

Polygenic risk score predicts risk of primary sclerosing cholangitis in inflammatory bowel disease

Ming-Hsi Wang ,^{1,2} Jessica J Friton,² Laura E Raffals,² Jonathan A Leighton,³ Shabana F Pasha,³ Michael F Picco,⁴ Kelly Monroe,⁵ Billy D Nix,⁵ Rodney D Newberry,⁵ William A Faubion^{2,3}

To cite: Wang M-H, Friton JJ, Raffals LE, *et al*. Polygenic risk score predicts risk of primary sclerosing cholangitis in inflammatory bowel disease. *BMJ Open Gastroenterol* 2023;**10**:e001141. doi:10.1136/bmjgast-2023-001141

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjgast-2023-001141>).

RDN and WAF contributed equally.

Received 28 February 2023
Accepted 8 September 2023



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Mayo Clinic, Mankato, Minnesota, USA

²Mayo Clinic, Rochester, Minnesota, USA

³Mayo Clinic Scottsdale, Scottsdale, Arizona, USA

⁴Mayo Clinic in Florida, Jacksonville, Florida, USA

⁵Washington University in St Louis, St Louis, Missouri, USA

Correspondence to

Dr Ming-Hsi Wang;
wang.ming-hsi@mayo.edu

ABSTRACT

Background Forty distinct primary sclerosing cholangitis (PSC) genomic loci have been identified through multiancestry meta-analyses. The polygenic risk score (PRS) could serve as a promising tool to discover unique disease behaviour, like PSC, underlying inflammatory bowel disease (IBD).

Aim To test whether PRS indicates PSC risk in patients with IBD.

Materials and methods Mayo Clinic and Washington University at St Louis IBD cohorts were used to test our hypothesis. PRS was modelled through the published PSC loci and weighted with their corresponding effect size. Logistic regression was applied to predict the PSC risk.

Results In total, 63 (5.6%) among 1130 patients with IBD of European ancestry had PSC. Among 381 ulcerative colitis (UC), 12% had PSC; in contrast to 1.4% in 761 Crohn disease (CD). Compared with IBD alone, IBD-PSC had significantly higher PRS (PSC risk: 3.0% at the lowest PRS quartile vs 7.2% at the highest PRS quartile, $P_{\text{trend}} = .03$). In IBD subphenotypes subgroup analysis, multivariate analysis shows that UC-PSC is associated with more extensive UC disease (OR, 5.60; $p=0.002$) and younger age at diagnosis ($p=0.02$). In CD, multivariate analysis suggests that CD-PSC is associated with colorectal cancer (OR, 50; $p=0.005$).

Conclusions We found evidence that patients with IBD with PSC presented with a clinical course difference from that of patients with IBD alone. PRS can influence PSC risk in patients with IBD. Once validated in an independent cohort, this may help identify patients with the highest likelihood of developing PSC.

INTRODUCTION

Primary sclerosing cholangitis (PSC) is characterised by chronic immune-mediated stricturing bile ducts, which eventually leads to liver failure and the need for liver transplantation.¹ PSC is highly associated with inflammatory bowel disease (IBD), especially ulcerative colitis (UC). The prevalence of UC in patients with PSC ranges from 60% to 70%.² Conversely, PSC affects about 7% of patients with UC.³ It is important for clinicians to identify which patients with

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Inflammatory bowel disease combined (IBD) with primary sclerosing cholangitis (PSC) imposes a different clinical course and outcome. Genetic links to help differentiate the subgroups of IBD associated with PSC from classic IBD remain to be identified.

WHAT THIS STUDY ADDS

⇒ In our large multicentre cohort of patients with IBD who had long-term follow-up (average >10 years), through a polygenic risk score (PRS) model derived from genome-wide association studies-PSC loci, patients with IBD combined with PSC had significantly higher PRS and provides evidence of IBD-PSC may be genetically distinct from classic patients with IBD.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ This may help guide an individualised IBD patients' management strategy and identify patients with the highest likelihood of developing PSC. Further research in discovering which specific molecular pathways influencing the differences is warranted.

