Title
Genetic and clinical analyses of psychosis spectrum symptoms in a large multiethnic youth cohort reveal significant link with ADHD.

Permalink
https://escholarship.org/uc/item/46n1c2q9

Journal
Translational psychiatry, 11(1)

ISSN
2158-3188

Authors
Olde Loohuis, Loes M
Mennigen, Eva
Ori, Anil PS
et al.

Publication Date
2021

DOI
10.1038/s41398-021-01203-2

Peer reviewed

Genetic and clinical analyses of psychosis spectrum symptoms in a large multiethnic youth cohort reveal significant link with ADHD

Loes M. Olde Loohuis [1], Eva Mennigen[1,2], Anil P. S. Ori[1], Diana Perkins[3], Elise Robinson[4,5,6], Jean Addington [7], Kristin S. Cadenhead[8], Barbara A. Cornblatt[9], Daniel H. Mathalon[10], Thomas H. McGlashan[11], Larry J. Seidman[12], Matcheri S. Keshavan[12], William S. Stone[12], Ming T. Tsuang [8], Elaine F. Walker[13], Scott W. Woods[11], Tyrone D. Cannon[14], Ruben C. Gur[15], Raquel E. Gur[15], Carrie E. Bearden [1,16] and Roel A. Ophoff [1,17,18]

Abstract

Psychotic symptoms are not only an important feature of severe neuropsychiatric disorders, but are also common in the general population, especially in youth. The genetic etiology of psychosis symptoms in youth remains poorly understood. To characterize genetic risk for psychosis spectrum symptoms (PS), we leverage a community-based multiethnic sample of children and adolescents aged 8–22 years, the Philadelphia Neurodevelopmental Cohort ($n =$ 7225, 20% PS). Using an elastic net regression model, we aim to classify PS status using polygenic scores (PGS) based on a range of heritable psychiatric and brain-related traits in a multi-PGS model. We also perform univariate PGS associations and evaluate age-specific effects. The multi-PGS analyses do not improve prediction of PS status over univariate models, but reveal that the attention deficit hyperactivity disorder (ADHD) PGS is robustly and uniquely associated with PS (OR 1.12 (1.05, 1.18) $P = 0.0003$). This association is driven by subjects of European ancestry (OR = 1.23 (1.14, 1.34), $P = 4.15 \times 10^{-7}$) but is not observed in African American subjects ($P = 0.65$). We find a significant interaction of ADHD PGS with age ($P = 0.01$), with a stronger association in younger children. The association is independent of phenotypic overlap between ADHD and PS, not indirectly driven by substance use or childhood trauma, and appears to be specific to PS rather than reflecting general psychopathology in youth. In an independent sample, we replicate an increased ADHD PGS in 328 youth at clinical high risk for psychosis, compared to 216 unaffected controls (OR 1.06, CI(1.01, 1.11), $P = 0.02$). Our findings suggest that PS in youth may reflect a different genetic etiology than psychotic symptoms in adulthood, one more akin to ADHD, and shed light on how genetic risk can be investigated across early disease trajectories.

Introduction

Psychotic symptoms, such as delusions and hallucinations, are an important feature of severe psychiatric disorders such as schizophrenia and bipolar disorder. They are, however, also common in the general population, and occur in ~5–10%[1] of adults; a prevalence much higher than that of clinical diagnoses of schizophrenia and bipolar disorder (about 0.5–3% each[2,3]) In children and adolescents, the prevalence of psychotic symptoms/psychotic-like experiences is even higher, as high as 20%[4]. Youth experiencing psychotic symptoms typically exhibit a multitude of other comorbid symptoms, such as increased mood, anxiety, and attention deficit hyperactivity disorder (ADHD) symptoms, as well as increased

Correspondence: Loes M. Olde Loohuis (LOldeLoohuis@mednet.ucla.edu) or Roel A. Ophoff (ophoff@ucla.edu)
[1]Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, CA, USA
[2]Department of Psychiatry and Psychotherapy, University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany
Full list of author information is available at the end of the article

substance use and impairments in global functioning[5–7]. While subclinical psychopathology poses a risk for later development of overt psychiatric illnesses[5,8–11], only a minority of youth reporting psychotic symptoms will convert to full-blown psychotic disorders.

With recent progress in psychiatric genetics, psychotic disorders are becoming well-characterized genetically[12–14]. In particular, the landmark genome-wide association study (GWAS) of schizophrenia provides aggregate risk conferred by variants identified, polygenic scores (PGS), which explain about 7–10% of variance in case-control status[12,15]. In individuals with bipolar disorder, both genetic risk for schizophrenia as well as for bipolar disorder have been associated with psychotic symptoms[16,17].

In the general population, however, the genetic etiology of psychotic symptoms across development is still largely unknown. The heritability of psychotic experiences has been estimated between 30 and 50% from twin studies[18,19] with the proportion of genetic variance explained by common variants (SNP-heritability) of 3–17% in adolescents[18,20]. Adults with psychotic symptoms harbor increased genetic liability for a broad spectrum of psychiatric disorders, including schizophrenia and other neuropsychiatric disorders[21]. While in adolescents some evidence suggests increased genetic risk for schizophrenia (and major depressive disorder) for specific features of psychosis[18], the reported effect sizes are very small, and these effects not very robust[22]. In pre-adolescent youth, the relationship between genetic risk for psychiatric traits and psychotic symptoms has not yet been explored. The genetic characterization of psychosis spectrum symptoms across development may increase our understanding of their etiological and pathological significance.

To study genetic risk for psychosis symptoms in a population sample of youth, we leverage a large well-characterized community-based sample of youth aged 8–22 years, the Philadelphia Neurodevelopmental Cohort (PNC, $n = 9498$ in total). The PNC is a multiethnic cohort, with the largest proportion of individuals of European (66%) and African American (26%) ancestry. In this community-based sample that is not ascertained for neuropsychiatric disorders, about 20% of the youth are classified as having psychotic spectrum symptoms (PS). In the PNC, having PS has been associated with structural and functional brain alterations[23–25], qualitatively similar to those present in overt psychotic disorders, as well as cognitive deficits[26]. These findings underscore both the increased vulnerability in youth experiencing psychosis spectrum symptoms and the importance of investigating psychosis risk as a dynamic developmental process[7].

In this well-characterized sample, we explore the genetic architecture of PS based on common variant liability for psychiatric illness. Specifically, we aim to understand whether PGS for psychiatric disorders can be used to classify psychotic symptoms in youth. To do so, we adopt a recently developed multi-PGS approach[27]. The method combines multiple summary statistics from different GWAS into a single predictive model, thus potentially increasing classification power. Given evidence suggesting PS increases risk for broader psychopathology[11,21] and the wide spectrum of genetic inter-correlations in adults with psychotic experiences[21], we include a range of heritable psychiatric and brain traits in our analyses, an analytic approach that has not been previously explored in this type of cohort. To shed light on, and generate hypotheses about, the nature of psychotic symptoms across development in an unbiased manner, we also perform univariate associations for all traits included in our multivariate model and assess phenotypic overlap between traits.

