Behavior Genetics (2023) 53:292–309
https://doi.org/10.1007/s10519-023-10139-w

ORIGINAL RESEARCH

Comparing Pruning and Thresholding with Continuous Shrinkage Polygenic Score Methods in a Large Sample of Ancestrally Diverse Adolescents from the ABCD Study®

Jonathan Ahern[1,2] · Wesley Thompson[3,4] · Chun Chieh Fan[4,5] · Robert Loughnan[1,2]

Received: 1 December 2022 / Accepted: 28 February 2023 / Published online: 5 April 2023

Abstract
Using individuals' genetic data researchers can generate Polygenic Scores (PS) that are able to predict risk for diseases, variability in different behaviors as well as anthropomorphic measures. This is achieved by leveraging models learned from previously published large Genome-Wide Association Studies (GWASs) associating locations in the genome with a phenotype of interest. Previous GWASs have predominantly been performed in European ancestry individuals. This is of concern as PS generated in samples with a different ancestry to the original training GWAS have been shown to have lower performance and limited portability, and many efforts are now underway to collect genetic databases on individuals of diverse ancestries. In this study, we compare multiple methods of generating PS, including pruning and thresholding and Bayesian continuous shrinkage models, to determine which of them is best able to overcome these limitations. To do this we use the ABCD Study, a longitudinal cohort with deep phenotyping on individuals of diverse ancestry. We generate PS for anthropometric and psychiatric phenotypes using previously published GWAS summary statistics and examine their performance in three subsamples of ABCD: African ancestry individuals (n = 811), European ancestry Individuals (n = 6703), and admixed ancestry individuals (n = 3664). We find that the single ancestry continuous shrinkage method, PRScs (CS), and the multi ancestry meta method, PRScsx Meta (CSx Meta), show the best performance across ancestries and phenotypes.

Keywords Polygenic score · ABCD · Diverse · Admixed · Adolescent · Continuous shrinkage

Introduction

The study of human genetics has established that the vast majority of complex traits that are genetically influenced are polygenic—with many small effects being distributed across the genome (Visscher et al. 2021). Polygenic Scores (PSs) are a class of prediction methods that can aggregate these polygenic effects for a given phenotype (e.g. height, BMI, psychiatric risk) to explain a substantial amount of variation in phenotypes using genetic data (Sugrue and Desikan 2019). While not useful as stand-alone diagnostic measures, previous research has shown that, when incorporated with other traditional risk measurements, PS can improve predictive model accuracy for common diseases like cancers (Jia et al. 2020; Kachuri et al. 2020), Coronary Artery Disease (Inouye et al. 2018; Klarin and Natarajan 2022), and Type 2 Diabetes (Ashenhurst et al. 2022; Ge et al. 2022).

A major challenge facing the application of PS is their diminished performance when being trained and deployed in different ancestry groups. Thus far the majority of

Handling Editor: John K Hewitt.

✉ Jonathan Ahern
jrahern@ucsd.edu

1 Department of Cognitive Science, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA

2 Center for Human Development, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92161, USA

3 Herbert Wertheim School of Public Health and Human Longevity Science, University of California San Diego, 9500 Gilman Drive, La Jolla, San Diego, CA 92161, USA

4 Center for Population Neuroscience and Genetics, Laureate Institute for Brain Research, Tulsa, OK 74103, USA

5 Department of Radiology, University of California, San Diego School of Medicine, 9500 Gilman Drive, La Jolla, CA 92037, USA

well-powered Genome-Wide Association Studies (GWASs) have been conducted on individuals of European ancestry (Bitarello and Mathieson 2020; Lewis and Green 2021; Martin et al. 2019; Peterson et al. 2019), limiting their utility in non-European ancestry groups. The over-representation of European ancestry individuals in genetic studies has led many to fear this may exacerbate health disparities (Martin et al. 2019). It is thought this drop in performance when deploying PS between ancestry groups is related to differences in linkage-disequilibrium (LD) and allele frequencies (Wang et al. 2020), which will likely be addressed by conducting GWAS in more diverse samples. While there have been efforts to improve the number of non-European GWAS samples by groups like the Hispanic/Latino Anthropometry (HISLA) Consortium (Fernández-Rhodes et al. 2022), the Population Architecture Using Genetics and Epidemiology (PAGE) Study (Matise et al. 2011), and the African Ancestry Anthropometry Genetics Consortium (AAAGC) (Ng et al. 2017) among others, European ancestry samples are usually the largest for a given phenotype and cross-ancestry PS methods are limited.

There exist multiple different methods for computing PS (Choi and O'Reilly 2019; Ge et al. 2019, 2022; Ruan et al. 2022) which each attempt to address two issues of GWAS: (1) distinguishing impactful and non-impactful variants and (2) LD correlations across the genome. The classic PS method, Pruning and Thresholding (P + T), addresses these issues by (a) grouping genomic variants according to LD correlations (pruning), then (b) restricting remaining variants to those meeting a p-value threshold (thresholding) (Choi et al. 2020; Choi and O'Reilly 2019; Marees et al. 2018). An alternative PS approach tackles these issues by using shrinkage techniques which have been shown to achieve superior performance (Ge et al. 2019; Privé et al. 2020). Finally, PS methods exist which are specifically designed to be deployed across multiple ancestries (Ge et al. 2022; Ruan et al. 2022). Benchmarking these methods in independent datasets of ancestral diverse individuals is of importance in evaluating their relative performance.

The Adolescent Brain Cognitive Development (ABCD) Study® is a longitudinal study with deep phenotyping of over 11,000 children from 9 to 11 years old, with wide sociodemographic and genetic diversity across the United States. This study provides an ideal opportunity to profile the performance of these recently developed PS methods in an ancestrally diverse cohort. In the present study, we deploy four PS methods: PRSice2, PRScs, PRScsx, and PRScsx Meta. For this analysis, we use PRSice2 as our P + T method (Choi and O'Reilly 2019). PRScs (Ge et al. 2019) has emerged as a particularly effective shrinkage PS method, and PRScsx and PRScsx Meta leverage information from multiple ancestries to improve PS generation trans-ancestrally (Ge et al. 2022; Ruan et al. 2022). We use each of these four methods to generate PS for four phenotypes in the ABCD sample: height, body mass index (BMI), schizophrenia risk, and depression risk. For each of these phenotypes, we utilize previous GWAS in five ancestral groups where possible: European (EUR), East Asian (EAS), South Asian (SAS), Hispanic (HIS), and African (AFR). The PSs from these methods are then evaluated, against relevant phenotypes, separately in European, African, and Mixed ancestry individuals in the ABCD cohort. We hope that the results from this analysis will help guide future work aiming to utilize PS in populations of diverse ancestries.

Methods

ABCD sample

Our sample consisted of 11,178 children from the 4.0 of the Adolescent Brain Cognitive Development (ABCD) study (https://doi.org/10.15154/1523041) with qualified genetic data. The ABCD cohort was recruited to ensure the sample was as close to nationally representative as possible, and therefore exhibits large sociodemographic diversity (Garavan et al. 2018). There is an embedded twin cohort and many siblings.

ABCD phenotypes

Dependent variables in this analysis come from the baseline visit from ABCD release 4.0. Two anthropomorphic traits were used in our analysis: height, calculated as the mean of 3 measurements, and body mass index (BMI), calculated as weight/height^2. Two behavioral metrics were used in our analysis: KSADS Total Symptoms as reported by the participant's caregiver and CBCL Total Problems based on an extensive battery of questionnaires and interviews. The Kiddie schedule for affective disorders and schizophrenia (KSADS) measure represents the combined outcome of a self-administered version of the K-SADS-5 assessment filled out by the caregiver of the child enrolled in the ABCD study. The KSADS assessment is a diagnostic tool commonly used to identify symptoms, behaviors, and impairments potentially related to psychiatric disorders, and our variable represents a normalized total of all responses to a participants K-SADS-5 assessment and is used in this analysis as a single metric of the level of presence or absence of psychiatric disorders (Kaufman et al. 1997). The KSADS variable used in our analysis was normalized using a rank-based inverse normal transformation. The K-SADS-5 assessment has been shown to be valid, reliable, and replicable for children of diverse cultural and national backgrounds (de la Peña et al. 2018; Dun et al. 2022; Kaufman et al. 1997; Kim et al. 2004; Nishiyama et al. 2020; Shahrivar et al. 2010).

The child behavioral checklist (CBCL) is a caregiver-reported assessment of their child on eight syndrome scales: anxious/depressed, withdrawn/depressed, somatic complaints, social problems, thought problems, attention problems, rule-breaking behavior, and aggressive behavior which are then combined into a single score for this analysis representing the normalized total number of behavioral problems reported in the CBCL assessment. The resulting combined score was normalized using a rank-based inverse normal transformation. The CBCL Total Symptoms score represents a single metric of a participants emotional functioning and well-being as evidenced through behavior (Achenbach and Rescorla 2004). The CBCL is a portion of the Achenbach System of Empirically Based Assessment (ASEBA) designed to be used on school-aged children between the ages of 6 and 18 (Achenbach and Rescorla 2004) and it has been found to be valid and reliable for children of diverse cultural and national backgrounds (Albores-Gallo et al. 2007; Dutra et al. 2004; Hartini et al. 2015; Leung et al. 2006).

Summary statistics data access

Our analyses were limited by the public availability of large and diverse GWAS of mental phenotypes. Furthermore, our trans-ancestry method required that our summary statistics be easily separable into distinct continental ancestries which prevented us from using some large trans-ancestry meta-analyses. The phenotypes height and BMI were selected to show the efficacy of these methods across some of the most well-powered anthropomorphic phenotypes. Depression was chosen because of its prevalence among adolescence (Goodwin et al. 2022) and because of its availability as summary statistics in multiple ancestry groups. Schizophrenia was chosen despite its low prevalence in an adolescent population sample because of its availability of summary statistics in non-European ancestries and because previous studies have found differential experiential and mental phenotypic manifestations in adolescents with high genetic load for schizophrenia before clinical manifestation of the disorder (Jones et al. 2016; Woolway et al. 2022). Data were collected from publicly available GWAS summary statistics found through GWAS Catalog (https://www.ebi.ac.uk/gwas/

Table 1 Size and demographic information of the ancestry subsamples used in this analysis

Ancestry group	African ancestry (AFR)	European ancestry (EUR)	Admixed ancestry (MIX)	All samples
Total N	811	6703	3664	11,178
	Mean (SD)			
Age in years	9.9 (0.59)	9.9 (0.63)	9.9 (0.63)	9.9 (0.92)
Height in inches	56.19 (3.62)	55.16 (3.19)	55.28 (3.39)	55.27 (3.30)
BMI	20.73 (5.16)	17.99 (3.41)	19.93 (4.84)	18.82 (4.21)
CBCL total problems	20.02 (21.21)	17.81 (17.14)	18.89 (18.75)	18.33 (18.01)
KSADS total problems	23.32 (30.25)	18.37 (24.03)	19.14 (26.69)	18.98 (25.45)
Sex	%			
Female	51.30%	47.30%	47.90%	47.80%
Parental education				
< HS diploma	10.10%	0.60%	10.60%	4.60%
HS diploma/GED	24.10%	3.70%	16.50%	9.40%
Some college	39.20%	19.80%	36%	26.50%
Bachelor's degree	12.50%	30.90%	18.70%	25.60%
Post graduate degree	14%	44.90%	18.20%	33.90%
Household income				
Less than $50,000	70.30%	13.70%	51.80%	29.40%
Between $50,000 and $100,000	20.10%	30.70%	26.10%	28.60%
Greater than $100,000	9.60%	55.60%	22.00%	42.00%
Race and ethnicity (self-report)				
Asian	0%	0.10%	2.50%	0.90%
Black	97.50%	0.10%	28.40%	16.30%
Hispanic	0.60%	7.40%	47.40%	20%
White	0.10%	94.50%	27.00%	65.50%

