Title

A polygenic risk score for alcohol-associated cirrhosis among heavy drinkers with European ancestry.

Journal

Hepatology Communications, 8(6)

Authors

Schwantes-An, Tae-Hwi

Whitfield, John

Aithal, Guruprasad

et al.

Publication Date

2024-06-01

DOI

10.1097/HC9.0000000000000431

Peer reviewed

Received: 10 July 2023 | Accepted: 1 November 2023

DOI: 10.1097/HC9.0000000000000431

ORIGINAL ARTICLE

OPEN

A polygenic risk score for alcohol-associated cirrhosis among heavy drinkers with European ancestry

Tae-Hwi Schwantes-An[1] | John B. Whitfield[2] | Guruprasad P. Aithal[3] | Stephen R. Atkinson[4] | Ramon Bataller[5] | Greg Botwin[6,7] | Naga P. Chalasani[8] | Heather J. Cordell[9] | Ann K. Daly[10] | Rebecca Darlay[9] | Christopher P. Day[11] | Florian Eyer[12] | Tatiana Foroud[1] | Samer Gawrieh[8] | Dermot Gleeson[13] | David Goldman[14] | Paul S. Haber[15,16] | Jean-Marc Jacquet[17] | Craig S. Lammert[8] | Tiebing Liang[8] | Suthat Liangpunsakul[18] | Steven Masson[10] | Philippe Mathurin[19] | Romain Moirand[20] | Andrew McQuillin[21] | Christophe Moreno[22] | Marsha Y. Morgan[23] | Sebastian Mueller[24] | Beat Müllhaupt[25] | Laura E. Nagy[26] | Pierre Nahon[27,28,29] | Bertrand Nalpas[17,30] | Sylvie Naveau[31] | Pascal Perney[32] | Munir Pirmohamed[33] | Helmut K. Seitz[24] | Michael Soyka[34] | Felix Stickel[25] | Andrew Thompson[33,35] | Mark R. Thursz[4] | Eric Trépo[22] | Timothy R. Morgan[36,37] | Devanshi Seth[15,16,38] | for the GenomALC Consortium

[1]Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis IN, USA

[2]Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Queensland 4029, Australia

[3]NIHR Nottingham Biomedical Research Centre, Nottingham University Hospitals and the University of Nottingham, Nottingham NG7 2UH, UK

[4]Department of Metabolism, Digestion & Reproduction, Imperial College London, UK

[5]Center for Liver Diseases, University of Pittsburgh Medical Center, 3471 Fifth Avenue, Pittsburgh, PA 15213, USA

[6]Department of Veterans Affairs, VA Long Beach Healthcare System, 5901 East Seventh Street, Long Beach, CA 90822, USA

[7]F. Widjaja Family Foundation Inflammatory Bowel and Immunobiology Research Institute, Cedars-Sinai Medical Center, Los Angeles, California CA 90048, USA

[8]Department of Medicine, Indiana University, Indianapolis, IN 46202-5175, USA

[9]Population Health Sciences Institute, Faculty of Medical Sciences, Newcastle University, International Centre for Life, Central Parkway, Newcastle upon Tyne NE1 3BZ, UK

[10]Faculty of Medical Sciences, Newcastle University Medical School, Framlington Place, Newcastle upon Tyne NE2 4HH, UK

[11]Newcastle University, Framlington Place, Newcastle upon Tyne NE2 4HH, UK

[12]Division of Clinical Toxicology, Department of Internal Medicine 2, Klinikum rechts der Isar, School of Medicine, Technical University of Munich, Ismaninger Str. 22, 81675 Munich, Germany

[13]Liver Unit, Sheffield Teaching Hospitals, AO Floor Robert Hadfield Building, Northern General Hospital, Sheffield S5 7AU, UK

[14]Office of the Clinical Director and Laboratory of Neurogenetics, NIAAA, Bethesda, MD 20952, USA

[15]Edith Collins Centre (Translational Research in Alcohol Drugs and Toxicology), Sydney Local Health District, Missenden Road, Camperdown, NSW 2050, Australia

[16]Faculty of Medicine and Health, the University of Sydney, Sydney, NSW 2006, Australia

[17]Service Addictologie, CHRU Caremeau, 30029 Nîmes, France

[18]Division of Gastroenterology and Hepatology, Department of Medicine, Indiana University and Roudebush Veterans Administration Medical Center, Indianapolis, USA

[19]CHRU de Lille, Hôpital Claude Huriez, Rue M. Polonovski CS 70001, 59 037 Lille Cedex, France

[20]Univ Rennes, INRA, INSERM, CHU Rennes, Institut NUMECAN (Nutrition Metabolisms and Cancer), F-35000 Rennes, France

[21]Molecular Psychiatry Laboratory, Division of Psychiatry, University College London, London WC1E 6DE, UK

[22]CUB Hôpital Erasme, Université Libre de Bruxelles, clinique d'Hépatologie, Brussels, Belgium; Laboratory of Experimental Gastroenterology, Université Libre de Bruxelles, Brussels, Belgium

[23]UCL Institute for Liver & Digestive Health, Division of Medicine, Royal Free Campus, University College London, London NW3 2PF, UK

[24]Department of Internal Medicine, Salem Medical Center and Center for Alcohol Research, University of Heidelberg, Zeppelinstraße 11-33, 69121 Heidelberg, Germany

[25]Department of Gastroenterology and Hepatology, University Hospital Zurich, Rämistrasse 100, CH-8901 Zurich, Switzerland

[26]Lerner Research Institute, 9500 Euclid Avenue, Cleveland, Ohio, OH 44195, USA

[27]Service d'Hépatologie, APHP Hôpital Avicenne et Université Paris 13, Bobigny, France

[28]University Paris 13, Bobigny, France

[29]Inserm U1162 Génomique fonctionnelle des tumeurs solides, Paris, France

[30]DISC, Inserm, 75013 Paris, France

[31]Hôpital Antoine-Béclère, 157 Rue de la Porte de Trivaux, 92140 Clamart, France

[32]Hôpital Universitaire Caremeau, Place du Pr. Robert Debre, 30029 Nîmes, France

[33]MRC Centre for Drug Safety Science, Liverpool Centre for Alcohol Research, University of Liverpool, The Royal Liverpool and Broadgreen University Hospitals NHS Trust, and Liverpool Health Partners, Liverpool, L69 3GL, UK

[34]Psychiatric Hospital University of Munich, Nussbaumsstr.7, 80336 Munich, Germany

[35]Health Analytics, Lane Clark & Peacock LLP, London, UK

[36]Department of Medicine, University of California, Irvine, USA

[37]Department of Veterans Affairs, VA Long Beach Healthcare System, 5901 East Seventh Street, Long Beach, CA 90822, USA

[38]Centenary Institute of Cancer Medicine and Cell Biology, the University of Sydney, Sydney, NSW 2006, Australia

Correspondence
Tae-Hwi Schwantes-An, HS 4023, 410 W, 10th Street, Indianapolis, Indiana, USA.
Email: tlschwan@iu.edu

Devanshi Seth, Centenary Institute of Cancer Medicine and Cell Biology, The University of Sydney, Sydney, NSW 2006, Australia.
Email: d.seth@sydney.edu.au

Abstract

Background: Polygenic Risk Scores (PRS) based on results from genome-wide association studies offer the prospect of risk stratification for many common and complex diseases. We developed a PRS for alcohol-associated cirrhosis by comparing single-nucleotide polymorphisms among patients with alcohol-associated cirrhosis (ALC) versus drinkers who did not have evidence of liver fibrosis/cirrhosis.

Methods: Using a data-driven approach, a PRS for ALC was generated using a meta-genome-wide association study of ALC (N=4305) and an independent cohort of heavy drinkers with ALC and without significant liver disease (N=3037). It was validated in 2 additional independent cohorts from the UK Biobank with diagnosed ALC (N=467) and high-risk drinking controls (N=8981) and participants in the Indiana Biobank Liver cohort with alcohol-associated liver disease (N=121) and controls without liver disease (N=3239).