We hypothesize that developmental changes in the expression of psychotic symptoms will be reflected via age-specific genetic etiology, with the genetic architecture underlying psychosis spectrum symptoms in older youth being more similar to psychotic disorders. We evaluate this hypothesis by testing whether the observed correlations and the interplay with phenotypic overlap changes across the development.

Methods

Cohort description

Data were obtained from the publicly available Philadelphia Neurodevelopmental Cohort (PNC, 1st release, phs000607.v1.p1, #7147) via the Database of Genotypes and Phenotypes (dbGap) platform. The PNC is a community-based sample consisting of 9498 genotyped youth (ages 8–22 years) who participated in clinical and neurocognitive assessment, with a subsample undergoing MRI, after providing written informed consent or assent with parental consent (youth under 18 years old). Psychiatric symptomatology was assessed using the GOAS-SESS interview[28] covering broad domains of psychopathology including mood, anxiety, phobias, psychosis, and externalizing behaviour[29]. Psychotic symptoms were specifically assessed with questions from the Kiddie Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS)[30], the Structured Interview for Prodromal Syndromes (SIPS)[31], and the PRIME Screen Revised (P-SR).

Definition of psychosis spectrum

Criteria to establish a group of individuals that experiences psychosis spectrum symptoms were defined as in prior PNC studies[7,24,32]. These criteria consider lifetime occurrence of positive psychotic symptoms such as hallucinations and persecutory thinking, negative/disorganized symptoms such as flattened affect, as well as age-appropriateness of these symptoms (Supplementary

Methods). Previous publications have described the clinical and functional significance of these criteria[7,24,32].

Since MRI data are available for only a subset of the PNC sample that is too small for genetic analyses, we do not include these in our analyses. Other phenotypes, such as PS domain scores of positive (PRIME) and negative/disorganized (SOPS) as well as symptoms of ADHD and additional phenotypes for follow-up analyses are described in the Supplementary Material.

Genotyping QC and imputation

Genotyping QC, imputation and selection of individuals of European and African ancestry are described in detail in the Supplementary Materials. In brief, imputation followed the standard Ricopili pipeline (see URLS) and best-guess genotypes of well-imputed variants (INFO > 0.8) were selected for further analysis. The imputed dataset included 7774 subjects, 7764 with phenotyping data. Individuals were assigned ancestry group based on estimates from ADMIXTURE[33] and related subjects (specifically 477 siblings) were removed within ancestry groups (Supplementary Note and Figs. S1–S3). After these final filtering steps, the total sample size is 7225, with a total of 1,937,561 included SNPs. We identified two subcohorts: one including individuals of European ancestry (EA, $n = 4852$) and one African American ancestry group (AA, $n = 1802$).

GWAS summary statistics

We selected GWAS summary statistics from LD hub[34], a centralized repository for summary statistics (accessed June 2018). Specifically, we included 23 GWAS summary statistics of psychiatric, brain traits and personality traits, that were either publicly available or obtained via personal correspondence. From these, a total of 12 had a linkage disequilibrium (LD) score heritability z-score >5, indicating good statistical power[27](which is a function of variance explained and sample size) and complete GWAS information available. If available, we replaced summary statistics in LD hub with more recent or powerful GWAS: For the Psychiatric Genomics Consortium (PGC), we replaced MDD, BIP, ADHD. Finally, we also included the 23andMe traits for Morningness and self-reported depression. For more details on the included GWAS see Supplementary Table 1.

After filtering, the 13 traits included in our analyses include psychiatric disorders (PGC GWAS for: ADHD[35], Autism[36], bipolar disorder[13], schizophrenia[12], cross-disorder, a joint analyses of severe mental illness[37], major depression[38]; other psychiatric GWAS for: self-reported depression from 23andMe[39] major depressive disorder from CONVERGE[40]) brain traits (ENIGMA Caudate volume and Putamen volume[41]) and behavioural traits (Morningness 23andMe[42], Neuroticism[43] and subjective well-being[43](SWB)).

Polygenic scoring

Polygenic scoring was performed in a standard clumping and thresholding fashion, based on a *p*-value threshold of 0.05 (see Supplementary Methods for details). Specifically, for analyses only involving the EA (or AA) cohort, we included ancestry-specific principal components (PCs) after exclusion of related samples. The standardized residuals were used for follow-up analyses.

In addition to the conventional approach of thresholding and clumping, which can lead to loss of information, especially in cases where the ancestry of the GWAS sample diverges from the target sample, we also computed polygenic scores using LDPRED for follow-up analysis[44]. As recommended, we used the target sample genotype data as the LD reference panel, performing scoring separately in the EA and AA ancestry samples. We used standard settings with an LD radius of 500 SNPs.

Multi-score PGS analyses

As has been successfully implemented previously[27], we used elastic net regularized regression to predict outcomes by selecting predictors and estimating their contribution to the prediction. Elastic net uses a linear combination of two regularization techniques, L2 regularization (used in ridge regression) and L1 regularization (used in LASSO), and has been shown to work particularly well in case of correlated predictors, as is expected in the context of highly correlated genetic traits[45].

Elastic net regularized regression employs two hyper-parameters, alpha and lambda. To achieve optimized balance between variance explained and minimum bias, we fit models to tune over both alpha and lambda parameter values in repeated 10-fold cross-validation, and used the minimal lambda for the prediction model. As a performance measure, we use area under the ROC curve (AUC). Models were trained on a random subset of 70% of the data and weights of the selected variables subsequently used to test their cumulative discriminatory power in predicting psychosis status in the remaining 30%. We performed model fit both in the entire sample as well as within EA and AA separately. To obtain an estimate of the robustness and range of the selected parameters we performed 1000 repetitions of the procedure. In addition, we also generated models for 1000 permutations of the phenotype (within the All, EA and AA cohorts). Comparing our models to those derived from permutation, we adopt the conservative approach to compare the mean of the 1000 repetitions in the actual sample to the distribution of the permutations.

Univariate PGS analyses

To estimate the effect of each PGS individually, we fit a series of logistic regression models for each of the

corrected polygenic scores including Age, Age[2] and Sex as covariates. Effect sizes are reported as odds ratio (OR) relative to one standard deviation increase in PGS. To account for multiple testing, we applied a Bonferroni correction, dividing the p-value by the total number of tests ($14 \times 3 = 42$ tests yields a *p*-value cutoff of 0.001). Interactions were tested by including the interaction term in the full model.