IBD are more likely to have PSC, since this could significantly determine IBD disease behaviour and clinical outcome and allow for a more personalised therapeutic approach.

The aetiology of PSC is not clearly understood. Studies of heritability have indicated a genetic component to the development of PSC. Recent genome-wide association studies (GWAS) have identified 40 regions of the genome, including the HLA locus at chromosome 6p21 and several regions located outside of previously known PSC loci, which underlies the risk of PSC and suggests the presence of shared genetic risk factors between PSC and UC.⁴⁻⁶ Among the established PSC loci, several have been found to be associated with UC.⁷ In the latest PSC GWAS, in combination with the International IBD Genetics Consortium data analysis to estimate the genetic

correlation (r_G) between PSC and subphenotypes of IBD, it was found that PSC is more genetically related to UC ($r_G=0.29$) than Crohn's disease (CD) ($r_G=0.04$) ($p=2.55E-15$) but is still much lower than the genetic correlation between UC and CD ($r_G=0.56$).⁵ In combination with previous study,⁴ these findings suggest that PSC-IBD is likely the result of a unique disease, which can be genetically distinct from classic IBD.

Genetic variants influencing complex traits—mostly single-nucleotide polymorphisms (SNPs) identified through GWAS—typically have little effects and limited predictive power. However, a larger proportion of heritability among complex traits can be explained by a polygenic risk score (PRS), which can be constructed through statistical models to evaluate the effects of all SNPs simultaneously.^{8,9} Recent GWAS studies have revealed 40 genomic regions related to the risk of developing PSC.⁵ However, the genetic biomarker to help differentiate the subgroups of IBD associated with PSC from classic IBD remains to be identified.

In the present study, we hypothesise that a PRS derived from the established PSC GWAS loci may help differentiate PSC-IBD from IBD without PSC through a large retrospective cohort.

MATERIALS AND METHODS

Cohort datasets

Multicentre GWAS data sets were used for this study: the Mayo Clinic IBD cohort (Cohort 1) (ImmunoChip custom genotyping array, Illumina) and Washington University IBD cohort (Cohort 2) (ImmunoChip custom genotyping array, Illumina). All data collection and study procedures were approved in both cohorts by the respective institutional review boards.

Both GWAS cohorts recruited adult patients (≥ 18 years) with a confirmed diagnosis of IBD. Clinical phenotype information was obtained through chart review, and biospecimens were collected for DNA sequencing. All participants gave written informed consent. Clinical factors included gender, age at diagnosis of UC, age at study enrolment, UC disease duration, disease location, tobacco use, family history of IBD and extraintestinal manifestations (EIMs). Disease location, behaviour and extent were classified according to the validated National Institutes of Diabetes and Digestive and Kidney Diseases IBD Genetics Consortium modification of the Montreal Classification, as previously described.¹⁰ EIMs included joint involvement (small joint, large joint, ankylosing spondylitis and sacroiliitis), eye involvement (iritis, uveitis, nonspecific ocular inflammation), skin involvement (erythema nodosum, pyoderma gangrenosum) and PSC.

Statistical analysis

We applied the following quality control criteria to exclude SNPs: (1) Hardy-Weinberg equilibrium test p values $< 1.0E-05$; (2) minor allele frequency less than 1%

and (3) call rate less than 95%. In total, 213 386 SNPs (ie, 84% of the originally genotyped SNPs) passed the quality control filters. Samples with cryptic relatedness (estimated identical-by-descent, PI-HAT more than 0.25) were excluded.

PRS was derived from the recently published PSC GWAS SNPs.⁵ Among the 40 published PSC loci, 36 were successfully genotyped in the cohort 1 and cohort 2. The PRS was calculated by computing the sum of risk alleles across the 36 SNPs that an individual has, weighted with their corresponding effect size (log-ORs) (online supplemental table 1).

ORs and 95% CI were calculated to estimate single locus effects for risk alleles and genotypes (online supplemental table 1, Excel spreadsheet). Two-sample t -tests were used for continuous variables, and nominal data were analysed using the χ^2 test or Fisher exact test. Age at diagnosis was analysed as a continuous variable. A two-tailed p value < 0.05 was considered significant. These analyses were conducted using the BlueSky Statistics V.7.40 and the SVS software suite V.8.3 (Golden Helix).

