HEPATOBILIARY

Differentiating pediatric autoimmune liver diseases by quantitative magnetic resonance cholangiopancreatography

Leah A. Gilligan[1] · Andrew T. Trout[1,2] · Simon Lam[3] · Ruchi Singh[3] · Jean A. Tkach[1,2] · Suraj D. Serai[1,4] · Alexander G. Miethke[3,5] · Jonathan R. Dillman[1,2]

Abstract

Purpose Autoimmune liver diseases (AILD), including primary sclerosing cholangitis (PSC), autoimmune sclerosing cholangitis (ASC), and autoimmune hepatitis (AIH), have overlapping clinical features but distinct management strategies and outcomes. The purpose of this study was to assess the diagnostic performance of quantitative magnetic resonance cholangiopancreatography (MRCP) parameters for distinguishing PSC/ASC from AIH in children and young adults.

Materials and methods This IRB-approved, cross-sectional study included participants from an institutional AILD registry that underwent baseline serum liver biochemistry testing and 3D fast spin-echo MRCP. The biliary tree was extracted and modeled from MRCP images using novel proprietary software (MRCP+™; Perspectum Diagnostics; Oxford, United Kingdom), and quantitative parameters were generated (e.g., biliary tree volume; number and length of bile ducts, strictures, and dilations; bile duct median/maximum diameters). Mann–Whitney U tests were performed to compare laboratory values and MRCP metrics between patient cohorts (clinical diagnosis of PSC/ASC versus AIH). Receiver operating characteristic (ROC) curves and multivariable logistic regression were used to assess diagnostic performance of serum biochemistry values and MRCP parameters for discriminating PSC/ASC from AIH.

Results Thirty percent (14/47) of MRCP exams failed post-processing due to motion artifact. The remaining 33 patients included 20 males and 13 females, with a mean age of 15.1 ± 3.9 years. Eighteen patients were assigned the clinical diagnosis of PSC or ASC and 15 of AIH. All but one quantitative MRCP parameter were significantly different between cohorts ($p < 0.05$) and predictive of diagnosis (ROC $p < 0.05$), including numbers of bile duct strictures (area under curve [AUC] = 0.86, $p < 0.0001$) and dilations (AUC = 0.87, $p < 0.0001$) and total length of dilated ducts (AUC = 0.89, $p < 0.0001$). Laboratory values were not significantly different between cohorts ($p > 0.05$). The best multivariable model for distinguishing PSC/ASC from AIH included total length of dilated ducts (odds ratio [OR], 1.08; 95% CI 1.02–1.14) and maximum left hepatic duct diameter (OR, 1.21; 95% CI 0.57–2.56) [AUC = 0.92].

Conclusion Quantitative MRCP parameters provide good discrimination of PSC/ASC from AIH.

Keywords Autoimmune hepatitis · Autoimmune sclerosing cholangitis · Biliary tree · Children · Magnetic resonance cholangiopancreatography · Primary sclerosing cholangitis · Quantitative imaging

✉ Leah A. Gilligan
leah.gilligan@cchmc.org

[1] Department of Radiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA

[2] Department of Radiology, University of Cincinnati College of Medicine, Cincinnati, OH, USA

[3] Division of Pediatric Gastroenterology, Hepatology and Nutrition, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA

[4] Department of Radiology, Children's Hospital of Philadelphia, Philadelphia, PA, USA

[5] Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, USA

Introduction

Autoimmune liver diseases (AILD) comprise a group of disorders, including autoimmune hepatitis (AIH), primary sclerosing cholangitis (PSC), and autoimmune sclerosing cholangitis (ASC) [1, 2]. ASC is an overlap syndrome with features of both AIH and PSC [3]. Notably, ASC is more prevalent in the pediatric population than in the adult population, comprising between 30 and 50% of pediatric AILD cases [3, 4]. Determining the correct AILD diagnosis in a pediatric patient can be challenging but has important implications for management strategies and clinical outcomes. AIH and ASC are typically treated with corticosteroids and/or azathioprine to induce remission of hepatocyte damage, whereas biliary injury in ASC and PSC may not respond to these immunosuppressive therapies [5–7]. To date, no medical treatment has been definitively shown to alter disease progression in PSC, and these patients may ultimately require liver transplantation [8, 9].

Magnetic resonance cholangiopancreatography (MRCP) is the preferred imaging modality for the diagnosis and follow-up of pediatric AILD patients, and it is routinely obtained in children with suspected AILD. MRCP provides a non-invasive, detailed anatomic assessment of the biliary tree, including both intra- and extra-hepatic bile ducts, without the need for intravascular contrast material. In addition, MRCP has been shown to have good sensitivity (84–86%) and specificity (94%) for identifying PSC in children and adults [10, 11]. However, interpretation of MRCP currently relies on qualitative evaluation which is limited by inter-observer disagreement [12, 13]. Furthermore, given the lack of objective data derived from MRCP examination, the potential of MRCP to provide imaging biomarkers for diagnosis, prediction of key outcomes, disease progression, and response to therapy in AILD has not yet been fully explored. Quantitative parameters derived from MRCP might serve as surrogate endpoints for clinical trials of novel therapies for AILD in the future. To this end, validating the performance of quantitative MRCP parameters as diagnostic biomarkers would be the first step before examining their predictive and dynamic properties.

