

Article

Prognostic Value and Inter-Reader Agreement of the ANALI_{NG} and DiStrict MRCP Scores in Primary Sclerosing Cholangitis

Thomas Goodwin ^{1,2,*} , Numan Kutaiba ^{3,4} , Mark Goodwin ^{3,4} , Ruth P. Lim ^{3,4} , Leon Winata ⁵, Bruno Di Muzio ^{6,7} , William Kemp ^{1,2} , Stuart Roberts ^{1,2} , Natassia Tan ^{1,2} and Ammar Majeed ^{1,2}

¹ Department of Gastroenterology and Hepatology, Alfred Health, Melbourne, VIC 3004, Australia

² School of Translational Medicine, Monash University, Melbourne, VIC 3004, Australia

³ Department of Radiology, Austin Health, Melbourne, VIC 3084, Australia

⁴ Department of Radiology, Melbourne Medical School, The University of Melbourne, Melbourne, VIC 3050, Australia

⁵ Department of Medical Imaging, St Vincent's Hospital, Melbourne, VIC 3065, Australia

⁶ Department of Radiology, Alfred Health, Melbourne, VIC 3004, Australia

⁷ Department of Surgery, Monash University, Melbourne, VIC 3168, Australia

* Correspondence: t.goodwin@alfred.org.au

Abstract

Background/Objectives: The ANALI and DiStrict scores are novel quantitative magnetic resonance cholangio-pancreatography (MRCP)-derived scoring systems designed to predict outcomes in primary sclerosing cholangitis (PSC). However, as their prognostic utility and inter-reader reproducibility are not widely validated, their utility in clinical practice is currently limited. This study aimed to assess the reproducibility and prognostic performance of the ANALI without gadolinium (ANALI_{NG}) and DiStrict scores in predicting liver-related outcomes in patients with PSC. **Methods:** We conducted a multicentre retrospective cohort study enrolling adult patients with large-duct PSC and at least one MRCP from the time of diagnosis. MRCPs were scored by five blinded senior abdominal radiologists with inter- and intra-reader agreement assessed via intraclass correlation coefficients. Multivariate Cox proportional models were used to evaluate the prognostic value of the ANALI_{NG} and DiStrict scores, with a composite endpoint of liver transplantation, hepatic decompensation, or liver-related death. **Results:** Eighty-nine patients with a median of 5.5 years (IQR 3.4–9.2) of follow-up were included. The ANALI_{NG} score showed higher intra-reader agreement (ICC 0.87, 95% CI 0.82–0.92 vs. ICC 0.64, 95% CI 0.36–0.79) and similar inter-reader agreement (ICC 0.71, 95% CI 0.63–0.78 vs. ICC 0.67, 95% CI 0.59–0.75) compared to the DiStrict score. Only the ANALI_{NG} score was associated with liver-related events, which remained significant when adjusted for age and sex (aHR 1.28, 95% CI 1.01–1.63) and MELD (aHR 1.69, 95% CI 1.22–2.34). **Conclusions:** The ANALI_{NG} score is a reproducible, MRI-derived scoring system that is independently associated with liver outcomes and has higher intra-reader agreement compared to the DiStrict score.

Keywords: primary sclerosing cholangitis; ANALI; DiStrict; prognosis; magnetic resonance imaging



Academic Editor: Ralf Weiskirchen

Received: 23 April 2026

Revised: 22 June 2026

Accepted: 24 June 2026

Published: 1 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Primary sclerosing cholangitis (PSC) is a progressive cholestatic liver disease characterised by chronic inflammation and fibrosis of the biliary tree, often culminating in cirrhosis, liver failure, or the need for liver transplantation [1]. Patients with PSC are also at

increased risk of hepatobiliary malignancy, contributing to high morbidity and mortality [2]. Despite advances in understanding the disease, prognostication remains challenging.

Several prognostic models including the Mayo Risk Score [3], UK-PSC score [4], the Primary Sclerosing Cholangitis Risk Estimate Tool (PREsTo) [5], and Amsterdam-Oxford model [6,7], which rely on a combination of clinical and biochemical parameters, have been developed to predict transplant-free survival. In addition, recent studies have demonstrated the utility of qualitative and quantitative scoring systems of magnetic resonance cholangiopancreatography (MRCP) as complementary prognostic tools [8–10]. As MRCP is already used for routine surveillance in patients with PSC, validation of available scoring systems may allow us to further refine risk stratification in both clinical and trial contexts for patients with PSC [11–13].

Imaging-based tools including the ANALI and DiStrict scores were originally developed to standardise MRCP assessment in PSC. By directly characterising structural liver and biliary changes, these MRI-based scoring systems may offer complementary prognostic insights beyond existing clinical and biochemical models, with the potential to refine risk stratification. The ANALI score incorporates both biliary and parenchymal assessment of the liver and has been validated in external cohorts to strongly predict liver-related outcomes [14,15]. In contrast, the DiStrict score focuses exclusively on cholangiographic changes, grading the severity and extent of intra- and extra-hepatic strictures and dilatations. Developed in a multicentre cohort, the score demonstrated good inter-reader reproducibility and was a strong predictor of liver transplant or liver-related death [16].