RESULTS

Clinical factors and PRS of PSC in patients with IBD

Among 1130 patients of European ancestry with IBD (cohort 1, $n=586$; cohort 2, $n=607$) who were successfully genotyped and had available detailed clinical information, in total, 63 (5.6%) were diagnosed with PSC (cohort 1, $n=27$ (4.6%); cohort 2, $n=36$ (5.9%)).

In univariate analysis, IBD-PSC was associated with more men (OR, 1.64; $p=0.06$), younger age at diagnosis of UC (OR, 0.97; $p=0.02$), longer IBD disease duration (18.9 years vs 13.3 years; $p<0.001$), less smoking (OR, 0.28; $p<0.001$) and less IBD-associated surgeries (OR, 0.51; $p=0.01$). Family history of IBD was not associated with PSC. Multivariate analysis showed that longer IBD disease duration ($p<0.001$), less smoking (OR, 0.34; $p=0.002$) and less IBD-associated surgeries (OR, 0.39; $p=0.05$) remained statistically significant (table 1).

Compared with IBD alone, IBD-PSC has higher PRS ($-2.58 \pm SE 0.08$ vs $-2.82 \pm SE 0.02$; $p=0.004$) (figure 1). The PSC risk among IBD was 3% at the lowest PRS quartile, compared with 7.2% at the highest PRS quartile ($P_{trend}=0.03$) (figure 2).

Clinical factors and PRS of PSC in patients with UC

Among a total of 381 patients with UC of European ancestry successfully genotyped and with available detailed clinical information, 46 (12.1%) were diagnosed with PSC. There was no significant difference of PRS between UC-PSC and UC alone ($-2.59 \pm SE 0.09$ vs $-2.73 \pm SE 0.03$; $p=0.14$).

In univariate analysis, UC-PSC was associated with younger age at diagnosis of UC (OR, 0.94; $p<0.001$), longer UC disease duration (21.6 years vs 11.7 years; $p<0.001$), more extensive disease (OR, 6.25; $p<0.001$) and less smoking (OR, 0.31; $p=0.007$). Family history of

Table 1 Associations between clinical factors and polygenic risk score between IBD with PSC and IBD without PSC in the cohort 1 and cohort 2

Predictors	Cohort 1 IBD with PSC/ without PSC (n=27/559)	Cohort 2 IBD with PSC/ without PSC (n=36/571)	Combined cohort IBD with PSC/without PSC (n=63/1130)	Univariate analysis OR (95% CI) P value	Multivariate analysis OR (95% CI) P value
Polygenic PSC risk score, mean (SE)	2.50 (0.12)/2.86 (0.03)	2.66 (0.10)/2.80 (0.03)	2.58 (0.08)/2.82 (0.02) (non-parametric Wilcoxon rank sum test p=0.004	1.78 (1.19 to 2.68) p=0.005	1.81 (1.16 to 2.84) p=0.009
Gender (male), number (%)	15 (62.5%)/236 (42.6%)	17 (47.2%)/226 (39.6%)	32 (53.3%)/462 (41.1%)	1.64 (0.97 to 2.77) p=0.06	1.26 (0.71 to 2.23) p=0.32
Family history of IBD (yes)	4 (14.8%)/87 (15.5%)	8 (22.2%)/81 (14.2%)	12 (19%)/168 (14.9%)	1.34 (0.67 to 2.49) p=0.38	NA
Age at diagnosis of IBD, mean (SE)	27.08 (2.92)/30.7 (0.61)	26.1 (2.40)/30.8 (0.59)	26.6 (1.89)/30.7 (0.42)	0.97 (0.95 to 0.99) p=0.02	0.99 (0.96 to 1.01) p=0.55
Disease duration, years (SE)	15.8 (11.1)/13.7 (11.1)	21.0 (13.7)/12.9 (8.6)	18.9 (12.9)/13.3 (9.9)	1.04 (1.02 to 1.06) p<0.001	1.05 (1.03 to 1.08) p<0.0001
Ever smoked (yes), number (%)	2 (7.4%)/145 (25.9%)	5 (13.8%)/201 (35.3%)	7 (11.1%)/346 (30.6%)	0.28 (0.12 to 0.59) p<0.001	0.34 (0.14 to 0.74) p=0.002
Bowel resection surgery (yes), number (%)	8 (29.6%)/247 (44.1%)	9 (25%)/225 (39.4%)	17 (27%)/471 (41.7%)	0.51 (0.28 to 0.89) p=0.01	0.39 (0.20 to 0.74) p=0.05

