for predicting disease outcome in primary sclerosing cholangitis. Histopathology. 2022;81(2):192–204. https://doi.org/10.1111/his.14677
12. Van Eyken P, Sciot R, Desmet VJ. A cytokeratin immunohistochemical study of cholestatic liver disease: evidence that hepatocytes can express 'bile duct-type' cytokeratins. Histopathology. 1989;15(2):125–35. https://doi.org/10.1111/j.1365-2559.1989.tb03060.x
13. Carpino G, Cardinale V, Folseraas T, Overi D, Floreani A, Franchitto A, et al. Hepatic stem/progenitor cell activation differs between primary sclerosing and primary biliary cholangitis. Am J Pathol. 2018;188(3):627–39. https://doi.org/10.1016/j.ajpath.2017.11.010
14. Karlsen TH, Folseraas T, Thorburn D, Vesterhus M. Primary sclerosing cholangitis - a comprehensive review. J Hepatol. 2017;67(6):1298–323. https://doi.org/10.1016/j.jhep.2017.07.022
15. Burt AD, MacSween RN. Bile duct proliferation--its true significance? Histopathology. 1993;23(6):599–602. https://doi.org/10.1111/j.1365-2559.1993.tb01258.x
16. Verdonk RC, Lozano MF, van den Berg AP, Gouw AS. Bile ductal injury and ductular reaction are frequent phenomena with different significance in autoimmune hepatitis. Liver Int. 2016;36(9):1362–9. https://doi.org/10.1111/liv.13083
17. Sakellariou S, Michaelides C, Voulgaris T, Vlachogiannakos J, Manesis E, Tiniakos DG, et al. Keratin 7 expression in hepatic cholestatic diseases. Virchows Arch. 2021;479(4):815–24. https://doi.org/10.1007/s00428-021-03152-z
18. Desmet VJ. Histopathology of cholestasis. Verh Dtsch Ges Pathol. 1995;79:233–40.
19. Starling C, Bhathal PS, Quaglia A. Do orcein-positive copper-binding protein deposits and cytokeratin 7 co-localise in periportal hepatocytes in chronic cholestasis? J Clin Pathol. 2018;71(6):563–4. https://doi.org/10.1136/jclinpath-2018-205139
20. Bateman AC, Hübscher SG. Cytokeratin expression as an aid to diagnosis in medical liver biopsies. Histopathology. 2010;56(4):415–25. https://doi.org/10.1111/j.1365-2559.2009.03391.x
21. Quaglia A, Bhathal PS. Copper, copper-binding protein and cytokeratin 7 in biliary disorders. Histopathology. 2017;71(6):1006–8. https://doi.org/10.1111/his.13314
22. Yabushita K, Yamamoto K, Ibuki N, Okano N, Matsumura S, Okamoto R, et al. Aberrant expression of cytokeratin 7 as a histological marker of progression in primary biliary cirrhosis. Liver. 2001;21(1):50–5. https://doi.org/10.1034/j.1600-0676.2001.210108.x
23. Barakauskienė A, Speičienė D, Liakina V, Semuchinienė T, Valantinas J. Expression of cytokeratin 7 as a histological marker of cholestasis and stages of primary biliary cirrhosis. Medicina. 2011;47(1):31–8. https://doi.org/10.3390/medicina47010005
24. Seki H, Ikeda F, Nanba S, Moritou Y, Takeuchi Y, Yasunaka T, et al. Aberrant expression of keratin 7 in hepatocytes as a predictive marker of rapid progression to hepatic failure in asymptomatic primary biliary cirrhosis. Acta Med Okayama. 2015;69:137–44.
25. Sjöblom N, Boyd S, Manninen A, Knuuttila A, Blom S, Farkkila M, et al. Chronic cholestasis detection by a novel tool: automated

analysis of cytokeratin 7-stained liver specimens. Diagn Pathol. 2021;16(1):41. https://doi.org/10.1186/s13000-021-01102-6

26. Saffioti F, Hall A, de Krijger M, Verheij J, Hubscher SG, Maurice J, et al. Collagen proportionate area correlates with histological stage and predicts clinical events in primary sclerosing cholangitis. Liver Int. 2021;41(11):2681–92. https://doi.org/10.1111/liv.14979

