This is the accepted manuscript of the article, which has been published in Journal of pediatric gastroenterology and nutrition
https://dx.doi.org/10.1097/MPG.0000000000002522
Journal of Pediatric Gastroenterology and Nutrition, Publish Ahead of Print

DOI : 10.1097/MPG.0000000000002522

Assessing the validity of adult-derived prognostic models for primary sclerosing cholangitis outcomes in children

Authors:

name	Institution	email
Mark R. Deneau	**University of Utah, Salt Lake City, UT**	**mark.deneau@hsc.utah.edu**
Pamela L. Valentino	Yale University School of Medicine, New Haven, CT	pamela.valentino@yale.edu
Cara Mack	University of Colorado School of Medicine, Aurora, CO	Cara.Mack@childrenscolorado.org
Khaled Alqoaer	King Salman Armed Forces Hospital, Tabuk, Saudi Arabia	dralquaer@hotmail.com
Mansi Amin	Phoenix Children's Hospital, Phoenix, AZ	mamin@phoenixchildrens.com
Achiya Z. Amir	The Dana-Dwek Children's Hospital, The Tel-Aviv Medical Center, Tel-Aviv University, Tel Aviv, Israel	achiyaa@tlvmc.gov.il
Madeleine Aumar	Lille University Hospital of Lille, Lille, France	Madeleine.AUMAR@chru-lille.fr
Marcus Auth	Alder Hey Children's Hospital, Liverpool, United Kingdom	Marcus.Auth@alderhey.nhs.uk
Annemarie Broderick	University College Dublin, Dublin, Ireland	annemarie.broderick@ucd.ie
Matthew DiGuglielmo	Nemours Alfred I duPont Hospital For Children, Wilmington, DE	Matthew.diguglielmo@nemours.org
Laura G. Draijer	Academic Medical Centre, Amsterdam, The Netherlands	l.g.draijer@amc.uva.nl
Wael El-Matary	University of Manitoba, Winnipeg, Manitoba, Canada	welmatary@exchange.hsc.mb.ca
Federica Ferrari	Sapienza University of Rome, Rome, Italy	federica.ferrari@uniroma1.it
Katryn N. Furuya	Mayo Clinic, Rochester, MN	Furuya.Katryn@mayo.edu
Frederic Gottrand	Lille University Hospital of Lille, Lille, France	Frederic.GOTTRAND@CHRU-LILLE.FR
Nitika Gupta	Emory University School of Medicine, Atlanta, GA	nitika.gupta@emory.edu
Matjaz Homan	University of Ljubljana, Ljubljana, Slovenia	matjaz.homan@guest.arnes.si
M. K. Jensen	University of Utah, Salt Lake City, UT	kyle.jensen@hsc.utah.edu
Binita M.	University of Toronto, Toronto,	binita.kamath@sickkids.ca

Kamath	Ontario, Canada	
Kyung Mo Kim	University of Ulsan, Seoul, South Korea	kmkim@amc.seoul.kr
Kaija-Leena Kolho	University of Helsinki, Helsinki, Finland	kaija-leena.kolho@helsinki.fi
Bart Koot	Academic Medical Centre, Amsterdam, The Netherlands	b.g.koot@amc.uva.nl
Raffaele Iorio	University of Naples Federico II, Naples, Italy	riorio@unina.it
Mercedes Martinez	Columbia University College of Physicians and Surgeons, New York, NY	Mm2479@cumc.columbia.edu
Tamir Miloh	Texas Children's Hospital, Houston, TX	tamiloh@texaschildrens.org
Parvathi Mohan	Children's National Medical Center, Washington, DC	pmohan@childrensnational.org
Sirish Palle	Emory University School of Medicine, Atlanta, GA	Sirish-Palle@ouhsc.edu
Alexandra Papadopoulou	University of Athens, Athens, Greece	Papadop5@otenet.gr
Amanda Ricciuto	University of Toronto, Toronto, Ontario, Canada	amanda.ricciuto@sickkids.ca
Lawrence Saubermann	University of Rochester Medical Center, Rochester, NY	Lawrence_Saubermann@URMC.Rochester.edu
Pushpa Sathya	Memorial University, St. John's, Newfoundland and Labrador, Canada	Pushpa.Sathya@easternhealth.ca
Eyal Shteyer	Shaare Zedek Medical Center, Jerusalem, Israel	eyals@szmc.org.il
Vratislav Smolka	Palacky University, Olomouc, Czech Republic	Vratislav.Smolka@fnol.cz
Atsushi Tanaka	Teikyo University School of Medicine, Tokyo, Japan	a-tanaka@med.teikyo-u.ac.jp
Raghu Varier	Northwest Pediatric Gastroenterology LLC, Portland, OR	r.u.varier@gmail.com
Veena Venkat	University of Pittsburgh Medical Center, Pittsburgh, PA	venkatvl@upmc.edu
Bernadette Vitola	Medical College of Wisconsin, Milwaukee, WI	bvitola@mcw.edu
Marek Woynarowski	Children's Health Memorial Institute, Warsaw, Poland	m.woynarowski@med-net.pl
Stephen Guthery	University of Utah, Salt Lake City, UT	Stephen.guthery@hsc.utah.edu