Recently, a novel post-processing and quantitative image analysis software tool (MRCP+™; Perspectum Diagnostics, Oxford, United Kingdom) received 510(k) clearance for general use from the United States Food and Drug Administration (FDA). This software provides quantitative metrics of the biliary tree derived from three-dimensional (3D) MRCP images using advanced computational techniques, including artificial intelligence. The purpose of our study was to assess the diagnostic performance of these quantitative MRCP parameters for distinguishing PSC/ASC from AIH in children and young adults. We hypothesized that one or more quantitative MRCP parameter(s) would allow clinically meaningful discrimination of PSC/ASC from AIH.

Materials and methods

The institutional review board at Cincinnati Children's Hospital Medical Center approved this single-center, cross-sectional, Health-Insurance Portability and Accountability Act (HIPAA)-compliant pilot study. Imaging data were prospectively collected as part of a longitudinal study of pediatric and young adult AILD. Written informed consent was obtained from either adult patients or from parents/guardians of patients less than 18 years of age. Assent was obtained from children between ages 11 and 18 years, as appropriate. In-kind research support was provided by Perspectum Diagnostics in the form of image analysis; no financial support was provided.

Imaging data were collected from children and young adults aged 6 through 25 years with known or suspected AILD who had been enrolled in the institutional AILD registry at Cincinnati Children's Hospital Medical Center. Patients were excluded from the registry if they had any of the following: (1) history of liver transplantation; (2) chronic hepatitis B or C infection; (3) pregnancy; (4) absolute contraindication to magnetic resonance imaging (MRI); (5) diagnosis of cystic fibrosis, biliary atresia, Langerhans cell histiocytosis, or other non-PSC cholangiopathies; (6) diagnosis of cardiac hepatopathy; or (7) diagnosis of Wilson's disease, alpha-1 antitrypsin deficiency, or glycogen storage disease. Data analyzed in the current study are from a subset of consecutively recruited registry participants.

MRCP protocol

All patients underwent research 3D fast spin-echo (FSE) MRCP imaging on a 1.5 T MRI scanner (Ingenia; Philips Healthcare, Best, The Netherlands) following at least four hours of fasting. Detailed acquisition parameters are presented in Table 1. MRCP images were acquired using a 16-channel phased-array anterior surface coil and respiratory triggering. Respiratory triggering was performed via a pneumatic sensor placed on the upper abdomen; data acquisition occurred during the quiescent portion of end expiration.
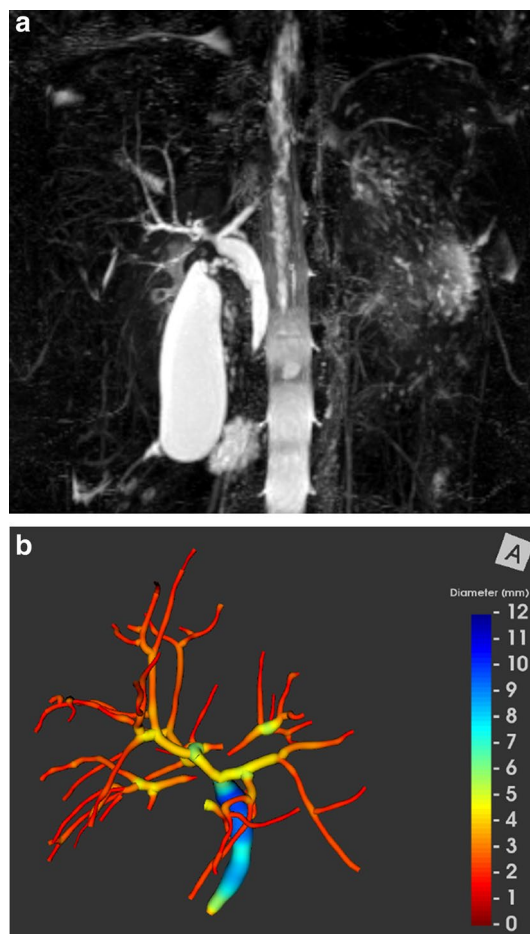
Image post-processing

3D MRCP exams were post-processed using MRCP+™ software (Perspectum Diagnostics) to extract and create a 3D model of the biliary tree and to derive related quantitative metrics. MRCP+™ analysis included the following steps: (1) tubular structures, including bile ducts, were enhanced

Table 1 MRCP pulse sequence parameters

Imaging platform	Ingenia (Philips Healthcare)
Field strength (T)	1.5
Sequence	3D fast spin-echo
Coil	16 element anterior torso coil; 12 element posterior array, integrated into table
Scan plane	Coronal
Frequency direction	S/I
FOV (mm)	256×256
Matrix	256×177
No. signals acquired	1
TE (ms)	600
Fat suppression	SPIR—strong
TR (ms)	1204
TSE factor	160
Refocusing flip angle (degrees)	80
Bandwidth (Hz)/pixel	278
No. sections	120
Slice thickness (mm)	1
In-plane voxel size (mm^2)	1×1
Acceleration	SENSE
Acceleration factors	2 in plane (R/L), 1.5 slice select (A/P)

MRCP magnetic resonance cholangiopancreatography, *T* tesla, *S/I* superior/inferior, *FOV* field of view, *TE* echo time, *SPIR* spectral presaturation inversion recovery, *TR* repetition time, *TSE* turbo spin echo, *SENSE* sensitivity encoding, *R/L* right/left, *A/P* anterior/posterior

using anisotropic diffusion followed by Frangi's multi-scale vessel enhancement filtering [14, 15]; (2) binarization was performed to identify connected components using a proprietary modification of Otsu's thresholding algorithm [16]; (3) pancreatobiliary components of interest were distinguished from uninteresting components (for example, gastrointestinal structures and blood vessels) by an expert operator, an MRI technologist with over 5 years of experience and trained in MRCP+ analysis, who was blinded to patients' assigned clinical diagnoses (see below); (4) an initial path through part of the biliary tree was determined by applying an intelligent path search algorithm using features from the Frangi analysis together with features from gradient vector flow and other information [17]; (5) the entire biliary tree was traversed by recursively following branches that arose from the initial path using the same methods as in (4); (6) biliary tree paths were refined with proprietary algorithms; (7) the diameter of each duct in perpendicular cross section was quantified at all points of the tree to achieve sub-voxel accuracy for the duct center lines and diameter measurements, and (8) ducts of interest (e.g., common bile duct [CBD], right hepatic bile duct [RHBD], and left hepatic bile duct [LHBD]) were identified by the expert operator. The