While both the above scoring systems have shown promise, concerns have been raised regarding their reproducibility, with prior studies demonstrating only poor to moderate inter-reader agreement [15,17–19]. Furthermore, no studies to date have directly compared the prognostic performance of ANALI and DiStrict within the same patient cohort, leaving uncertainty about their relative strengths and clinical utility.

In Australia, the vast majority of MRCP performed for PSC are without contrast; hence, we aimed to validate the prognostic performance of the ANALI score without gadolinium (ANALI_{NG}) and the DiStrict score MRI-based scoring systems and evaluate the reproducibility of each scoring system in an Australian setting.

2. Materials and Methods

2.1. Study Design and Population

We conducted a multicentre retrospective cohort study across three tertiary hospitals in Melbourne, Australia, including one liver transplant centre. Patients with a confirmed diagnosis of large-duct PSC who were diagnosed between 1 January 2005 and 30 March 2021 were identified from institutional hepatology databases at the three participating centres. This study represents a sub-cohort of a larger Australian PSC population previously described [20]. Follow-up data were collected where available up to 31 December 2023.

The study was approved by The Austin Ethics Committee (project: HREC/98594/Austin-2024). All relevant ethical and regulatory approvals were obtained in accordance with the guidelines laid down by the NHMRC of Australia, the Helsinki Declaration, and their subsequent amendments.

2.2. Inclusion and Exclusion Criteria

We included all adult patients with a confirmed diagnosis of large-duct PSC, defined by characteristic cholangiographic features on MRCP or endoscopic retrograde cholangiopancreatography in combination with a cholestatic biochemical profile [13]. Patients were required to have an adequate quality index MRCP—or the earliest available surveillance scan—suitable for radiologic scoring and sufficient longitudinal clinical follow-up.

Patients were excluded if they had small-duct PSC, prior liver transplantation, or an episode of hepatic decompensation (defined as an episode of ascites, hepatic encephalopathy, or variceal bleeding) before the index MRCP. Patients with inadequate MRCP quality, such as severe motion artefact, were also excluded from analysis.

2.3. Imaging Evaluation

MRCP examinations were performed between 2005 and 2021 using 1.5-T and 3-T scanners from multiple vendors. Assessments were based on three-dimensional MRCP images, together with axial T1 and T2, and coronal T2-weighted sequences to evaluate the biliary tree, liver morphology, splenomegaly, and portosystemic shunts. All MRCPs were performed as part of routine surveillance.

Five abdominal radiologists, with a range of 2 to 19 years of experience, independently reviewed all MRCP examinations while blinded to clinical and laboratory data. Before scoring, readers participated in a calibration session involving ten training cases of varying severity (not included in the study cohort) to ensure consistency in interpretation.

Each examination was scored according to both the ANALI_{NG} and DiStrict scoring systems. The ANALI_{NG} score quantifies intrahepatic bile duct dilatation, liver dysmorphism, and features of portal hypertension to generate a composite score ranging from 0 to 5 [15]. The DiStrict score assesses intra- and extra-hepatic ductal involvement, grading the extent and severity of strictures and upstream dilatation on a 0–4 scale for each domain (total 0–8) [16]. Full scoring criteria and imaging examples are provided in the Supplementary Materials (Tables S1 and S2, Figures S3 and S4). Of note, our analyses used the non-contrast ANALI_{NG} score as contrast-enhanced MRCP is rarely used in Australia for surveillance [21].

2.4. Outcomes

The primary endpoint was a composite of liver transplantation, hepatic decompensation, or liver-related death. Decompensation was defined as encephalopathy, ascites, or oesophageal variceal bleeding. Hepatobiliary cancers were not considered an endpoint, in keeping with prior studies [9,14–16]. Deaths attributable to liver cancer were included as liver-related mortality, whereas patients with incident malignancy who remained alive at study completion were censored.

2.5. Statistical Analysis

Continuous variables were reported as medians with interquartile ranges and categorical variables as absolute numbers and percentages.

Associations between imaging scores and clinical outcomes were evaluated using Cox proportional hazards regression, with hazard ratios (HRs) and 95% confidence intervals (CIs) estimated from both univariate and multivariate models adjusted for age, sex, and MELD score. The Kaplan–Meier method with log-rank analysis was used to assess differences in time to composite endpoint between patients stratified by ANALI_{NG} scores of 0–2 versus 3–5 and DiStrict scores of 0–4 versus 5–8, as per the original studies. Statistical significance was set at $p < 0.05$.

The inter-reader and intra-reader reproducibility of both scoring systems were assessed using the intraclass correlation coefficient (ICC) with 95% confidence intervals. To assess intra-reader reliability, one radiologist (anonymised for review; six years of abdominal imaging experience) re-evaluated all MRCP studies two months after the initial assessment. For survival analysis, this radiologist's initial assessment was used. Agreement was interpreted using the following conventional thresholds: ICC < 0.50 (poor), 0.50–0.75 (moderate), 0.75–0.90 (good), and >0.90 (excellent) [22].

Analyses were conducted using SPSS statistical software version 29.