Results: A 20-single-nucleotide polymorphisms PRS for ALC (PRSALC)

Abbreviations: ALC, alcohol-associated liver cirrhosis; ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; GWAS, Genome-wide Association Study; *HSD17B13*, hydroxysteroid 17-beta dehydrogenase 13 gene; IB-Liver, Indiana Biobank Liver; ICD, International Classification of Diseases; *KLHL8*, Kelch-like family member 8 gene; LNG, Laboratory for Neurogenetics; MASLD, metabolic dysfunction–associated steatotic liver disease; *MBOAT7*, membrane-bound O-acyltransferase domain containing 7; NIAAA, National Institute on Alcohol Abuse and Alcoholism; NPV, negative predictive value; *PDE7B*, phosphodiesterase 7B gene; *PNPLA3*, patatin-like phospholipase domain-containing 3 gene, PPV, positive predictive value; PRS, Polygenic Risk Score; R-sq, R-squared; *SERPINA1*, serpin family A member 1 gene; SNP, single-nucleotide polymorphism; *SUGP1*, SURP and G-patch domain containing 1; *TM6SF2*, transmembrane 6 superfamily member 2 gene; UKB, UK Biobank.

Timothy R. Morgan and Devanshi Seth are equal senior authors.

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.hepcommjournal.com.

was generated that stratified risk for ALC comparing the top and bottom deciles of PRS in the 2 validation cohorts (ORs: 2.83 [95% CI: 1.82 -4.39] in UK Biobank; 4.40 [1.56 -12.44] in Indiana Biobank Liver cohort). Furthermore, PRSALC improved the prediction of ALC risk when added to the models of clinically known predictors of ALC risk. It also stratified the risk for metabolic dysfunction -associated steatotic liver disease -cirrhosis (3.94 [2.23 -6.95]) in the Indiana Biobank Liver cohort -based exploratory analysis. **Conclusions:** PRS_{ALC} incorporates 20 single-nucleotide polymorphisms, predicts increased risk for ALC, and improves risk stratification for ALC compared with the models that only include clinical risk factors. This new score has the potential for early detection of heavy drinking patients who are at high risk for ALC.

INTRODUCTION

Polygenic risk scores (PRS) offer a tool for stratifying disease risk and initiating action in the presymptomatic phase for a wide range of common diseases that do not currently have useful biochemical or physiological risk markers.[1–5] PRSs are generated from genome-wide association studies[1,6,7] and increasingly, from sequencing data, which can provide information on less common but large-effect variants.[8] In either case, large cohort-based studies are needed to derive unbiased estimates of effect sizes, which are often small to moderate, as hundreds of thousands of variants are likely to be involved in the genetic underpinnings of common and complex diseases.

Because the genetic architectures of common diseases (ie, the number and effect sizes of loci contributing appreciably to the overall risk) vary considerably, PRS needs to be evaluated on a case-by-case basis using ORs and measures of sensitivity and specificity. In the context of risk stratification, sensitivity does not necessarily need to meet the high levels required for a diagnostic test, but it is essential that people in the top decile or quintile for PRS have a substantially elevated risk compared to the general population or those in the lower-PRS deciles. Specificity also needs to be considered; does the PRS identify risk for only one disease, for a group of diseases that have a genetic component in common, or for diseases that have previously been viewed as unrelated? Lack of specificity might not be a disadvantage, provided that its impact on any PRS-based diagnostic or management pathway is recognized.

We have recently developed and evaluated a 3 single-nucleotide polymorphism (SNP)-based risk score for alcohol-associated liver cirrhosis (ALC).[9] This 3-SNP–based genetic risk score included well-known genetic risk loci for both alcohol-associated and nonalcohol-associated liver diseases, using rs738409 for patatin-like phospholipase domain-containing protein 3 (*PNPLA3*),[10] rs6834314 for hydroxysteroid 17-beta dehydrogenase 13 (*HSD17B13*),[11] and rs10401969 for Transmembrane 6 superfamily member 2 (*TM6SF2*).[12] The score identified a 3-fold difference in risk (for cirrhosis among high-risk drinkers) between people in the top and bottom quintiles, which was also recently confirmed in an independent cohort.[13] The score performed well for risk stratification in ALC, comparably to PRSs for other common diseases with larger sample sizes ($N \geq 100{,}000$) and based on hundreds of thousands of variants.[6,14]

Two questions arise from our 3-SNP genetic risk score. First, can the inclusion of additional loci (SNPs), and hypothesis-free generation of a risk score by an algorithmic process, improve the estimate of genetic risk for developing alcohol-associated cirrhosis? Second, can the PRS for alcohol-associated cirrhosis also stratify risk for other related liver diseases? To answer these questions, we generated a new *a priori* PRS using a data-driven approach (ie, PRSice-2)[7] using results from our and other Genome-wide Association Study (GWAS)[10,15] on ALC risk and from an additional GenomALC-2 cohort including patients with ALC and heavy drinking controls. We also validated the new PRS in 2 independent cohorts, the UK Biobank (UKB) and the Indiana Biobank Liver (IB-Liver) cohort.

METHODS

Principles of generating PRS for ALC

Three independent data sets are required to generate and validate a PRS;

(1) Base GWAS that contains variants' *p*-values and allelic effect estimates
(2) Target cohort that contains individual-level genotypes and phenotype data

(3) Validation cohort(s) that also contains individual-level data but independent of the base and target cohorts.

The *p*-values from the base GWAS are used to filter SNPs based on significance. At each *p*-value threshold, betas from SNPs that pass the threshold are multiplied by the genotype values in the target data to calculate a PRS. This is done across multiple *p*-value thresholds to generate many PRSs. The PRS that maximizes the variance explained in the target data is chosen as the best PRS. It is then validated in an independent cohort.

In the present study, combined GWAS data from the GenomALC-1 and the Laboratory for Neurogenetics (LNG) cohorts and the summary statistics reported by Buch et al[10] were meta-analyzed to provide the *Base* GWAS. Data from the *Target* GenomALC-2 cohort was used to generate and identify the PRS that maximized the variance explained. Data from the UKB and the IB-Liver cohort were used as independent Validation cohorts. Details of the individual cohorts are described below and summarized in Figure 1.

Base GWAS participants

GenomALC-1

The GenomALC study participants were recruited at clinical sites in Australia, France, Germany, Switzerland, the United Kingdom, and the United States using our previously published enrollment protocol.[16] Briefly, chronic heavy drinkers who reported consuming ≥ 80 g of alcohol/day (male) or 50 g/d (female) for ≥ 10 years with cirrhosis (cases) or without significant alcohol-associated liver disease (controls) were recruited. Controls (N = 764) were defined as having normal bilirubin, aspartate aminotransferase (AST), and alanine aminotransferase levels at the time of heavy alcohol use and no prior evidence of liver injury. Cases (N = 1128) were defined by at least one of the following: (1) clinically evident portal hypertension or decompensated cirrhosis (eg, ascites, esophageal varices, HE), (2) FibroScan stiffness >22 kPa if AST <100 IU/L/>32 kPa if AST 100–200 IU/L/exclude if AST >200 IU/L, or 3) liver histology data (Metavir score of F4) if available. Patients with HIV, viral hepatitis, hemochromatosis, or other liver diseases were excluded.[16] The study was approved by