IBD, inflammatory bowel disease; PSC, primary sclerosing cholangitis.

IBD, male gender and intestinal resection were not associated with PSC in UC. Multivariate analysis showed that younger age at diagnosis (OR, 0.97; p=0.02), longer UC disease duration (p<0.0001) and extensive (pancolitis) disease (OR, 5.60; p=0.002) remained associated with PSC in UC but not ever-smoking (OR, 0.47; p=0.15) (table 2).

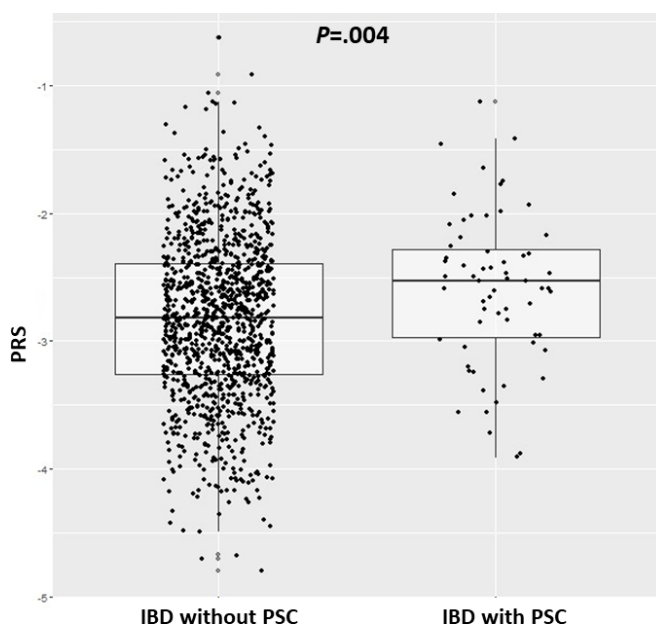


Figure 1 Boxplots of polygenic risk scores between PSC and non-PSC in IBD patients. IBD, inflammatory bowel disease; PSC, primary sclerosing cholangitis.

Clinical factors and PRS of PSC in patients with CD

Among a total of 711 CD patients of European ancestry successfully genotyped, and with available detailed clinical information, 11 patients (1.4%) were diagnosed with PSC. There was no significant difference of PRS between CD-PSC and CD alone ($-2.53 \pm SE 0.19$ vs $-2.87 \pm SE 0.02$; p=0.06).

In univariate analysis, CD-PSC was associated with a suggestive higher PRS (p=0.09), shorter disease duration (7.45 years vs 13.98 years; p=0.009), less complex CD behaviour (stricturing or penetrating) (OR, 0.17; p=0.009) and higher risk of colorectal cancer (OR, 18.9;

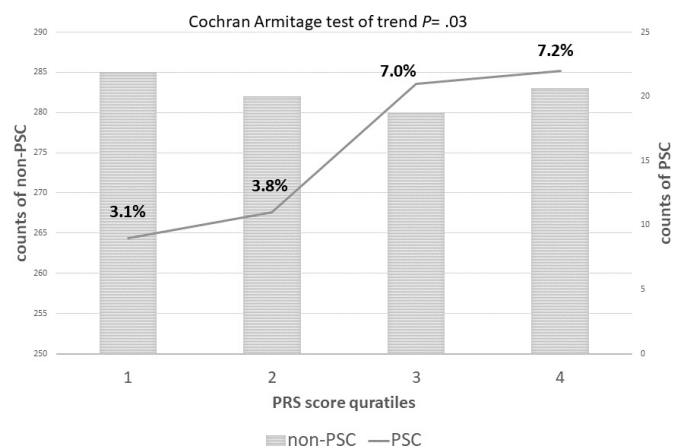


Figure 2 Plot of polygenic risk score quartiles and PSC status in IBD. IBD, inflammatory bowel disease; PSC, primary sclerosing cholangitis.