Additional graphics showing the distributions of our phenotypes of interest can be found in Supplementary Figs. 2–5

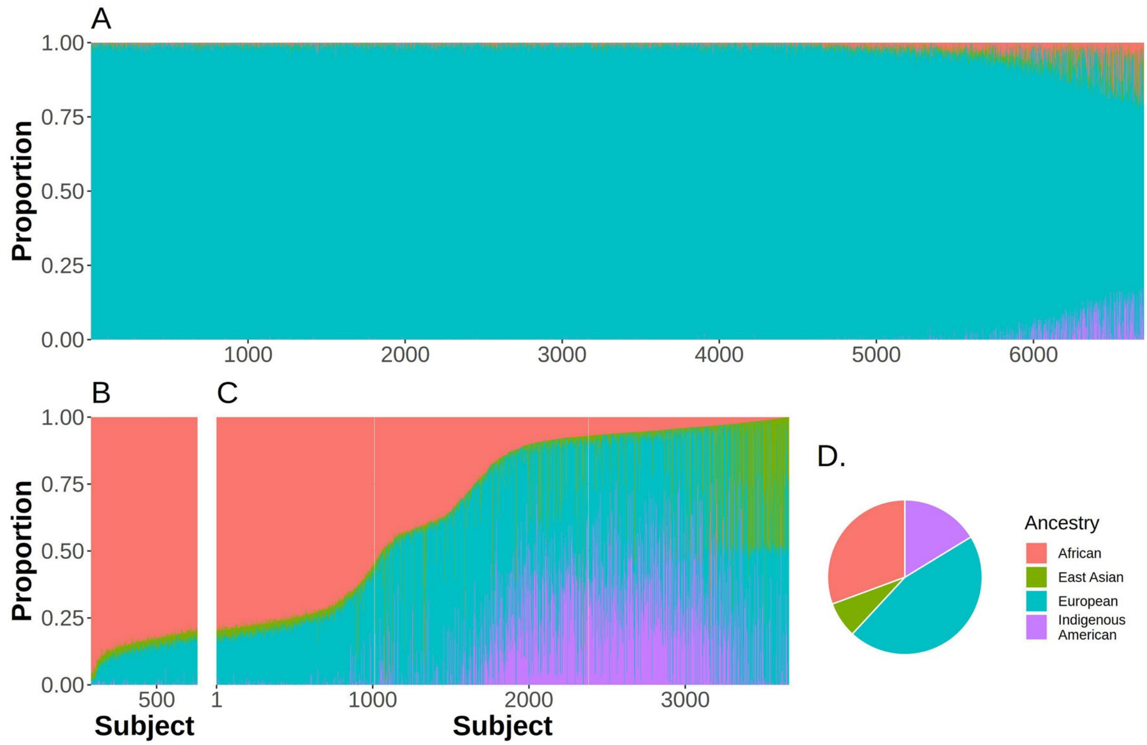


Fig. 1 Genetic ancestry proportion for individuals of each ancestry subpopulation in ABCD categorized into African, East Asian, European, and Indigenous North and South American ancestry portions. **A**–**C** represent European, African, and Admixed ancestry populations respectively. **D** shows the mean genetic ancestry components of the admixed population

home) and Google scholar. Additional information regarding existing and available GWAS summary statistics was also gathered from the GWAS catalog (https://www.ebi.ac.uk/gwas/home). A full list of data sets used can be found in Supplementary Table 1. All summary statistics were aligned to genome build GRCh38.

Genetic data

Genetic data was collected using blood or saliva samples from participants of the ABCD study (Uban et al. 2018). 656,247 genomic markers were measured using the Smokescreen array (Baurley et al. 2016). Genetic principal components were calculated from these genetic data using PC-Air (Conomos et al. 2015) with default settings. We calculated participants' continental genetic ancestry as calculated using SNPweights (Chen et al. 2013) and precompiled external genomic reference panels from the 1000 Genomes Project (Auton et al. 2015), and Indigenous reference panels (Reich et al. 2012). Participants were categorized into groups depending on if they had an inferred genetic ancestry at least 80% consistent with a continental reference panel (African continental ancestry, East Asian continental ancestry, European continental ancestry, or indigenous North and South American ancestry) or admixed meaning that their genetic ancestry did not meet the 80% threshold for any of the ancestries due to a genetic admixture of two or more of the aforementioned continental ancestral components. We chose to use inferred genetic ancestry as opposed to other methods as this enabled us to not only define individuals of continental ancestries but also to define an admixed ancestry group (i.e., those that did not meet criteria for a single continental ancestry). Genetic PCA plots of our sample (see Supplementary Fig. 1) show that inferred ancestries labels individuals as would be expected. Due to the relatively small portion of the participants of genomic indigenous ancestry (n = 53) and participants of continental East Asian ancestry (n = 158) were left out of our total testing sample. Our total testing sample consisted of an African ancestry group (AFR, n = 811), a European ancestry group (EUR, n = 6703), and an admixed ancestry group (MIX, n = 3664).

To increase the overlap of genetic variants in ABCD with summary statistics from previous GWAS we imputed markers measured from the Smokescreen array using the TOPMED imputation server (Taliun et al. 2021). These imputed variants were fractional dosages that were converted to an integer number of alleles using a best guess threshold of 0.9. This resulted in 280,850,795 imputed variants aligned to genome build GRCh38. After imputation target genetic data was restricted to only autosomal variants with a minor allele frequency of 1% (0.01) or greater leaving just under 11 million Single Nucleotide Polymorphisms (SNPs) in the target data.

Polygenic score methods

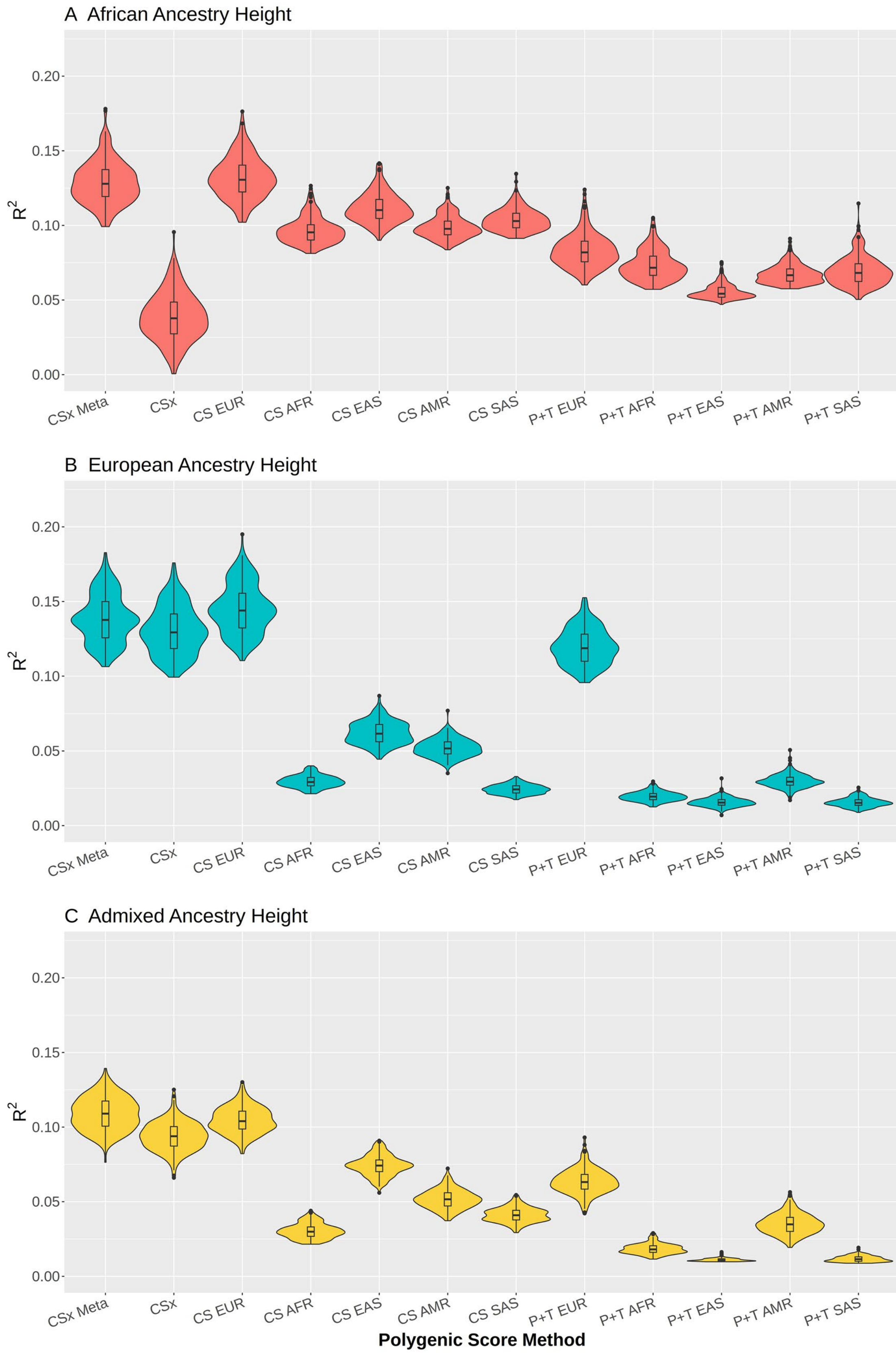
As some methods required hyperparameter tuning we generated 100 50/50 cross-validation splits within each ancestry group. We ensured that family members were not split across training and testing folds (with families being defined using the 'rel_family_id' variable). This provided training folds, for hyperparameter tuning, and testing folds for the evaluation of PS performance.