3. Results

3.1. Clinical Characteristics

A total of 94 patients were initially identified from the retrospective cohort. Two were excluded due to a diagnosis of small-duct PSC, and three were excluded because of prior decompensation. The final cohort comprised 89 patients with large-duct PSC, with a median follow-up of 5.5 (IQR 3.4–9.2) years from the time of index MRCP, with a total follow-up time of 576 person-years. An additional 11 patients lacked up-to-date biochemical data within six months of the index MRCP and were therefore excluded from the sub-analysis evaluating the biochemical predictors of outcome.

The median age at MRCP was 39 years, and 55% were male. At the end of the follow-up period, 31 patients (35%) met the composite primary outcome: 17 (19%) underwent liver transplantation, 19 (21%) developed liver decompensation, and 4 (5%) died from end-stage liver disease. Of those who underwent transplantation, five experienced hepatic decompensation during follow-up prior to transplantation, and all four patients who died from liver-related causes experienced hepatic decompensation prior to death. The median time to the composite outcome was 5.8 (IQR 2.6–7.4) years. Baseline demographic, laboratory, and clinical characteristics are summarised in Table 1.

Table 1. Demographic and laboratory data of the study population.

	Study Population
Number of patients	89
Liver transplant centre	61 (68%)
Non-transplant centre	28 (32%)
Female/male	40/49
Age, years	39 (27–55)
	Laboratory Values
ALT U/L	71 (35–165)
AST U/L	64 (38–102)
ALP U/L	217 (123–401)
Albumin g/L	36 (32–40)
Bilirubin mcmol/L	16 (8–34)
INR	1 (1–1.2)
Platelets 10 ⁹ /L	240 (178–319)
AFP mcg/L	2 (1.2–3.8)
MELD Score	8 (6.5–11.5)
Follow-up time from MRI (years)	5.5 (3.4–9.2)
Patients who met composite endpoint	31
Total events	40
Liver transplant	17
Liver decompensation	19
Liver-related death	4
HCC/CCA	2/1

All values are expressed as medians (inter-quartile range). ALT, alanine aminotransferase; AST, aspartate transaminase; ALP, alkaline phosphatase; INR, international normalised ratio; HCC, hepatocellular carcinoma; CCA, cholangiocarcinoma; AFP, alpha-fetoprotein; MELD, Model for End-Stage Liver Disease.

There was one patient with an endobiliary stent, one with a prior right hemihepatectomy, and fourteen patients with a prior cholecystectomy who were included within the study. These patients were included in the final analysis to best reflect a real-world cohort.

3.2. Assessment of Inter-Reader and Intra-Reader Agreement

Inter-reader and intra-reader reproducibility of both scoring systems are shown in Table 2. The ANALI_{NG} score demonstrated moderate inter-reader agreement (ICC = 0.71, 95% CI 0.63–0.78) and good intra-reader agreement (ICC = 0.87, 95% CI 0.82–0.92). The

DiStrict score showed moderate inter-reader agreement (ICC = 0.67, 95% CI 0.59–0.75) and moderate intra-reader agreement (ICC = 0.64, 95% CI 0.36–0.79).

Table 2. Intraclass correlation coefficient (ICC) with 95% confidence intervals (CI).

Score	Inter-Reader ICC (95% CI)	Intra-Reader ICC (95% CI)
ANALI _{NG} score	0.71 (0.63–0.78)	0.87 (0.82–0.92)
Intrahepatic ducts	0.75 (0.68–0.81)	0.80 (0.71–0.86)
Liver dysmorphism	0.60 (0.50–0.69)	0.77 (0.67–0.84)
Porto-systemic shunts	0.69 (0.60–0.76)	0.83 (0.75–0.88)
DiStrict score	0.67 (0.59–0.75)	0.64 (0.36–0.79)
Intrahepatic ducts	0.62 (0.53–0.71)	0.62 (0.31–0.78)
Extrahepatic ducts	0.56 (0.47–0.66)	0.52 (0.35–0.66)

Among the ANALI_{NG} subcomponents, inter-reader agreement was good for intra-hepatic ducts (ICC, 0.75; 95% CI, 0.68–0.81) and moderate for porto-systemic shunts (0.69; 95% CI 0.60–0.76) and liver dysmorphism (0.60; 95% CI 0.50–0.69). Intra-reader agreement for these features was good across all scores (ICCs, 0.77–0.83).

For DiStrict subcomponents, inter-reader agreement was moderate for intrahepatic ducts (0.62; 95%CI 0.53–0.71) and for extrahepatic ducts (0.56; 95% CI 0.47–0.66), with corresponding intra-reader ICCs of 0.62 (95% CI 0.31–0.78) and 0.52 (95% CI 0.35–0.66).

3.3. Prognostic Performance

Higher ANALI_{NG} scores were significantly associated with meeting the composite primary outcome (HR 1.28, 95% CI 1.02–1.61, $p = 0.033$), and this association remained significant after adjustment for age and sex (aHR 1.28, 95% CI 1.01–1.63, $p = 0.045$) and adjustment for MELD score (aHR 1.69, 95% CI 1.22–2.34, $p = 0.002$; Table 3). Kaplan–Meier analysis demonstrated significantly lower outcome-free survival among patients with ANALI_{NG} scores of 3–5 compared to those with scores of 0–2 (Figure 1).