27. Janowczyk A, Madabhushi A. Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. J Pathol Inf. 2016;7(1):29. https://doi.org/10.4103/2153-3539.186902

28. Chang HY, Jung CK, Woo JI, Lee S, Cho J, Kim SW, et al. Artificial intelligence in pathology. J Pathol Transl Med. 2019;53:1–12. https://doi.org/10.4132/jptm.2018.12.16

29. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology. Nat Rev Clin Oncol. 2019;16(11):703–15. https://doi.org/10.1038/s41571-019-0252-y

30. O'Mahony N, Campbell S, Carvalho A, Harapanahalli S, Velasco-Hernandez G, Krpalkova L. et al., editors. Deep learning vs. traditional computer vision 2020. Cham: Springer International Publishing. p. 128–44.

31. Wang S, Yang DM, Rong R, Zhan X, Xiao G. Pathology image analysis using segmentation deep learning algorithms. Am J Pathol. 2019;189(9):1686–98. https://doi.org/10.1016/j.ajpath.2019.05.007

32. Boyd S, Tenca A, Jokelainen K, Mustonen H, Krogerus L, Arola J, et al. Screening primary sclerosing cholangitis and biliary dysplasia with endoscopic retrograde cholangiography and brush cytology: risk factors for biliary neoplasia. Endoscopy. 2016;48(05):432–9. https://doi.org/10.1055/s-0041-110792

33. Nakanuma Y, Zen Y, Harada K, Sasaki M, Nonomura A, Uehara T, et al. Application of a new histological staging and grading system for primary biliary cirrhosis to liver biopsy specimens: interobserver agreement. Pathol Int. 2010;60(3):167–74. https://doi.org/10.1111/j.1440-1827.2009.02500.x

34. Bergquist A, Ekbom A, Olsson R, Kornfeldt D, Loof L, Danielsson A, et al. Hepatic and extrahepatic malignancies in primary sclerosing cholangitis. J Hepatol. 2002;36(3):321–7. https://doi.org/10.1016/s0168-8278(01)00288-4

35. Team RC. R: a language and environment for statistical computing. MSOR Connect. 2014;1.

36. Terry MT. A package for survival analysis in R [R package version 3.4-0]. New York: Springer; 2022.

37. Terry M, Therneau PMG. Modeling survival data: extending the Cox model. New York: Springer; 2000.

38. Alboukadel Kassambara MK, Biecek P. CRAN2021. p. Contains the function 'ggsurvplot()' for drawing easily beautiful and 'ready-to-publish' survival curves with the 'number at risk' table and 'censoring count plot'. Other functions are also.available to plot adjusted curves for `Cox` model and to visually examine 'Cox' model assumptions. Scheipl Fabian Survminer Drawing Surviv Curves Using 'ggplot2'. 0.4.9 edn.

39. Thiele C, Hirschfeld G. cutpointr: improved Estimation and Validation of Optimal cutpoints in R. J Stat Software. 2021;98(11):1–27. https://doi.org/10.18637/jss.v098.i11

40. de Jong IEM, Matton APM, van Praagh JB, van Haaften WT, Wiersema-Buist J, van Wijk LA, et al. Peribiliary glands are key in regeneration of the human biliary epithelium after severe bile duct injury. Hepatology. 2019;69(4):1719–34. https://doi.org/10.1002/hep.30365

41. Scrushy M, O'Brien A, Glaser S. Recent advances in understanding bile duct remodeling and fibrosis. F1000Res. 2018;7:F1000Res. https://doi.org/10.12688/f1000research.14578.1

42. Wang N, Kong R, Han W, Lu J. Wnt/β-catenin signalling controls bile duct regeneration by regulating differentiation of ductular reaction cells. J Cell Mol Med. 2020;24(23):14050–8. https://doi.org/10.1111/jcmm.16017

43. Sato K, Marzioni M, Meng F, Francis H, Glaser S, Alpini G. Ductular reaction in liver diseases: pathological mechanisms and translational significances. Hepatology. 2019;69(1):420–30. https://doi.org/10.1002/hep.30150

44. European Association for the Study of the Liver. EASL Clinical Practice Guidelines: management of cholestatic liver diseases. J Hepatol. 2009;51(2):237–67. https://doi.org/10.1016/j.jhep.2009.04.009

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

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.