Corresponding Author:

Mark Deneau MD, MS
Associate Professor of Pediatrics
Division of Gastroenterology, Hepatology and Nutrition
University of Utah
Intermountain Healthcare / Primary Children's Hospital
81 N. Mario Capecchi Dr.
Salt Lake City, UT 84113
(Phone) 801-213-3599
(Fax) 801-587-7539
mark.deneau@hsc.utah.edu

Keywords:

PSC, pediatric, risk stratification, natural history, prognosis

Potential Conflicts of interest:

Dr. Mark Deneau has consulted for HighTide Biopharmaceuticals LLC
Dr. Binita Kamath is a consultant for Retrophin
Dr. Tamir Miloh consults, advises and is on the speaker board for Alexion
Dr. Parvathi Mohan received grants from Gilead

Financial Support:

Research reported in this publication was supported by PSC Partners Seeking A Cure, the Primary Children's Hospital Foundation, the National Center for Advancing Translational Sciences of the National Institutes of Health under Award Numbers KL2TR001065 and 8UL1TR000105 (formerly UL1RR025764). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health

Acknowledgement:

The authors thank Drs. Jason Yap and Reham Abdou for assistance with data collection.

Author's Contributions:

All authors participated in study design, data abstraction, and authoring of the manuscript. Dr. Deneau additionally performed the statistical analysis.

Abstract:

Background:

Natural history models for primary sclerosing cholangitis (PSC) are derived from adult patient data, but have never been validated in children. It is unclear how accurate such models are for children with PSC.

Methods:

We utilized the pediatric PSC consortium database to assess the Revised Mayo Clinic, Amsterdam-Oxford and Boberg models. We calculated the risk stratum and predicted survival for each patient within each model using patient data at PSC diagnosis, and compared it to observed survival. We evaluated model fit using the c-statistic.

Results:

Model fit was good at one year (c-statistics 0.93, 0.87, 0.82) and fair at ten years (0.78, 0.75, 0.69) in the Mayo, Boberg and Amsterdam-Oxford models, respectively. The Mayo model correctly classified most children as low risk, whereas the Amsterdam-Oxford model incorrectly classified most as high risk. All of the models underestimated survival of patients classified as high risk. Albumin, bilirubin, AST and platelets were most associated with outcomes. Autoimmune hepatitis was more prevalent in higher risk groups, and over-weighting of AST in these patients accounted for the observed vs. predicted survival

Conclusion:

All three models offered good short-term discrimination of outcomes but only fair long-term discrimination. None of the models account for the high prevalence of features of autoimmune hepatitis overlap in children and the associated elevated aminotransferases. A pediatric-specific model is needed. AST, bilirubin, albumin and platelets will be important predictors, but must be weighted to account for the unique features of PSC in children.

Introduction:

Several natural history models have been derived from clinical data in adult populations with primary sclerosing cholangitis (PSC) (1-10). No consensus exists regarding the optimal model (11), and none have been validated for use in children. Important clinical differences exist between pediatric and adult-onset PSC patients. At PSC diagnosis, dominant strictures are present in 4% of children (12, 13), compared to 45% of adults (14). Similarly, cholangiocarcinoma is rare in pediatric-onset PSC, occurring in 1% of children by 10 years (12, 13), compared to at least 7-13% of adults (15-17). A small duct phenotype is present in 20% of children (12, 13), but only 10% of adults (18, 19). Features of autoimmune hepatitis overlap with PSC are present in over 33% of children (12, 13), but only 7% of adults with PSC (20, 21). With these clinical differences, it is unclear how well risk models derived from adult patient data are generalizable to children.

The most widely-used model to estimate transplant-free patient survival is the Revised Natural History Model for PSC, from a group at the Mayo Clinic (the "Mayo model") (5). It estimates survival with native liver for up to four years, and is available as an online calculator tool (22). A subsequent risk model from five European centers was created by Boberg et al. to more accurately estimate one-year survival to inform immediate transplant listing decisions (the "Boberg model") (6). The most recent Amsterdam-Oxford model (the "A-O model") included the largest model creation and validation cohorts to date, and had an added strength of originating from population-based data (10). It estimates survival with native liver out to 15 years, and is also available online (23). Characteristics of these models and their creation and validation cohorts are described and are compared to the Pediatric PSC Consortium in **Table 1**. We aimed to test the predictive utility of the Mayo, Boberg and A-O prognostic models for PSC using data from the Pediatric PSC Consortium, a large, multicenter cohort of children with PSC (12).