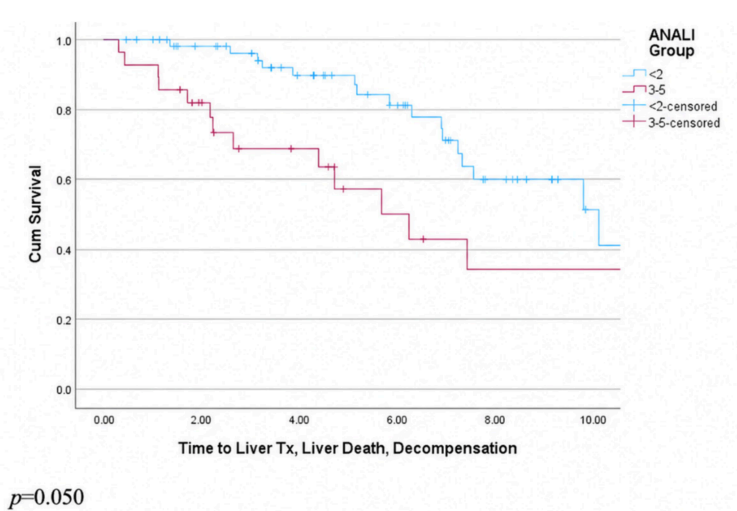
In contrast, DiStrict score was not significantly associated with outcomes in unadjusted analysis (HR 1.03, 95% CI 0.87–1.23, $p = 0.71$), and the association remained non-significant after adjustment for age and sex (aHR 1.025, 95% CI 0.86–1.22, $p = 0.078$) and MELD (aHR 1.04, 95% CI 0.85–1.28, $p = 0.69$; Table 3). Kaplan–Meier analysis similarly showed no significant difference in outcome-free survival between patients with DiStrict scores of 5–8 and those with scores of 0–4 (Figure 2).

Table 3. Risk estimates for clinical outcomes adjusted for age and sex, and MELD score.

Variable	Univariable Analysis		Multivariable Analysis (Age and Sex)		Multivariable Analysis (MELD)	
	HR (95% CI)	p -Value	aHR (95% CI)	p -Value	aHR (95% CI)	p -Value
ANALI _{NG} score	1.28 (1.02–1.61)	0.033	1.28 (1.01–1.63)	0.045	1.69 (1.22–2.34)	0.002
DiStrict score	1.033 (0.87–1.23)	0.712	1.025 (0.86–1.22)	0.078	1.04 (0.85–1.28)	0.693

Analysis of the individual ANALI_{NG} score components demonstrated that liver dysmorphism (aHR 1.70, 95% CI 1.14–2.52, $p = 0.009$) and portosystemic shunts (aHR 4.32, 95% CI 1.41–13.26, $p = 0.011$) were independently associated with outcomes while intrahepatic ductal changes alone were not (Supplementary Table S3). In the DiStrict score subgroup analysis, intrahepatic ductal changes were independently associated with outcomes (aHR

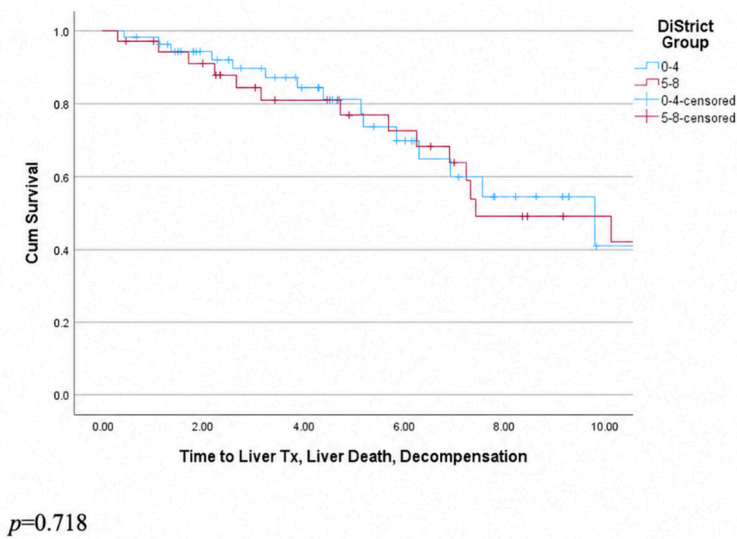
1.52, 95% CI 1.05–2.20, $p = 0.026$), while extrahepatic ductal changes showed no association with outcome (Supplementary Table S4).



Number at risk

Time (years)	0	2	4	6	8	10
ANALING 0-2	60	51	38	26	14	5
ANALING 3-5	29	19	14	12	4	4

Figure 1. Kaplan–Meier curve for outcome-free survival for ANALING scores of 0–2 and 3–5.