Analyses of phenotyping data

Symptom overlap was tested using Fisher's exact test (pairs of binary traits), Wilcoxon Rank-sum test (binary vs quantitative) and Spearman rank correlation (pairs of quantitative traits).

Results

Ancestry

After quality control (see Supplementary methods), the PNC cohort consists of a total of $n = 7225$ youth ages 8–22 with both phenotypic and imputed data available with two ancestry groups: European ancestry (EA, $n = 4852$) and African American ancestry (AA, $n = 1802$). From the total sample, 1369 youth (19%) are classified as having psychosis spectrum symptoms (PS). Figure 1 displays the demographics of the cohorts. Common variant heritability of the PS phenotype was estimated using GCTA, but due to lack of power, we were unable to obtain an accurate estimate (SNP-$h^2 = 0.11$, se $= 0.21$, $P = 0.3$, Supplementary Methods).

Multi-PGS prediction

Our multi-PGS models classified PS marginally better than chance, both in the whole sample and in the EA cohort alone (average AUC All $= 0.53$ (sd 0.01) $P = 0.009$, average AUC EA $= 0.55$ (sd 0.02) $P < 0.001$). Within the AA cohort, however, the multivariate prediction was not different from chance (average AUC AA $= 0.51$ (sd 0.02) $P = 0.35$). In EA, the highest weight was consistently assigned to the PGS of ADHD, with an average standardized coefficient of 0.09 (sem 0.0007) corresponding to an OR of 1.10 (Fig. 2), after correcting for all other selected variables. This effect was driven by the EA cohort (OR = 1.18; Fig. 2 and Supplementary Fig. 1) but absent in AA (OR = 1.00). As expected, permutation of case-control status separately within the ancestry groups did not highlight any single trait (Fig. S4).

Univariate association

In-line with the multivariate model, univariate logistic regression yielded a modest but significant association between PS and ADHD liability (OR 1.12 (1.05, 1.18) $P = 0.0003$, Fig. 3, Table 1, Fig. S5). This effect was driven by youth with EA ancestry (OR = 1.23 (1.14, 1.34), $P = 4.15 \times 10^{-7}$), and not observed in the AA cohort (OR =

0.98 (0.88,1.08) $P = 0.65$). Genetic liability for all other neuropsychiatric traits, including schizophrenia, was not associated with PS in either ancestry group. Since the AA cohort is of a different ancestry than the majority of GWAS cohorts, with different allele frequencies, LD patterns, and effect sizes[46,47], we also performed PGS computation using LDPRED[44], a method that explicitly models LD. No other traits were significantly associated using this method either. To test whether the multivariate predictor outperformed ADHD PGS alone, we performed the same classification procedure including only the ADHD GWAS. That is, we estimated the regression coefficient on a training dataset including 70% of data and tested on the subset that was left out. Our multivariate model did not outperform a univariate predictor. In fact, in the case of EA, the univariate predictor even performed slightly better ($P = 0.004$, Wilcoxon rank-sum test, Fig. S6), indicating the PGS for additional traits introduce more noise than signal to the classifier.

Since the EA cohort drives the observed genetic association between PS and ADHD, we next explored the nature and robustness of the association in the EA cohort only.

Comparing the extremes of the distribution, the top decile of PGS was associated with a nearly 2.5-fold increased risk for being assigned to the PS group compared to the lowest decile (OR = 2.43 (1.71, 3.51) $P = 1.25e\text{-}06$, Fig. 3b). Globally, within EA the ADHD PGS explains about 1% of variation in case-control status, as measured by reduction in Nagelkerke R2. Moreover, the association is (i) robust across *P*-value thresholds (ii) not driven by subtle population stratification within the EA cohort, (ii) extends to the European-only version of the ADHD GWAS (see Supplementary Results and Figs. S7, S8, for details).
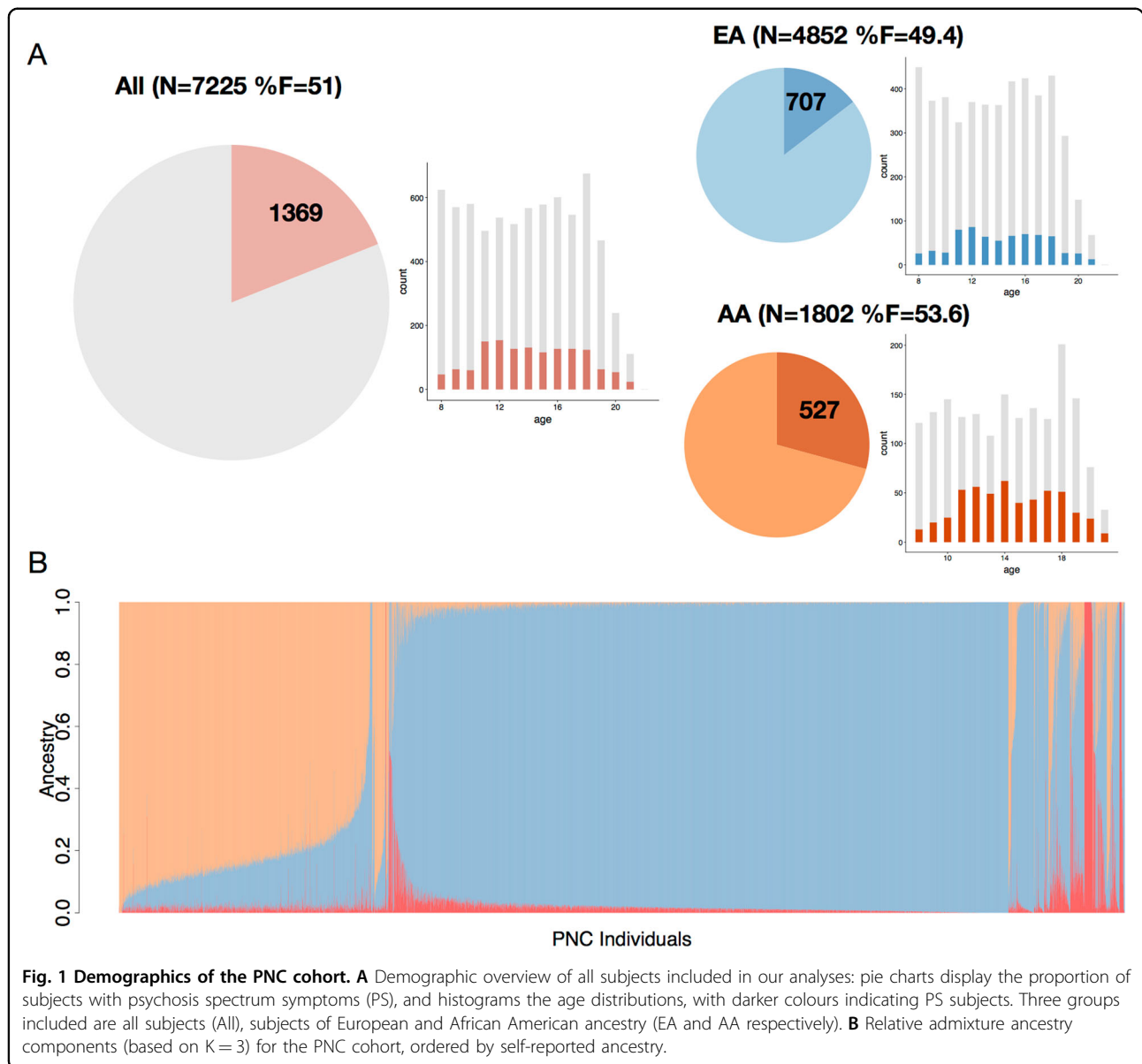
Developmental effects