Table 2 Associations between clinical factors and polygenic risk score between UC with PSC and UC without PSC

Predictors	Combined cohort UC with PSC/without PSC (n=46/335)	Univariate analysis OR (95% CI) P value	Multivariate analysis OR (95% CI) P value
Polygenic PSC risk score, mean (SE)	-2.59 (0.09)/-2.73 (0.03) (non-parametric) Wilcoxon rank sum test p=0.14	1.48 (0.88 to 2.53) p=0.13	1.41 (0.80 to 2.50) p=0.23
Gender (male), number (%)	23 (50%)/167 (49.8%)	1.27 (0.67 to 2.48) p=0.46	NA
Family history of IBD (yes)	8 (17.4%)/38 (11.3%)	1.64 (0.67 to 3.63) p=0.25	NA
Age at diagnosis of UC, mean (SE)	25.0 (2.11)/34.3 (0.78)	0.94 (0.92 to 0.97) p<0.001	0.97 (0.94 to 0.99) p=0.02
Disease duration, years (SE)	21.6 (1.36)/11.7 (0.54)	1.08 (1.05 to 1.11) p<0.001	1.07 (1.04 to 1.10) p<0.0001
Ever smoked (yes), number (%)	5 (10.8%)/94 (28%)	0.31 (0.11 to 0.75) p=0.007	0.47 (0.17 to 1.33) p=0.15
Extended UC (yes), number (%)	42 (91.3%)/210 (62.6%)	6.25 (2.45 to 21.1) p<0.001	5.60 (1.88 to 16.6) p=0.002
Colorectal cancer (yes), number (%)	2 (4.3%)/5 (1.5%)	3.00 (0.42 to 14.3) p=0.23	NA
Bowel resection surgery (yes), number (%)	13 (28.3%)/66 (19.7%)	1.51 (0.28 to 0.89) p=0.19	NA

IBD, inflammatory bowel disease; PSC, primary sclerosing cholangitis; UC, ulcerative colitis.

p=0.05). Family history of IBD, age at diagnosis of CD, male gender, ever smoking and intestinal resection were not associated with PSC in CD. Interestingly, multivariate analysis showed shorter disease duration (p=0.04), and CRC (OR, 50; p=0.005) remained significantly associated with PSC among CD (table 3).

DISCUSSION

In a large multicentre cohort of patients with IBD who had long-term follow-up (average >10 years), 12.3% of patients with UC also had PSC, in contrast to 1.4% of patients with CD with PSC. Patients who had both IBD and PSC presented with a clinical course that differed from that of patients with IBD alone. UC patients with PSC were younger at the age of diagnosis and with more extensive disease, whereas CD patients with PSC had less complex CD behaviour (stricturing/penetrating). Through a PRS model from GWAS PSC loci, compared with IBD alone, IBD-PSC had significantly higher PRS (PSC risk: 3.0% at the lowest PRS quartile vs 7.2% at the highest PRS quartile, $P_{\text{trend}} = 0.03$). Our findings provide evidence of PSC in IBD, presenting a unique clinical course, and the PRS influencing the PSC risk among IBD.

In patients with UC, PSC is characterised by younger age at onset¹¹; milder but more extensive intestinal inflammation, rectal sparing and backwash ileitis^{12 13}; increased a threefold to fivefold higher risk of colonic carcinoma¹⁴; and higher risk for development of