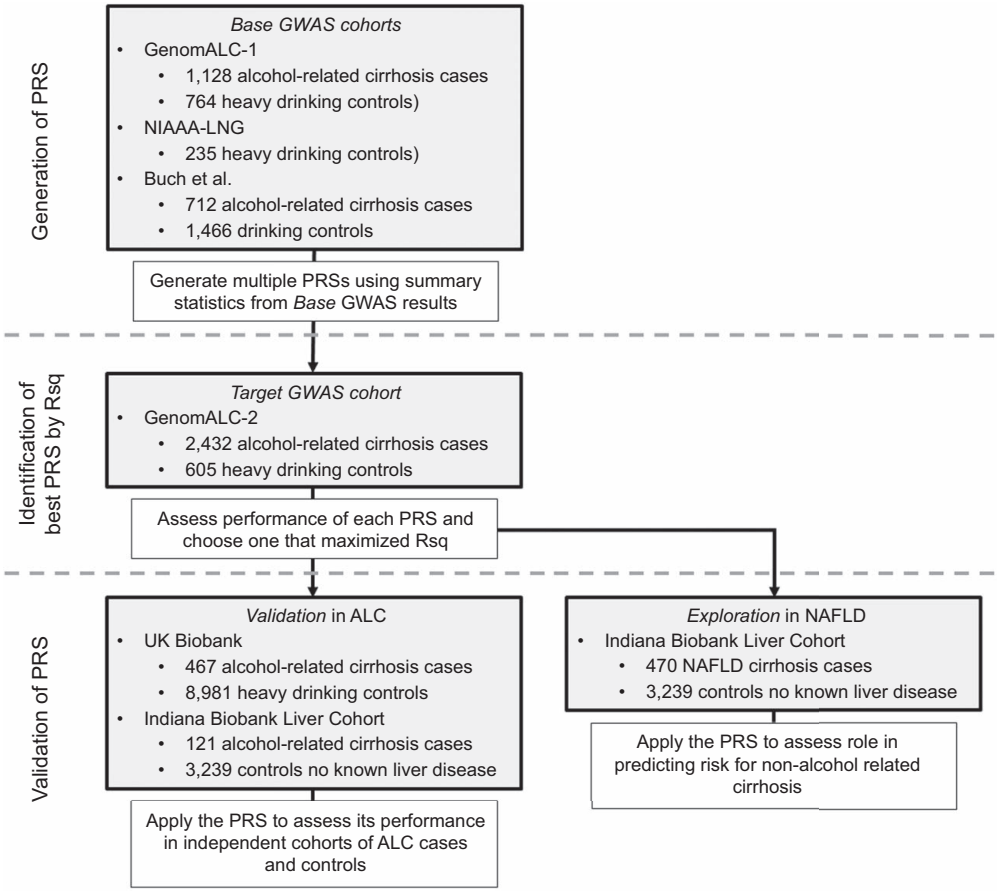

FIGURE 1 Flowchart of cohorts. Flowchart depicting the overall study design and number of participants included in each component. Abbreviations: ALC, alcohol-associated liver cirrhosis; GWAS, Genome-wide Association Study; NIAAA, National Institute on Alcohol Abuse and Alcoholism; PRS, Polygenic Risk Score; R-sq, R-squared.

appropriate Ethics Committees or Institutional Review Boards at each site and conformed to the ethical guidelines of the 1975 Declaration of Helsinki. Genotyping was performed using the Global Screening Array v1.0 (Illumina, 5200 Illumina Way, San Diego, CA 92122). Genotyped SNPs were filtered for call rate, violation of the Hardy-Weinberg equilibrium, and minor allele frequency. Samples were checked for genotyping rate, sex, and unexpected relatedness. Genetic ancestry was determined using SNPRelate package using the 1000 Genomes Project as reference. Additional genotypes were imputed using the Michigan Imputation Server.[17] Additional details on GWAS data cleaning are available.[15]

LNG cohort

LNG and Office of the Clinical Director at the National Institutes on Alcohol Abuse and Alcoholism (NIAAA, Rockville, MD) provided 235 heavy drinking controls (per GenomALC-1 definitions) under protocols (98-AA-009; 05-AA-0121). Participants had no known liver disease or serious medical issues requiring ongoing treatment. Genotype data (Illumina OmniExpress Bead Chip) were cleaned and imputed in the same manner as GenomALC-1 data as described above. This additional control cohort was used in a combined GWAS analysis with the GenomALC-1 data.

Buch and colleagues' cohort

The summary statistics from GWAS of 712 ALC cases and 1466 drinking controls were obtained from http://gengastro.med.tu-dresden.de/suppl/alc_cirrhosis/ to be meta-analyzed with GenomALC-1 plus NIAAA cohort's GWAS results. Detailed information on the Buch et al[10] cohort is available.

Target cohort participants

GenomALC-2 cohort

The GenomALC-2 DNA samples and project-specific data were donated by research groups who had independently collected them for prior genetic studies of alcohol-associated liver disease. Patients gave informed consent, and the studies were approved by the appropriate Ethics Review Boards. Patients enrolled in GenomALC-2 were from hospitals and clinics in the countries in GenomALC-1 and from Belgium. Cases and controls were defined similarly to GenomALC-1. Some GenomALC-2 samples that overlapped with the Buch et al[10] GWAS were removed from this analysis, leaving 2432 cases and 605 controls in the target cohort. Genotype data were cleaned and imputed using the same pipeline as for cleaning GenomALC-1 and LNG cohorts. Detailed information on this cohort is available.[9]

Validation cohorts participants

UKB cohort

Data from the UKB were accessed under approval number 18870. All participants gave informed consent, consistent with the UK Biobank Ethics and Governance Framework. Cirrhosis cases (n = 467) were defined as having International Classification of Diseases (ICD)10 code K70.3; "alcohol-associated cirrhosis of liver', or ICD-9 code 571.2, "Cirrhosis, liver, alcohol-associated or ICD10 code K70.1; "alcoholic hepatitis without ascites" or ICD-9 code 571.1, "acute alcoholic hepatitis." Controls (n = 8981) were defined as having (1) reported alcohol intake of $\geq$80 g/d (males) and $\geq$50 g/d (females) and/or (2) ICD10 diagnosis of F10.2 (mental and behavioral disorders due to alcohol) but with no recorded diagnosis of any liver disease as defined previously.[18] Genotype data were cleaned and imputed by the UKB investigators. Additional details are available in our previous report.[15]

The Indiana Biobank (IB) liver cohort (IB-Liver)

The IB-liver participants were recruited through the Indiana Biobank, which is a centralized statewide biobanking initiative at Indiana University, Purdue University, and the University of Notre Dame. Enrolled participants' medical information is obtained from the Indiana Network for Patient Care research database. Indiana Biobank cohort participants were recruited in-person from participating clinics and consented to research under IB protocol approved by Institutional Review Boards at respective institutes.

Chronic liver disease cases were identified using ICD9/10 liver disease–specific codes and subsequently confirmed via electronic medical chart review (N = 1033). Among chronic liver disease cases, cirrhosis was defined by liver histology (Metavir stage IV fibrosis), imaging (liver suggestive of cirrhosis or ascites), or a clinical diagnosis of ascites, esophageal varices, or HE (N = 591). To further specify the cause of cirrhosis, alcohol-associated versus metabolic dysfunction–associated steatotic liver disease (MASLD), clinical notes were reviewed for indications of alcohol use disorder history and ICD codes for alcohol-associated cirrhosis. After chart review, 121 participants were identified as having ALC and 470 patients with MASLD-cirrhosis.

Liver disease–free controls from the Indiana Biobank (IB-Liver cohort) were identified as those without a liver-specific ICD9/10 code or laboratory data, suggesting a risk for liver injury or disease (ie, abnormal AST, alanine aminotransferase tests, positive HBsAg, or hepatitis C antibody). In total, 3239 IB participants were identified as liver disease–free controls.

Genotyping was performed using the Global Screening Array v1.0 (Illumina, 5200 Illumina Way, San Diego, CA) by Regeneron. Genotype data cleaning and imputation were done using the pipeline published previously for GenomALC[15] and as described above.