We hypothesized that the association between PS and ADHD PGS would be strongest in the younger children and weaker closer to the typical age of onset of schizophrenia and other psychotic disorders. As predicted, we observe a significant interaction of ADHD PGS with age, with a stronger association for younger children (age 12 or younger), weakening in late adolescence ($P = 0.02$ for the interaction term in the full model, Fig. 3c). Given this decrease of association across age, we hypothesized an association with schizophrenia PGS in the older age group, but this was not the case (Supplementary Results).

We observed no interaction with sex ($P = 0.85$): the association between males and females are near-identical, with the OR in males-only 1.23 (1.10,1.37) and females-only 1.25 (1.10,1.41).
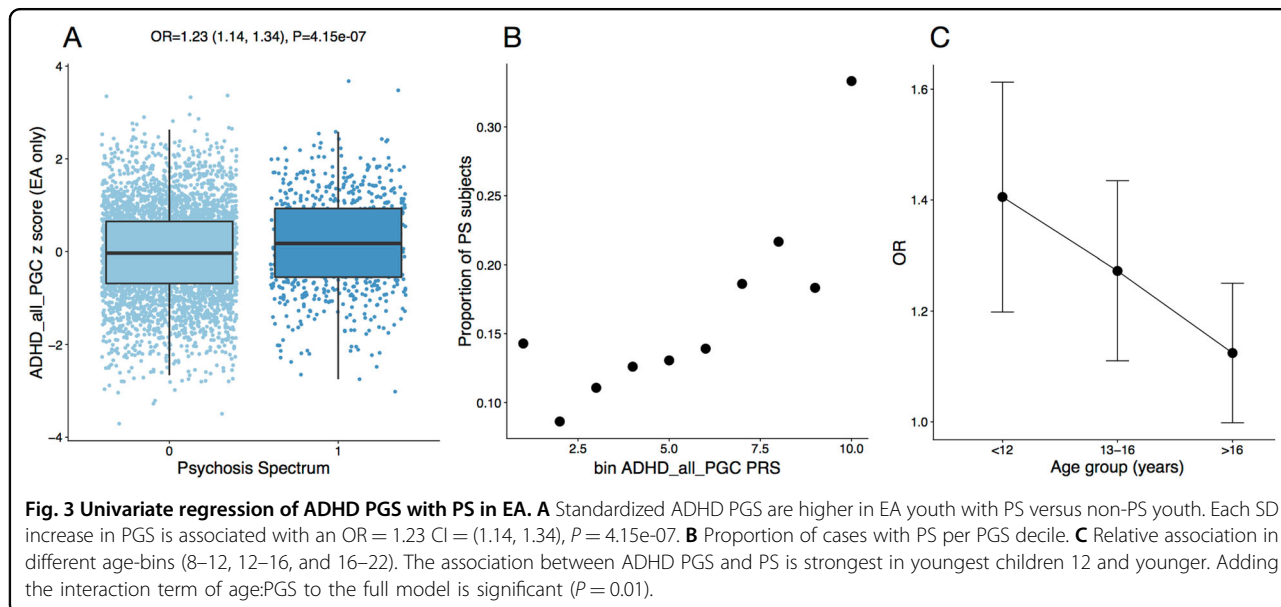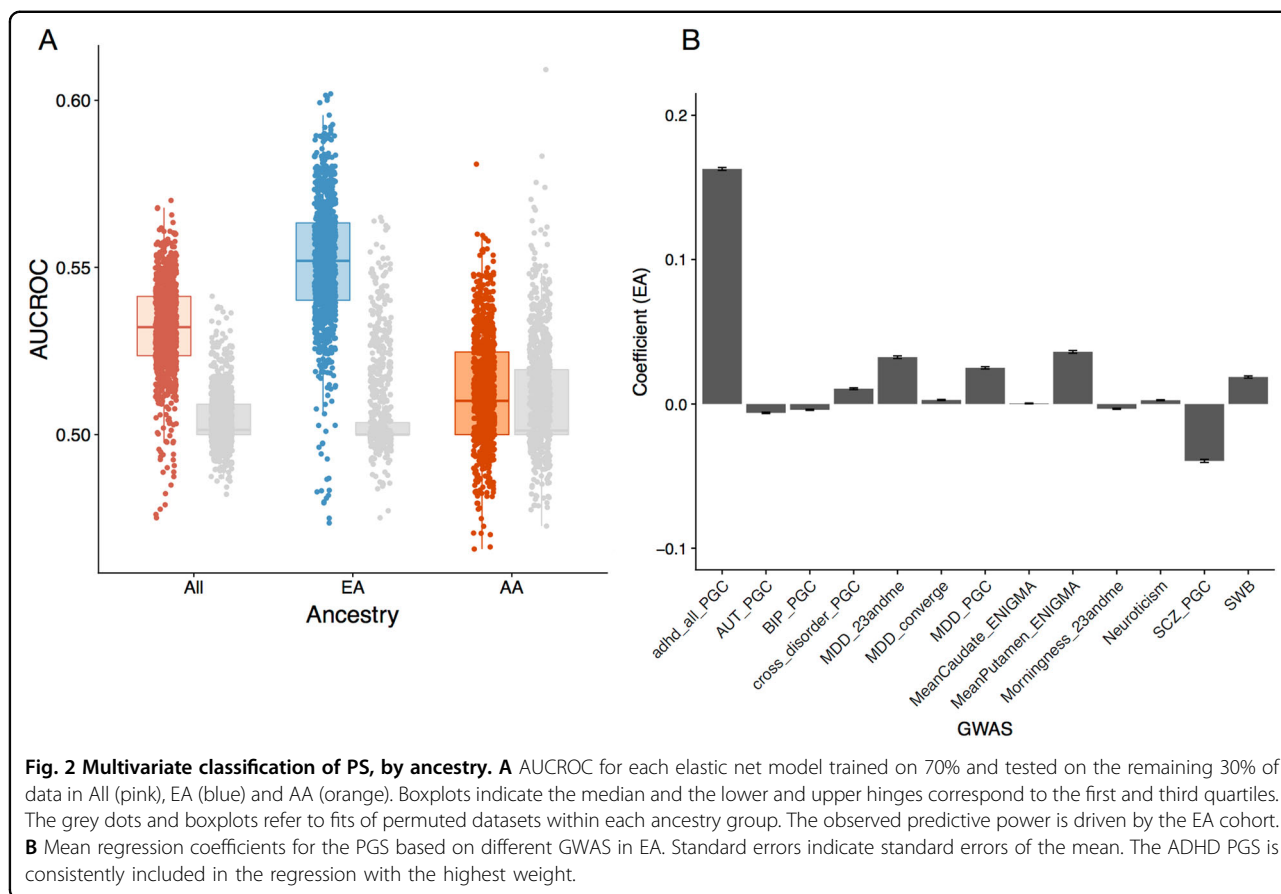
Phenotypic overlap