We provide a brief description of each PS method used in this analysis:

1. P + T: PRSice2
    a. PRSice2 is an LD-informed pruning and P-value thresholding method meaning that it groups and thins SNPs according to LD and P-value and then limits these SNPs to only those that exceed a given P-Threshold (Choi and O'Reilly 2019). For our analysis, we used default parameters with a range of P-value thresholds (0, 0.001, 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 0.9 and 1) where we applied the threshold maximizing R^2 in the training fold to the test fold. Reported R^2 values were calculated in the testing fold.
2. CS: PRScs
    a. PRScs is a Bayesian calculation method that uses GWAS summary statistics and LD reference information as well as a continuous shrinkage prior to infer posterior SNP effect sizes which is found to be both reliable and computationally efficient in varying genetic architectures (Ge et al. 2019). For our analysis, we used base parameters except with regards to MCMC and burnin where we used a slightly higher threshold of 10,000 MCMC iterations and 5000 burnin. These values were chosen as they have been shown to help increase stability of posterior effects between runs without being too computationally intensive (Schultz et al. 2022). LD references were based on data from 1000 Genomes Phase 3 to match respective summary statistics (AFR, AMR, EAS, EUR, or SAS). PLINK 2.0 was used to generate Polygenic Risk scores from PRScs posterior effect sizes (Purcell et al. 2007).
3. CSx: PRScsx
    a. PRScsx is a Bayesian polygenic modeling method that integrates GWAS and LD information from multiple ancestrally diverse populations to improve the estimation of posterior SNP effects (Ruan et al. 2022). For our analysis, we used recommended parameters (apart from MCMC iterations and burnin) and the appropriate provided LD references from 1000 Genomes Phase 3 for a given summary statistic (AFR, AMR, EAS, EUR, or SAS). PLINK 2.0 was used to generate ancestry-specific PSs from PRScsx posterior effect sizes (Purcell et al. 2007). We used 10,000 MCMC iterations and 5000 burnin to increase stability (Schultz et al. 2022) and to be constant with other methods. Ancestry-specific PSs were combined using a linear combination of ancestry-specific PSs as advised by the authors of the method. Weights were learned in training folds using linear regression to predict the phenotype of interest as:

$$\mathrm{PRS} = w_{\mathrm{Ancestry1}}\mathrm{PRS}_{\mathrm{Ancestry1}} + w_{\mathrm{Ancestry2}}\mathrm{PRS}_{\mathrm{Ancestry2}} \dots + w_{\mathrm{AncestryN}}\mathrm{PRS}_{\mathrm{AncestryN}},$$

    where w represents the relative weight. R^2 values were reported in the validation fold after applying weights learned in training folds.
4. CSx Meta: PRScsx Meta
    a. PRScsx Meta uses the same Bayesian polygenic model as PRScsx, but instead of needing hyperparameter tuning, it uses an inverse-variance-weighted meta-analysis to produce a single set of posterior effects (Ge et al. 2022). For our analysis, we used recommended parameters (apart from MCMC iterations and burnin) and the appropriate provided LD references from 1000 Genomes Phase 3 for a given summary statistic (AFR, AMR, EAS, EUR, or SAS). We used 10,000 MCMC iterations and 5000 burnin to increase stability (Schultz et al. 2022) and to be constant with other continuous shrinkage methods. PLINK 2.0 was used to generate Polygenic Risk scores from PRScsx posterior effect sizes (Purcell et al. 2007).


A African Ancestry Height
B European Ancestry Height
C Admixed Ancestry Height
R^2
0.00
0.05
0.10
0.15
0.20
CSx Meta
CSx
CS EUR
CS AFR
CS EAS
CS AMR
CS SAS
P+T EUR
P+T AFR
P+T EAS
P+T AMR
P+T SAS
Polygenic Score Method

◂**Fig. 2** Variance explained of PS methods in predicting height across 100 folds of 50/50 cross-validation. **A**–**C** Represent performance in AFR, EUR, and MIX ancestry populations respectively. Methods marked 'CSx Meta' and 'CSx' represent the meta-analysis and hyperparameter-weighted trans-ancestry outputs from the continuous shrinkage method PRScsx. Methods marked 'CS' represent the single ancestry outputs of the continuous shrinkage method PRScs (both LD reference and summary statistics from a single ancestry). Methods marked 'P+T' represent the single ancestry outputs from the linear pruning and thresholding method PRSice2. The three letter abbreviations included in some methods represents the sample ancestries as follows: European ancestry (EUR), African ancestry (AFR), East Asian ancestry (EAS), North or South American ancestry (AMR), and South Asian ancestry (SAS). Additional information about the summary statistics can be found in Supplementary Tables 1 and additional phenotype data can be found in Supplementary Fig. 2

Statistical analysis

To assess the association between each PS method and the relevant dependent variable Generalized Additive Mixed Models (GAMMS) were fitted using the gamm4 package in R (Wood and Scheipl 2022) in each test fold. Each model predicted a different mental or anthropomorphic feature, depending on the PS of interest. Each model was corrected for participants' age in months, sex, ABCD study site, and the top ten generic principal components calculated using PC-AIR (Conomos et al. 2015)as fixed effects and a participants' family id as a random effect. Nagelkerke R^2 (Nagelkerke 1991) values were calculated between reduced (covariates only) and full (covariates + PS) models. Supplementary Tables were generated using the t-tests functions and False Discovery Rate (FDR) correction from the R package 'stats' (R Core Team 2022).

Genome-wide complex trait analysis (GCTA)

For some traits we unexpectedly found higher performance in non-European ancestry cohorts. In attempt to understand this further we conducted GCTA (Yang et al. 2011) each ancestry subsample separately to estimate snp-heritability (h^2_{snp}) for each analyzed measure in ABCD to quantify the genetic variability contributing to each trait—independent of previous GWAS summary statistics or any specific polygenic score method. For this we constructed a GRM (genetic relatedness matrix) using 'gcta –make-grm' for each ancestry individually. We then filtered this GRM to unrelated individuals using a threshold of --grm-cutoff = 0.025 as recommended. This resulted in 5133, 683 and 414 individuals in European, African and Admixed ancestry cohorts, respectively. We then performed GCTA on this pruned GRM, using covariates described above to obtain point estimates of h^2_{snp}. With the small sample sizes for this analysis we observe large error bars and so use point estimates as an indication of differences and interpret with caution.

Results

The results in this analysis utilize the full baseline visit data from ABCD data release 4.0, an ancestrally diverse longitudinal cohort of children from 21 different data acquisition sites around the United States (https://doi.org/10.15154/1523041). We restricted our analysis to the three largest ancestral groups of the ABCD sample: African (AFR), European (EUR), and Admixed (MIX). Membership within each ancestry group was defined as an individual having greater than 80% inferred continental ancestry for the given group. We inferred participants' continental genetic ancestry as calculated using SNPweights (Chen et al. 2013) and precompiled external genomic reference panels from the 1000 Genomes Project (Auton et al. 2015), and Indigenous reference panels (Reich et al. 2012). Genetic ancestry estimates for each group are shown in Fig. 1. Assignment into each of these three ancestral groups covers 98% of the full ABCD baseline sample. Sample sizes and demographic information of these groups are presented in Table 1.

Height polygenic score

Figure 2 shows the predictive performance of different PS models in predicting height. The single ancestry continuous shrinkage method based on European ancestry LD and summary statistics from a large European GWAS (Yengo et al. 2022), CS EUR, outperformed other methods in the European and African ancestry groups with a mean variance explained of 14.5% and 13.2% respectively. In the admixed ancestry group the best performing method was the trans-ancestry, continuous shrinkage, meta-analysis method, CSx Meta, which accounted for a mean variance explained of 10.9%. CS EUR was roughly comparable to CSx Meta ($R_{CS\ EUR}^2/R_{CSx\ Meta}^2$ = EUR:1.05, AFR:1.03, MIX:0.96) and was an improvement on CSx ($R_{CS\ EUR}^2/R_{CSx}^2$ = EUR:1.11, AFR:3.42, MIX:1.11) and the best performing P + T Method, P + T EUR ($R_{CS\ EUR}^2/R_{P+T\ EUR}^2$ = EUR:1.21, AFR:1.59, MIX:1.65).

Although CSx achieved worse but comparable performance in EUR and MIX groups, it showed particularly low performance in the AFR group (see Fig. 2). This may be explained by the wide variability in weightings of ancestry-specific PS across training folds used to calculate CSx for the AFR group: compare panel A with panels B and C in Supplementary Fig. 6. P + T (using PRSice2) showed the lowest performance across all ancestry groups except the African ancestry sample where PRScsx showed the lowest performance. Numeric summaries of these results are available in Supplemental Tables 2–4.

BMI polygenic score

When comparing the performance of PS methods in predicting Body Mass Index (BMI), we found that the trans-ancestry, continuous shrinkage, meta-analysis method, CSx Meta, performed best in all ancestries. CSx Meta accounted for a men variance explained of 11.7% in the European ancestry sample, 11.9% in the African ancestry sample, and 9.9% in the admixed ancestry sample. Numeric summaries of these results are available in Supplemental Tables 5–7.

Once again, we observed particularly low performance of the CSx method in the AFR and observed wide variability in ancestry specific PS weights in training for this method (Supplementary Fig. 7). Once again, the P + T methods (PRSice2) exhibited the lowest performance across ancestry groups, except in the AFR group where the CSx method had the lowest performance (Fig. 3).

Depression polygenic score

In predicting CBCL total problems from depression GWAS, we found the CS EAS method trained on an East Asian GWAS (Giannakopoulou et al. 2021) achieved the highest performance in the African ancestry sample, the CS EUR methods trained on a European GWAS (Howard et al. 2019) achieved the highest performance in the European ancestry sample, and CSx Meta performed best in the admixed ancestry sample—see Fig. 4. The mean variance explained by CS EAS was 10.1% in the African ancestry sample, 1.1% in the European ancestry sample, and 2.0% in the admixed ancestry sample. The mean variance explained for the CS EUR method was: 9.8% in the African ancestry sample, 1.7% in the European ancestry sample, and 2.2% in the admixed sample. The mean variance explained by CSx Meta was 10.0% in the African ancestry sample, 1.6% in the European ancestry sample, and 2.2% in admixed ancestry sample. Numeric summaries of these results are available in supplemental Tables 8–10.

In predicting KSADs total problems from depression GWAS, we found the CS EAS method trained on an East Asian GWAS (Giannakopoulou et al. 2021) achieved the highest performance in the African ancestry sample, the CS EUR methods trained on a European GWAS (Howard et al. 2019) achieved the highest performance in the European ancestry sample, the CSx Meta method achieved the highest performance in the admixed ancestry sample—see Fig. 4. Mean variance explained for the CS EAS method was 10.7% in the African ancestry sample, 1.1% in the European ancestry sample, and 2.1% in the admixed ancestry sample. Mean variance explained for the CS EUR method was 10.1% in the African ancestry sample, 1.5% in the European ancestry sample, and 2.3% in the admixed ancestry sample. Mean variance explained for the CSx Meta method was 10.2% in the African ancestry sample, 1.4% in the European ancestry sample, and 2.3% in the admixed ancestry sample.

We observe low performance of the CSx method in predicting both CBCL total problems and KSADS total problems across all ancestry groups, this once again may be due to unstable weightings of ancestry-specific PS that make up the CSx method across cross-validation folds (Supplementary Figs. 8–9 Numeric summaries of these results are available in supplemental Tables 11–13 (Fig. 5).

Schizophrenia polygenic score

In predicting CBCL total problems from schizophrenia GWAS, we find that the CS EAS method trained on the large East Asian GWAS (Trubetskoy et al. 2022) performed best in the African and admixed ancestry samples and the CS EUR method trained on the large European GWAS (Trubetskoy et al. 2022) performed best in the European ancestry sample. Mean variance explained for the CS EAS method was 9.5% in the African ancestry sample, 1.0% in the European ancestry sample, and 1.9% in the admixed ancestry sample. Mean variance explained for the CS EUR method was 9.2% in the African ancestry sample, 1.0% in the European ancestry sample, and 1.8% in the admixed ancestry sample. There is not a significant difference between the results from CS EAS, CS EUR, and CSx Meta which all accounted for a mean variance of approximately 1.0% ($p_{\text{CS EAS}-\text{CS EUR}} = 0.39$, $p_{\text{CS EAS}-\text{CSx Meta}} = 0.88$, $p_{\text{CS EUR}-\text{CSx Meta}} = 0.08$). A graphic summary of these results can be found in Fig 6. Numeric summaries of these results are available in Supplemental Tables 14–16.