Generating PRS for ALC

Base GWAS

Base GWAS results used to generate PRS were from meta-analysis of the combined GenomALC-1 and LNG cohorts GWAS, which comprises 1128 cases and 999 controls (adjusting for age, sex, and the first 10 principal components) and summary statistics reported by Buch et al.[10] (including 712 case and 1466 controls). We removed ambiguous SNPs (having A/T or C/G allele combination, where allele flips cannot be resolved) to avoid meta-analyzing opposite alleles together, and InDels. In total, meta-analysis included 4,882,650 SNPs and was conducted using METAL[19] with a genomic control option.

Generation of PRS

First, PRSice-2[7] was used to define a set of independent SNPs that represent all associated loci from the base GWAS. A total of 19,271 independent SNPs were identified based on linkage disequilibrium patterns in the target data (GenomALC-2) using an R-squared (R-sq) value a threshold of 0.2. Using the corresponding betas and *p*-values of the independent SNPs from the base GWAS, PRSs were iteratively generated across *p*-value thresholds between 5×10^{-8} to 5×10^{-2} by steps of 5×10^{-8} for SNP inclusion criteria for PRS. At each *p*-value threshold, a PRS score was calculated as the sum of (genotype value [effect allele count] multiplied by the beta) across all independent SNPs with *p*-values meeting the threshold for each individual in the target dataset (GenomALC-2). The score was then tested against the ALC status in the target data to calculate R-sq (the proportion of variance explained) as a measure of its performance. After calculating through all *p*-value thresholds and obtaining R-sq values at each threshold, the PRS that maximized R-sq was chosen to be the best-performing PRS of ALC (PRS_{ALC}).

Validation of PRS_{ALC} in 2 independent cohorts

As a validation of PRS_{ALC}, we tested its ability to stratify risk for ALC in 2 independent cohorts, UKB (N = 9,448) and IB-Liver (N = 3,360). In each cohort, we calculated PRS_{ALC}, conducted a logistic regression analysis between ALC versus control status and PRS_{ALC}, and calculated performance metrics (detailed below) to assess PRS_{ALC}'s ability to stratify risk for ALC in each cohort.

Exploratory analysis of PRS_{ALC} in MASLD-cirrhosis

To assess the genetic overlap between ALC and MASLD-cirrhosis using PRS_{ALC}, we tested for the association between PRS_{ALC}-related and MASLD-related liver cirrhosis. A total of 3709 participants from IB-Liver were included in the exploratory analyses. There were 470 patients with MASLD-cirrhosis diagnosis and 3239 controls who did not have a known history of liver diseases.

Metrics for PRS_{ALC} performance

The performance of the PRS_{ALC} was assessed by calculating the ORs, AUC, positive predictive value (PPV), negative predictive value (NPV), sensitivity (true positive rate), and specificity (true negative rate). ORs were calculated by comparing individuals in each PRS_{ALC} score decile with those in the lowest decile. Available covariates were included in the calculation of ORs and adjusted for in each cohort: GenomALC-2 (age, sex); IB-Liver (age, sex, BMI), and the UKB (age, sex, BMI, diabetes). The 80th percentile score of PRS_{ALC} from GenomALC-2 cohort (−2.75) was used as a threshold for dichotomizing samples in all 3 cohorts into high versus low-risk groups for calculation of AUC, PPV, NPV, sensitivity, and specificity. Finally, following the steps outlined by Hao et al[20], a secondary cutoff for PRS_{ALC} was calculated to stratify samples into high and low risk for ALC. To calculate the cutoff, the beta for PRS_{ALC} was obtained using standardized PRS_{ALC} score after adjusting for age and sex in GenomALC-2 cohort. Then, the ratio of log[2] and our beta was calculated as the threshold (1.778) for identifying those at higher risk of developing ALC. All calculations were done in R 4.1.1.[21]

TABLE 1 List of variants included in PRS_{ALC}

Chromosome	BP	rsID	Effect allele	Other allele	OR	SE	*p*	Gene
3	65,137,329	rs1482588	A	G	0.57	0.12	4.99E-06	—
4	7,166,787	rs199571141	A	G	1.35	0.07	1.01E-05	*SORCS2*
4	88,169,536	rs67373430	A	G	1.51	0.06	3.38E-11	*HSD17B13-KLHL8*
5	116,751,170	rs13184765	A	G	0.62	0.11	6.67E-06	*LINC00992*
6	136,244,663	rs9494417	A	G	1.44	0.08	3.77E-06	*PDE7B*
10	9,928,512	rs4749900	T	C	0.65	0.09	3.72E-06	*LINC02663*
14	66,591,630	rs115335372	T	C	0.43	0.18	2.03E-06	*LINC02290*
14	94,844,947	rs28929474	T	C	2.40	0.20	1.20E-05	*SERPINA1*
16	12,652,801	rs12162092	A	G	0.72	0.07	3.50E-06	*SNX29*
18	23,918,945	rs1676988	A	C	0.73	0.06	1.46E-06	*TAF4B*
19	19,393,714	rs8100204	A	G	1.63	0.08	1.06E-10	*SUGP1*
19	19,643,028	rs10422819	A	C	1.28	0.05	4.34E-06	*YJEFN3*
19	54,676,763	rs641738	T	C	1.31	0.05	1.20E-07	*MBOAT7*
22	44,339,055	rs117772800	A	G	0.53	0.10	6.58E-11	*PNPLA3*
22	44,340,904	rs2294915	T	C	2.25	0.06	2.06E-47	*PNPLA3*
22	44,344,885	rs16991199	T	G	0.46	0.15	3.85E-07	*PNPLA3*
22	44,348,116	rs11912828	A	G	0.66	0.06	3.85E-11	*SAMM50*
22	44,390,366	rs2235779	A	G	0.74	0.06	3.30E-07	*SAMM50*
22	44,391,588	rs2179642	T	C	0.74	0.05	4.99E-09	*SAMM50*
22	44,400,149	rs13055235	A	G	2.23	0.13	2.92E-10	*PARVB*

Note: BP, base pair location, rsID: dbSNP ID, other allele: noneffect (reference) allele for the tested variant, OR: effect for each copy of the effect allele, Gene: Name of associated or nearby gene for each variant. *p*-value, OR, and SE come from the base GWAS.

RESULTS

The overall design of the study is depicted in Figure 1. Across the tested PRS models, the one with 20 SNPs (*p*-value threshold was 1.205×10^{-5}) maximized the R-sq for ALC versus heavy drinking controls in the GenomALC-2 (target) cohort (Supplemental Figure S1, http://links.lww.com/HC9/A864). The list of included SNPs, their betas, *p*-values, associated genes are shown in Table 1. In GenomALC-2, the top decile (highest PRS_{ALC} group) had an OR of 2.9 compared to the bottom decile (lowest PRS_{ALC} group) after adjusting for age and sex ($p = 1.91\times10^{-6}$) (Figure 2). Furthermore, the fourth and higher deciles (40%–100%) compared to the bottom decile all showed an increased risk for ALC with *p*-values $<9.33\times10^{-4}$ as shown in Table 2. As comparison in the validation cohorts, seventh decile and above showed increased risk for ALC in the UKB (ORs ranging between 1.7 and 2.8) and IB-Liver cohort for ALC versus healthy controls (ORs ranging between 2.8 and 4.4) (Figure 2).

The PRS_{ALC} explained $>3.5\%$ of the variance in ALC risk for the GenomALC-2 cohort ($p = 1.12\times10-^{15}$). In the UKB validation cohort, the PRS_{ALC} was significantly associated with ALC status ($p = 1.03\times10^{-15}$) with an R-sq of 2.0%. In IB-Liver, PRS_{ALC} was also significantly associated with ALC status (R-sq of 2.9%, *p*-value 1.89×10^{-7}) (Supplemental Table S1, http://links.lww.com/HC9/A864).