While evidence for shared genetic risk for (categorically defined) SCZ and ADHD is minimal (genetic correlation

Fig. 1 Demographics of the PNC cohort. A Demographic overview of all subjects included in our analyses: pie charts display the proportion of subjects with psychosis spectrum symptoms (PS), and histograms the age distributions, with darker colours indicating PS subjects. Three groups included are all subjects (All), subjects of European and African American ancestry (EA and AA respectively). **B** Relative admixture ancestry components (based on K = 3) for the PNC cohort, ordered by self-reported ancestry.

is $r = 0.11$, se $= 0.04$, $P = 0.001$, LD-score regression[17]), ADHD and psychosis symptoms are known to co-occur in youth[5,6]. Indeed, across ancestries, we observe a strong phenotypic overlap between PS and ADHD symptoms in the PNC (see Supplementary Materials). In the total sample, 5% of youth satisfy the DSM criteria for lifetime ADHD ($n = 384$; $n = 222$ and $n = 114$ within EA and AA, respectively). A large fraction of ADHD cases also endorse psychosis spectrum symptoms, and the majority of subjects in the PS group endorse ADHD symptoms (OR 2–4.7, Table 1).

We additionally explored the overlap between ADHD and PS using a variety of phenotypic constructs, and observed strong phenotypic overlap across domains, and across ancestry groups (Supplementary Methods and

Results). For example, we observe strong positive correlations between ADHD and PS domain scores (For all correlations, spearman rho is $0.25 < r < 0.69$ and $P < 10^{-16}$). Both higher inattention and hyperactivity scores are associated with an increased probability of being classified as PS, while at the same time higher PRIME and SOPS scores are associated with increased probability of meeting ADHD criteria, across ancestry groups (Supplementary Results and Figs. S9–S12). The age interaction effect we observed at the genetic level is not consistently observed at the phenotypic level; e.g., the effect of answering "yes" to any of the ADHD screener questions on the probability of being classified as PS does not change during development ($P = 0.41$ in EA, Supplementary results).

Fig. 2 Multivariate classification of PS, by ancestry. A AUCROC for each elastic net model trained on 70% and tested on the remaining 30% of data in All (pink), EA (blue) and AA (orange). Boxplots indicate the median and the lower and upper hinges correspond to the first and third quartiles. The grey dots and boxplots refer to fits of permuted datasets within each ancestry group. The observed predictive power is driven by the EA cohort. **B** Mean regression coefficients for the PGS based on different GWAS in EA. Standard errors indicate standard errors of the mean. The ADHD PGS is consistently included in the regression with the highest weight.

Fig. 3 Univariate regression of ADHD PGS with PS in EA. A Standardized ADHD PGS are higher in EA youth with PS versus non-PS youth. Each SD increase in PGS is associated with an OR = 1.23 CI = (1.14, 1.34), $P = 4.15e\text{-}07$. **B** Proportion of cases with PS per PGS decile. **C** Relative association in different age-bins (8–12, 12–16, and 16–22). The association between ADHD PGS and PS is strongest in youngest children 12 and younger. Adding the interaction term of age:PGS to the full model is significant ($P = 0.01$).

Phenotypic overlap and genetic liability

Given the strong phenotypic overlap between PS and ADHD symptoms, we performed a series of follow-up analyses to determine the extent to which the observed genetic association is driven by this overlap. We tested the association after removing ADHD cases. We also removed

Table 1   Univariate association of ADHD PGS with PS across ancestry groups.

Group	N	%PS	OR	2.5%	97.5%	P
All	7008	19%	1.12	1.05	1.18	0.0003
EA	4790	15%	1.23	1.14	1.34	4.15×10^{-7}
AA	1746	29%	0.98	0.88	1.06	0.65
EA_noADHD	4349	13%	1.26	1.16	1.38	2.63×10^{-7}
EA_noSymptoms	2497	7%	1.23	1.05	1.44	0.01

EA_noADHD and EA_noSymptoms denote the EA group after removing ADHD cases and subjects that endorse any of the symptoms in the ADHD screener, respectively. The 95% confidence interval of the association is represented by the 2.5% and 97.5% columns. *N* denotes the number of subjects included in each model, i.e., with complete information for all variables.

Table 2   Clinical overlap between PS and ADHD across ancestry groups.

	PS_ALL		PS_EA		PS_AA	
	1	0	1	0	1	0
ADHD						
1	151	233	83	139	50	64
0	1115	5220	584	3765	426	1080
Overlap statistic	OR 3.0 (2.4–3.7), $P < 2.2 \times 10^{-16}$		OR 3.8 (2.8–5.1), $P < 2.2 \times 10^{-16}$		OR 1.9 (1.3–2.9) $P = 0.0008$	
ADHD_Symptoms						
1	1073	2515	536	1694	427	621
0	284	3120	169	2328	91	573
Overlap statistic	OR 4.7 (4.1–5.4), $P < 2.2 \times 10^{-16}$		OR 4.6 (3.6–5.3), $P < 2.2 \times 10^{-16}$		OR 4.3 (3.3–5.6), $P < 2.2 \times 10^{-16}$	

Confusion matrices for overlap between PS and subjects meeting DSM criteria for ADHD (ADHD) and subjects that endorse any of the symptoms in the ADHD screener (ADHD_Symptoms), respectively. Only subjects with complete information for all variables were included in each overlap analysis.

the subset of youth that endorsed any ADHD symptoms from the screener. The latter analysis leaves only 2497 EA subjects and reduces the percentage of PS cases from 15 to 7%, with a total $n = 169$ PS subjects. In both cases, despite the smaller sample sizes and the relative depletion of the number of cases, the effect of ADHD PGS on PS risk remains stable (Table 2, Supplementary results, Table S2).