In predicting KSADS total problems from schizophrenia GWAS, we find that the CS EAS method trained on the large East Asian GWAS (Trubetskoy et al. 2022) performed best in the African ancestry sample, the CSx Meta method performed best in the European ancestry sample, and the CS EUR method trained on the large European GWAS (Trubetskoy et al. 2022) performed best in the admixed ancestry sample. Mean variance explained for the CS EAS method was 9.5% in the African ancestry sample, 1.0% in the European ancestry sample, and 1.9% in the admixed ancestry sample. Mean variance explained for the CSx Meta method was 9.0% in the African ancestry sample, 1.0% in the European ancestry sample, and 1.8% in the admixed ancestry sample. Mean variance explained for the CS EUR method was 9.3% in the African ancestry sample, 1.0% in the European ancestry sample, and 1.9% in the admixed ancestry sample. It is worth noting that the results of CS EAS and CS EUR are not significantly different in the MIX group

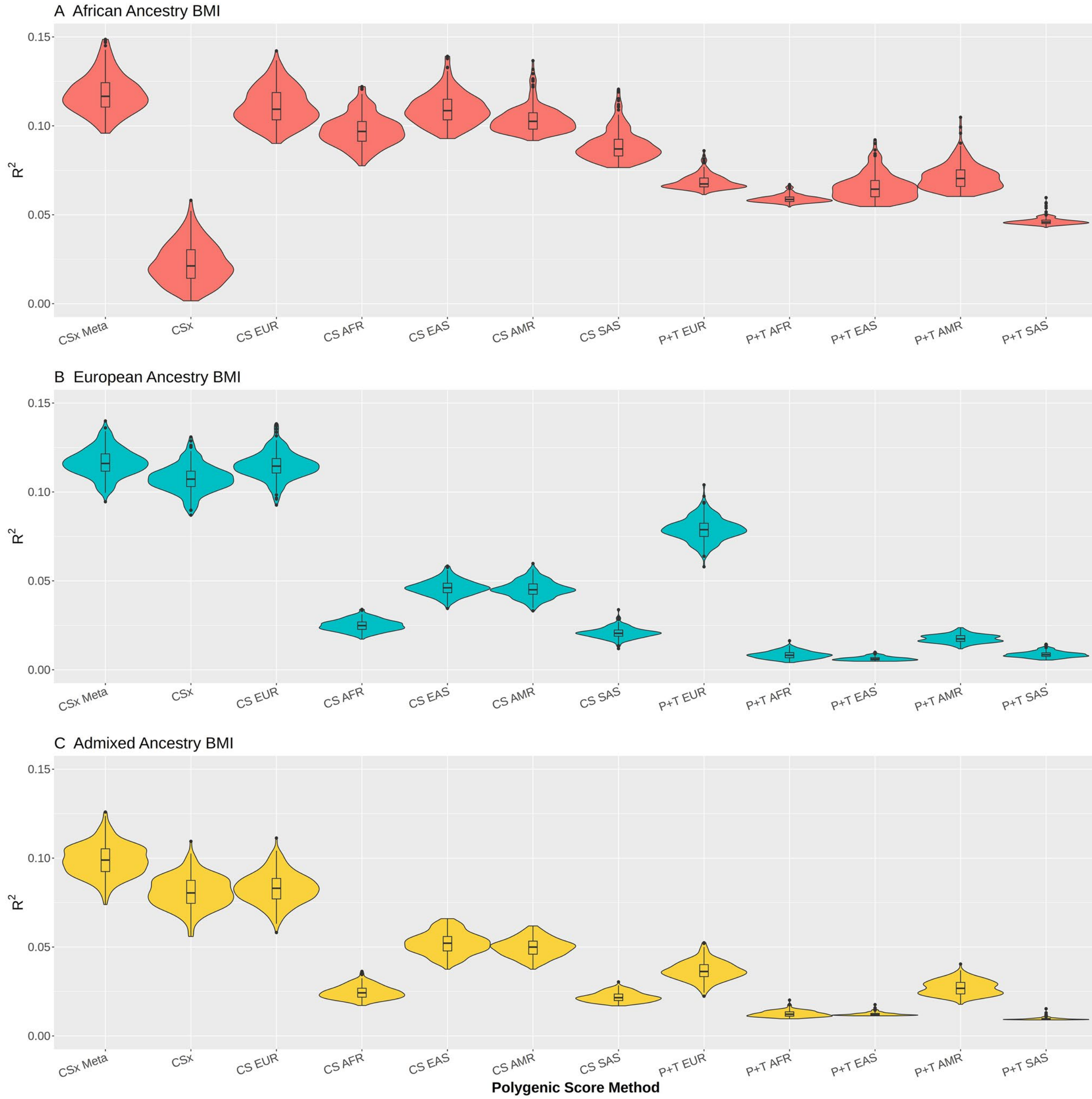


Fig. 3 Variance explained of PS methods in predicting BMI across 100 folds of 50/50 cross-validation. Panels **A**, **B**, and **C** represent performance in AFR, EUR, and MIX ancestry populations respectively. Methods marked 'CSx Meta' and 'CSx' represent the meta-analysis and hyperparameter-weighted trans-ancestry outputs from the continuous shrinkage method PRScsx. Methods marked 'CS' represent the single ancestry outputs of the continuous shrinkage method PRScs (both LD reference and summary statistics from a single ancestry). Methods marked 'P+T' represent the single ancestry outputs from the linear pruning and thresholding method PRSice2. The three letter abbreviations included in some methods represents the sample ancestries as follows: European ancestry (EUR), African ancestry (AFR), East Asian ancestry (EAS), North or South American ancestry (AMR), and South Asian ancestry (SAS). Additional information about the summary statistics can be found in Supplementary Tables 1 and additional phenotype data can be found in Supplementary Fig. 3

($p=0.87$). A graphic summary of these results can be found in Fig. 7. Numeric summaries of these results are available in Supplemental Tables 17–19.As with depression PSs, schizophrenia PSs predicting CBCL total problems and KSADS total problems, we observe particularly poor performance of the CSx method; this may be due to unstable

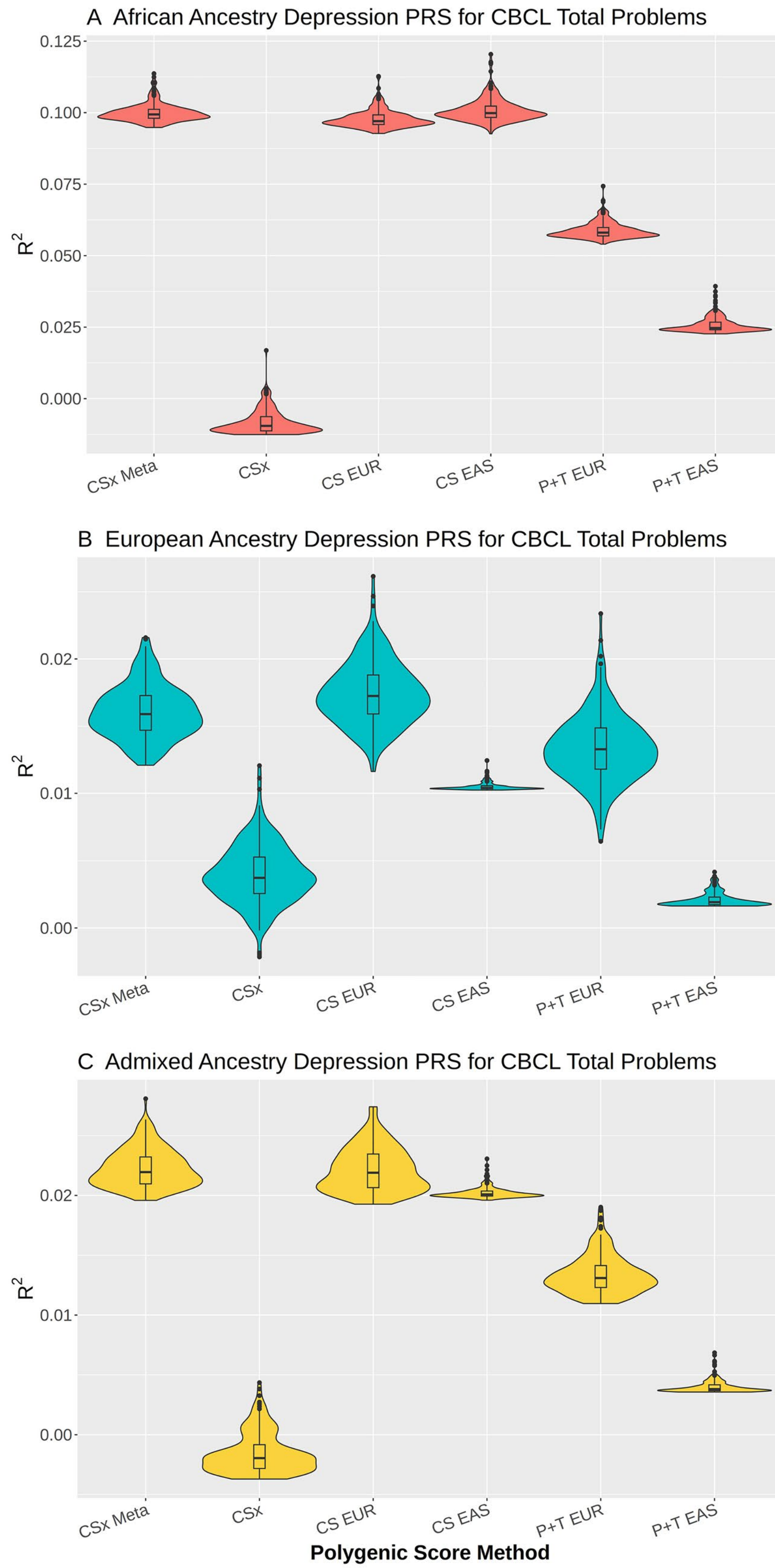


Fig. 4 Variance explained of PS methods applied to depression in predicting CBCL total problems scores across 100 folds of 50/50 cross-validation. **A**–**C** represent performance in AFR, EUR, and MIX ancestry populations respectively. Methods marked 'CSx Meta' and 'CSx' represent the meta-analysis and hyperparameter-weighted trans-ancestry outputs from the continuous shrinkage method, PRScsx. Methods marked 'CS' represent the single ancestry outputs of the continuous shrinkage method PRScs (both LD reference and summary statistics from a single ancestry). Methods marked 'P + T' represent the single ancestry outputs from the linear pruning and thresholding method PRSice2. The three letter abbreviations included in some methods represents the sample ancestries as follows: European ancestry (EUR) and East Asian ancestry (EAS). Additional information about the summary statistics can be found in Supplementary Tables 1 and additional phenotype data can be found in Supplementary Fig. 4