Methods:

We previously reviewed medical records on all known PSC patients at 36 different institutions throughout Europe, North America, the Middle East, and Asia (12). The PSC diagnosis was based on a cholestatic laboratory profile and either cholangiography showing multifocal stricturing and segmental dilations of the biliary tree and/or liver biopsy showing periductal, concentric fibrosis, fibro-obliterative cholangitis, or primary ductular involvement (11). Patients with abnormal cholangiograms were labeled as large duct PSC. Patients with normal cholangiograms but abnormal liver biopsy were labeled as small duct PSC. Autoimmune hepatitis (AIH) was diagnosed in patients who met a 'probable' or 'definite' score on the simplified AIH criteria that have been validated in children (24). We collected demographics, laboratory, histopathology, cholangiography and endoscopy data at liver disease diagnosis, as well as the presence of an esophageal variceal bleeding history. Alkaline phosphatase values were normalized for age. Complete data was present in 670/781 patients (86%). To account for missing data, we performed multivariate imputation using iteratively chained equations, combining the results of 10 imputed data sets. We validated the models using this imputed data set.

We calculated survival probabilities for each child using the equations derived from the Mayo (5), Boberg (6) and A-O (10) risk models **(Appendix, Supplemental Digital Content, http://links.lww.com/MPG/B735).** We did not validate other models because they necessitated access to original histopathology (1, 4), full images from cholangiography studies (7, 9), or included subjective assessments of organomegaly (2, 3, 8) that were not included in our dataset, and none are widely used. To generate observed survival probabilities, we created a retrospective cohort of all patients and followed them from time of PSC diagnosis to endpoints of liver transplantation or death from liver disease. Person-time was censored at the date of the last known clinical encounter. We used the Kaplan-

Meier method to calculate rates of survival each year after diagnosis. The endpoints of each model were somewhat different, with the Mayo model derived to predict only a risk of dying with a horizon of 4 years, and with liver transplant treated as if the patient would die within one year. The Boberg model was designed to predict one-year transplant-free survival, and the A-O model offered predictions for 10+ years. For uniformity in assessing multiple models, and to extrapolate longer-term prediction capability, we kept patients in their initial risk strata and observed survival out to ten years regardless of each model's original intent.

We evaluated the ability of each model to yield accurate survival probabilities for a given patient graphically, by comparing overlaid plots of observed and calculated survival probabilities. We plotted the Kaplan-Meier curve of observed outcomes alongside the annual predicted probabilities of survival for each risk group. For the plots of predicted survival, we calculated the median of the annual survival probabilities of each patient within each risk group, and connected these with straight lines (25, 26). The utility of risk score cutoffs specified by the adult models to stratify patients into distinct groups (e.g. "low" and "high" risk) with distinct observed survival probabilities was assessed using the logrank test. The logrank test is used to test the null hypothesis that there is no difference between the risk groups in the probability of an event (transplant or death) at any time point (27). Discriminatory ability of the models was assessed with the concordance statistic (c-statistic). The c-statistic was calculated by comparing observed and expected survival between every possible pairing of two of the 781 patients in the cohort (1 vs. 2, 1 vs. 3, … , 780 vs. 781). The c-statistic is the percentage of all 609,180 of these possible pairings that the model "guessed" correctly (assigned a worse predicted survival to the patient with the worst observed survival) (28). The c-statistic ranges from 0.5 (no discrimination, e.g. random risk stratification using a coin toss) to 1.0 (perfect discrimination), with values of 0.8 or higher generally regarded as "good discrimination" (29). We created time-truncated datasets to each

of one through ten years of follow-up and assessed the c-statistic for each time point to follow the accuracy of each model out to longer and longer prediction windows.

We broke down the median risk score in each risk group for the Mayo and A-O models and calculated the proportion of the risk score attributable to each individual predictor. We compared three or more groups of continuous variables using the Kruskal-Wallis test. All calculations were done using Stata version 13.0 (StataCorp, College Station, TX). The protocol of the study was approved by the institutional review and/or research ethics board of each collaborating institution.

Results:

The Revised Mayo Clinic Model:

The Mayo model was designed to report four-year outcomes. Overall, the Mayo model offered good discrimination of four-year outcomes with a c-statistic of 0.83. Predicted vs. observed survival with native liver (SNL) was similar in low, medium and high risk groups at one-year (99 vs. 99, 97 vs. 98, and 80 vs. 79%, respectively), but more disparate at four years (98 vs. 96, 89 vs. 79, and 33 vs. 47%, respectively). The low, medium and high risk cutoffs created three distinct populations of patients with progressively worse outcomes, logrank $p<0.001$ between all groups as shown in **Figure 1**. Most children were correctly stratified into the low risk group.

Serum albumin and aspartate aminotransferase levels made up the majority of the risk score for each patient, whereas total bilirubin, patient age, and variceal hemorrhage history contributed very little to the risk score, as shown in **Supplemental Figure 1 (Supplemental Digital Content, http://links.lww.com/MPG/B735)**. Each of the predictor variables varied significantly between groups as shown in **Supplemental Table 1 (Supplemental Digital Content, http://links.lww.com/MPG/B735)**. Inflammatory bowel disease was most

prevalent in low vs. medium and high risk groups: 80 vs. 73 vs. 52%, while autoimmune hepatitis was least prevalent in low vs. medium and high risk groups: 29 vs. 39 vs. 52%, respectively (both $p<0.001$). Large duct disease was distributed evenly among risk groups.