Thus, the association between increased ADHD liability and PS in youth does not appear to be driven by symptom overlap. Moreover, ADHD genetic risk is not only associated with PS as a categorical variable, but also similarly with the psychosis severity scales, measured quantitatively (beta = 0.47, $P = 0.0002$ for PRIME and beta = 0.21, $P = 1.68 \times 10^{-8}$ for SOPS), and significantly increased in the subset of PS subjects that endorse hallucinations ($n = 258$; OR = 1.15 (1.02,1.30), $P = 0.02$). As expected, the ADHD PGS is associated with ADHD status (OR 1.18, CI(1.04, 1.36), $P = 0.01$). Conversely, and contrary to existing evidence[48], the schizophrenia PGS is not associated with ADHD status ($P = 0.51$).

We investigate the relationship of ADHD and psychosis phenotypes to substance use and childhood trauma. In the PNC EA cohort, substance use information was available for 43% of the sample. In this sample, PS overlapped the use of alcohol, tobacco and marijuana only nominally ($P > 0001$), and not cocaine or over the counter substance use ($P > 0.05$). Including these as covariates in the model did not alter the estimated effect sizes (See Table S2). ADHD PGS has been associated with environmental adversity, including childhood maltreatment[49], which indirectly could contribute to the observed association with PS. While we replicate this association between childhood trauma and AHDH PGS in the EA cohort (OR:1.13 CI (1.05,1.21) $P = 0.001$), when we exclude subjects endorsing traumatic events the association remains stable (Table S2).

Finally, we evaluate whether ADHD PGS may be more broadly associated with other domains of psychopathology,

including specific symptoms related to depression, mania and anxiety. From these three domains, only depression and mania symptoms are nominally associated with ADHD PGS (depression OR:1.10 CI(1.02,1.18) $P = 0.01$, mania OR:1.10 CI(1.01,1.18) $P = 0.02$). However, these phenotypes both overlap strongly with PS (Table S3), and the association does not survive the exclusion of PS subjects ($P > 0.4$). The opposite, however, does hold: the association between ADHD PGS and PS remains consistent even after excluding subjects who endorsing either of these three phenotypes (Table S2).

Thus, the association between PS and ADHD PGS is not driven by symptom overlap, or indirectly driven by substance use or childhood trauma. Moreover, it appears to be specific to PS, and not a measure of general psychopathology.

Replication in an independent cohort

We sought to replicate our finding of increased polygenic liability for ADHD in youth with psychotic spectrum symptoms in the North American Prodrome Longitudinal Study, Phase 2 (NAPLS2) cohort[50,51]. NAPLS2 is an eight-site longitudinal study of predictors and mechanisms of conversion to psychosis, and includes help-seeking adolescents and young adults at clinical high risk (CHR) for psychosis (ages 12–35, with a median age of 18; $n = 328$) as well as unaffected control subjects ($n = 216$)[52].

As in our discovery sample, we observe increased ADHD risk in CHR youth compared to controls (OR = 1.06 (1.01, 1.11) $P = 0.02$; in EA, OR = 1.09 (1.01, 1.18) $P = 0.03$; $n = 124$ and $n = 70$, respectively). These effect sizes are similar to the ones observed in the >16 age group in the PNC EA cohort (Fig. 2). However, there is no difference in ADHD PGS between CHR subjects who subsequently convert to psychosis versus those who do not ($n = 80$ converters, $n = 248$ non-converters; $P = 0.62$).

Discussion

Leveraging a large community-based sample, we sought to characterize genetic risk profiles for psychotic spectrum symptoms across childhood and adolescence. We applied multi-PGS prediction models as well univariate statistical tests, based on GWAS of multiple brain and behavioural traits. Notably, we observed a modest but robust association between broadly defined psychosis symptomatology and genetic liability for ADHD, but not for schizophrenia or any other psychiatric traits. This effect was only observed in participants of European ancestry, for whom those within the highest decile of ADHD genetic risk have an almost 2.5-fold increased likelihood of being in the PS group, compared to those with lowest ADHD polygenic scores. This association is strongest in children 12 years or younger, and diminishes closer to typical age of onset of schizophrenia. In addition to being associated with the dichotomous PS phenotype, the ADHD PGS is associated with the quantitative symptom scales as well, generally improving power[53]. To our knowledge, this is the first investigation of ADHD polygenic scores for association with symptoms of psychosis in youth. We therefore replicated our finding in an independent cohort of subjects at clinical high risk for psychosis.

Contrary to recent genetic evidence based on psychotic experiences in adulthood[21], psychosis spectrum symptoms in youth did not yield a general association with multiple psychiatric illnesses, but with ADHD specifically. While an association between schizophrenia genetic risk and psychosis symptoms has been reported in adults[18,21] and older adolescents aged 15–19[18], the lack of such an association in our study is consistent with literature in young adolescents (in a population sample of >5k genotyped youth aged 12–18, no such association was observed[22]). Rather than testing a specific hypothesis, we test and correct for multiple genetic risk scores in a single cohort. This is an unbiased way of establishing the genetic architecture of a trait and reduces the risk of identifying false positive associations through performing single tests.

The association between PS and ADHD PGS, while modest, holds even when excluding all subjects that endorse any ADHD symptoms. Moreover, while we observed some association between ADHD PGS and the experience of trauma, replicating earlier evidence[49], the association with PS is not substantially mediated by trauma, or substance use.

One important question is whether the association between ADHD PGS and PS is specific to psychosis, or reflect more general liability for psychopathology. For example, a weak association between ADHD PGS and early onset depression has been reported prevoiusly[54]. Despite observing strong phenotypic overlap between PS and symptoms of depression, mania and anxiety domains,

we did not observe a robust association for these phenotypes with ADHD PGS in the PNC cohort.

In our study, the multivariate approach did not improve classification accuracy above a single trait association of ADHD. However, future efforts to improve risk scoring methodology and especially more powerful GWAS of related traits are likely to improve prediction as well.

As the discovery GWAS of ADHD included 55,374 children and adults (20,183 ADHD cases and 35,191 control subjects), ADHD genetic risk across all developmental ranges is captured in the downstream polygenic scores. An important avenue for future work will be to investigate if the association with general psychotic symptomatology is driven more by genetic liability present in children versus adults diagnosed with ADHD.