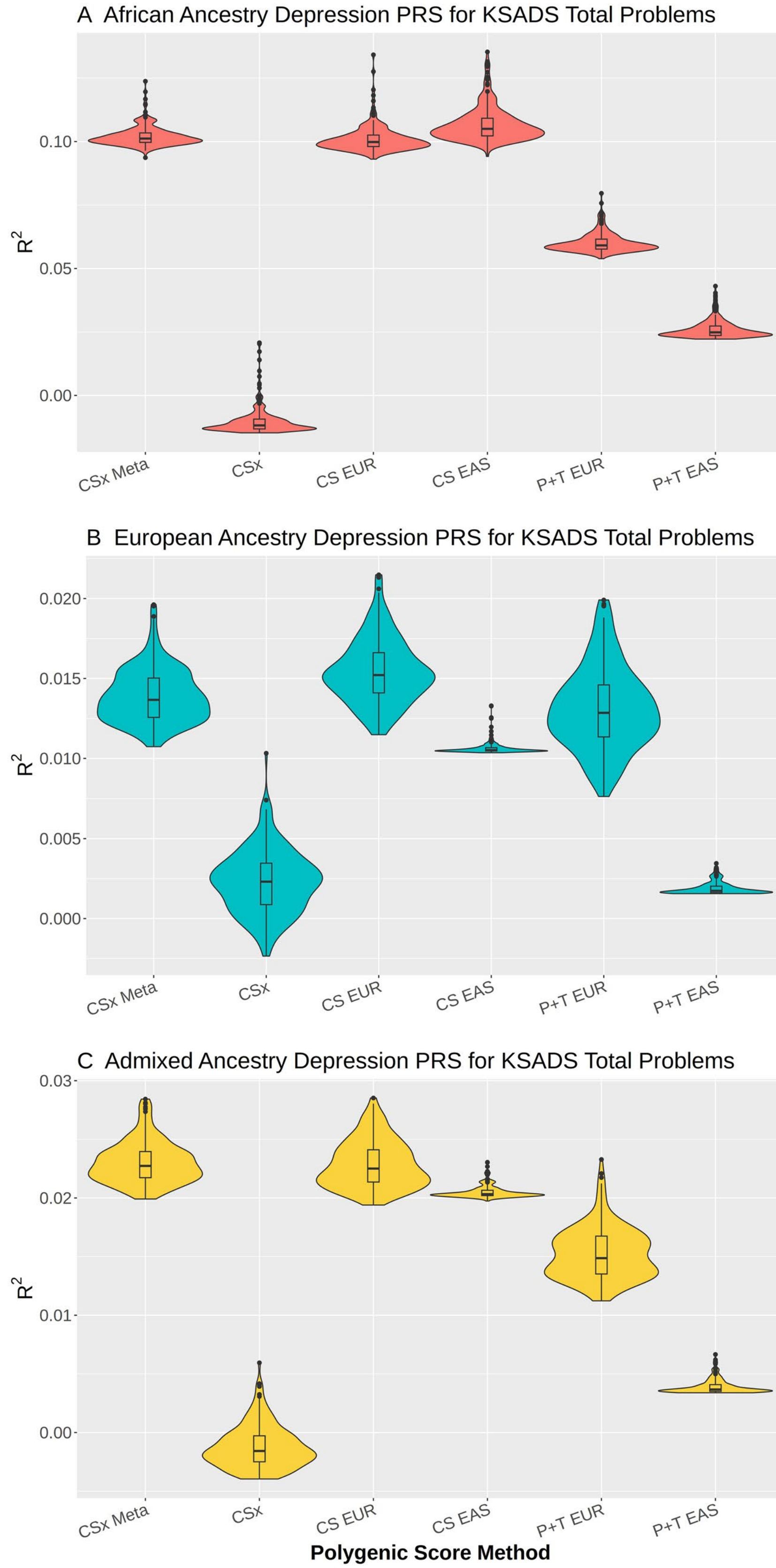
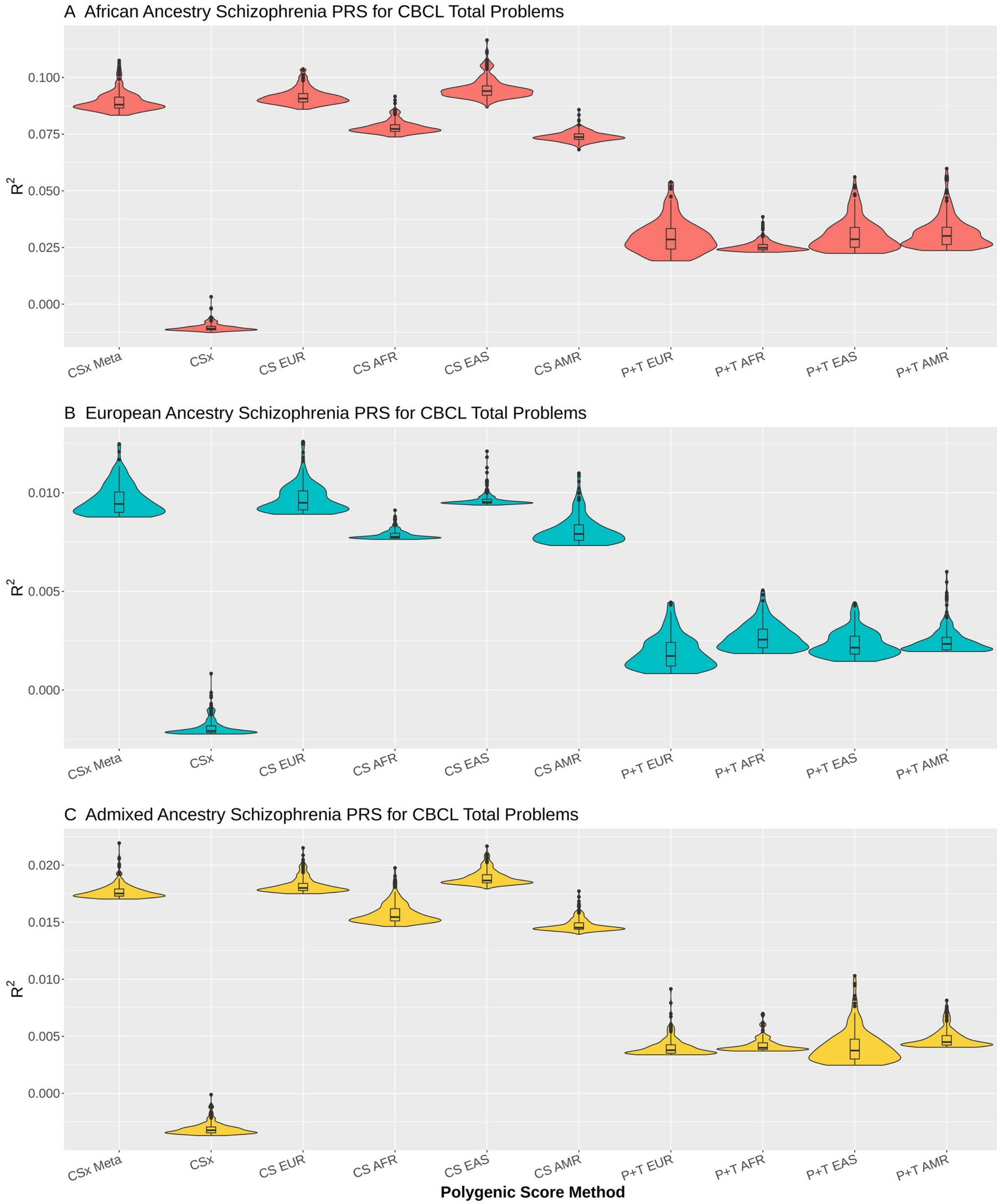


Fig. 5 Variance explained of PS methods applied to depression in predicting KSADs total problems scores across 100 folds of 50/50 cross-validation. **A–C** represent performance in AFR, EUR, and MIX ancestry populations respectively. Methods marked 'CSx Meta' and 'CSx' represent the meta-analysis and hyperparameter-weighted trans-ancestry outputs from the continuous shrinkage method PRScsx. Methods marked 'CS' represent the single ancestry outputs of the continuous shrinkage method PRScs (both LD reference and summary statistics from a single ancestry). Methods marked 'P+T' represent the single ancestry outputs from the linear pruning and thresholding method PRSice2. The three letter abbreviations included in some methods represents the sample ancestries as follows: European ancestry (EUR) and East Asian ancestry (EAS). Additional information about the summary statistics can be found in Supplementary Tables 1 and additional phenotype data can be found in Supplementary Fig. 5



weightings of ancestry-specific PS that make up the CSx method across cross-validation folds (Supplementary Figs. 9–10).

GCTA analysis

For psychological traits the polygenic prediction in the African ancestry group showed surprisingly high performance.

◂**Fig. 6** Variance explained of PS methods applied to schizophrenia in predicting CBCL total problems scores across 100 folds of 50/50 cross-validation. **A**–**C** represent performance in AFR, EUR, and MIX ancestry populations respectively. Methods marked 'CSx Meta' and 'CSx' represent the meta-analysis and hyperparameter-weighted trans-ancestry outputs from the continuous shrinkage method PRScsx. Methods marked 'CS' represent the single ancestry outputs of the continuous shrinkage method PRScs (both LD reference and summary statistics from a single ancestry). Methods marked 'P+T' represent the single ancestry outputs from the linear pruning and thresholding method PRSice2. The three letter abbreviations included in some methods represents the sample ancestries as follows: European ancestry (EUR), African ancestry (AFR), East Asian ancestry (EAS), and North or South American ancestry (AMR). Additional information about the summary statistics can be found in Supplementary Tables 1 and additional phenotype data can be found in Supplementary Fig. 4