Number at risk

Time (years)	0	2	4	6	8	10
DiStrict 0-4	54	42	30	16	8	2
DiStrict 5-8	35	30	22	17	10	7

Figure 2. Kaplan–Meier curve for outcome-free survival for DiStrict scores of 0–4 and 5–8.

A comparison with established biochemical predictors was conducted in a subgroup with complete biochemical data (79 patients). This showed that MELD, Mayo Risk Score,

serum bilirubin, ALP, and albumin all retained prognostic value in multivariate models (Supplementary Table S5).

4. Discussion

In this multicentre retrospective study, we evaluated the prognostic performance and reproducibility of the ANALI_{NG} and DiStrict MRCP-based scoring systems in patients with PSC. We found that, among the two scoring systems, ANALI_{NG} demonstrated stronger prognostic performance, remaining independently associated with liver transplantation, hepatic decompensation, or liver-related death after adjustment for age, sex, and MELD score. In contrast, the DiStrict score was not associated with outcomes in our cohort, although the intrahepatic stricturing component within its scoring framework was associated with clinical outcomes, highlighting the prognostic relevance of intrahepatic disease burden.

These findings support the findings by Grigoriadis et al. [15,16], who described and validated both ANALI_{NG} and DiStrict scores as MRI-based tools for PSC prognostication. Previous studies have demonstrated that ANALI_{NG} is strongly associated with outcomes but have raised concerns regarding poor-to-moderate inter-reader agreement, particularly for subjective features such as liver dysmorphism and parenchymal heterogeneity [14,15,19]. This variability has been consistently reported in the PSC literature, especially for MRI-based assessments, and appears to persist across different cohorts and reader experience levels [17,18,23].

Importantly, we found inter- and intra-reader agreement for ANALI_{NG} to be moderate-to-good, indicating acceptable reproducibility. The inter-reader agreement for ANALI_{NG} (ICC 0.71) was higher than that reported in prior work (ICC 0.56), which may in part reflect the greater reproducibility of the liver dysmorphism component in our cohort (ICC 0.60 vs. 0.28). Notably, our cohort demonstrated improved agreement across all three ANALI_{NG} domains [15]. This improved reader agreement may be the result of the pre-assessment training session, suggesting that enhancing inter-reader reliability is feasible.

For DiStrict score, both inter- and intra-reader agreement were moderate, indicating greater variability in scoring compared with ANALI_{NG}. Agreement was higher for intrahepatic than extrahepatic ductal assessment (ICC 0.62 vs. 0.56), although, to our knowledge, this comparison has not been previously studied. MRCP studies in our cohort were acquired over a 16-year period using evolving protocols and variable image quality, which may have contributed to lower agreement in the assessment of biliary strictures and ductal dilatation.

In contrast to the ANALI_{NG} score and earlier European studies, the DiStrict score was not independently associated with outcomes in our patient cohort. In the original Swedish validation study by Grigoriadis et al. [16], higher DiStrict scores (≥ 5) were strongly associated with adverse outcomes, whereas this association was not observed in our study, either in the univariate analysis or after the multivariable adjustment. This discrepancy may be partially attributable to differences in disease severity between the cohorts, with our patients demonstrating lower DiStrict scores at baseline (median 3; mean 3.1) relative to the Swedish cohort (median 4; mean 4.1) [16]. The distribution of scores in the study population is shown in Supplementary Figures S1 and S2.

This discrepancy may also reflect differences in sample size and event rates, as our cohort was smaller and included fewer total endpoints likely resulting in type II error. As the DiStrict score captures only the burden of biliary stricturing, rather than markers of disease progression such as fibrosis or portal hypertension and given that stricturing typically precedes cirrhosis and hepatic decompensation, the relatively shorter median follow-up of 5.5 years in our cohort may have been insufficient to capture later stage

events. Longer observation may therefore be required to fully assess its prognostic value for advanced outcomes. DiStrict score reproducibility was also lower than previously reported, with only moderate inter-reader and intra-reader agreement, suggesting that ductal scoring may be more susceptible to variability both in interpretation and image acquisition, although we did not formally assess image quality in this study.

An analysis of individual score components showed that liver dysmorphism and the presence of portosystemic shunts, a surrogate for portal hypertension, were strongly associated with outcomes, whereas intrahepatic ductal dilatation was not. In contrast, the DiStrict score identified intrahepatic strictures as independently prognostic. This difference may reflect how intrahepatic disease is captured differently across the two scores. ANALI_{NG} evaluates only the largest measurable duct, grading dilatation on a 0–2 scale without accounting for the extent or distribution of stricturing, whereas DiStrict employs a more detailed 0–4 scale that incorporates both the proportion of liver affected (<25% vs. >25%) and the presence of associated biliary dilatation. As a result, the DiStrict score may more sensitively capture diffuse intrahepatic involvement that better reflects risk of progressive fibrosis rather than isolated ductal dilatation.

Other novel tools, such as MRCP+, also quantify biliary changes using more granular metrics and have shown positive correlations with outcomes [8]. Validation studies integrating MRCP+ parameters with biochemical markers have demonstrated superior prognostic performance compared with traditional scores such as the Mayo risk score [24], suggesting that future iterations of ANALI_{NG} and DiStrict may benefit from the incorporation of complementary non-invasive markers. The lack of association between intrahepatic dilatation and outcome in ANALI_{NG} suggests that duct calibre alone may be an imprecise marker of disease severity.