pouchitis after proctocolectomy with ileoanal-pouch anastomosis¹⁵ compared with patients with UC without PSC. In this study, univariate analysis showed that patients with UC with PSC are more likely to be younger at UC diagnosis (25 vs 34.3 years old; p<0.001), have longer disease duration (21.6 years vs 11.7 years; p<0.001), more extensive disease (pancolitis) (OR, 6.25 (2.45 to 21.1); p<0.001), and less tobacco exposure (OR, 0.31 (0.11 to 0.75); p=0.007) when compared with patients with UC without PSC. Compared with UC alone, colorectal cancer was found more common in UC-PSC (4.3% vs 1.5%) but not statistically significant (p=0.23). Family history of first-degree relatives with IBD, male gender and intestinal surgical resection were not associated with PSC in those with UC. Multivariate analysis showed that younger age at diagnosis (p=0.02), longer UC disease duration (p<0.0001) and extensive (pancolitis) disease (OR, 5.60 (1.88 to 16.6); p=0.002) remained associated with PSC in patients with UC but not tobacco exposure (OR, 0.47 (0.17 to 1.33); p=0.15) (table 2). This suggests that UC-PSC is a unique subgroup of patients with UC.

Among patients with CD, univariate analysis showed that CD-PSC are more likely to have shorter disease duration (7.45 years vs 13.9 years; p<0.009), and less complex CD behaviour (stricturing or penetrating) (OR, 0.17 (0.03 to 0.67); p<0.009). Compared with CD alone, more colorectal cancer was observed in CD-PSC patients (9% vs 0.5%; OR, 18.9; p=0.05). This is consistent with

Table 3 Associations between clinical factors and polygenic risk score between CD with PSC and CD without PSC

Predictors	Combined cohort CD with PSC/without PSC (n=11/760)	Univariate analysis OR (95% CI) P value	Multivariate analysis OR (95% CI) P value
Polygenic PSC risk score, mean (SE)	-2.53 (0.19)/-2.87 (0.02) (non-parametric) Wilcoxon rank sum test p=0.06	2.16 (0.87 to 5.37) p=0.09	2.27 (0.89 to 5.76) p=0.08
Gender (male), number (%)	4 (35.4%)/277 (36.4%)	0.99 (0.30 to 3.90) p=0.98	NA
Family history of IBD (yes)	2 (18.2%)/125 (16.4%)	1.12 (0.17 to 4.45) p=0.88	NA
Age at diagnosis of IBD, mean (SE)	32.8 (5.1)/29.1 (0.5)	1.01 (0.98 to 1.06) p=0.39	NA
Disease duration, yes (SE)	7.45 (1.74)/13.9 (0.37)	0.88 (0.76 to 0.97) p=0.009	0.88 (0.78 to 0.99) p=0.04
Ever smoked (yes), number (%)	2 (18.2%)/245 (32.2%)	0.47 (0.07 to 1.83) p=0.30	NA
Complex phenotype (stricturing or penetrating), (yes) number (%)	2 (18.1%)/429 (56.4%)	0.17 (0.03 to 0.67) p=0.009	0.26 (0.05 to 1.27) p=0.09
Colorectal cancer, number (%)	1 (9%)/4 (0.5%)	18.9 (0.92 to 143) p=0.05	50.3 (3.08 to 820) p=0.005
Bowel resection surgery (yes), number (%)	4 (36.3%)/399 (52.5%)	0.52 (0.13 to 1.73) p=0.28	NA

CD, Crohn's disease; IBD, inflammatory bowel disease; PSC, primary sclerosing cholangitis.

the suggestive finding among UC. Family history of first-degree relatives with IBD, age at diagnosis of CD, male gender, ever smoking and intestinal resection were not associated with PSC in CD. Multivariate analysis showed that shorter disease duration ($p=0.04$) and CRC (OR, 50; $p=0.005$) remained associated with PSC among CD. This suggests that CD-PSC is a unique subgroup of CD patients. In addition, prior studies revealed that UC-PSC is associated with milder UC disease activity,^{11–13} which corresponds with our suggestive findings of less complex CD behaviour (stricturing/penetrating) in CD-PSC patients (table 3).