Fig. 1 Selected images from a 17-year-old girl with PSC including a maximum intensity projection 3D MRCP image (**a**) and the corresponding 3D biliary tree model derived from these data after postprocessing with MRCP+™ (strictures=3, dilations=9, length of dilated ducts=59.3 mm) (**b**). Colors correspond to duct diameter according to the scale in image (**b**). Note that areas of bile duct discontinuity represent artifact (e.g., motion, crossing blood vessel)

minimum threshold for the detection of a pancreatobiliary structure (e.g., biliary duct) was set by the acquisition resolution (1 mm isotropic), and no pancreatobiliary structures were excluded from the 3D model. On average, this postprocessing takes 15–20 min per examination.

From the constructed 3D model, presented as a 3D color-coded rendering of the biliary tree (Fig. 1), quantitative metrics were derived, including: (1) biliary tree volume; (2) median CBD diameter; (3) maximum CBD diameter; (4) median RHBD diameter; (5) maximum RHBD diameter; (6) median LHBD diameter; (7) maximum LHBD diameter; (8) number of modeled bile ducts; (9) total length of biliary tree; (10) number of biliary duct strictures (defined as local minima that were more than 30% narrower than neighboring maxima, after a proprietary algorithm is used to identify significant extrema);

(11) total length of biliary duct strictures; (12) number of biliary duct dilations (defined as local maxima that were more than 30% wider than neighboring minima); and (13) total length of biliary duct dilations. All length and diameter measurements were reported in millimeters.

Clinical and laboratory data

Serum liver biochemistries were obtained as part of routine clinical care at the time of the research MRCP. The following values were recorded for each patient: (1) total bilirubin (mg/dL); (2) alkaline phosphatase (ALP, U/L); (3) gamma-glutamyl transferase (GGT, U/L); (4) albumin (g/dL); (5) aspartate transaminase (AST, U/L); and (6) alanine aminotransferase (ALT, U/L). The registry also was used to establish patient age (at time of MRCP), sex, height, weight, and time between diagnosis and MRCP.

Assignment of clinical diagnosis

A team of pediatric hepatologists at Cincinnati Children's Hospital Medical Center assigned each registry participant a diagnosis of either AIH, PSC, or ASC based on established guidelines [1, 18, 19]. The assigned clinical diagnosis was used to place patients into one of the two cohorts for our study: (1) AIH and (2) PSC or ASC. PSC and ASC were grouped together given similarities in MRCP findings and clinical outcomes [1].

A clinical diagnosis of PSC was assigned on the basis of clinical history, biochemical features of cholestasis, radiologic findings compatible with cholangiopathy (from clinical MRCP, not the research MRCP examinations included in the current study), and/or histopathologic findings from liver biopsy with typical findings of PSC [18]. Patients were classified as having large duct PSC (or ASC) based on the presence of strictures and/or dilations from a clinical MRCP examination in which the interpreting radiologist had clinical information available to her/him. Patients with small duct PSC (or ASC) had the presence of histopathologic features of PSC in the setting of a normal clinical MRCP examination.

Patients were assigned the diagnosis of AIH if they met the international autoimmune hepatitis working group simplified criteria, including elevation of serum gamma globulin levels, autoantibodies (antinuclear, smooth muscle, and liver–kidney–microsomal), and liver histopathology compatible with AIH without radiologic or histopathologic evidence for cholangiopathy [19]. Patients with features of both AIH and PSC were classified as having ASC [1].

Statistical analysis

Continuous data were summarized either as means and standard deviations (SD) or medians and interquartile ranges (IQR); categorical data were summarized as counts and percentages. Student t test (two-sided) or Mann–Whitney U tests were performed to compare continuous variables (age, laboratory values, and quantitative MRCP parameters) between patient cohorts. Receiver operating characteristic (ROC) curves were generated to assess the diagnostic performance of age, laboratory values, and quantitative MRCP parameters in differentiating PSC/ASC from AIH. Area under the ROC curve (AUROC), sensitivity, and specificity were calculated for each laboratory test and quantitative MRCP parameter. Youden index was used to select the optimal cut-off value for each laboratory value and quantitative MRCP parameter [20].

Finally, multivariable logistic regression models were created to assess the diagnostic performance of combinations of covariates (i.e., age, laboratory values, and quantitative MRCP parameters) for discriminating the two patient cohorts. The best performing multivariable model was chosen based on the lowest Akaike Information Criterion (AIC) while avoiding covariate collinearity.

A p value of < 0.05 was considered statistically significant for all inference testing. 95% confidence intervals (CI) were generated, as appropriate. Statistical analyses were performed using MedCalc Statistical Software version 18.11.3 (MedCalc Software bvba, Ostend, Belgium; https://www.medcalc.org; 2019) and SAS, version 9.4 (SAS Institute, Inc, Cary, NC).