A key strength of this study is the multicentre design across three tertiary institutions, including a liver transplant centre, which enhances generalisability within the Australian PSC population. Reader blinding and independent image review strengthened reproducibility estimates, and outcomes were verified through longitudinal clinical follow-up.

However, several limitations should be acknowledged. This was a retrospective single-country study conducted across three tertiary centres in Melbourne, Australia, introducing risks of selection and referral bias. Additionally, imaging protocols were not standardised across sites or over time, with MRCPs performed between 2005 and 2019 on scanners from different vendors. This, however, reflects real-world practice where patients undergo imaging from external sites and grading would have to be performed with differing technical factors.

Notably, outcome definitions differed between the original ANALI_{NG} and DiStrict derivation studies, with hepatic decompensation being included only with the ANALI_{NG} score [14,16]. We chose to include hepatic decompensation in the composite outcome to capture clinically relevant disease progression, improve statistical power and to allow for direct comparison between the scores. Because ANALI_{NG} incorporates radiologic features of cirrhosis whereas DiStrict does not, this outcome choice likely contributed to the comparatively poorer performance of the DiStrict score. This highlights the challenges of comparing prognostic models with heterogeneous outcome frameworks and underscores the need for harmonised endpoints in future validation studies.

A further limitation is that survival analyses were based on assessments from a single radiologist without consensus review. Although reproducibility was moderate, this approach may limit generalisability and does not fully account for inter-observer variability. In addition, readers relied on available sequences for biliary tree assessment which included axial T1 and T2 and coronal T2 sequences as well as 3D-MRCP images. The DiStrict score was developed using 3D-MRCP images and relied on axial T2 sequence for

anatomical correlation only [17]. Furthermore, the follow-up period was relatively short and not all patients were observed to a definitive endpoint, and censoring may have led to an underestimation of risk associations. The number of outcome events, while adequate for primary analyses, was modest and may have limited power for multivariable and subgroup analyses.

Finally, hepatobiliary malignancies were not included in the primary endpoint, in keeping with prior studies, but this excludes an important cause of morbidity and mortality in PSC.

5. Conclusions

Our multicentre study demonstrates that the ANALI_{NG} score provides stronger prognostic performance and more consistent reproducibility than the DiStrict score for predicting liver-related outcomes in PSC. Although the DiStrict score itself was not independently associated with outcomes, intrahepatic strictures within its framework were, emphasising the importance of incorporating parenchymal, portal hypertension, and intrahepatic ductal features into MRI-based risk models. Validation in larger prospective cohorts, ideally alongside non-invasive markers of fibrosis, will be essential to refine these imaging tools and establish standardised MRI-based prognostic systems for PSC.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/livers6040059/s1>, Figure S1: Distribution of ANALI score without gadolinium in the study population; Figure S2: Distribution of DiStrict score in the study population; Figure S3: Example MRCP image, Axial T1 sequence; Figure S4: Example MRCP image of biliary stricture; Table S1: ANALI score without gadolinium; Table S2: DiStrict score; Table S3: ANALI score prognostic value by individual parameter; Table S4: DiStrict score prognostic value by individual parameter; Table S5: Risk estimates for clinical outcomes adjusted for age and sex; Table S6: Demographic breakdowns by score.

Author Contributions: Conceptualization, T.G., N.K., N.T. and A.M.; methodology, T.G., N.K., N.T. and A.M.; formal analysis, T.G., N.K., N.T. and A.M.; MRCP analysis, N.K., M.G., R.P.L., L.W. and B.D.M.; writing—original draft preparation, T.G.; writing—review and editing, T.G., N.K., S.R., N.T., A.M., M.G., R.P.L., L.W. and B.D.M.; supervision, N.T., S.R., W.K. and A.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Institutional Review Board Statement: This study was approved by the Austin Hospital Human Research Ethics Committee (approval code HREC/98594/Austin-2024, approval date 26 August 2024). The study was conducted in accordance with the National Health and Medical Research Council (NHMRC) National Statement on Ethical Conduct in Human Research.

Informed Consent Statement: Patient consent was waived due to study being approved by the local ethics institution as a low-risk retrospective study (approval code HREC/98594/Austin-2024, approval date 26 August 2024).

Data Availability Statement: The datasets presented in this article are not readily available because the data are part of an ongoing study. Requests to access the datasets should be directed to the corresponding author.

Acknowledgments: During the preparation of the graphical abstract, the authors used ChatGPT-5.5 for the purposes of image generation only. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The authors declare no conflicts of interest relevant to this work.