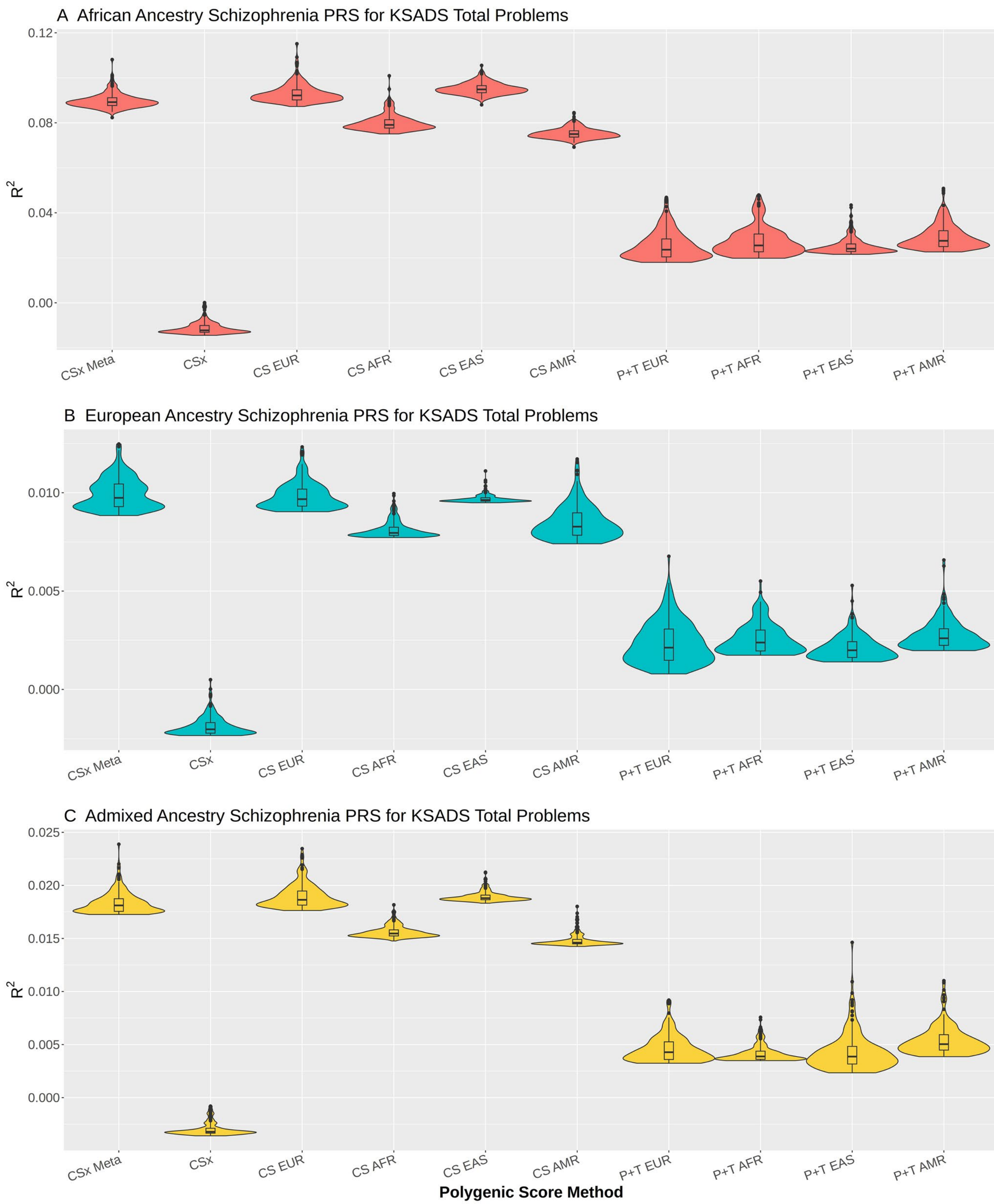
To further explore these results we performed Genome-wide Complex Trait Analyses (GCTAs) to see if there were significant differences in SNP heritability (h^2_{snp}) estimates for our mental health traits between our different samples (Yang et al. 2011).

We observed higher point estimates for h^2_{snp} the African ancestry cohort vs. European and admixed ancestry cohorts for CBCL Total Problems (AFR: h^2_{snp} = 0.34, SE = 0.88; EUR: h^2_{snp} = 0.19, SE = 0.08; MIX: h^2_{snp} = 0.06, SE = 0.82) and for KSADS Total Problems (AFR: h^2_{snp} = 1.00, SE = 0.93; EUR: h^2_{snp} = 0.00, SE = 0.08; MIX: h^2_{snp} = 0.00, SE = 0.09). This is consistent with greater genetic variability in this trait for African ancestry individuals – which may explain our unexpected result of higher polygenic prediction in this sample. However, with small sample sizes we observe wide error bars and so we advise caution in overinterpreting this result. Full results of GCTA Analysis can be found in Supplementary Tables 20–25.

Discussion

In ABCD, we find PRScsx Meta (CSx Meta) and PRScs (CS), especially those run on large European ancestry sample, provide improved polygenic prediction over pruning and thresholding methods. Also, the addition of multiple non-European ancestry reference panels and summary statistics seemed to provide greater predictive utility to models in African and Admixed populations than to European Populations.

Our results showed that PS for height and BMI explained between 9.8 and 14.5% of variance across all ancestry groups. For anthropometric traits, the methods used in this analysis perform best in the European ancestry sample. Differences in performance are likely due to differences in GWAS sample sizes for different ancestries (Karunamuni et al. 2020; Wu et al. 2022). Additionally, for the admixed cohort proportion of European ancestry may be a factor; previous research has shown that in ancestrally diverse populations the predictive validity of PS increases linearly with the individual's proportion of European genetic ancestry (Bitarello and Mathieson 2020). Unexpectedly, despite PRScsx being a trans-ancestry method it didn't show as strong of performance in this sample compared to a previous study (Ruan et al. 2022). This relatively poor performance could be due to a lack of homogeneity in the United States (Adhikari et al. 2017) or the high degree of genetic diversity within continental ancestry groups (Adhikari et al. 2016; Campbell and Tishkoff 2008). Indeed, the weighting of the different ancestral components making up PRScsx appeared unstable in our analysis, particularly for psychiatric PS, (see Supplementary Figs. 6–11) which may be driven by these factors and likely explains this method's low performance in our sample. Additionally, the two best-performing methods (PRScs and PRScsx-Meta) do not require hyperparameter tuning which negates the requirement for cross-validation in the target population making them much easier to deploy. In mental health traits we found uncharacteristically high performance in our African ancestry sample. GCTA analysis showed there were differences in heritability between our samples that could potentially account for the improved performance; however, the wide-error bars of the results prevent us from drawing any definitive conclusions. Additional research is needed to understand the potential factors that underlie this result. There was also surprising predictive power of the CS methods based on GWAS in East Asian Populations (Giannakopoulou et al. 2021; Trubetskoy et al. 2022) on our mental health phenotypes, despite their smaller discovery sample sizes. While the true difference between the results analysis is small and sometimes even not significant, we propose a potential explanation. First there may be differences in environmental factors or confounding diagnostic practices for schizophrenia or depression in different regions in which the discovery GWAS were conducted. Previous research has highlighted geographic differences in diagnostic criteria of psychiatric conditions (Mitchell et al. 2011; Saito et al. 2022). Because we are associating polygenic scores with measures of behavior and symptomology instead of formal diagnoses, it is possible that diagnostic criteria for schizophrenia and depression used in East Asia happens to be more related to variability in CBCL total problems and KSADS total problems than criteria for these same disorders in Europe. Such effects may explain differences in genetic correlations between eastern and western countries observed for psychiatric disorders (Saito et al. 2022; THE BRAINSTORM CONSORTIUM et al. 2018). In any case, given the generally low performance of these models in the majority of our participants we again caution the reader from over interpreting our results. Additional, more



targeted research is needed to fully understand the association between genetics and any behavioral or psychiatric outcome.

There is potential that PS performance on anthropometrics traits in our analysis may be low due to the age of our sample. Previous research has shown that genetic factors exert a particularly strong influence on body height between

◂**Fig. 7** Variance explained of PS methods applied to schizophrenia in predicting KSADs total problems scores across 100 folds of 50/50 cross-validation. **A**–**C** represent performance in AFR, EUR, and MIX ancestry populations respectively. Methods marked 'CSx Meta' and 'CSx' represent the meta-analysis and hyperparameter-weighted trans-ancestry outputs from the continuous shrinkage method PRScsx. Methods marked 'CS' represent the single ancestry outputs of the continuous shrinkage method PRScs (both LD reference and summary statistics from a single ancestry). Methods marked 'P+T' represent the single ancestry outputs from the linear pruning and thresholding method PRSice2. The three letter abbreviations included in some methods represents the sample ancestries as follows: European ancestry (EUR), African ancestry (AFR), East Asian ancestry (EAS), and North or South American ancestry (AMR). Additional information about the summary statistics can be found in Supplementary Tables 1 and additional phenotype data can be found in Supplementary Fig. 5

the ages of 14 and 18 (Silventoinen et al. 2008) and that PRSs of BMI increase in efficacy as an individual's age passes adolescence into adulthood (Sanz-de-Galdeano et al. 2020). Additionally, age may play a role in the efficacy of our polygenic models based on GWAS of schizophrenia as, in most cases, schizophrenia is diagnosed in late adolescence and early adulthood (Walker et al. 2004), and premorbid declines in cognition are normally assessed during the onset of puberty between the ages of 13 and 16 years (Fuller et al. 2002). Because symptoms related to schizophrenia may not have begun to fully manifest in our population it is possible that our estimates of the genetic variance explained are underestimates of associations at later time points.

This paper also represents an important step in the use of PSs in admixed individuals. There has been research into particular GWAS methods in admixed populations that have had some success in improving PSs (Bitarello and Mathieson 2020; Hou et al. 2022) these analyses often only look at two-way admixture between African ancestry and European ancestry. Our analysis of our admixed population shows encouraging performance in our admixed population agnostic of the degree of admixture and the component ancestry pieces.

While the precise causal mechanisms that underlie these interactions are yet poorly understood, and many traits, especially cognitive and psychiatric traits, are strongly influenced by environmental factors, this article shows the current potential that Bayesian continuous shrinkage PS methods display higher prediction of complex traits across ancestries. If PSs are to achieve clinical validity it is important that they show equitable performance across diverse populations and within populations of individuals that do not fit categorically into continental ancestries (Lewis and Green 2021). More work must continue to be done to improve the power and validity of diverse GWAS summary statistics, as well as to develop genomic datasets in individuals of diverse ancestry.

Limitations

It must be acknowledged that PSs only account for genetic influences. Any sort of modeling based on PS should account for other known non-genetic causal factors. For many traits, genetic factors only account for a portion of the variance in trait outcomes. It is especially important to acknowledge that, for essentially all cognitive and psychiatric traits, environmental factors and gene-environment interactions will be major components in outcomes. ABCD has very rich phenotyping, but this analysis was limited by the public availability of GWAS summary statistics in non-European populations, as the majority of large public GWAS studies have been performed on European populations (Martin et al. 2019). Moreover, while ABCD is a diverse cohort, the majority of participants are of European ancestry.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10519-023-10139-w.

Acknowledgements The authors wish to thank the youth and families participating in the Adolescent Brain Cognitive Development (ABCD) Study and all ABCD staff. Data used in the preparation of this article were obtained from the Adolescent Brain Cognitive Development (ABCD) Study (https://abcdstudy.org), held in the NIMH Data Archive (NDA). This is a multisite, longitudinal study designed to recruit more than 10,000 children age 9–10 and follow them over 10 years into early adulthood. The ABCD Study is supported by the National Institutes of Health and additional federal partners under award numbers U01DA041022, U01DA041028, U01DA041048, U01DA041089, U01DA041106, U01DA041117, U01DA041120, U01DA041134, U01DA041148, U01DA041156, U01DA041174, U24DA041123, U24DA041147, U01DA041093, and U01DA041025. A full list of supporters is available at https://abcdstudy.org/federal-partners.html. A listing of participating sites and a complete listing of the study investigators can be found at https://abcdstudy.org/Consortium_Members.pdf. ABCD consortium investigators designed and implemented the study and/or provided data but did not all necessarily participate in the analysis or writing of this report. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators. The ABCD data repository grows and changes over time. The data were downloaded from the NIMH Data Archive ABCD Collection Release 2.0.1 (https://doi.org/10.15154/1504041).

Author contributions All authors have contributed equally.

Funding This work was supported by Grant R01MH122688 and RF1MH120025 funded by the National Institute for Mental Health.

Data availability Information about data availability can be found in Supplementary Table 1.