Results

Forty-seven consecutive registry participants that underwent a research MRCP were included in our study. Fourteen of 47 (30%) of the MRCP exams failed post-processing with MRCP+™ due to motion artifact. The remaining 33 patients were included in all analyses with one exception: due to variant anatomy (presence of a right accessory bile duct and absence of LHBD, the LHBD was not quantified) in one patient with PSC (thus, $n = 32$). Detailed demographic information is presented in Table 2.

Median age of included patients was 16 years (IQR 10–19 years). This was not statistically significantly different from the age of the 14 patients excluded for motion artifacts (median age: 15 years; IQR 13–17 years, $p = 0.64$). Twenty of 33 (61%) patients were male. Median time between research MRCP and diagnosis of AILD was 1.6 and 3.1 years for the PSC/ASC and AIH cohorts, respectively. Fifteen (45%) patients had AIH; 10 (30%) had PSC; and 8 (24%) had ASC. Three of the eight (38%)

Table 2 Patient demographic and laboratory data per AILD cohort (AIH vs. PSC/ASC)

Variable	AIH ($n=15$)	PSC/ASC ($n=18$)	p value
Patient demographics			
Age (years)	16.0 (10.0–19.0)	16.0 (13.8–17.1)	0.82
No. male sex	10	10	0.72**
Weight (kg)*	63.4 ± 26.7	70.3 ± 26.9	0.47***
Height (cm)*	161.0 ± 19.4	164.8 ± 16.6	0.55***
BMI (kg/m^2)*	23.5 ± 7.4	25.4 ± 8.5	0.50***
Time from diagnosis to MRI (years)	3.1 (0.7–4.1)	1.6 (0.5–4.1)	0.63
Liver biochemistry			
Alkaline phosphatase (units/L)	108.0 (89.0–192.0)	138.5 (88.3–240.0)	0.73
Total bilirubin (mg/dL)	0.5 (0.3–1.0)	0.5 (0.3–0.5)	0.43
GGT (units/L)	38.0 (13.0–68.0)	63.0 (45.5–240.3)	0.09
ALT (units/L)	32.5 (19.3–57.8)	54.0 (29.3–89.5)	0.12
AST (units/L)	33.0 (18.0–39.0)	25.5 (19.0–44.3)	0.98
Albumin (g/dL)	4.0 (3.5–4.3)	4.0 (3.5–4.2)	0.42

Median and interquartile ranges (IQR) are reported unless otherwise specified. p values derived from Mann–Whitney U tests unless otherwise noted

AILD autoimmune liver disease, *AIH* autoimmune hepatitis, *PSC/ASC* primary sclerosing cholangitis/autoimmune sclerosing cholangitis, *BMI* body mass index, *MRI* magnetic resonance imaging, *GGT* gamma-glutamyltransferase, *ALT* alanine aminotransferase, *AST* aspartate aminotransferase

*Mean and standard deviation reported

**p value derived from Fisher's exact test, two sided

***p values derived from student t test, two sided

patients with ASC demonstrated small duct involvement on clinical registry data, while no patients with PSC demonstrated isolated small duct involvement. Finally, all but two patients (both with PSC) had a liver biopsy obtained for their clinical AILD workup.

There were no significant differences in patient age or serum biochemistry values between the AIH and PSC/ASC

Table 3 Quantitative MRCP data per AILD cohort (AIH vs. PSC/ASC)

Variable	AIH ($n=15$)	PSC/ASC ($n=18$)	p value
Biliary tree volume and diameter parameters			
Biliary tree volume (mL)	2.9 (1.3–4.7)	8.4 (3.9–15.5)	0.004
Median CBD diameter (mm)	3.6 (3.1–4.4)	4.5 (3.6–6.0)	0.049
Maximum CBD diameter (mm)	5.0 (4.3–6.2)	6.2 (4.7–9.8)	0.034
Median right hepatic duct diameter (mm)	3.0 (2.6–3.5)	3.4 (2.5–4.6)	0.37
Maximum right hepatic duct diameter (mm)	4.0 (3.2–4.3)	4.7 (3.9–6.9)	0.017
Median left hepatic duct diameter (mm)	2.9 (2.6–4.0)	4.3 (3.1–5.5)*	0.047
Maximum left hepatic duct diameter (mm)	4.3 (3.4–4.5)	5.0 (4.5–6.6)*	0.012
Biliary tree narrowing/dilation parameters			
Number of visualized bile ducts	15 (9–38)	45 (27.8–73.3)	0.006
Total length of biliary tree (mm)	327.7 (187.3–801.8)	1010 (552.6–2263)	0.005
Number of strictures	1 (0–3)	5 (3–12)	0.0002
Total length of strictures (mm)	7.4 (0–28.5)	36.4 (22.9–108.2)	0.001
Number of dilations	1 (1–4)	9 (4.8–13.3)	0.0001
Total length of dilations (mm)	9.7 (3.5–23.9)	56.4 (31.8–89.6)	<0.0001