Abbreviations

The following abbreviations are used in this manuscript:

MRCP	Magnetic resonance cholangio-pancreatography
PSC	Primary sclerosing cholangitis
ANALI _{NG}	ANALI with no gadolinium
MELD	Model for end stage liver disease
IQR	Inter-quartile range
ICC	Intraclass correlation coefficient
CI	Confidence interval

References

- Boonstra, K.; Weersma, R.K.; van Erpecum, K.J.; Rauws, E.A.; Spanier, B.M.; Poen, A.C.; van Nieuwkerk, K.M.; Drenth, J.P.; Witteman, B.J.; Tuynman, H.A. Population-based epidemiology, malignancy risk, and outcome of primary sclerosing cholangitis. *Hepatology* **2013**, *58*, 2045–2055. [\[CrossRef\]](#) [\[PubMed\]](#)
- Bowlus, C.L.; Arrivé, L.; Bergquist, A.; Deneau, M.; Forman, L.; Ilyas, S.I.; Lunsford, K.E.; Martinez, M.; Sapisochin, G.; Shroff, R. AASLD practice guidance on primary sclerosing cholangitis and cholangiocarcinoma. *Hepatology* **2023**, *77*, 659–702. [\[PubMed\]](#)
- Kim, W.R.; Therneau, T.M.; Wiesner, R.H.; Poterucha, J.J.; Benson, J.T.; Malinchoc, M.; Larusso, N.F.; Lindor, K.D.; Dickson, E.R. A Revised Natural History Model for Primary Sclerosing Cholangitis. *Mayo Clin. Proc.* **2000**, *75*, 688–694. [\[CrossRef\]](#) [\[PubMed\]](#)
- Goode, E.C.; Clark, A.B.; Mells, G.F.; Srivastava, B.; Spiess, K.; Gelson, W.T.; Trivedi, P.J.; Lynch, K.D.; Castren, E.; Vesterhus, M.N. Factors associated with outcomes of patients with primary sclerosing cholangitis and development and validation of a risk scoring system. *Hepatology* **2019**, *69*, 2120–2135. [\[PubMed\]](#)
- Eaton, J.E.; Vesterhus, M.; McCauley, B.M.; Atkinson, E.J.; Schlicht, E.M.; Juran, B.D.; Gossard, A.A.; LaRusso, N.F.; Gores, G.J.; Karlsen, T.H. Primary sclerosing cholangitis risk estimate tool (PREsTo) predicts outcomes of the disease: A derivation and validation study using machine learning. *Hepatology* **2020**, *71*, 214–224. [\[PubMed\]](#)
- de Vries, E.M.; Wang, J.; Williamson, K.D.; Leeflang, M.M.; Boonstra, K.; Weersma, R.K.; Beuers, U.; Chapman, R.W.; Gekus, R.B.; Ponsioen, C.Y. A novel prognostic model for transplant-free survival in primary sclerosing cholangitis. *Gut* **2018**, *67*, 1864–1869. [\[PubMed\]](#)
- Goet, J.C.; Floreani, A.; Verhelst, X.; Cazzagon, N.; Perini, L.; Lammers, W.J.; de Vries, A.C.; van der Meer, A.J.; van Buuren, H.R.; Hansen, B.E. Validation, clinical utility and limitations of the Amsterdam-Oxford model for primary sclerosing cholangitis. *J. Hepatol.* **2019**, *71*, 992–999. [\[CrossRef\]](#) [\[PubMed\]](#)
- Cristoferi, L.; Porta, M.; Bernasconi, D.P.; Leonardi, F.; Gerussi, A.; Mulinacci, G.; Palermo, A.; Gallo, C.; Scaravaglio, M.; Stucchi, E.; et al. A quantitative MRCP-derived score for medium-term outcome prediction in primary sclerosing cholangitis. *Dig. Liver Dis.* **2023**, *55*, 373–380. [\[CrossRef\]](#) [\[PubMed\]](#)
- Mandea, M.; Iacob, S.M.; Grasu, M.C.; Anghel, C.; Iacob, R.A.; Ghioca, M.C.; Gheorghe, C.; Gheorghe, L.S. Quantitative Magnetic Resonance Cholangiopancreatography Scoring and Its Predictive Value for Outcomes in Adults with Primary Sclerosing Cholangitis. *J. Clin. Med.* **2024**, *13*, 4548. [\[CrossRef\]](#) [\[PubMed\]](#)
- Trivedi, P.J.; Arndtz, K.; Abbas, N.; Telford, A.; Young, L.; Banerjee, R.; Eddowes, P.; Jhaveri, K.S.; Hirschfield, G.M. Quantitative MRCP and metrics of bile duct disease over time in patients with primary sclerosing cholangitis: A prospective study. *Aliment. Pharmacol. Ther.* **2024**, *59*, 1366–1375. [\[CrossRef\]](#) [\[PubMed\]](#)
- Chapman, R.; Fevery, J.; Kalloo, A.; Nagorney, D.M.; Boberg, K.M.; Shneider, B.; Gores, G.J. Diagnosis and management of primary sclerosing cholangitis. *Hepatology* **2010**, *51*, 660–678. [\[CrossRef\]](#) [\[PubMed\]](#)
- Chazouilleres, O.; Beuers, U.; Bergquist, A.; Karlsen, T.H.; Levy, C.; Samyn, M.; Schramm, C.; Trauner, M. EASL Clinical Practice Guidelines on sclerosing cholangitis. *J. Hepatol.* **2022**, *77*, 761–806. [\[CrossRef\]](#)
- Schramm, C.; Eaton, J.; Ringe, K.I.; Venkatesh, S.; Yamamura, J.; MRI Working Group of the International PSC Study Group (IPSCSG). Recommendations on the use of magnetic resonance imaging in PSC-A position statement from the International PSC Study Group. *Hepatology* **2017**, *66*, 1675–1688. [\[PubMed\]](#)
- Lemoine, S.; Cazzagon, N.; El Mouhadi, S.; Trivedi, P.J.; Dohan, A.; Kemgang, A.; Belkacem, K.B.; Housset, C.; Chretien, Y.; Corpechot, C. Simple magnetic resonance scores associate with outcomes of patients with primary sclerosing cholangitis. *Clin. Gastroenterol. Hepatol.* **2019**, *17*, 2785–2792.e3. [\[CrossRef\]](#) [\[PubMed\]](#)
- Grigoriadis, A.; Ringe, K.I.; Andersson, M.; Kartalis, N.; Bergquist, A. Assessment of prognostic value and interreader agreement of ANALI scores in patients with primary sclerosing cholangitis. *Eur. J. Radiol.* **2021**, *142*, 109884. [\[CrossRef\]](#) [\[PubMed\]](#)