The PRS_{ALC} was applied to MASLD-cirrhosis in the IB-Liver cohort to explore whether it was associated with MASLD-related liver cirrhosis. We observed significant associations with the risk for MASLD comparing the top and bottom deciles of PRS_{ALC} (OR of 3.94, *p*-value 2.35×10^{-6}) (Table 3). AUC for ROC models with and without PRS_{ALC} was calculated to assess the improvement over typical risk predictors (age, sex, BMI, diabetes) for alcohol-associated cirrhosis (Table 4). In GenomALC-2, adding PRS_{ALC} to age and sex improved AUC from 0.64 to 0.68 or 5.5% relative increase. In the UKB, adding PRS_{ALC} to age, sex, and BMI improved AUC by 8.3% (from 0.57 to 0.62), and to age, sex, and diabetes improved AUC by 7.6% (from 0.62 to 0.67). In IB-Liver, adding PRS_{ALC} to age, sex, and BMI increased AUC by 4.4% (from 0.70 to 0.73). Furthermore, in the IB-Liver cohort, prediction of MASLD-cirrhosis improved by 7.8% (from 0.60 to 0.65) by adding PRS_{ALC} respectively (calibration plots in Supplemental Figure S2, http://links.lww.com/HC9/A864). The PRS_{ALC} achieved PPV $>80\%$ in GenomALC-2 cohort and NPV $>95\%$ in both the UKB and IB-Liver cohorts (Table 5). Furthermore, specificity (true negative rate) was approximately 87% in all 3 cohorts.

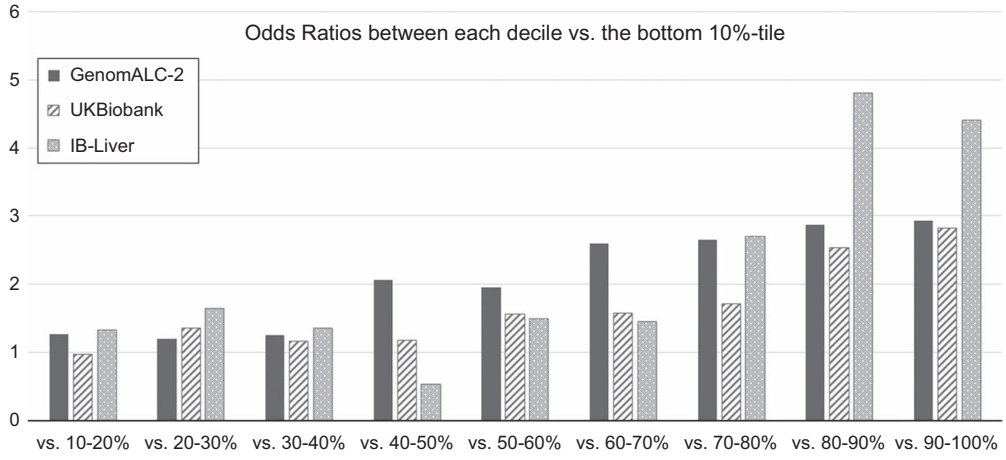

FIGURE 2 Plot of adjusted ORs across 3 cohorts Bar plot of covariates adjusted ORs between the bottom decile (0%–10% group) and each decile of PRSALC in the target and 2 validation cohorts. Abbreviations: ALC, alcohol-associated liver cirrhosis; IB-Liver: Indiana Biobank Liver.

TABLE 2 Adjusted ORs for PRS_{ALC} deciles in GenomALC-2, UK Biobank, and IB-liver cohorts

Cohort	Adjusted covariates	Deciles (vs. lowest 0%–10% decile)	OR (95% CI)	*p*
GenomALC-2	Age, sex, 10 PCs	10–20	1.26 (0.86, 1.84)	2.30E-01
		20–30	1.20 (0.83, 1.75)	3.30E-01
		30–40	1.25 (0.85, 1.82)	2.60E-01
		40–50	2.07 (1.37, 3.11)	5.35E-04
		50–60	1.95 (1.31, 2.90)	9.33E-04
		60–70	2.60 (1.70, 3.97)	1.01E-05
		70–80	2.66 (1.74, 4.05)	5.46E-06
		80–90	2.87 (1.85, 4.47)	2.64E-06
		90–100	2.93 (1.88, 4.56)	1.91E-06
UK Biobank	Age, sex, 10 PCs	10–20	0.91 (0.54, 1.53)	7.20E-01
		20–30	1.26 (0.78, 2.04)	3.40E-01
		30–40	1.11 (0.68, 1.83)	6.74E-01
		40–50	1.14 (0.70, 1.87)	6.02E-01
		50–60	1.51 (0.94, 2.40)	8.61E-02
		60–70	1.59 (1.00, 2.53)	5.05E-02
		70–80	1.66 (1.04, 2.64)	3.32E-02
		80–90	2.49 (1.61, 3.84)	3.74E-05
		90–100	2.75 (1.79, 4.21)	3.49E-06
IB-liver	Age, sex, 10 PCs	10–20	1.64 (0.47, 5.71)	4.40E-01
		20–30	1.77 (0.61, 5.15)	2.95E-01
		30–40	1.35 (0.41, 4.41)	6.20E-01
		40–50	0.53 (0.12, 2.34)	4.02E-01
		50–60	1.5 (0.49, 4.64)	4.79E-01
		60–70	1.45 (0.47, 4.44)	5.16E-01
		70–80	2.7 (0.95, 7.66)	6.20E-02
		80–90	4.81 (1.86, 12.44)	1.19E-03
		90–100	4.4 (1.56, 12.44)	5.18E-03

Note: Adjusted covariates: List of covariates that were adjusted for to obtain the ORs, Deciles: tested decile for the obtained ORs, the reference group is always bottom decile or 0%–10% of PRSALC distribution in each cohort. Estimate: log-OR estimate for the compared deciles.

TABLE 3 Adjusted ORs for PRS_{ALC} deciles in IB-liver cohort for MASLD-cirrhosis and liver disease–free participants

Phenotype	Adjusted covariates	Deciles (vs. lowest 0%–10% decile)	OR (95% CI)	*p*
MASLD-related cirrhosis vs. Healthy controls	Age, sex, BMI, 10 PCs	10–20	1.41 (0.74, 2.70)	2.91E-01
	—	20–30	1.26 (0.67, 2.37)	4.73E-01
	—	30–40	1.42 (0.75, 2.70)	2.83E-01
	—	40–50	1.65 (0.89, 3.05)	1.12E-01
	—	50–60	2.02 (1.12, 3.65)	1.95E-02
	—	60–70	2.00 (1.10, 3.61)	2.23E-02
	—	70–80	2.56 (1.44, 4.55)	1.40E-03
	—	80–90	3.29 (1.85, 5.82)	4.65E-05
	—	90–100	3.94 (2.23, 6.95)	2.35E-06

Note: Adjusted covariates: list of covariates that were adjusted for to obtain the ORs, Deciles: tested decile for the obtained ORs, the reference group is always bottom decile or 0%–10% of PRSALC distribution in each cohort. Estimate: log-OR estimate for the compared deciles.
CI, 95% confidence interval for the OR; OR, OR, 95%.

Finally, identifying the proportion of subjects in UKB and IB-Liver cohort that were determined to be high risk (OR > 2) by methods outlined by Hao et al,[22] 5.0% (498 out of 9,884) in the UKB and 5.4% of the participants in IB-Liver cohort (232 out of 4274) were found to be at high risk for ALC (Table 6, Supplemental Figure S3, http://links.lww.com/HC9/A864).