Median and IQRs reported. p values derived from Mann–Whitney U tests

MRCP magnetic resonance cholangiopancreatography, *AILD* autoimmune liver disease, *AIH* autoimmune hepatitis, *PSC/ASC* primary sclerosing cholangitis/autoimmune sclerosing cholangitis, *IQR* interquartile range, *CBD* common bile duct
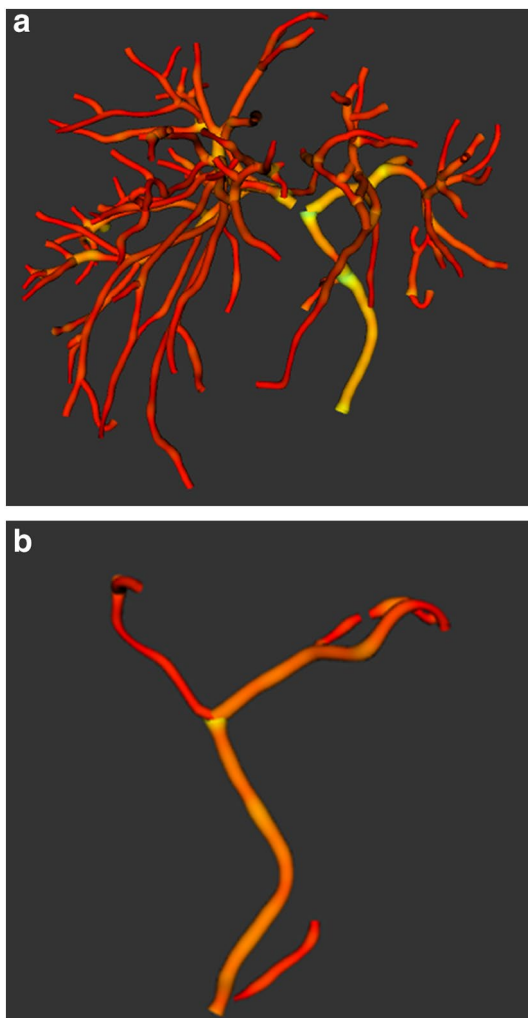
*$n=17$

cohorts (Table 2). All quantitative MRCP+ parameters except for median right hepatic duct diameter were statistically significantly different between the AIH and PSC/ASC cohorts (Table 3). Representative post-processed 3D biliary tree models from a patient from each diagnostic cohort (PSC/ASC, AIH) are provided in Fig. 2.

Assessment of the diagnostic performance of patient age, serum biochemistry, and quantitative MRCP+ parameters using ROC curve analyses is summarized in Table 4. No clinical parameter (age or biochemistry value) was significantly predictive of AILD diagnosis (all ROC p values > 0.05), while all quantitative MRCP+ parameters except median right hepatic duct diameter were significantly predictive of AILD diagnosis (ROC p values < 0.05). The most

discriminative MRCP+ parameters for distinguishing AIH from PSC/ASC included number of strictures (AUROC, 0.86; 95% CI AUROC, 0.69-0.95; sensitivity, 72%; specificity, 80%), number of dilations (AUROC, 0.87; 95% CI AUROC, 0.71–0.96; sensitivity, 89%; specificity, 73%), and total length of dilations (AUROC, 0.89; 95% CI AUROC, 0.73–0.97; sensitivity, 83%; specificity, 87%) (Table 4, Fig. 3).

Using multivariable logistic regression, the best model for discriminating PSC/ASC from AIH included two quantitative MRCP variables and achieved an AUC of 0.92: total length of biliary tree dilations (OR 1.08; 95% CI 1.02–1.14; $p = 0.01$) and maximum LHBD diameter (OR 1.21; 95% CI 0.57–2.56; $p = 0.62$) (Fig. 4). The sensitivity and specificity of this model were 88.2% and 73.3%, respectively.

Discussion

MRCP provides a non-invasive anatomic assessment of the intra- and extra-hepatic biliary tree and plays a key role in diagnosing pediatric AILD. In combination with clinical history, serum biochemistry, and liver histopathology, MRCP is utilized to reach a specific diagnosis of AIH, PSC, or ASC. The characteristic appearance of large duct PSC on MRCP is a random distribution of intra-hepatic multifocal strictures and associated segmental dilated upstream bile ducts that produce a "beaded" appearance of the biliary tree [18, 21]. However, MRCP findings in PSC patients can be variable (e.g., isolated intra- or extra-hepatic disease or dominant extra-hepatic stricture). Furthermore, in small duct PSC, characteristic findings of PSC are observed only on histopathology and not on MRCP [18]. The biliary tree of patients with AIH is typically normal without strictures or dilations; however, in cases of suspected AIH, MRCP is obtained to rule out ASC, which has concomitant findings of PSC on MRCP and/or histopathologic evaluation [1].

While it has been shown that radiologists are able to diagnose PSC by MRCP with high sensitivity (84–86%) and specificity (94%), interpretation of MRCP remains subjective [10, 11]. One study demonstrated the percentage of agreement between two observers for detecting the presence and location of ductal unit strictures on MRCP in patients with PSC to range from 45 to 81% [12]. Another study also demonstrated poor inter-observer agreement ($n = 44$ observers, kappa values of 0.13–0.19) for the following MRCP findings in patients with PSC: overall interpretation (typical, compatible, or atypical for PSC), bile duct changes and location(s), and dominant stricture and location(s) [13]. Such disagreements in interpretation are more likely to occur in clinical practice than in research studies with expert readers and may be even more problematic in interpretation of MRCP studies from children.