For the past two decades, GWAS studies have provided a great shift in our capacity to understand the genetic basis of human disease and have generated new insights into disease biology that have supported clinical translation. Genetic heterogeneity underpins the variable clinical natural history of IBD and the continuum of IBD subphenotypes. EIMs in patients with IBD are more frequently observed in those with the HLA-A2, HLA-DR1, HLA-DQw5¹⁶ and PTPN22 variants.¹⁷ Compared with UC alone, UC-PSC patients were found to be associated with two genetic loci: rs38904 (chr7, in the region of WNT2, CFTR; allele G vs A, OR 2.78; $p=0.01$) and rs11209026 (chr1, in IL23R; allele A vs G, OR 4.08; $p=0.02$)¹⁸; however, a larger cohort to validate these findings is warranted. In recent PSC GWAS studies, 40 genomic regions, including the HLA loci at chromosome 6p21 and additional 24 loci located outside of previously known PSC loci, related to

the risk of developing PSC, have been discovered⁵ and provide a fundamental tool to understand the underlying pathogenesis of PSC. The individual effect sizes of GWAS loci are mostly modest; however, examples such as variants linked to changes in integrin expression underscore the value of the regulatory consequences of GWAS variants.¹⁹

To our knowledge, the application of genetic risk score to IBD EIMs, especially PSC, is yet to be defined in the literature. In this study, we applied the PSC genetic risk score in a large multicentre cohort of patients with IBD who had long-term follow-up (average >10 years). Compared with IBD alone, IBD-PSC has a significantly higher PRS (-2.58 vs -2.82 ; $p=0.004$) and a significant trend of PSC risk (highest PSC risk 7.2% at the highest PRS quartile vs 3.0% at the lowest PRS quartile, $P_{\text{trend}}=0.03$). Our findings provide evidence that genetic risk score influences PSC risk in patients with IBD. This suggests that IBD-PSC may be genetically distinct from IBD alone.

Some limitations of our study warrant consideration. First, this is a retrospective study and the number of patients with IBD diagnosed with PSC is generally modest. The study power of comparing the PRS within IBD subphenotypes is further limited by the smaller number of samples. Second, this study was not designed for formal replication because there was no stringent statistical significance imposed separately in the two cohorts. Third, the exact date of PSC diagnosis was not available, and we were unable to determine the time interval



between IBD diagnosis and PSC diagnosis. Fourth, liver transplantation status and posttransplant immunosuppressive agents used in PSC patients were not available, which could affect the clinical course and disease extent of IBD; however, we do not think this would significantly confound our results.

Our study also has several strengths. We have long-term follow-up IBD cohorts and focused on the clinically relevant characteristics to identify the predictors of PSC risk in patients with IBD. In addition, the included studies were high quality according to the methodologic quality assessment in IBD clinical behaviour and phenotyping and genotyping data.

In conclusion, patients with IBD with PSC presented with a clinical course difference from that of patients with IBD alone. Among patients with UC, those with UC-PSC were younger at age of diagnosis and with more extensive disease. A similar trend of less complex CD behaviour— inflammatory only—was observed in CD-PSC than in CD. Through a genetic risk score (PRS) modelled from GWAS PSC loci, compared with IBD alone, IBD-PSC has a significantly higher PRS. Our findings provide evidence of PSC in IBD and that the PRS indicates PSC risk in patients with IBD, which suggests that PSC in IBD may be genetically distinct from most patients with IBD. Future work of replicating our findings in an independent and larger cohort is warranted. Once our result is validated, this may help guide an individualised IBD management strategy and identify patients with the highest likelihood of developing PSC.

Acknowledgements The Scientific Publications staff at Mayo Clinic provided copyediting support.

Contributors M-HW contributed to the conception and design, analysis and interpretation of the data, drafting and critical revision of the article, and generation/collection of figures. JJF contributed to the experiments and collection of data. LER, JAL, SFP, MFP, RDN, WAF contributed to the conception and design and critical revision of the article. KM contributed to the experiments and collection of data. BDN contributed to the experiments and collection of data. M-HW acts as guarantor.

Funding This study was supported by the Mayo Clinic Health System Intramural Grant (Research Grant: Company 301 PAU 47931).