Code availability The code used in this analysis can be found on the following websites: PRScsx: https://github.com/getian107/PRScsx. PRScs: https://github.com/getian107/PRScs. PRSice2: https://choishingwan.github.io/PRS-Tutorial/prsice/#:~:text=PRSice%2D2%20is%20one%20of,the%20standard%20 C%2BT%20method. TOPMED imputation scripts: https://github.com/robloughnan/TOPMED_Imputation_Scripts. PC-AIR scripts: https://github.com/robloughnan/ABCD_GeneticPCs_and_Relatedness. Plink 2.0: https://www.cog-genomics.org/plink/2.0/. SNPweights: https://mybiosoftware.com/tag/

snpweights. Gamm4: https://cran.r-project.org/web/packages/gamm4/index.html.

Declarations

Competing interests Jonathan Ahern, Wesley Thompson, Chun Chieh Fan and Robert Loughnan declare no competing interests.

Ethics Approval ABCD data is collected by the ABCD consortium with appropriate IRB approval and informed consent.

References

Achenbach TM, Rescorla LA (2004) The Achenbach System of empirically based Assessment (ASEBA) for ages 1.5 to 18 years. In: The use of psychological testing for treatment planning and outcomes assessment, 3rd edn. Routledge, London

Adhikari K, Mendoza-Revilla J, Chacón-Duque JC, Fuentes-Guajardo M, Ruiz-Linares A (2016) Admixture in Latin America. Curr Opin Genet Dev 41:106–114. https://doi.org/10.1016/j.gde.2016.09.003

Adhikari K, Chacón-Duque JC, Mendoza-Revilla J, Fuentes-Guajardo M, Ruiz-Linares A (2017) The genetic diversity of the Americas. Annu Rev Genom Hum Genet 18(1):277–296. https://doi.org/10.1146/annurev-genom-083115-022331

Albores-Gallo L, Lara-Muñoz C, Esperón-Vargas C, Zetina JAC, Soriano AMP, Colin GV (2007) Validity and reliability of the CBCL/6–18. Includes DSM scales. Actas Españolas de Psiquiatría 35:393–399

Ashenhurst JR, Sazonova OV, Svrchek O, Detweiler S, Kita R, Babalola L, McIntyre M, Aslibekyan S, Fontanillas P, Shringarpure S, 23andMe Research Team, Pollard JD, Koelsch BL (2022) A polygenic score for type 2 diabetes improves risk stratification beyond current clinical screening factors in an ancestrally diverse sample. Front Genet. https://doi.org/10.3389/fgene.2022.871260

Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, Chakravarti A, Clark AG, Donnelly P, Eichler EE, Flicek P, Gabriel SB, Gibbs RA, Green ED, Hurles ME, Knoppers BM, Korbel JO, Lander ES, Lee C et al (2015) A global reference for human genetic variation. Nature. https://doi.org/10.1038/nature15393

Baurley JW, Edlund CK, Pardamean CI, Conti DV, Bergen AW (2016) Smokescreen: a targeted genotyping array for addiction research. BMC Genomics 17(1):145. https://doi.org/10.1186/s12864-016-2495-7

Bitarello BD, Mathieson I (2020) Polygenic scores for height in admixed populations. G3 10(11):4027–4036. https://doi.org/10.1534/g3.120.401658

Campbell MC, Tishkoff SA (2008) AFRICAN GENETIC DIVERSITY: implications for human demographic history, modern human origins, and complex disease mapping. Annu Rev Genom Hum Genet 9:403–433. https://doi.org/10.1146/annurev.genom.9.081307.164258

Chen C-Y, Pollack S, Hunter DJ, Hirschhorn JN, Kraft P, Price AL (2013) Improved ancestry inference using weights from external reference panels. Bioinformatics 29(11):1399–1406. https://doi.org/10.1093/bioinformatics/btt144

Choi SW, O'Reilly PF (2019) PRSice-2: polygenic risk score software for biobank-scale data. GigaScience 8(7):giz082. https://doi.org/10.1093/gigascience/giz082

Choi SW, Mak TS-H, O'Reilly PF (2020) Tutorial: a guide to performing polygenic risk score analyses. Nat Protoc. https://doi.org/10.1038/s41596-020-0353-1

Conomos MP, Miller M, Thornton T (2015) Robust inference of population structure for ancestry prediction and correction of stratification in the presence of relatedness. Genet Epidemiol 39(4):276–293. https://doi.org/10.1002/gepi.21896

de la Peña FR, Villavicencio LR, Palacio JD, Félix FJ, Larraguibel M, Viola L, Ortiz S, Rosetti M, Abadi A, Montiel C, Mayer PA, Fernández S, Jaimes A, Feria M, Sosa L, Rodríguez A, Zavaleta P, Uribe D, Galicia F eta l (2018) Validity and reliability of the Kiddie-Schedule for Affective Disorders and Schizophrenia-Present and Lifetime Version DSM-5 (K-SADS-PL-5) Spanish version. BMC Psychiatry 18:193. https://doi.org/10.1186/s12888-018-1773-0

Dun Y, Li Q-R, Yu H, Bai Y, Song Z, Lei C, Li H-H, Gong J, Mo Y, Li Y, Pei X-Y, Yuan J, Li N, Xu C-Y, Lai Q-Y, Fu Z, Zhang K-F, Song J-Y, Kang S-M et al (2022) Reliability and validity of the Chinese version of the Kiddie-Schedule for Affective Disorders and Schizophrenia-Present and Lifetime Version DSM-5 (K-SADS-PL-C DSM-5). J Affect Disord 317:72–78. https://doi.org/10.1016/j.jad.2022.08.062

Dutra L, Campbell L, Westen D (2004) Quantifying clinical judgment in the assessment of adolescent psychopathology: reliability, validity, and factor structure of the Child Behavior Checklist for clinician report. J Clin Psychol 60(1):65–85. https://doi.org/10.1002/jclp.10234

Fernández-Rhodes L, Graff M, Buchanan VL, Justice AE, Highland HM, Guo X, Zhu W, Chen H-H, Young KL, Adhikari K, Palmer ND, Below JE, Bradfield J, Pereira AC, Glover L, Kim D, Lilly AG, Shrestha P, Thomas AG et al (2022) Ancestral diversity improves discovery and fine-mapping of genetic loci for anthropometric traits—the Hispanic/Latino Anthropometry Consortium. Hum Genet Genomics Adv 3(2):100099. https://doi.org/10.1016/j.xhgg.2022.100099

Fuller R, Nopoulos P, Arndt S, O'Leary D, Ho B-C, Andreasen NC (2002) Longitudinal assessment of premorbid cognitive functioning in patients with schizophrenia through examination of standardized scholastic test performance. Am J Psychiatry 159(7):1183–1189. https://doi.org/10.1176/appi.ajp.159.7.1183

Garavan H, Bartsch H, Conway K, Decastro A, Goldstein RZ, Heeringa S, Jernigan T, Potter A, Thompson W, Zahs D (2018) Recruiting the ABCD sample: design considerations and procedures. Dev Cogn Neurosci 32:16–22. https://doi.org/10.1016/j.dcn.2018.04.004

Ge T, Chen C-Y, Ni Y, Feng Y-CA, Smoller JW (2019) Polygenic prediction via Bayesian regression and continuous shrinkage priors. Nat Commun. https://doi.org/10.1038/s41467-019-09718-5

Ge T, Irvin MR, Patki A, Srinivasasainagendra V, Lin Y-F, Tiwari HK, Armstrong ND, Benoit B, Chen C-Y, Choi KW, Cimino JJ, Davis BH, Dikilitas O, Etheridge B, Feng Y-CA, Gainer V, Huang H, Jarvik GP, Kachulis C et al (2022) Development and validation of a trans-ancestry polygenic risk score for type 2 diabetes in diverse populations. Genome Med 14(1):70. https://doi.org/10.1186/s13073-022-01074-2

Giannakopoulou O, Lin K, Meng X, Su M-H, Kuo P-H, Peterson RE, Awasthi S, Moscati A, Coleman JRI, Bass N, Millwood IY, Chen Y, Chen Z, Chen H-C, Lu M-L, Huang M-C, Chen C-H, Stahl EA, Loos RJF, Kuchenbaecker K (2021) The genetic architecture of depression in individuals of east Asian ancestry. JAMA Psychiatry 78(11):1–12. https://doi.org/10.1001/jamapsychiatry.2021.2099

Goodwin RD, Dierker LC, Wu M, Galea S, Hoven CW, Weinberger AH (2022) Trends in U.S. depression prevalence from 2015 to 2020: the widening treatment gap. Am J Prev Med 63(5):726–733. https://doi.org/10.1016/j.amepre.2022.05.014

Hartini S, Hapsara S, Herini SE, Takada S (2015) Verifying the Indonesian version of the Child Behavior Checklist. Pediatr Int 57(5):936–941. https://doi.org/10.1111/ped.12669

Hou K, Ding Y, Xu Z, Wu Y, Bhattacharya A, Mester R, Belbin G, Conti D, Darst BF, Fornage M, Gignoux C, Guo X, Haiman C, Kenny E, Kim M, Kooperberg C, Lange L, Manichaikul A, North KE et al (2022) Causal effects on complex traits are similar across segments of different continental ancestries within admixed individuals. medRxiv. https://doi.org/10.1101/2022.08.16.22278868

Howard DM, Adams MJ, Clarke T-K, Hafferty JD, Gibson J, Shirali M, Coleman JRI, Hagenaars SP, Ward J, Wigmore EM, Alloza C, Shen X, Barbu MC, Xu EY, Whalley HC, Marioni RE, Porteous DJ, Davies G, Deary IJ et al (2019) Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. Nat Neurosci. https://doi.org/10.1038/s41593-018-0326-7

Inouye M, Abraham G, Nelson CP, Wood AM, Sweeting MJ, Dudbridge F, Lai FY, Kaptoge S, Brozynska M, Wang T, Ye S, Webb TR, Rutter MK, Tzoulaki I, Patel RS, Loos RJF, Keavney B, Hemingway H, Thompson J et al (2018) Genomic risk prediction of coronary artery disease in 480,000 adults: implications for primary Prevention. J Am Coll Cardiol 72(16):1883–1893. https://doi.org/10.1016/j.jacc.2018.07.079

Jia G, Lu Y, Wen W, Long J, Liu Y, Tao R, Li B, Denny JC, Shu X-O, Zheng W (2020) Evaluating the utility of polygenic risk scores in identifying high-risk individuals for eight common cancers. JNCI Cancer Spectr 4(3):pkaa021. https://doi.org/10.1093/jncics/pkaa021

Jones HJ, Stergiakouli E, Tansey KE, Hubbard L, Heron J, Cannon M, Holmans P, Lewis G, Linden DEJ, Jones PB, Davey Smith G, O'Donovan MC, Owen MJ, Walters JT, Zammit S (2016) Phenotypic manifestation of genetic risk for schizophrenia during adolescence in the general population. JAMA Psychiatry 73(3):221–228. https://doi.org/10.1001/jamapsychiatry.2015.3058

Kachuri L, Graff RE, Smith-Byrne K, Meyers TJ, Rashkin SR, Ziv E, Witte JS, Johansson M (2020) Pan-cancer analysis demonstrates that integrating polygenic risk scores with modifiable risk factors improves risk prediction. Nat Commun. https://doi.org/10.1038/s41467-020-19600-4