Despite similar phenotypic correlations across ancestry groups, the absence of any genetic association in youth of African ancestry highlights the need for increasing ethnic diversity in GWA studies. Because allele frequencies, linkage disequilibrium patterns, and effect sizes of common polymorphisms vary with ancestry, current common variant genetic findings do not translate well across populations[46,47,55]. Our study offers further evidence that polygenic scores, at this point, have limited predictive power in non-European ethnic groups. As PGS scores are approaching clinical utility[56], the crucial equity issue that arises as a result from this discrepancy should not be taken lightly. Novel tools to generate risk scores across ancestries, such as by scoring only segments of the genome matching the GWAS population in admixed populations, may improve applicability of the risk scores across ancestries. Most importantly, however, larger samples from different ancestries are needed to begin to close this gap.

Based on a follow-up study of PNC youth, about half of youth experiencing psychotic symptoms, symptoms persist or get worse over a 2-year follow-up, while even those whose psychotic symptoms remit exhibited comparatively higher symptom levels and lower functioning than typically developing youths[7]. An important follow-up question is whether youth with psychotic symptoms that ultimately develop a psychotic disorder have different genetic characteristics than those who do not. Consistent with our present findings in the PNC, we recently found that subjects meeting psychosis risk syndrome criteria that do not develop psychosis in a 2-year follow-up in the NAPLS cohort do not have increased genetic risk for schizophrenia compared to controls. However, polygenic liability for schizophrenia is a predictor for conversion to overt psychosis[6]. We now show, in the same NAPLS cohort, that ADHD PGS is also increased in CHR youth, but was not associated with conversion. While there are important clinical differences between PNC PS and NAPLS CHR cohorts (importantly, the latter consists of help-seeking youth, whereas PNC is a community

sample), the consistency of our findings and similar effect sizes across cohorts confirms the robustness of the ADHD PGS association with psychosis symptoms.

Our findings shed light on the genetic architecture of psychosis symptoms in a population-based youth cohort and suggest that broad psychosis spectrum symptoms in youth may reflect a different genetic etiology than psychotic symptoms in adulthood, or those that convert to psychosis. Rather, the genetic etiology of psychosis symptoms in youth seems more akin to ADHD. Findings indicate that genetic risk can be investigated across early symptom trajectories, and in non-help-seeking populations, to improve our understanding of disease risk factors and psychiatric comorbidities.

Author details
[1]Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, CA, USA. [2]Department of Psychiatry and Psychotherapy, University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany. [3]Department of Psychiatry, University of North Carolina, Chapel Hill, NC, USA. [4]Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA. [5]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA. [6]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA. [7]Department of Psychiatry, Hotchkiss Brain Institute, Calgary, AB, Canada. [8]Department of Psychiatry, UCSD, San Diego, CA, USA. [9]Department of Psychiatry, Zucker Hillside Hospital, Long Island, NY, USA. [10]Department of Psychiatry, UCSF, and SFVA Medical Center, San Francisco, CA, USA. [11]Department of Psychiatry, Yale University, New Haven, CT, USA. [12]Department of Psychiatry, Harvard Medical School at Beth Israel Deaconess Medical Center, Boston, MA, USA. [13]Departments of Psychology and Psychiatry, Emory University, Atlanta, GA, USA. [14]Department of Psychology, Yale University, New Haven, CT, USA. [15]Department of Psychiatry, University of Pennsylvania School of Medicine and the Penn-CHOP Lifespan Brain Institute, Philadelphia, PA, USA. [16]Department of Psychology, University of California, Los Angeles, CA, USA. [17]Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, CA, USA. [18]Department of Psychiatry, Erasmus University Medical Center, Rotterdam, The Netherlands

Conflict of interest
The authors declare that they have no conflict of interest.

Publisher's note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41398-021-01203-2.

References
1. McGrath, J. J. et al. Psychotic experiences in the general population: a cross-national analysis based on 31,261 respondents from 18 countries. *JAMA Psychiatry* **72**, 697–705 (2015).
2. Saha, S., Chant, D., Welham, J. & McGrath, J. A systematic review of the prevalence of schizophrenia. *PLoS Med.* **2**, e141 (2005).
3. Kessler, R. C., Chiu, W. T., Demler, O., Merikangas, K. R. & Walters, E. E. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch. Gen. Psychiatry* **62**, 617–627 (2005).
4. Schimmelmann, B. G., Michel, C., Martz-Irngartinger, A., Linder, C. & Schultze-Lutter, F. Age matters in the prevalence and clinical significance of ultra-high-risk for psychosis symptoms and criteria in the general population: Findings from the BEAR and BEARS-kid studies. *World Psychiatry* **14**, 189–197 (2015).
5. Lee, T. Y., Lee, J., Kim, M., Choe, E. & Kwon, J. S. Can we predict psychosis outside the clinical high-risk state? A systematic review of non-psychotic risk syndromes for mental disorders. *Schizophr. Bull.* **44**, 276–285 (2018).
6. Addington, J. et al. Comorbid diagnoses for youth at clinical high risk of psychosis. *Schizophr. Res.* **190**, 90–95 (2017).
7. Calkins, M. E. et al. Persistence of psychosis spectrum symptoms in the Philadelphia Neurodevelopmental Cohort: a prospective two-year follow-up. *World Psychiatry* **16**, 62–76 (2017).
8. Caspi, A. & Moffitt, T. E. All for one and one for all: mental disorders in one dimension. *Am. J. Psychiatry* **175**, 831–844 (2018).
9. Hartmann, J. A., Nelson, B., Ratheesh, A., Treen, D. & McGorry, P. D. At-risk studies and clinical antecedents of psychosis, bipolar disorder and depression: a scoping review in the context of clinical staging. *Psychol. Med.* **49**, 177–189, https://doi.org/10.1017/S0033291718001435 (2019).
10. Bufferd, S. J., Dougherty, L. R., Carlson, G. A., Rose, S. & Klein, D. N. Psychiatric disorders in preschoolers: continuity from ages 3 to 6. *Am. J. Psychiatry* **169**, 1157–1164 (2012).
11. Fisher, H. L. et al. Specificity of childhood psychotic symptoms for predicting schizophrenia by 38 years of age: a birth cohort study. *Psychol. Med.* **43**, 2077–2086 (2013).
12. Schizophrenia Working Group of the Psychiatric Genomics, C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).
13. Stahl, E. A. et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat. Genet.* **51**, 793–803 (2019).
14. Maier, R. M., Visscher, P. M., Robinson, M. R. & Wray, N. R. Embracing polygenicity: a review of methods and tools for psychiatric genetics research. *Psychol. Med.* **48**, 1055–1067 (2018).
15. Bergen, S. E. et al. Joint contributions of rare copy number variants and common SNPs to risk for schizophrenia. *Am. J. Psychiatry* **176**, 29–35 (2019).
16. Allardyce, J. et al. Association between schizophrenia-related polygenic liability and the occurrence and level of mood-incongruent psychotic symptoms in bipolar disorder. *JAMA Psychiatry* **75**, 28–35 (2018).
17. Bulik-Sullivan, B. K. et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
18. Pain, O. et al. Genome-wide analysis of adolescent psychotic-like experiences shows genetic overlap with psychiatric disorders. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **177**, 416–425 (2018).
19. Zavos, H. M. et al. Consistent etiology of severe, frequent psychotic experiences and milder, less frequent manifestations: a twin study of specific psychotic experiences in adolescence. *JAMA Psychiatry* **71**, 1049–1057 (2014).
20. Sieradzka, D. et al. Heritability of individual psychotic experiences captured by common genetic variants in a community sample of adolescents. *Behav. Genet.* **45**, 493–502 (2015).
21. Legge, S. E. et al. Association of genetic liability to psychotic experiences with neuropsychotic disorders and traits. *JAMA Psychiatry* https://doi.org/10.1001/jamapsychiatry.2019.2508 (2019).
22. Jones, H. J. et al. Phenotypic manifestation of genetic risk for schizophrenia during adolescence in the general population. *JAMA Psychiatry* **73**, 221–228 (2016).
23. Satterthwaite, T. D. et al. Connectome-wide network analysis of youth with Psychosis-Spectrum symptoms. *Mol. Psychiatry* **20**, 1508–1515 (2015).
24. Satterthwaite, T. D. et al. Structural brain abnormalities in youth with psychosis spectrum symptoms. *JAMA Psychiatry* **73**, 515–524 (2016).
25. Jalbrzikowski, M. et al. Structural brain alterations in youth with psychosis and bipolar spectrum symptoms. *J. Am. Acad. Child Adolesc. Psychiatry* https://doi.org/10.1016/j.jaac.2018.11.012 (2019).