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by Mayo Clinic IRB # ID: 15-000295. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information. Not applicable.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Ming-Hsi Wang <http://orcid.org/0000-0002-1188-8258>

REFERENCES

- 1 Ponsioen CY, Vrouenraets SME, Prawirodirdjo W, *et al*. Natural history of primary sclerosing cholangitis and prognostic value of cholangiography in a Dutch population. *Gut* 2002;51:562–6.
- 2 Boonstra K, van Erpecum KJ, van Nieuwkerk KMJ, *et al*. Primary sclerosing cholangitis is associated with a distinct phenotype of inflammatory bowel disease. *Inflamm Bowel Dis* 2012;18:2270–6.
- 3 Fausa O, Schrumph E, Elgjo K. Relationship of inflammatory bowel disease and primary sclerosing cholangitis. *Semin Liver Dis* 1991;11:31–9.
- 4 Ellinghaus D, Jostins L, Spain SL, *et al*. Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared Loci. *Nat Genet* 2016;48:510–8.
- 5 Ji S-G, Juran BD, Mucha S, *et al*. Genome-wide Association study of primary sclerosing cholangitis identifies new risk Loci and quantifies the genetic relationship with inflammatory bowel disease. *Nat Genet* 2017;49:269–73.
- 6 Liu JZ, Hov JR, Folseraas T, *et al*. Dense genotyping of immune-related disease regions identifies nine new risk Loci for primary sclerosing cholangitis. *Nat Genet* 2013;45:670–5.
- 7 Ellinghaus D, Folseraas T, Holm K, *et al*. Genome-wide Association analysis in primary sclerosing cholangitis and ulcerative colitis identifies risk Loci at GPR35 and TCF4. *Hepatology* 2013;58:1074–83.
- 8 Palla L, Dudbridge F. A fast method that uses polygenic scores to estimate the variance explained by genome-wide marker panels and the proportion of variants affecting a trait. *Am J Hum Genet* 2015;97:250–9.
- 9 Yang J, Benyamin B, McEvoy BP, *et al*. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* 2010;42:565–9.
- 10 Dassopoulos T, Nguyen GC, Bitton A, *et al*. Assessment of reliability and validity of IBD phenotyping within the National Institutes of diabetes and digestive and kidney diseases (NIDDK) IBD Genetics consortium (IBDGC). *Inflamm Bowel Dis* 2007;13:975–83.
- 11 Wang M-H, Mousa OY, Friton JJ, *et al*. Unique phenotypic characteristics and clinical course in patients with ulcerative colitis and primary sclerosing cholangitis: a multicenter US experience. *Inflamm Bowel Dis* 2020;26:774–9.
- 12 Loftus EV, Harewood GC, Loftus CG, *et al*. PSC-IBD: a unique form of inflammatory bowel disease associated with primary sclerosing cholangitis. *Gut* 2005;54:91–6.
- 13 Sano H, Nakazawa T, Ando T, *et al*. Clinical characteristics of inflammatory bowel disease associated with primary sclerosing cholangitis. *J Hepatobiliary Pancreat Sci* 2011;18:154–61.
- 14 Soetikno RM, Lin OS, Heidenreich PA, *et al*. Increased risk of colorectal neoplasia in patients with primary sclerosing cholangitis and ulcerative colitis: a meta-analysis. *Gastrointest Endosc* 2002;56:48–54.
- 15 de Vries AB, Janse M, Blokzijl H, *et al*. Distinctive inflammatory bowel disease phenotype in primary sclerosing cholangitis. *World J Gastroenterol* 2015;21:1956–71.
- 16 Orchard TR, Chua CN, Ahmad T, *et al*. Uveitis and erythema nodosum in inflammatory bowel disease: clinical features and the role of HLA genes. *Gastroenterology* 2002;123:714–8.
- 17 Cleynen I, González JR, Figueroa C, *et al*. Genetic factors conferring an increased susceptibility to develop Crohn's disease also influence disease phenotype: results from the Ibdchip European project. *Gut* 2013;62:1556–65.
- 18 Wang MH, Fiocchi C, Zhu X, *et al*. Gene-gene and gene-environment interactions in ulcerative colitis. *Hum Genet* 2014;133:547–58.
- 19 de Lange KM, Moutsianas L, Lee JC, *et al*. Genome-wide Association study Implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat Genet* 2017;49:256–61.