The Amsterdam-Oxford Model:

The A-O model was designed to report fifteen-year outcomes, but we had inadequate pediatric follow-up data to this time point and so followed it to a maximum of ten years. Overall, the A-O model offered fair discrimination of ten-year outcomes with a c-statistic of 0.69. Predicted vs. observed SNL was similar in low, low-intermediate, and medium risk groups, but disparate in the high risk group at one-year (100 vs. 99, 100 vs. 98, 100 vs. 97, 96 vs. 90%, respectively), five years (97 vs. 97, 96 vs. 94, 94 vs. 89, 83 vs. 66%, respectively) and ten years (88 vs. 93, 84 vs. 84, 76 vs. 74, 61 vs. 34%, respectively). The low, low-intermediate, medium and high risk cutoffs created four distinct populations of patients with progressively worse outcomes, log-rank $p<0.001$ between all groups as shown in **Figure 2**. The original model stratified 16, 34, 34 and 16% of adult patients as low, low-intermediate, medium and high risk, respectively. Children were stratified with 19, 9, 14 and 57% falling into these respective groups, over-classifying most as high risk.

Serum aspartate aminotransferase levels and platelet count made up the majority of the risk score for each patient, whereas total bilirubin, alkaline phosphatase, and albumin contributed little to the risk score, as shown in **Supplemental Figure 2 (Supplemental Digital Content, http://links.lww.com/MPG/B735)**. Age and large duct phenotype were similar in all risk groups, whereas all of the laboratory-based predictors were significantly different as shown in **Supplemental Table 2 (Supplemental Digital Content, http://links.lww.com/MPG/B735)**. Inflammatory bowel disease was equally prevalent in

lower risk groups: 84 vs. 77 vs. 80 vs. 74%, respectively, while autoimmune hepatitis was more prevalent in higher risk groups: 17 vs. 31 vs. 37% vs. 38%, respectively.

The Boberg model:

The Boberg model was designed to report one-year outcomes. The Boberg model provided excellent discrimination of one-year outcomes, with a c-statistic of 0.87, making it generally accurate at predicting if an individual patient would require liver transplantation or not on the basis of his or her laboratory studies. The patient's bilirubin (median 0.6 [IQR 0.4-1.2]) made up the majority of the prognostic score, accounting for 80%. Serum albumin (median 4 [IQR 3.6-4.4]) and patient age (median 12y [IQR 8-15]) accounted for 10% each. The model was overly pessimistic in predicting SNL for the group however. We observed 24 deaths or liver transplants in the first year after diagnosis, whereas the Boberg model predicted that over 170 would have occurred. The observed vs. predicted SNL at one year was 98% vs. 78%, respectively.

Model Comparison:

We assessed the performance of each model to discriminate outcomes at each of one to ten years after diagnosis, even though this was beyond the intended window for the Mayo and Boberg models. This is shown in **Figure 3**. The Mayo model was excellent at predicting need for transplant at one year, outperforming the other models (c-statistic 0.93 vs. 0.87 vs. 0.82 for the Mayo, Boberg and A-O models, respectively). Despite the Mayo score being designed for outcomes up to four years, and the Boberg model designed for outcomes at one year, use of either score as a predictor outperformed the A-O model at every time point cutoff through ten years. AST, platelet count, bilirubin and albumin were most associated with outcomes and accounted for the bulk of each risk score, where used in each model. Overall the Mayo model provided the best discrimination at all points in follow-up.

Discussion:

We used a large dataset of pediatric-onset PSC cases to assess the validity of prognostic and risk stratification tools created for adult PSC patients. We showed that the Mayo model offered the best discrimination of outcomes up to ten years. The Mayo and A-O models accurately estimated SNL in patients for 4-5 years after diagnosis. The Mayo model provided the best stratification to low, medium and high risk groups. A large source of inaccuracy of the models appeared to be weighting of AST that did not take into account the high prevalence of autoimmune hepatitis in children.

AST level contributed the largest variance explained in calculating risk scores in the Mayo and A-O model, and in stratifying patients into higher risk groups. AST rises with extensive fibrosis and cirrhosis. Indeed, the AST to Platelet ratio index (APRI) is a useful surrogate marker of hepatic fibrosis in many liver diseases (30-32), including PSC (33, 34). While an important predictor of disease progression, the Mayo and A-O models do not take into account the high prevalence of features of AIH overlap in children. At least one third of children with PSC are affected with AIH (12) compared to 0-5% of the adult cohorts (5, 6, 10) used to create these models. The median AST at diagnosis in children with PSC-AIH overlap was 290 U/L, yet most of these children had an uncomplicated clinical course, with a five-year SNL of 90% (12). The large number of children with marked elevations of AST that are unrelated to fibrosis, and which do not imply a negative prognosis, is the largest source of inaccuracy in prediction and risk stratification in these models.