DISCUSSION

In the present study, we developed a 20-SNP–based PRS for stratifying the risk for ALC utilizing a data-driven approach in meta-analyzed GWAS applied to an independent cohort (GenomALC-2). In GenomALC-2, participants in the top decile of PRS_{ALC} are at approximately 3-fold higher odds of developing cirrhosis

TABLE 4 Area Under Curve (AUC) comparisons between available clinical risk factors and PRS_{ALC}

Study	Outcome	Covariates	N_Case	N_Ctrl	AUC (CI)	Delta AUC
GenomALC-2	ALC	Age + sex + 10 PCs	2244	602	0.6438 (0.6182–0.6695)	
	ALC	Age + sex + 10 PCs + PRS_{ALC}	2244	602	0.6791 (0.6551–0.7032)	5.5%
	ALC	Age + sex + 10 PCs + 3-SNP score	2244	602	0.6710 (0.6464–0.6956)	4.2%
UK Biobank	ALC	Age + sex + BMI + 10 PCs	462	8931	0.5717 (0.5445–0.5988)	
	ALC	Age + sex + BMI + 10 PCs + PRS_{ALC}	462	8931	0.6194 (0.5923–0.6465)	8.3%
	ALC	Age + sex + BMI + 10 PCs + 3-SNP score	462	8931	0.6356 (0.6085–0.6626)	11.2%
UK Biobank	ALC	Age + sex + diabetes + 10 PCs	467	8981	0.6205 (0.5931–0.6480)	
	ALC	Age + sex + diabetes + 10 PCs + PRS_{ALC}	467	8981	0.6645 (0.6376–0.6914)	7.1%
	ALC	Age + sex + diabetes + 10 PCs + 3-SNP score	467	8981	0.6797 (0.5637–0.7057)	9.5%
IB-Liver	ALC	Age + sex + BMI + 10 PCs	101	2648	0.7029 (0.6556–0.7502)	—
	ALC	Age + sex + BMI + 10 PCs + PRS_{ALC}	101	2648	0.7339 (0.6847–0.7830)	4.4%
	ALC	Age + sex + BMI + 10 PCs + 3-SNP score	101	2648	0.7295 (0.6787–0.7802)	3.8%
IB-Liver	MASLD-Cirrhosis	age + sex + BMI + 10 PCs	343	2648	0.6028 (0.5729–0.6327)	
	MASLD-Cirrhosis	age + sex + BMI + 10 PCs + PRS_{ALC}	343	2,648	0.6496 (0.6188–0.6803)	7.8%

Note: Outcome: Tested outcomes for each AUC calculation. ALC: alcohol-related liver cirrhosis outcome with cirrhosis cases and noncirrhosis controls, MASLD-cirrhosis: metabolic dysfunction–associated steatotic liver disease cirrhosis outcome with cirrhosis cases and liver disease history negative controls.
Abbreviations: Age, age at enrollment, BMI, body mass index; Delta AUC, differences in AUC between models with and without genetic risk scores (PRSALC and 3-SNP score); N_Case, Number of cases; N_Ctrl, Number of controls; PCs, principal components of genetic distance among participants in; sex, biological sex.

TABLE 5 Table of PRS_{ALC} performance statistics

GenomALC-2						
PRS_{ALC} > −2.75	Below	Above	PPV	NPV	TPR	TNR
Ctrl	527	78	87.1%	21.7%	21.7%	87.1%
Case	1905	527	—	—	—	—
3-SNP score > 0.7	Below	Above	PPV	NPV	TPR	TNR
Ctrl	440	165	85.5%	23.2%	40.1%	72.7%
Case	1458	974	—	—	—	—
UK Biobank						
PRS_{ALC} > -2.75	Below	Above	PPV	NPV	TPR	TNR
Ctrl	7833	1148	8.5	95.6	22.9	87.2
Case	360	107	—	—	—	—
3-SNP score > 0.7	Below	Above	PPV	NPV	TPR	TNR
Ctrl	6935	2046	8.4%	96.1%	40.3%	77.2%
Case	279	188	—	—	—	—
IB-liver cohort						
PRS_{ALC} > -2.75	Below	Above	PPV	NPV	TPR	TNR
Ctrl	2845	394	7.5%	96.9%	26.0%	87.8%
Case	91	32	—	—	—	—
3-SNP score > 0.7	Below	Above	PPV	NPV	TPR	TNR
Ctrl	2493	746	7.4%	97.5%	48.8%	77.0%
Case	63	60	—	—	—	—

Note: Below: Number of participants whose genetic risk score is below the threshold value (−2.75), above: number of participants whose genetic risk score is above the threshold value (−2.75).
Abbreviations: NPV, negative predictive value; PPV, positive predictive value; TNR, true negative rate (specificity); TPR, true positive rate (sensitivity).

compared to those in the bottom decile. We then validated the PRS_{ALC} in 2 additional independent cohorts of ALC. Compared to GenomALC-2, we observe a similar or higher-fold increase in risk for ALC (UKB OR = 2.8, IB-Liver OR = 4.4). Furthermore, our PRS_{ALC} also adds to known risk factors such as age, sex, BMI, and diabetes status in risk stratification for alcohol-associated liver cirrhosis. Adding PRS_{ALC} to risk models increases AUC for ALC prediction by 4% to 8% across all 3 cohorts.

Our data-driven PRS_{ALC} included a total of 20 SNPs in regions containing 12 genes, 3 long noncoding RNAs, and 1 intergenic region. This included variants from the genes that were previously reported to be associated with alcohol-associated cirrhosis (*PNPLA3*, *HSD17B13*, SURP and G-patch domain containing 1 [*SUGP1*], membrane-bound O-Acyltransferase domain containing 7 [*MBOAT7*], serpin family A member 1 [*SERPINA1*], and phosphodiesterase 7B [*PDE7B*][15]). *PNPLA3-SAMM50-PARVB* region was represented by 7 SNPs that were in very weak linkage disequilibrium with each other. In lieu of rs738409, which is a G/C SNP that was excluded from our analysis to avoid ambiguous alleles, we included rs2294915 which is in complete linkage disequilibrium with rs738409. Other SNPs had been associated with liver injury/diseases, although not specifically with alcohol-associated cirrhosis. *KLHL8* (Kelch-like family member 8 gene) on chromosome 4, which was nearby rs67373430 in our PRS_{ALC}, is reported to be associated with total cholesterol level,[23] alanine aminotransferase levels,[24] AST levels,[25] and Fib-4 score among heavy drinkers,[26] and phosphodiesterase 7B gene (*PDE7B*) is associated with AST.[27]

TABLE 6 Table of proportions of participant with standardized PRS_{ALC} above OR > 2 threshold

	Below	Above	% at high risk
GenomALC-2	2888	149	4.9
IB-liver cohort	4042	232	5.4
UK Biobank	9386	498	5.0

Note: The number of participants with standardized PRS_{ALC} score below 1.778 (threshold for OR > 2) in each cohort. Above: The number of participants with standardized PRS_{ALC} score above 1.778 (threshold for OR > 2). % At high risk: the proportion of the participants with standardized PRS_{ALC} score above the threshold in each cohort.

Compared to our previous 3-SNP–based score, our new PRS_{ALC} slightly improved the AUC in GenomALC-2 and IB-Liver cohort. This suggests that the genetic underpinning for ALC is not completely driven by the three well-known loci, but those additional genes and variants may further contribute to the development of liver cirrhosis in heavy alcohol drinking environment. Moreover, we show that our PRS_{ALC} also stratifies risk for MASLD-cirrhosis, showing a similar increase in risk (OR = 3.9) for developing MASLD-cirrhosis in the IB-

Liver cohort when compared with people without known liver disease. The ability of PRS_{ALC} to stratify risk for MASLD-cirrhosis yet again suggests shared genetic etiology between the 2 diseases, which is supported by significant allelic associations seen in or near *PNPLA3*, *TM6SF2*, and *HSD17B13*[28] for both ALC and MASLD-cirrhosis, and by the development of PRS for other liver diseases based on similar genetic loci.[29–31]

Across GenomALC-2, UKB, and IB-Liver cohorts, we observed similar sensitivity levels (ie, identification of drinkers at high risk of ALC) of 22%, 23%, and 26%, respectively, and specificity levels of approximately 87% for all 3 cohorts (ie, identification of patients with low risk who did not develop cirrhosis). Low sensitivity values suggest that even with a high genetic risk for developing ALC, other factors have effects on the overall risk of developing ALC. On the other hand, PPV and NPV showed various values depending on the prevalence of ALC in the cohorts. GenomALC-2, a cohort of preselected patients with ALC and a smaller number of heavy drinkers free of known alcohol-associated liver injury, had an ALC prevalence $>80\%$ (87% PPV and 22% NPV), while the population-based UKB and IB-Liver cohort had a prevalence of approximately 5% and 4% (PPV ~ 8% and NPV $>95\%$ in both cohorts), respectively. This difference in PPV and NPV values based on prevalence highlights the importance of clinical settings in interpreting PRS such as our PRS_{ALC}, and how it can be implemented to identify those at increased or decreased risk based on prevalence.