Fig. 2 Representative biliary tree models from a 17-year-old boy with ASC (strictures = 4, dilations = 4, length of dilated ducts = 38.8 mm) (**a**) and a 10-year-old boy with AIH (strictures = 0, dilations = 1, length of dilated ducts = 5.5 mm) (**b**). Note that areas of bile duct discontinuity represent artifact (e.g., motion, crossing blood vessel) and/or strictures, necessitating a review of source MRCP and anatomic images

Table 4 Assessment of the diagnostic performance of clinical and imaging parameters for discriminating PSC/ASC from AIH using receiver operating characteristic (ROC) curve analyses

Variable	AUROC	95% CI	p value	Cutoff	Sensitivity (%)	Specificity (%)
Age (years)	0.52	0.34–0.70	0.82	≤ 18	94	27
Liver biochemistry						
Alkaline phosphatase (units/L)	0.54	0.36–0.71	0.72	> 108	67	53
Total bilirubin (mg/dL)	0.58	0.40–0.75	0.43	≤ 0.5	78	47
GGT (units/L)	0.67	0.49–0.83	0.07	> 78	44	93
ALT (units/L)	0.61	0.43–0.77	0.29	> 20	94	27
AST (units/L)	0.50	0.33–0.68	0.97	> 13	100	20
Albumin (g/dL)	0.58	0.40–0.75	0.41	≤ 3.9	50	73
Biliary tree parameters						
Biliary tree volume (mL)	0.79	0.61–0.91	0.005	> 6.6	56	93
Median CBD diameter (mm)	0.70	0.52–0.85	0.030	> 4.5	50	93
Maximum CBD diameter (mm)	0.72	0.53–0.86	0.016	> 6.5	44	93
Median right hepatic duct diameter (mm)	0.59	0.41–0.76	0.36	> 4	39	93
Maximum right hepatic duct diameter (mm)	0.74	0.56–0.88	0.007	> 4.5	56	93
Median left hepatic duct diameter (mm)	0.71	0.52–0.85	0.034	> 4.2	53	93
Maximum left hepatic duct diameter (mm)	0.76	0.57–0.89	0.005	> 4.6	71	87
Number of visualized bile ducts	0.78	0.60–0.90	0.002	> 18	89	67
Total length of biliary tree (mm)	0.79	0.61–0.91	0.0009	> 526.9	83	83
Number of strictures	0.86	0.69–0.95	< 0.0001	> 3	72	80
Total length of strictures (mm)	0.82	0.65–0.93	< 0.0001	> 11.7	94	60
Number of dilations	0.87	0.71–0.96	< 0.0001	> 3	89	73
Total length of dilations (mm)	0.89	0.73–0.97	< 0.0001	> 31.0	83	87

Area under the curve (AUROC) and corresponding 95% confidence intervals (CI) are reported

GGT gamma-glutamyltransferase, *ALT* alanine aminotransferase, *AST* aspartate aminotransferase, *CBD* common bile duct

In the current study, we report the diagnostic performance of serum biochemistry and quantitative MRCP parameters derived from a novel post-processing algorithm for the discrimination of types of AILD: AIH versus PSC or ASC. Our results suggest that quantitative biliary tree parameters derived from 3D MRCP examinations provide good discrimination of AIH versus PSC/ASC in children and young adults. All but a single quantitative MRCP parameter were significantly different between patient cohorts. Similarly, all but a single MRCP parameter allowed significant discrimination between cohorts. The single best metric, total length of biliary tree dilations, was able to discriminate between cohorts with an AUROC of 0.89. The ability to distinguish between cohorts was slightly improved to AUC of 0.92 by the addition of another MRCP parameter, maximum diameter of the LHBD, to the model. Conversely, serum biochemistry values, including ALP, GGT, AST, and ALT, provided no discriminative ability to distinguish AIH versus PSC/ASC in the same population. This lack of discrimination between cohorts by serum biochemistry could in part be explained by the amount of time between diagnosis and laboratory testing and that patients with PSC/ASC are typically treated with ursodiol at our institution, which may lower ALP and GGT [22].

Our results demonstrate satisfactory diagnostic performance of quantitative MRCP for the discrimination of AILD and provide a foundation upon which further research into the diagnostic and prognostic capability of quantitative MRCP can be based. There is little doubt that radiologists can distinguish normal biliary trees from cases of marked cholangiopathic change. However, we hypothesize that quantitative MRCP parameters may be more sensitive and/or specific for cases of subtle cholangiopathic change. For example, a small number of mild strictures (e.g., more than 3 areas of ductal narrowing just greater than 30%) could go undetected in clinical practice. Our results support the use of quantitative MRCP as a biomarker of AILD at diagnosis and suggest that it might be applied to assessment of disease progression/change over time, establishing prognosis, and predicting treatment response, for which such radiologic biomarkers are currently lacking.