16. Grigoriadis, A.; Ringe, K.I.; Bengtsson, J.; Baubeta, E.; Forsman, C.; Korsavidou-Hult, N.; Rorsman, F.; Nilsson, E.; Kartalis, N.; Bergquist, A. Development of a prognostic MRCP-score (DiStrict) for individuals with large-duct primary sclerosing cholangitis. *JHEP Rep.* **2022**, *4*, 100595. [[PubMed](#)]
17. Grigoriadis, A.; Morsbach, F.; Voulgarakis, N.; Said, K.; Bergquist, A.; Kartalis, N. Inter-reader agreement of interpretation of radiological course of bile duct changes between serial follow-up magnetic resonance imaging/3D magnetic resonance cholangiopancreatography of patients with primary sclerosing cholangitis. *Scand. J. Gastroenterol.* **2020**, *55*, 228–235. [[CrossRef](#)] [[PubMed](#)]
18. Zenouzi, R.; Liwinski, T.; Yamamura, J.; Weiler-Normann, C.; Sebode, M.; Keller, S.; Lohse, A.W.; Schramm, C.; Aabakken, L.; Arrive, L.; et al. Follow-up magnetic resonance imaging/3D-magnetic resonance cholangiopancreatography in patients with primary sclerosing cholangitis: Challenging for experts to interpret. *Aliment. Pharmacol. Ther.* **2018**, *48*, 169–178. [[CrossRef](#)] [[PubMed](#)]
19. Poetter-Lang, S.; Ba-Ssalamah, A.; Messner, A.; Bastati, N.; Ambros, R.; Kristic, A.; Kittinger, J.; Pocheptnia, S.; Ba-Ssalamah, S.A.; Hodge, J.C. Disease severity prognostication in primary sclerosing cholangitis: A validation of the Anali scores and comparison with the potential functional stricture. *Eur. Radiol.* **2024**, *34*, 7632–7644. [[CrossRef](#)] [[PubMed](#)]
20. Tan, N.; Ngu, N.; Worland, T.; Lee, T.; Abrahams, T.; Pandya, K.; Freeman, E.; Hannah, N.; Gazelakis, K.; Madden, R.G.; et al. Epidemiology and outcomes of primary sclerosing cholangitis: An Australian multicentre retrospective cohort study. *Hepatol. Int.* **2022**, *16*, 1094–1104. [[CrossRef](#)] [[PubMed](#)]
21. Tan, N.; Ngu, N.; Worland, T.; Lee, T.; Abrahams, T.; Freeman, E.; Hannah, N.; Gazelakis, K.; Madden, R.G.; Lynch, K.D.; et al. Surveillance MRI is associated with improved survival in patients with primary sclerosing cholangitis. *Hepatol. Commun.* **2024**, *8*, e0442. [[CrossRef](#)] [[PubMed](#)]
22. Koo, T.K.; Li, M.Y. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* **2016**, *15*, 155–163. [[CrossRef](#)] [[PubMed](#)]
23. Venkatesh, S.K.; Welle, C.L.; Miller, F.H.; Jhaveri, K.; Ringe, K.I.; Eaton, J.E.; Bungay, H.; Arrivé, L.; Ba-Ssalamah, A.; Grigoriadis, A. Reporting standards for primary sclerosing cholangitis using MRI and MR cholangiopancreatography: Guidelines from MR Working Group of the International Primary Sclerosing Cholangitis Study Group. *Eur. Radiol.* **2022**, *32*, 923–937. [[PubMed](#)]
24. Vuppalanchi, R.; Are, V.; Telford, A.; Young, L.; Mouchti, S.; Ferreira, C.; Kettler, C.; Gromski, M.; Akisik, F.; Chalasani, N. A composite score using quantitative magnetic resonance cholangiopancreatography predicts clinical outcomes in primary sclerosing cholangitis. *JHEP Rep.* **2023**, *5*, 100834. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.