Our expanded 20-SNP PRS (PRS_{ALC}), when applied to GenomALC-2, UKB, and IB-Liver cohort, predicted 4.91%, 5.04%, and 5.43% of the participants to be at higher risk of ALC ($OR>2$), respectively. Across different drinking settings, ranging from heavy (GenomALC-2) to general population levels, our 20-SNP genetic risk score consistently predicted that approximately 5% of the population would be at double the average risk for developing liver cirrhosis if their alcohol exposure was sufficient. This proportion is similar to the estimated prevalence of alcohol-associated liver disease[32,33] among heavy drinkers.

The PRS_{ALC} has potential clinical utility. Although PRS may provide minimal improvement in risk prediction when added to clinical scores such as aminotransferase-to-platelet ratio index and FIB-4 (fibrosis-4 index) as demonstrated by Innes et al,[32] the key utility of PRS is that it can be calculated before the onset of a disease. This allows early identification of subjects at higher risk for developing a disease of interest and the potential to deploy strategies that would prevent disease development. Thus, PRS could be used early in the clinical course while clinical risk score could be used after sufficient exposure to the risk (eg, alcohol use for alcohol-associated cirrhosis).

There are limitations to this study. First, even though our PRS was based on one of the largest cohorts of patients with ALC and validated using independent cohorts (more than 20,000 total participants), this still was a modest size for current-day genetic studies. This may be reflected in our significant but still modest amount of variance explained in each cohort by PRS_{ALC}. Conducting larger GWAS would improve the accuracy and potential clinical utility of our score by discovering additional variants that underpin risk of ALC. Such efforts could incorporate resources such as Million Veterans Program[33] and All of Us (National Institutes of Health)[34] to better estimate the true effects of genetic variants in the at-risk population, making the overall PRS more reflective of true genetic risk for ALC. Second, our PRS_{ALC} is limited to European genetic ancestry participants due to the availability of ALC GWAS data. While we understand the importance of studying non-European ancestry samples, our current access precludes this. Future studies in large non-European cohorts are needed to generate cross-population PRSs that incorporate GWAS results from multiple populations.[35] Lastly, our new 20-SNP–based PRS only modestly improved the R-sq compared to our 3-SNP genetic risk score. Especially in UKB cohort, the 3-SNP score explained more variance than our new PRS_{ALC}. This incremental improvement of our PRS may reflect the need for much larger sample sizes to identify additional risk genes for ALC but also may reflect the prominent roles *PNPLA3*, *HSD17B13*, and *TM6SF2* play in the genetic underpinning of ALC.

In conclusion, a PRS was developed for alcohol-associated cirrhosis risk, based on a data-driven and hypothesis-free approach. The resulting 20-SNP–based PRS_{ALC} stratifies risk for developing ALC in a drinking population. Also, our new PRS_{ALC} stratifies risk for MASLD-related cirrhosis, emphasizing the common genetic underpinnings of the 2 traditionally distinct (alcohol-associated and fatty liver related) diseases.

AUTHOR CONTRIBUTIONS

Tae-Hwi Schwantes-An, John B. Whitfield, Devanshi Seth, and Timothy R. Morgan conceived and designed the study. Recruitment and data acquisition was done for GenomALC-1 by Guruprasad P. Aithal, Florian Eyer, Jean-Marc Jacquet, Suthat Liangpunsakul, Steven Masson, Philippe Mathurin, Romain Moirand, Timothy R. Morgan, Sebastian Mueller, Beat Müllhaupt, Pierre Nahon, Bertrand Nalpas, Sylvie Naveau, Munir Pirmohamed, Helmut K. Seitz, Devanshi Seth, Felix Stickel; for GenomALC-2 by Guruprasad P. Aithal, Stephen R. Atkinson, Ramon Bataller, Naga P. Chalasani, Ann K. Daly, Florian Eyer, David Goldman, Steven Masson, Philippe Mathurin, Christophe Moreno, Andrew McQuillin, Marsha Y. Morgan, Timothy R. Morgan, Laura E. Nagy, Devanshi Seth, Felix Stickel, Mark R. Thursz, Eric Trépo; and for IB-Liver cohort by Naga P.

Chalasani, Craig S. Lammert, Samer Gawrieh. All analyses were performed by Tae-Hwi Schwantes-An, John B. Whitfield. Tiebing Liang facilitated DNA processing for genotyping. Tae-Hwi Schwantes-An, Devanshi Seth, John B. Whitfield, and Timothy R. Morgan led the writing of the manuscript. All authors read, critically reviewed, and approved the final version. Devanshi Seth and Timothy R. Morgan are the guarantors.

ACKNOWLEDGMENTS

The authors thank Ms Donna Sheedy, Ms Julia Stevens, and Prof Jillian Krill for providing access to brain tissue (included in GenomALC-2) from the New South Wales Brain Tissue Resource Centre at the University of Sydney, NSW 2006, Australia (collection of tissues reported in this publication was supported by the National Institute of Alcohol Abuse and Alcoholism of the National Institutes of Health under Award Number R28AA012725). The content is solely the responsibility of the authors and does not represent the official views of the National Institutes of Health). The authors thank 4 Alcoholic Hepatitis Consortia TREAT, InTEAM, DASH, and SCAHC (U01 AA021886) from the United States for providing DNA/biospecimens contributing toward GenomALC-2.

FUNDING INFORMATION

Funding for this study was provided by NIH/NIAAA UO1AA018389 and RO1AA018389 for data collection, analysis, interpretation, and patient recruitment. Suthat Liangpunsakul: U01 grant from the NIAAA and R01 AA025208, U01 AA026917, and 1I01CX000361. Mark R. Thursz: The UK Medical Research Council Stratified Medicine Award (Ref MR/R014019/1), NIHR Imperial Biomedical Research Centre and NIHR Senior Investigator Award (NIHR 200153).

CONFLICTS OF INTEREST

Naga P. Chalasani has a number of consulting agreements with and research grants from the pharmaceutical industry, but they are not significantly or directly related to this paper. Munir Pirmohamed receives research funding from various organizations, including the MRC and NIHR. He has also received partnership funding for the following: MRC Clinical Pharmacology Training Scheme (co-funded by MRC and Roche, UCB, Eli Lilly and Novartis); a PhD studentship jointly funded by EPSRC and Astra Zeneca; and grant funding from Vistagen Therapeutics. He has also unrestricted educational grant support for the UK Pharmacogenetics and Stratified Medicine Network from Bristol-Myers Squibb and UCB. He has developed an HLA genotyping panel with MC Diagnostics but does not benefit financially from this. He is part of the IMI Consortium ARDAT (www.ardat.org). None of these funding sources were deployed in the undertaking of this study. Timothy R. Morgan has conducted clinical research with AbbVie, Genfit, Gilead, and Merck but none of these are related to this manuscript. Tae-Hwi Schwantes-An is a paid consultant for Target RWE. Guruprasad P. Aithal consults and advises Clinipace, Servier, Pfizer, GSK, AstraZeneca, DNDi, BenevolentAI, Abbott Product Operations AG, Albireo, PureTech and Merck. RB is on the speakers' bureau for AbbVie. Samer Gawrieh consults for TransMedics and Pfizer and received grants from Zydus, Viking, and SonicIncytes. Paul S. Haber received grants from Woke Pharmaceuticals. S Masson consults and is on the speakers' bureau for Norgine and is on the speakers' bureau for Dr. Falk and Sandoz. Christophe Moreno consults and received grants from AbbVie and Gilead. He consults for Surrozen, Intercept, Echosens, Novartis, Bayer, and Astellas. Beat Müllhaupt consults and received grants from Gilead and consults for Ipsen. Pierre Nahon consults and received grants from AstraZeneca, Bristol-Myers Squibb, and Eisai. He consults for Ipsen and Roche. Eric Trépo received grants from Gilead and Julius Clinical.