It may seem remarkable that the models provide reasonable discrimination of outcomes at all, given the derivation and validation cohorts range in median age from 36-45 years old, and the median child in our cohort is only 12. Despite differing prevalence of complications at diagnosis of PSC, disease progression to new adverse liver events is similar between children and adults, occurring consistently in approximately 4% of patients each year. Cholangiocarcinoma is more common in adult patients however, who may have decades of disease duration and potential for hepatobiliary inflammation to progress to dysplasia and cancer. The higher rate of cholangiocarcinoma in adults (and their associated high mortality) is likely a large source of inaccuracy when pediatric data are entered into these models. There are no known differences in the underlying pathogenesis of PSC in children as compared to adults. Other than patient age, the laboratory markers and phenotypic features included in each of the adult models have generally been shown to be useful predictors in children (12,

13). It is likely that an optimized pediatric-specific model will include many of the same predictors, but will apply different weights to each. Bilirubin, platelet count and serum albumin are strong candidates for a pediatric model.

The strength of this study was the large size of the validation cohort we utilized. The Pediatric PSC Consortium is the largest cohort of pediatric-onset PSC patients, and includes a diverse mix of secondary and tertiary referral centers. The weakness of the study is the retrospective nature of the Pediatric PSC Consortium data. This prevented a standardized diagnostic and therapeutic algorithm for each patient, and misclassification bias may be present. While we were able to evaluate the most popular and user-friendly risk stratification models, were unable to evaluate all existing prognostic models due to lack of original histopathology and cholangiography data, and lack of subjective assessments of organomegaly in all patients.

In conclusion, we used the Pediatric PSC Consortium dataset to evaluate the validity of adult-derived prognostic models to predict clinical outcomes in children. The best discrimination, prediction and risk-stratification was provided by the Mayo model. None of the models accounted for the high prevalence of features of autoimmune hepatitis overlap in children and the associated elevations of aminotransferase levels that are unrelated to cirrhosis. Total bilirubin, albumin and platelet count are strong candidates for inclusion into a future pediatric-specific model. Weighting of predictors to account for the unique biochemical profile of children, is likely to yield more useful and accurate predictions and risk-stratification for pediatric-onset PSC.

References:

1. Wiesner RH, Grambsch PM, Dickson ER, Ludwig J, MacCarty RL, Hunter EB, Fleming TR, et al. Primary sclerosing cholangitis: natural history, prognostic factors and survival analysis. Hepatology 1989;10:430-436.
2. Farrant JM, Hayllar KM, Wilkinson ML, Karani J, Portmann BC, Westaby D, Williams R. Natural history and prognostic variables in primary sclerosing cholangitis. Gastroenterology 1991;100:1710-1717.
3. Dickson ER, Murtaugh PA, Wiesner RH, Grambsch PM, Fleming TR, Ludwig J, LaRusso NF, et al. Primary sclerosing cholangitis: refinement and validation of survival models. Gastroenterology 1992;103:1893-1901.
4. Broome U, Olsson R, Loof L, Bodemar G, Hultcrantz R, Danielsson A, Prytz H, et al. Natural history and prognostic factors in 305 Swedish patients with primary sclerosing cholangitis. Gut 1996;38:610-615.
5. Kim WR, Therneau TM, Wiesner RH, Poterucha JJ, Benson JT, Malinchoc M, LaRusso NF, et al. A revised natural history model for primary sclerosing cholangitis. Mayo Clin Proc 2000;75:688-694.
6. Boberg KM, Rocca G, Egeland T, Bergquist A, Broome U, Caballeria L, Chapman R, et al. Time-dependent Cox regression model is superior in prediction of prognosis in primary sclerosing cholangitis. Hepatology 2002;35:652-657.
7. Ponsioen CY, Vrouenraets SM, Prawirodirdjo W, Rajaram R, Rauws EA, Mulder CJ, Reitsma JB, et al. Natural history of primary sclerosing cholangitis and prognostic value of cholangiography in a Dutch population. Gut 2002;51:562-566.
8. Tischendorf JJ, Hecker H, Kruger M, Manns MP, Meier PN. Characterization, outcome, and prognosis in 273 patients with primary sclerosing cholangitis: A single center study. Am J Gastroenterol 2007;102:107-114.
9. Ponsioen CY, Reitsma JB, Boberg KM, Aabakken L, Rauws EA, Schrumpf E. Validation of a cholangiographic prognostic model in primary sclerosing cholangitis. Endoscopy 2010;42:742-747.
10. de Vries EM, Wang J, Williamson KD, Leeflang MM, Boonstra K, Weersma RK, Beuers U, et al. A novel prognostic model for transplant-free survival in primary sclerosing cholangitis. Gut 2017.
11. Chapman R, Fevery J, Kalloo A, Nagorney DM, Boberg KM, Shneider B, Gores GJ, et al. Diagnosis and management of primary sclerosing cholangitis. Hepatology 2010;51:660-678.
12. Deneau MR, El-Matary W, Valentino PL, Abdou R, Alqoaer K, Amin M, Amir AZ, et al. The natural history of primary sclerosing cholangitis in 781 children: A multicenter, international collaboration. Hepatology 2017;66:518-527.
13. Valentino PL, Wiggins S, Harney S, Raza R, Lee CK, Jonas MM. The Natural History of Primary Sclerosing Cholangitis in Children: A Large Single-Center Longitudinal Cohort Study. J Pediatr Gastroenterol Nutr 2016;63:603-609.
14. Bjornsson E, Lindqvist-Ottosson J, Asztely M, Olsson R. Dominant strictures in patients with primary sclerosing cholangitis. Am J Gastroenterol 2004;99:502-508.
15. Burak K, Angulo P, Pasha TM, Egan K, Petz J, Lindor KD. Incidence and risk factors for cholangiocarcinoma in primary sclerosing cholangitis. Am J Gastroenterol 2004;99:523-526.
16. Kornfeld D, Ekbom A, Ihre T. Survival and risk of cholangiocarcinoma in patients with primary sclerosing cholangitis. A population-based study. Scand J Gastroenterol 1997;32:1042-1045.