One important finding in our study is the relatively high percentage (~ 30%) of MRCP examinations that were unable to be post-processed due to motion artifact. Traditional 3D
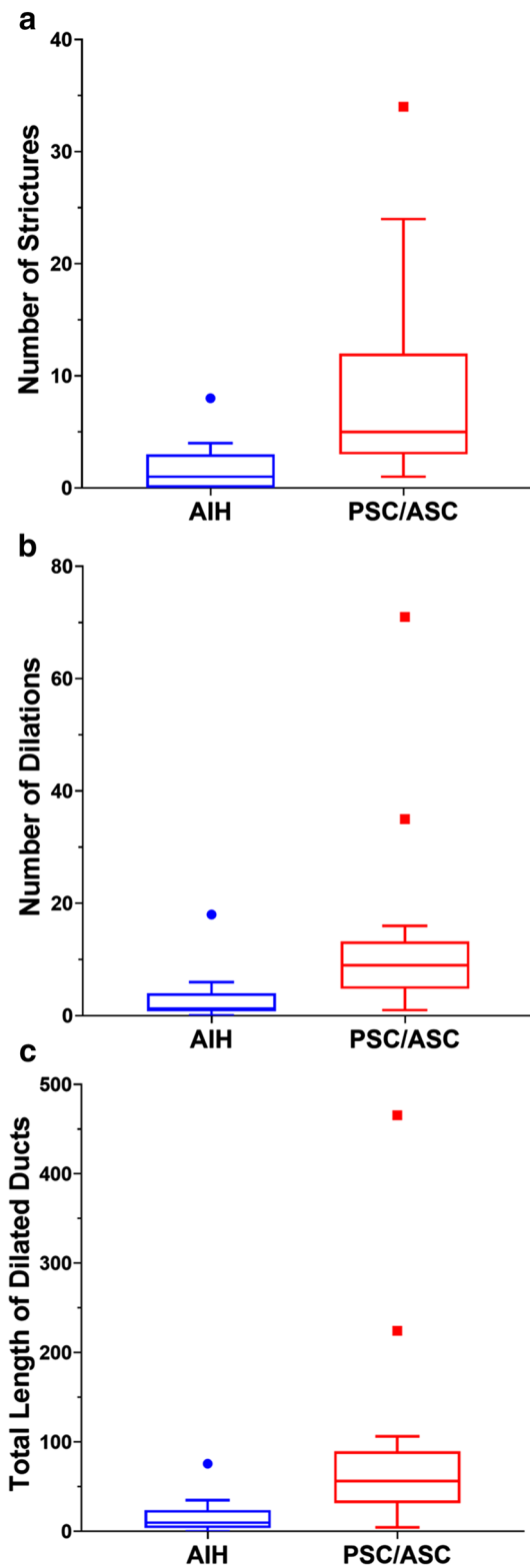
Fig. 3 Tukey box and whisker plots demonstrating the number of strictures (**a**), number of dilations (**b**), and total length of dilations [mm] (**c**) for patients with AIH and PSC/ASC
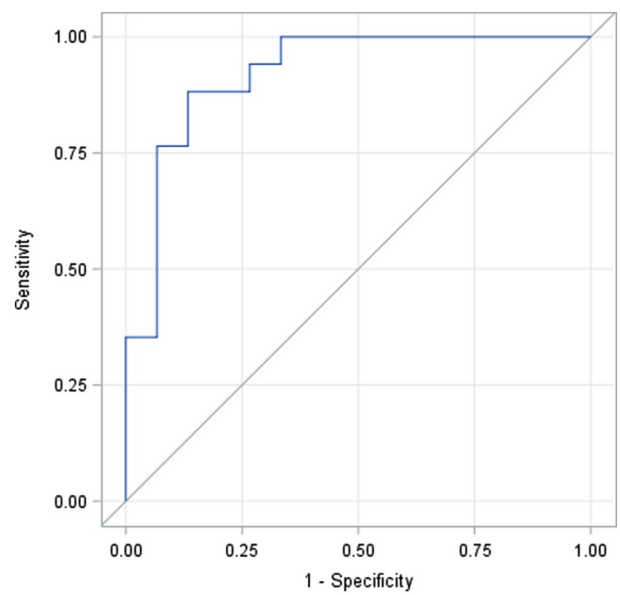
Fig. 4 ROC curve of the best performing multivariable logistic regression model for discrimination of PSC/ASC from AIH. Model has an AUROC of 0.92 and includes two quantitative MRCP parameters: total length of biliary tree dilations (mm) [OR 1.08] and maximum diameter of the left hepatic bile duct (mm) [OR 1.21]

MRCP is notoriously limited by imaging artifacts due to its length of acquisition (e.g., 4–8 min) despite the use of respiratory-triggering or navigator-gating, and this technical challenge is highlighted by our study. Importantly, the observed rate of examinations that were unable to be post-processed might limit the application of quantitative MRCP in clinical practice. This deserves further investigation, including the application of accelerated and sparse sampling (i.e., compressed sensing) MRI techniques.

Our study has several limitations. First, we had a relatively small-analyzed sample size of 33 patients. Second, the post-processing algorithm used in our study requires further research to better characterize its accuracy and repeatability within and across scanner platforms. In addition, this algorithm is not yet widely available and requires expert input, potentially limiting its clinical application. Third, we did not directly compare qualitative versus quantitative interpretation of the registry baseline MRCP examinations but instead relied upon clinical diagnosis as the reference standard, which incorporated clinical MRCP as well as other clinical data, such as laboratory values and histopathology. Finally, we included patients with small duct PSC/ASC, even though, by definition, these patients do not have qualitative findings on MRCP and could represent false negative cases in our MRCP analyses, thus slightly lowering the AUROCs of the various quantitative MRI parameters. However, it is conversely possible that subtle cholangiopathy in small duct PSC/ASC may be detected by quantitative,

and not qualitative, MRCP; thus, further investigations are warranted.

In conclusion, quantitative biliary MRCP parameters provide good discrimination of PSC/ASC from AIH. Our results suggest that quantitative MRCP has the potential to provide numerous imaging biomarkers of AILD, although there is need for further studies to determine if this technique is sensitive to change over time and if it is associated with, or predictive of, important clinical outcomes.