Karunamuni RA, Huynh-Le M-P, Fan CC, Eeles RA, Easton DF, Kote-Jarai Z, Al Olama A, Garcia ABenlloch, Muir S, Gronberg K, Wiklund H, Aly F, Schleutker M, Sipeky J, Tammela C, Nordestgaard TLJ, Key BG, Travis TJ, Neal RC, Seibert DE, T. M (2020) The effect of sample size on polygenic hazard models for prostate cancer. Eur J Hum Genet 28(10):1467–1475. https://doi.org/10.1038/s41431-020-0664-2

Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, Williamson D, Ryan N, Version L, K-SADS-PL (1997) Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): initial reliability and validity data. J Am Acad Child Adolesc Psychiatry 36(7):980–988. https://doi.org/10.1097/00004583-199707000-00021

Kim YS, Cheon KA, Kim BN, Chang SA, Yoo HJ, Kim JW, Cho SC, Seo DH, Bae MO, So YK, Noh JS, Koh YJ, McBurnett K, Leventhal B (2004) The reliability and validity of Kiddie-Schedule for Affective Disorders and Schizophrenia-Present and Lifetime Version-Korean version (K-SADS-PL-K). Yonsei Med J 45(1):81–89

Klarin D, Natarajan P (2022) Clinical utility of polygenic risk scores for coronary artery disease. Nat Rev Cardiol. https://doi.org/10.1038/s41569-021-00638-w

Leung PWL, Kwong SL, Tang CP, Ho TP, Hung SF, Lee CC, Hong SL, Chiu CM, Liu WS (2006) Test–retest reliability and criterion validity of the Chinese version of CBCL, TRF, and YSR. J Child Psychol Psychiatry 47(9):970–973. https://doi.org/10.1111/j.1469-7610.2005.01570.x

Lewis ACF, Green RC (2021) Polygenic risk scores in the clinic: new perspectives needed on familiar ethical issues. Genome Med 13(1):14. https://doi.org/10.1186/s13073-021-00829-7

Marees AT, de Kluiver H, Stringer S, Vorspan F, Curis E, Marie-Claire C, Derks EM (2018) A tutorial on conducting genome-wide association studies: quality control and statistical analysis. Int J Methods Psychiatr Res 27(2):e1608. https://doi.org/10.1002/mpr.1608

Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ (2019) Clinical use of current polygenic risk scores may exacerbate health disparities. Nat Genet. https://doi.org/10.1038/s41588-019-0379-x

Matise TC, Ambite JL, Buyske S, Carlson CS, Cole SA, Crawford DC, Haiman CA, Heiss G, Kooperberg C, Marchand LL, Manolio TA, North KE, Peters U, Ritchie MD, Hindorff LA, Haines JL, for the PAGE Study (2011) The Next PAGE in understanding complex traits: design for the analysis of Population Architecture Using Genetics and Epidemiology (PAGE) study. Am J Epidemiol 174(7):849–859. https://doi.org/10.1093/aje/kwr160

Mitchell AJ, Rao S, Vaze A (2011) International comparison of clinicians' ability to identify depression in primary care: meta-analysis and meta-regression of predictors. Br J Gen Pract 61(583):e72–e80. https://doi.org/10.3399/bjgp11X556227

Nagelkerke NJD (1991) A note on a general definition of the coefficient of determination. Biometrika 78(3):691–692. https://doi.org/10.1093/biomet/78.3.691

Ng MCY, Graff M, Lu Y, Justice AE, Mudgal P, Liu C-T, Young K, Yanek LR, Feitosa MF, Wojczynski MK, Rand K, Brody JA, Cade BE, Dimitrov L, Duan Q, Guo X, Lange LA, Nalls MA, Okut H et al (2017) Discovery and fine-mapping of adiposity loci using high density imputation of genome-wide association studies in individuals of African ancestry: African Ancestry Anthropometry Genetics Consortium. PLoS Genet 13(4):e1006719. https://doi.org/10.1371/journal.pgen.1006719

Nishiyama T, Sumi S, Watanabe H, Suzuki F, Kuru Y, Shiino T, Kimura T, Wang C, Lin Y, Ichiyanagi M, Hirai K (2020) The Kiddie Schedule for Affective Disorders and Schizophrenia Present and Lifetime Version (K-SADS-PL) for DSM-5: a validation for neurodevelopmental disorders in Japanese outpatients. Compr Psychiatry 96:152148. https://doi.org/10.1016/j.comppsych.2019.152148

Peterson RE, Kuchenbaecker K, Walters RK, Chen C-Y, Popejoy AB, Periyasamy S, Lam M, Iyegbe C, Strawbridge RJ, Brick L, Carey CE, Martin AR, Meyers JL, Su J, Chen J, Edwards AC, Kalungi A, Koen N, Majara L et al (2019) Genome-wide association studies in ancestrally diverse populations: opportunities, methods, pitfalls, and recommendations. Cell 179(3):589–603. https://doi.org/10.1016/j.cell.2019.08.051

Privé F, Arbel J, Vilhjálmsson BJ (2020) LDpred2: better, faster, stronger. Bioinformatics 36(22–23):5424–5431. https://doi.org/10.1093/bioinformatics/btaa1029

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, Sham PC (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81(3):559–575

R Core Team (2022) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org/

Reich D, Patterson N, Campbell D, Tandon A, Mazieres S, Ray N, Parra MV, Rojas W, Duque C, Mesa N, García LF, Triana O, Blair S, Maestre A, Dib JC, Bravi CM, Bailliet G, Corach D, Hünemeier T et al (2012) Reconstructing native American population history. Nature. https://doi.org/10.1038/nature11258

Ruan Y, Lin Y-F, Feng Y-CA, Chen C-Y, Lam M, Guo Z, He L, Sawa A, Martin AR, Qin S, Huang H, Ge T (2022) Improving polygenic

prediction in ancestrally diverse populations. Nat Genet. https://doi.org/10.1038/s41588-022-01054-7
Saito T, Ikeda M, Terao C, Ashizawa T, Miyata M, Tanaka S, Kanazawa T, Kato T, Kishi T, Iwata N (2022) Differential genetic correlations across major psychiatric disorders between eastern and western countries. J Neuropsychiatry Clin Neurosci 77(2):118–119. https://doi.org/10.1111/pcn.13498
Sanz-de-Galdeano A, Terskaya A, Upegui A (2020) Association of a genetic risk score with BMI along the life-cycle: evidence from several US cohorts. PLoS ONE 15(9):e0239067. https://doi.org/10.1371/journal.pone.0239067
Schultz LM, Merikangas AK, Ruparel K, Jacquemont S, Glahn DC, Gur RE, Barzilay R, Almasy L (2022) Stability of polygenic scores across discovery genome-wide association studies. Hum Genet Genomics Adv 3(2):100091. https://doi.org/10.1016/j.xhgg.2022.100091
Shahrivar Z, Kousha M, Moallemi S, Tehrani-Doost M, Alaghband-Rad J (2010) The reliability and validity of kiddie-schedule for affective disorders and schizophrenia—present and life-time version—Persian version. Child Adolesc Mental Health 15(2):97–102. https://doi.org/10.1111/j.1475-3588.2008.00518.x
Silventoinen K, Pietiläinen KH, Tynelius P, Sørensen TIA, Kaprio J, Rasmussen F (2008) Genetic regulation of growth from birth to 18 years of age: the Swedish young male twins study. Am J Hum Biol 20(3):292–298. https://doi.org/10.1002/ajhb.20717
Sugrue LP, Desikan RS (2019) What are polygenic scores and why are they important? JAMA 321(18):1820–1821. https://doi.org/10.1001/jama.2019.3893
Taliun D, Harris DN, Kessler MD, Carlson J, Szpiech ZA, Torres R, Taliun SAG, Corvelo A, Gogarten SM, Kang HM, Pitsillides AN, LeFaive J, Lee S, Tian X, Browning BL, Das S, Emde A-K, Clarke WE, Loesch DP et al (2021) Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. Nature. https://doi.org/10.1038/s41586-021-03205-y
THE BRAINSTORM CONSORTIUM, Anttila V, Bulik-Sullivan B, Finucane HK, Walters RK, Bras J, Duncan L, Escott-Price V, Falcone GJ, Gormley P, Malik R, Patsopoulos NA, Ripke S, Wei Z, Yu D, Lee PH, Turley P, Grenier-Boley B, Chouraki V et al (2018) Analysis of shared heritability in common disorders of the brain. Science 360(6395):eaap8757. https://doi.org/10.1126/science.aap8757
Trubetskoy V, Pardiñas AF, Qi T, Panagiotaropoulou G, Awasthi S, Bigdeli TB, Bryois J, Chen C-Y, Dennison CA, Hall LS, Lam M, Watanabe K, Frei O, Ge T, Harwood JC, Koopmans F, Magnusson S, Richards AL, Sidorenko J et al (2022) Mapping genomic loci implicates genes and synaptic biology in schizophrenia. Nature. https://doi.org/10.1038/s41586-022-04434-5
Uban KA, Horton MK, Jacobus J, Heyser C, Thompson WK, Tapert SF, Madden PAF, Sowell ER, Adolescent Brain Cognitive Development Study (2018) Biospecimens and the ABCD study: rationale, methods of collection, measurement and early data. Dev Cogn Neurosci 32:97–106. https://doi.org/10.1016/j.dcn.2018.03.005
Visscher PM, Yengo L, Cox NJ, Wray NR (2021) Discovery and implications of polygenicity of common diseases. Science 373(6562):1468–1473. https://doi.org/10.1126/science.abi8206
Walker E, Kestler L, Bollini A, Hochman KM (2004) Schizophrenia: etiology and course. Ann Rev Psychol 55(1):401–430. https://doi.org/10.1146/annurev.psych.55.090902.141950
Wang Y, Guo J, Ni G, Yang J, Visscher PM, Yengo L (2020) Theoretical and empirical quantification of the accuracy of polygenic scores in ancestry divergent populations. Nat Commun. https://doi.org/10.1038/s41467-020-17719-y
Wood S, Scheipl F (2022) gamm4: generalized additive mixed models using 'mgcv' and 'lme4'. Version 0.2-6 R package. https://cran.r-project.org/web/packages/gamm4/gamm4.pdf
Woolway GE, Smart SE, Lynham AJ, Lloyd JL, Owen MJ, Jones IR, Walters JTR, Legge SE (2022) Schizophrenia polygenic risk and experiences of childhood adversity: a systematic review and meta-analysis. Schizophr Bull 48(5):967–980. https://doi.org/10.1093/schbul/sbac049
Wu T, Liu Z, Mak TSH, Sham PC (2022) Polygenic power calculator: statistical power and polygenic prediction accuracy of genome-wide association studies of complex traits. Front Genet. https://doi.org/10.3389/fgene.2022.989639
Yang J, Lee SH, Goddard ME, Visscher PM (2011) GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 88(1):76–82. https://doi.org/10.1016/j.ajhg.2010.11.011
Yengo L, Vedantam S, Marouli E, Sidorenko J, Bartell E, Sakaue S, Graff M, Eliasen AU, Jiang Y, Raghavan S, Miao J, Arias JD, Graham SE, Mukamel RE, Spracklen CN, Yin X, Chen S-H, Ferreira T, Highland HH et al (2022) A saturated map of common genetic variants associated with human height. Nature. https://doi.org/10.1038/s41586-022-05275-y