17. Bergquist A, Ekbom A, Olsson R, Kornfeldt D, Loof L, Danielsson A, Hultcrantz R, et al. Hepatic and extrahepatic malignancies in primary sclerosing cholangitis. J Hepatol 2002;36:321-327.
18. Angulo P, Maor-Kendler Y, Lindor KD. Small-duct primary sclerosing cholangitis: a long-term follow-up study. Hepatology 2002;35:1494-1500.
19. Bjornsson E, Boberg KM, Cullen S, Fleming K, Clausen OP, Fausa O, Schrumpf E, et al. Patients with small duct primary sclerosing cholangitis have a favourable long term prognosis. Gut 2002;51:731-735.
20. van Buuren HR, van Hoogstraten HJE, Terkivatan T, Schalm SW, Vleggaar FP. High prevalence of autoimmune hepatitis among patients with primary sclerosing cholangitis. J Hepatol 2000;33:543-548.
21. Kaya M, Angulo P, Lindor KD. Overlap of autoimmune hepatitis and primary sclerosing cholangitis: an evaluation of a modified scoring system. J Hepatol 2000;33:537-542.
22. Mayo Foundation for Medical Education and Research. The Revised Natural History Model for Primary Sclerosing Cholangitis (https://www.mayoclinic.org/medical-professionals/model-end-stage-liver-disease/revised-natural-history-model-for-primary-sclerosing-chonalgitis). In; 2000.
23. PSC Expertise Centrum - Academic Medical Centre. Amsterdam-Oxford PSC Score calculator: (https://www.amc.nl/web/leren/research-62/research/amsterdam-oxford-psc-score-calculator.htm). In. Amsterdam, NL; 2017.
24. Mileti E, Rosenthal P, Peters MG. Validation and modification of simplified diagnostic criteria for autoimmune hepatitis in children. Clin Gastroenterol Hepatol 2012;10:417-421 e411-412.
25. Arjas E. A Graphical Method for Assessing Goodness of Fit in Cox's Proportional Hazards Model. Journal of the American Statistical Association 1988;83:204-212.
26. Bland JM, Altman DG. Survival probabilities (the Kaplan-Meier method). BMJ 1998;317:1572.
27. Bland JM, Altman DG. The logrank test. BMJ 2004;328:1073.
28. Harrell FE, Jr., Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. JAMA 1982;247:2543-2546.
29. Caetano SJ, Sonpavde G, Pond GR. C-statistic: A brief explanation of its construction, interpretation and limitations. Eur J Cancer 2018;90:130-132.
30. D'Souza RS, Neves Souza L, Isted A, Fitzpatrick E, Vimalesvaran S, Cotoi C, Amin S, et al. AST-to-platelet ratio index in non-invasive assessment of long-term graft fibrosis following pediatric liver transplantation. Pediatr Transplant 2016;20:222-226.
31. Joshita S, Umemura T, Ota M, Tanaka E. AST/platelet ratio index associates with progression to hepatic failure and correlates with histological fibrosis stage in Japanese patients with primary biliary cirrhosis. J Hepatol 2014;61:1443-1445.
32. McGoogan KE, Smith PB, Choi SS, Berman W, Jhaveri R. Performance of the AST-to-platelet ratio index as a noninvasive marker of fibrosis in pediatric patients with chronic viral hepatitis. J Pediatr Gastroenterol Nutr 2010;50:344-346.
33. Vesterhus M, Hov JR, Holm A, Schrumpf E, Nygard S, Godang K, Andersen IM, et al. Enhanced liver fibrosis score predicts transplant-free survival in primary sclerosing cholangitis. Hepatology 2015;62:188-197.
34. de Vries EMG, Farkkila M, Milkiewicz P, Hov JR, Eksteen B, Thorburn D, Chazouilleres O, et al. Enhanced liver fibrosis test predicts transplant-free survival in primary sclerosing cholangitis, a multi-centre study. Liver Int 2017;37:1554-1561.
35. Talwalkar JA, Lindor KD. Natural history and prognostic models in primary sclerosing cholangitis. Best Pract Res Clin Gastroenterol 2001;15:563-575.