Compliance with ethical standards

References

1. Deneau M, Jensen MK, Holmen J, Williams MS, Book LS, Guthery SL (2013) Primary sclerosing cholangitis, autoimmune hepatitis, and overlap in Utah children: epidemiology and natural history. Hepatology 58 (4):1392-1400. https://doi.org/10.1002/hep.26454

2. Singh H, Balouch F, Noble C, Lewindon P (2018) Evolving Practice and Changing Phenotype in Pediatric Autoimmune Liver Disease: Outcomes From an Australian Center. J Pediatr Gastroenterol Nutr 67 (1):80-85. https://doi.org/10.1097/mpg.0000000000001927

3. Gregorio GV, Portmann B, Karani J, Harrison P, Donaldson PT, Vergani D, Mieli-Vergani G (2001) Autoimmune hepatitis/sclerosing cholangitis overlap syndrome in childhood: a 16-year prospective study. Hepatology 33 (3):544-553. https://doi.org/10.1053/jhep.2001.22131

4. Wilschanski M, Chait P, Wade JA, Davis L, Corey M, St Louis P, Griffiths AM, Blendis LM, Moroz SP, Scully L, et al. (1995) Primary sclerosing cholangitis in 32 children: clinical, laboratory, and radiographic features, with survival analysis. Hepatology 22 (5):1415-1422

5. Della Corte C, Sartorelli MR, Sindoni CD, Girolami E, Giovannelli L, Comparcola D, Nobili V (2012) Autoimmune hepatitis in children: an overview of the disease focusing on current therapies. Eur J Gastroenterol Hepatol 24 (7):739-746. https://doi.org/10.1097/meg.0b013e328353750c

6. Manns MP, Czaja AJ, Gorham JD, Krawitt EL, Mieli-Vergani G, Vergani D, Vierling JM, American Association for the Study of Liver D (2010) Diagnosis and management of autoimmune hepatitis. Hepatology 51 (6):2193-2213. https://doi.org/10.1002/hep.23584

7. Angulo P, Batts KP, Jorgensen RA, LaRusso NA, Lindor KD (2000) Oral budesonide in the treatment of primary sclerosing cholangitis. Am J Gastroenterol 95 (9):2333-2337. https://doi.org/10.1111/j.1572-0241.2000.02323.x

8. Tischendorf JJ, Hecker H, Kruger M, Manns MP, Meier PN (2007) Characterization, outcome, and prognosis in 273 patients with primary sclerosing cholangitis: A single center study. Am J Gastroenterol 102 (1):107-114. https://doi.org/10.1111/j.1572-0241.2006.00872.x

9. Wiesner RH, Grambsch PM, Dickson ER, Ludwig J, MacCarty RL, Hunter EB, Fleming TR, Fisher LD, Beaver SJ, LaRusso NF (1989) Primary sclerosing cholangitis: natural history, prognostic factors and survival analysis. Hepatology 10 (4):430-436

10. Chavhan GB, Roberts E, Moineddin R, Babyn PS, Manson DE (2008) Primary sclerosing cholangitis in children: utility of magnetic resonance cholangiopancreatography. Pediatr Radiol 38 (8):868-873. https://doi.org/10.1007/s00247-008-0918-6

11. Dave M, Elmunzer BJ, Dwamena BA, Higgins PD (2010) Primary sclerosing cholangitis: meta-analysis of diagnostic performance of MR cholangiopancreatography. Radiology 256 (2):387-396. https://doi.org/10.1148/radiol.10091953

12. Vitellas KM, El-Dieb A, Vaswani KK, Bennett WF, Tzalonikou M, Mabee C, Kirkpatrick R, Bova JG (2002) MR cholangiopancreatography in patients with primary sclerosing cholangitis: interobserver variability and comparison with endoscopic retrograde cholangiopancreatography. AJR Am J Roentgenol 179 (2):399-407. https://doi.org/10.2214/ajr.179.2.1790399

13. Zenouzi R, Liwinski T, Yamamura J, Weiler-Normann C, Sebode M, Keller S, Lohse AW, Schramm C, IPSCSG IPSG (2018) Follow-up magnetic resonance imaging/3D-magnetic resonance cholangiopancreatography in patients with primary sclerosing cholangitis: challenging for experts to interpret. Aliment Pharm Ther 48 (2):169-178. https://doi.org/10.1111/apt.14797

14. Frangi AF, Niessen WJ, Vincken KL, Viergever MA (1998) Multiscale vessel enhancement filtering. Lect Notes Comput Sc 1496 (1496):130-137

15. Perona P MJ (1990) Scale-space and edge detection using anisotropic diffusion IEEE Transactions on Pattern Analysis and Machine Intelligence 12 (7):629-639

16. Otsu N (1979) A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man and Cybernetics 9 (1):62-66

17. Xu C, Prince JL (1998) Snakes, shapes, and gradient vector flow. IEEE Trans Image Process 7 (3):359-369. https://doi.org/10.1109/83.661186

18. Chapman R, Fevery J, Kalloo A, Nagorney DM, Boberg KM, Shneider B, Gores GJ, American Association for the Study of Liver D (2010) Diagnosis and management of primary sclerosing cholangitis. Hepatology 51 (2):660-678. https://doi.org/10.1002/hep.23294

19. Mileti E, Rosenthal P, Peters MG (2012) Validation and modification of simplified diagnostic criteria for autoimmune hepatitis in children. Clin Gastroenterol Hepatol 10 (4):417-421 e411-412. https://doi.org/10.1016/j.cgh.2011.11.030

20. Schisterman EF, Perkins NJ, Liu A, Bondell H (2005) Optimal cut-point and its corresponding Youden Index to discriminate individuals using pooled blood samples. Epidemiology 16 (1):73-81

21. Vitellas KM, Keogan MT, Freed KS, Enns RA, Spritzer CE, Baillie JM, Nelson RC (2000) Radiologic manifestations of sclerosing cholangitis with emphasis on MR cholangiopancreatography. Radiographics 20 (4):959-975; quiz 1108-1109, 1112. https://doi.org/10.1148/radiographics.20.4.g00jl04959

22. Lindor KD (1997) Ursodiol for primary sclerosing cholangitis. Mayo Primary Sclerosing Cholangitis-Ursodeoxycholic Acid Study Group. N Engl J Med 336 (10):691-695. https://doi.org/10.1056/nejm199703063361003