26. Gur, R. C. et al. Neurocognitive growth charting in psychosis spectrum youths. *JAMA Psychiatry* **71**, 366–374 (2014).

27. Krapohl, E. et al. Multi-polygenic score approach to trait prediction. *Mol. Psychiatry* **23**, 1368–1374 (2018).

28. Satterthwaite, T. D. et al. The Philadelphia Neurodevelopmental Cohort: a publicly available resource for the study of normal and abnormal brain development in youth. *Neuroimage* **124**, 1115–1119 (2016).

29. Kobayashi, H. et al. A self-reported instrument for prodromal symptoms of psychosis: testing the clinical validity of the PRIME Screen-Revised (PS-R) in a Japanese population. *Schizophr. Res.* **106**, 356–362 (2008).

30. Kaufman, J. et al. 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**, 980–988 (1997).

31. Miller, T. J. et al. Prodromal assessment with the structured interview for prodromal syndromes and the scale of prodromal symptoms: predictive validity, interrater reliability, and training to reliability. *Schizophr. Bull.* **29**, 703–715 (2003).

32. Calkins, M. E. et al. The psychosis spectrum in a young U.S. community sample: findings from the Philadelphia Neurodevelopmental Cohort. *World Psychiatry* **13**, 296–305 (2014).

33. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).

34. Zheng, J. et al. LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* **33**, 272–279 (2017).

35. Demontis, D. et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat. Genet.* **51**, 63–75 (2019).

36. Autism Spectrum Disorders Working Group of The Psychiatric Genomics, C. Meta-analysis of GWAS of over 16,000 individuals with autism spectrum disorder highlights a novel locus at 10q24.32 and a significant overlap with schizophrenia. *Mol. Autism* **8**, 21 (2017).

37. Cross-Disorder Group of the Psychiatric Genomics, C. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* **381**, 1371–1379 (2013).

38. Wray, N. R. et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* **50**, 668–681 (2018).

39. Hyde, C. L. et al. Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nat. Genet.* **48**, 1031–1036 (2016).

40. consortium, C. Sparse whole-genome sequencing identifies two loci for major depressive disorder. *Nature* **523**, 588–591 (2015).

41. Hibar, D. P. et al. Common genetic variants influence human subcortical brain structures. *Nature* **520**, 224–229 (2015).

42. Hu, Y. et al. GWAS of 89,283 individuals identifies genetic variants associated with self-reporting of being a morning person. *Nat. Commun.* **7**, 10448 (2016).

43. Okbay, A. et al. Corrigendum: Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nat. Genet.* **48**, 1591 (2016).

44. Vilhjalmsson, B. J. et al. Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *Am. J. Hum. Genet.* **97**, 576–592 (2015).

45. Zou, H. & Hastie, T. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **67**, 301–320 (2005).

46. Mills, M. C. & Rahal, C. A scientometric review of genome-wide association studies. *Commun. Biol.* **2**, 9 (2019).

47. Martin, A. R. et al. Human demographic history impacts genetic risk prediction across diverse populations. *Am. J. Hum. Genet.* **100**, 635–649 (2017).

48. Nivard, M. G. et al. Genetic overlap between schizophrenia and developmental psychopathology: longitudinal and multivariate polygenic risk prediction of common psychiatric traits during development. *Schizophr. Bull.* **43**, 1197–1207 (2017).

49. Zwicker, A. et al. Neurodevelopmental and genetic determinants of exposure to adversity among youth at risk for mental illness. *J. Child Psychol. Psychiatry* **61**, 536–544 (2020).

50. Cannon, T. D. et al. An individualized risk calculator for research in prodromal psychosis. *Am. J. Psychiatry* **173**, 980–988 (2016).

51. Addington, J. et al. North American Prodrome Longitudinal Study (NAPLS 2): The Prodromal Symptoms. *J. Nerv. Ment. Dis.* **203**, 328–335 (2015).

52. Perkins, D. O. et al. Polygenic risk score contribution to psychosis prediction in a target population of persons at clinical high risk. *Am. J. Psychiatry* https://doi.org/10.1176/appi.ajp.2019.18060721 (2019).

53. van der Sluis, S., Posthuma, D., Nivard, M. G., Verhage, M. & Dolan, C. V. Power in GWAS: lifting the curse of the clinical cut-off. *Mol. Psychiatry* **18**, 2–3 (2013).

54. Rice, F. et al. Characterizing developmental trajectories and the role of neuropsychiatric genetic risk variants in early-onset depression. *JAMA Psychiatry* **76**, 306–313 (2019).

55. Martin, A. R. et al. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat. Genet.* **51**, 584–591 (2019).

56. Khera, A. V. et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat. Genet.* **50**, 1219–1224 (2018).