Figure 1. Observed vs. Predicted survival with native liver by risk group in the Mayo model

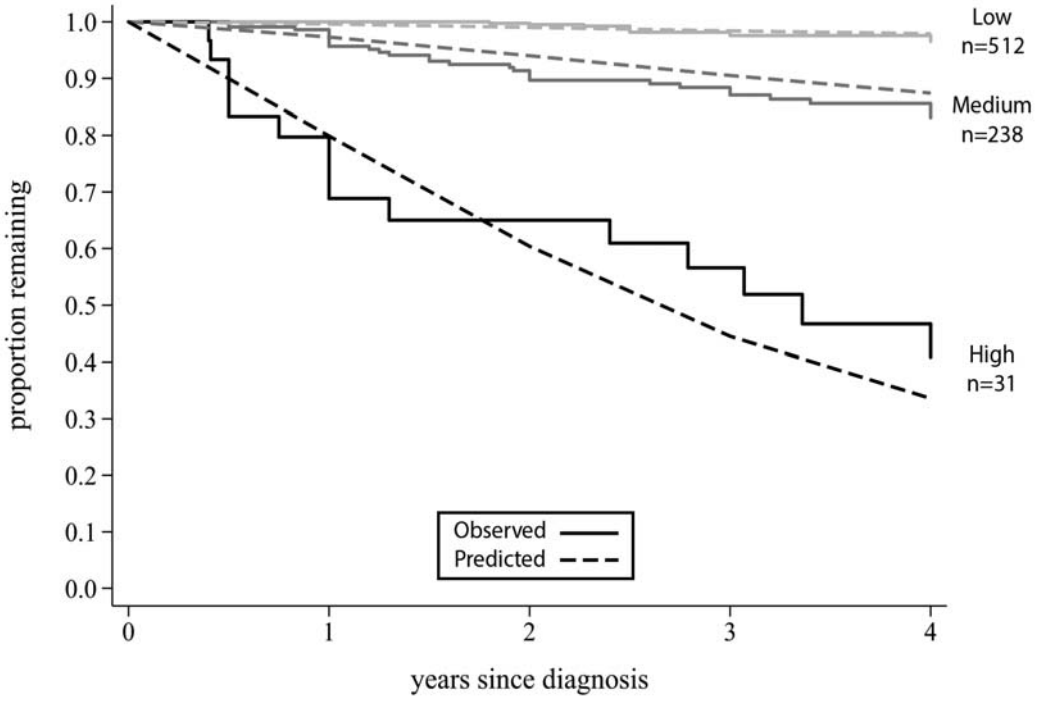

Figure 2: Observed vs. Predicted survival with native liver by risk group in the Amsterdam-Oxford model