REFERENCES

1. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nat Genet. 2018;50:1219–24.
2. Lewis CM, Vassos E. Polygenic risk scores: From research tools to clinical instruments. Genome Medicine. 2020;12:44.
3. Torkamani A, Wineinger NE, Topol EJ. The personal and clinical utility of polygenic risk scores. Nature Reviews Genetics. 2018; 19:581–90.
4. Klarin D, Natarajan P. Clinical utility of polygenic risk scores for coronary artery disease. Nat Rev Cardiol. 2022;19:291–301.
5. Lambert SA, Abraham G, Inouye M. Towards clinical utility of polygenic risk scores. Hum Mol Genet. 2019;28:R133–42.
6. Choi SW, Mak TS-H, O'Reilly PF. Tutorial: A guide to performing polygenic risk score analyses. Nature Protocols. 2020;15: 2759–72.
7. Choi SW, O'Reilly PF. PRSice-2: Polygenic Risk Score software for biobank-scale data. GigaScience. 2019;8:giz082.
8. Backman JD, Li AH, Marcketta A, Sun D, Mbatchou J, Kessler MD, et al. Exome sequencing and analysis of 454,787 UK Biobank participants. Nature. 2021;599:628–34.
9. Whitfield JB, Schwantes-An TH, Darlay R, Aithal GP, Atkinson SR, Bataller R, et al. A genetic risk score and diabetes predict development of alcohol-related cirrhosis in drinkers. J Hepatol. 2022;76:275–82.
10. Buch S, Stickel F, Trépo E, Way M, Herrmann A, Nischalke HD, et al. A genome-wide association study confirms PNPLA3 and identifies TM6SF2 and MBOAT7 as risk loci for alcohol-related cirrhosis. Nature Genetics. 2015;47:1443.
11. Abul-Husn NS, Cheng X, Li AH, Xin Y, Schurmann C, Stevis P, et al. A protein-truncating HSD17B13 variant and protection from chronic liver disease. N Engl J Med. 2018;378: 1096–106.
12. Liu Y-L, Reeves HL, Burt AD, Tiniakos D, McPherson S, Leathart JBS, et al. TM6SF2 rs58542926 influences hepatic fibrosis progression in patients with non-alcoholic fatty liver disease. Nat Commun. 2014;5:4309.
13. Johansen S, Thiele M, Juel HB, Hansen T, Krag A. External validation of a genetic risk score that predicts development of alcohol-related cirrhosis. J Hepatol. 2022.

14. Mars N, Koskela JT, Ripatti P, Kiiskinen TTJ, Havulinna AS, Lindbohm JV, et al. Polygenic and clinical risk scores and their impact on age at onset and prediction of cardiometabolic diseases and common cancers. Nat Med. 2020;26:549–57.
15. Schwantes-An TH, Darlay R, Mathurin P, Masson S, Liangpunsakul S, Mueller S, et al. GenomALC Consortium. Genome-wide association study and meta-analysis on alcohol-related liver cirrhosis identifies novel genetic risk factors. Hepatology. 2021; 73:1920–31.
16. Whitfield JB, Rahman K, Haber PS, Day CP, Masson S, Daly AK, et al. Brief report: Genetics of alcoholic cirrhosis—GenomALC Multinational Study. Alcohol Clin Exp Res. 2015; 39:836–42.
17. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. Nature Genetics. 2016;48:1284.
18. Whitfield JB, Seth D, Morgan TR, Consortium G. All-cause and liver-related mortality risk factors in excessive drinkers: Analysis of data from the UK biobank. Alcohol Clin Exp Res. 2022;46:2245–57.
19. Willer CJ, Li Y, Abecasis GR. METAL: Fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010;26:2190–1.
20. Hao L, Kraft P, Berriz GF, Hynes ED, Koch C, Kortegere V Kumar P, et al. Development of a clinical polygenic risk score assay and reporting workflow. Nature Medicine. 2022;28:1006–3.
21. Team RC. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; 2019.
22. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016;536:285.
23. Klarin D, Damrauer SM, Cho K, Sun YV, Teslovich TM, Honerlaw J, et al. Genetics of blood lipids among ~300,000 multi-ethnic participants of the Million Veteran Program. Nat Genet. 2018;50:1514–23.
24. Nielsen JB, Rom O, Surakka I, Graham SE, Zhou W, Roychowdhury T, et al. Loss-of-function genomic variants highlight potential therapeutic targets for cardiovascular disease. Nat Commun. 2020;11:6417.
25. Zhou W, Brumpton B, Kabil O, Gudmundsson J, Thorleifsson G, Weinstock J, et al. GWAS of thyroid stimulating hormone highlights pleiotropic effects and inverse association with thyroid cancer. Nat Commun. 2020;11:3981.
26. Innes H, Buch S, Hutchinson S, Guha IN, Morling JR, Barnes E, et al. Genome-wide association study for alcohol-related cirrhosis identifies risk loci in MARC1 and HNRNPUL1. Gastroenterology. 2020;159:1276–89 e1277.
27. Young KA, Palmer ND, Fingerlin TE, Langefeld CD, Norris JM, Wang N, et al. Genome-wide association study identifies loci for liver enzyme concentrations in Mexican Americans: the GUARDIAN consortium. Obesity (Silver Spring). 2019;27: 1331–7.
28. Scott E, Anstee QM. Genetics of alcoholic liver disease and non-alcoholic steatohepatitis. Clin Med (Lond). 2018;18:s54–9.
29. Bianco C, Jamialahmadi O, Pelusi S, Baselli G, Dongiovanni P, Zanoni I, et al. Non-invasive stratification of hepatocellular carcinoma risk in non-alcoholic fatty liver using polygenic risk scores. J Hepatol. 2021;74:775–82.
30. Bianco C, Tavaglione F, Romeo S, Valenti L. Genetic risk scores and personalization of care in fatty liver disease. Curr Opin Pharmacol. 2021;61:6–11.
31. De Vincentis A, Tavaglione F, Jamialahmadi O, Picardi A, Antonelli Incalzi R, Valenti L, et al. A polygenic risk score to refine risk stratification and prediction for severe liver disease by Clinical Fibrosis Scores. Clin Gastroenterol Hepatol. 2022;20: 658–73.
32. Innes H, Morling JR, Buch S, Hamill V, Stickel F, Guha IN. Performance of routine risk scores for predicting cirrhosis-related morbidity in the community. J Hepatol. 2022;77:365–76.
33. Gaziano JM, Concato J, Brophy M, Fiore L, Pyarajan S, Breeling J, et al. Million Veteran Program: A mega-biobank to study genetic influences on health and disease. J Clin Epidemiol. 2016;70:214–3.
34. Denny JC, Rutter JL, Goldstein DB, Philippakis A, Smoller JW, Jenkins G, et al. The “All of Us” Research Program. N Engl J Med. 2019;381:668–76.
35. Bitarello BD, Mathieson I. Polygenic scores for height in admixed populations. G3 (Bethesda). 2020;10:4027–36.

How to cite this article: Schwantes-An T-H, Whitfield JB, Aithal GP, Atkinson SR, Bataller R, Botwin G, et al. A polygenic risk score for alcohol-associated cirrhosis among heavy drinkers with European ancestry. Hepatol Commun. 2024;8: e0431. https://doi.org/10.1097/HC9.0000000000000431