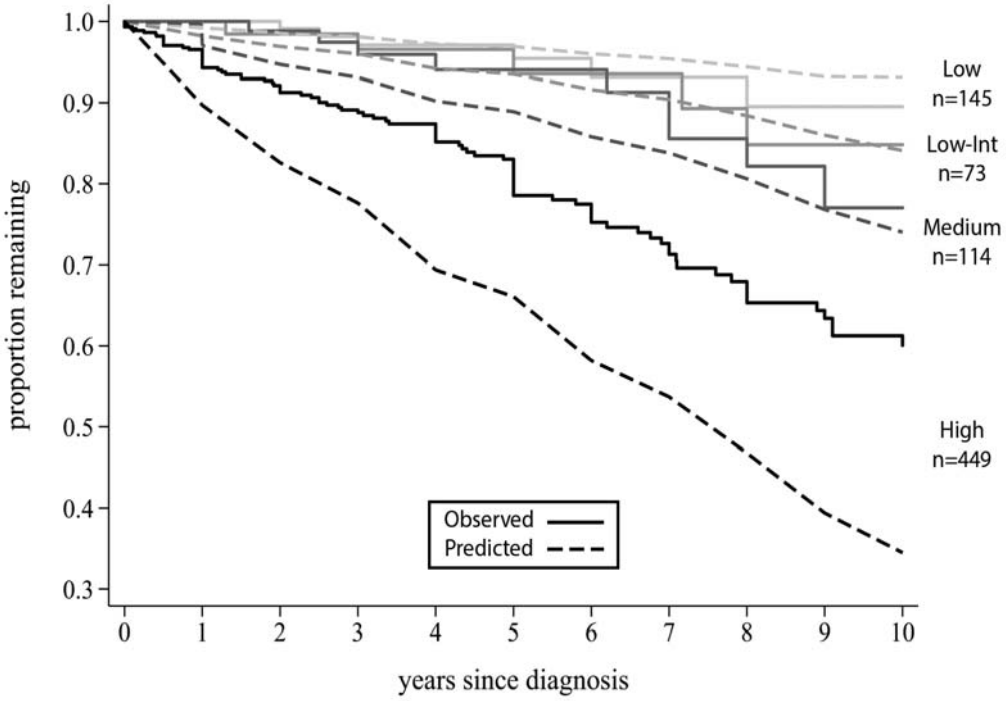

Figure 3: Comparison of model fit at annual time points after PSC diagnosis, extrapolated to 10 years

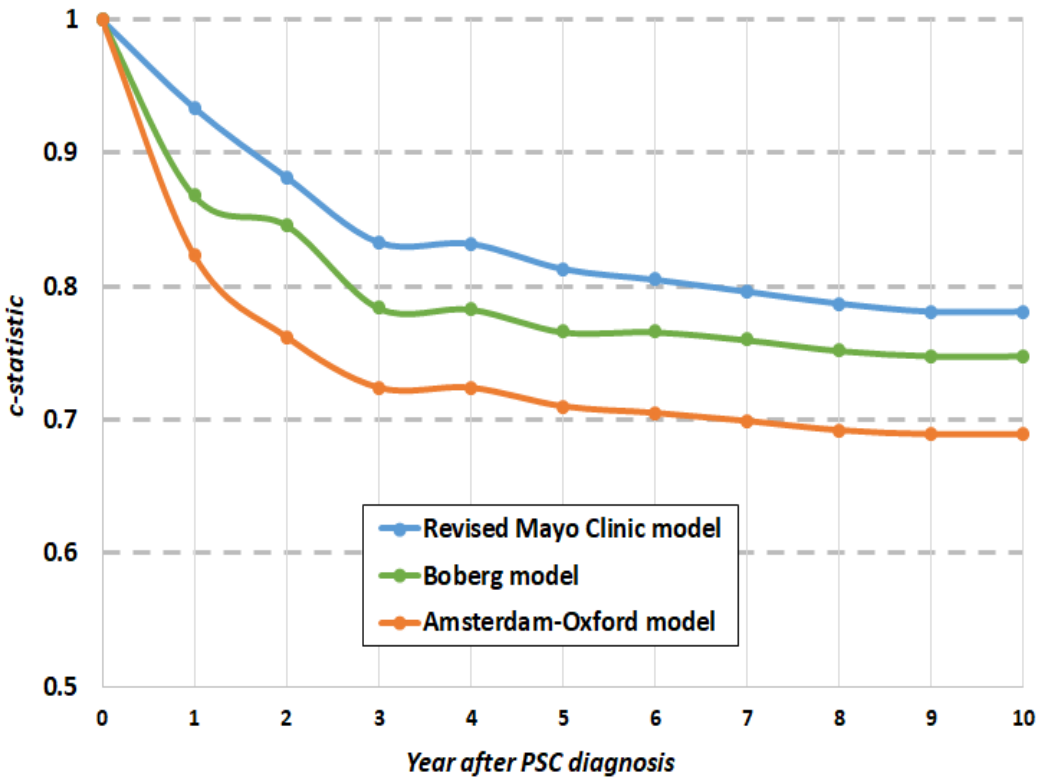

Table 1: Characteristics of Adult Prognostic Models for Primary Sclerosing Cholangitis

	Mayo Model Kim et al. 2000 (5)	**A-O Model** de Vries et al. 2017 (10)	**Boberg Model** Boberg et al. 2002 (6)	**Pediatric PSC Consortium** Deneau et al. 2017 (12)
Creation Cohort				
Location(s)	four United States referral centers	44 Dutch centers (population-based)	five European referral centers	36 North American, European, Middle East and Asian centers
n	405	692	330	781
age (years)	42	37 [IQR 27-49]	37 (range 13-82)	12
% female	33%	35%	32%	39%
% with IBD	74%	70%	83%	76%
% with AIH	0%	5%	0%	33%
median survival with native liver	not reported	20 years	11.7 years	16 years
Externally Validated?	Yes	Yes	No	
Validation Cohort				
location	King's College hospital London, UK	John Radcliffe hospital Oxford, UK		
n	124	264		
age (median)	36 years	45 years		
% with IBD	71%	74%		
% with AIH	0%	2%		
median survival with native liver	12 years	23 years		
Variables	age at diagnosis albumin AST bilirubin variceal bleeding history	age at diagnosis albumin alkaline phosphatase AST bilirubin large duct phenotype platelets	age at diagnosis albumin bilirubin	
Survival with native liver estimates:	1-4 years	1-15 years	1 year	