

## ARTICLE OPEN



# Polygenic risk of social isolation behavior and its influence on psychopathology and personality

Adam J. Socrates<sup>1</sup> , Niamh Mullins<sup>1,2,3</sup> , Ruben C. Gur<sup>4</sup> , Raquel E. Gur<sup>4</sup>, Eli Stahl<sup>2,10</sup> , Paul F. O'Reilly<sup>2</sup> , Abraham Reichenberg<sup>1</sup>, Hannah Jones<sup>5,6</sup> , Stanley Zammit<sup>6,7,8</sup> and Eva Velthorst<sup>1,9</sup>

© Crown 2024

Social isolation has been linked to a range of psychiatric issues, but the behavioral component that drives it is not well understood. Here, a genome-wide associations study (GWAS) was carried out to identify genetic variants that contribute specifically to social isolation behavior (SIB) in up to 449,609 participants from the UK Biobank. 17 loci were identified at genome-wide significance, contributing to a 4% SNP-based heritability estimate. Using the SIB GWAS, polygenic risk scores (PRS) were derived in ALSPAC, an independent, developmental cohort, and used to test for association with self-reported friendship scores, comprising items related to friendship quality and quantity, at age 12 and 18 to determine whether genetic predisposition manifests during childhood development. At age 18, friendship scores were associated with the SIB PRS, demonstrating that the genetic factors can predict related social traits in late adolescence. Linkage disequilibrium (LD) score correlation using the SIB GWAS demonstrated genetic correlations with autism spectrum disorder (ASD), schizophrenia, major depressive disorder (MDD), educational attainment, extraversion, and loneliness. However, no evidence of causality was found using a conservative Mendelian randomization approach between SIB and any of the traits in either direction. Genomic Structural Equation Modeling (SEM) revealed a common factor contributing to SIB, neuroticism, loneliness, MDD, and ASD, weakly correlated with a second common factor that contributes to psychiatric and psychotic traits. Our results show that SIB contributes a small heritable component, which is associated genetically with other social traits such as friendship as well as psychiatric disorders.

*Molecular Psychiatry*; <https://doi.org/10.1038/s41380-024-02617-2>

## INTRODUCTION

Social contact is essential for surviving and thriving in human societies [1]. As such, having limited contact with other people, or social isolation, can have detrimental effects on both physical and mental health. There is evidence that lack of social contact is associated with schizophrenia [2, 3], autism spectrum disorder [3], and depression [4], as well as with medical conditions such as cardiovascular disease [5] and diabetes [6]. Longitudinal studies indicate that Social-isolation can predate mental health issues and have a strong causal effect on poor mental health outcomes [4, 7, 8]. These issues have been acutely brought to light in the context of the Covid-19 pandemic, in which forced social isolation has had a substantial negative effect on mental health [9]. Social isolation has been strongly associated with the development of psychosis, and it has been hypothesized that this contribution may be due to isolated individuals with negative, delusional, or paranoid thoughts not having the opportunity to apply or test these beliefs in real world situations, and therefore not being

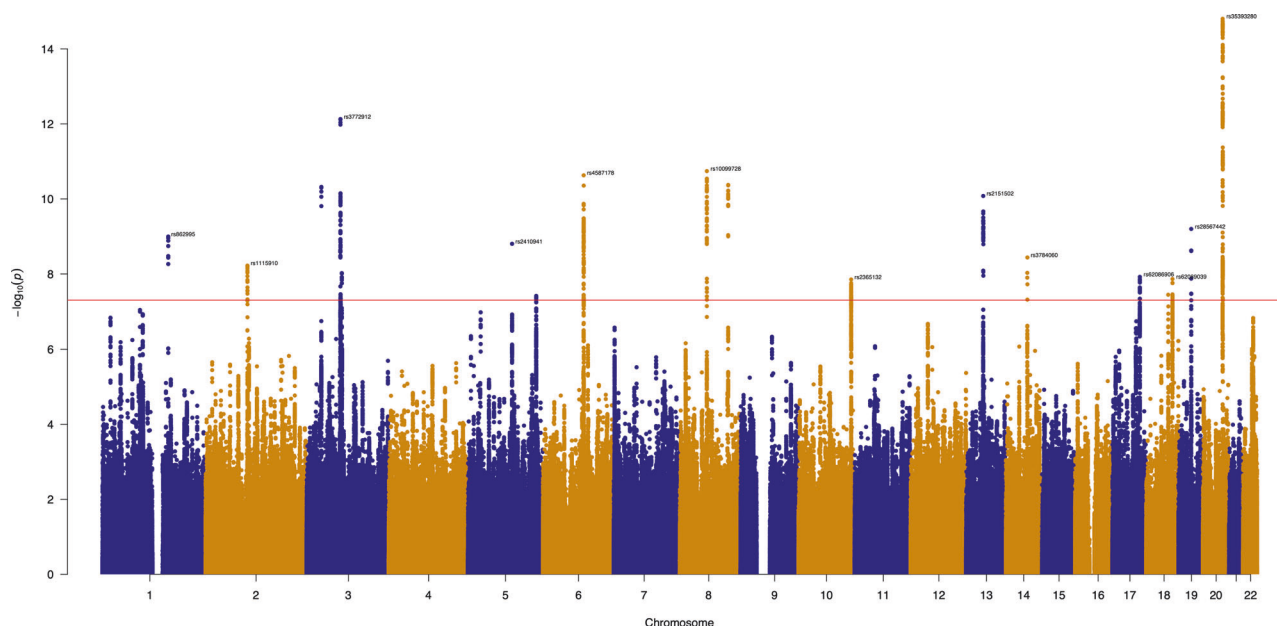
challenged and not having these delusions corrected by actual social interactions [10, 11].

Despite the impact of social isolation on mental and physical health, it remains among the least studied factors in psychiatric disorders, limiting understanding of etiology and causality with regards to psychiatric disorders [12–15]. Associations between genetics and traits related to social contact such as feelings of loneliness (feelings of distress or discomfort from being alone) and sociability (the ability to connect and socialize with others) have been noted [16]. More recently, Bralten et al. investigated the genetic underpinnings of a sociability phenotype combining behavioral and internalizing traits [17]. However, the existence and influence of an exclusive genetic predisposition towards social isolation *behaviors* specifically (SIB), i.e., action that leads to isolation, as distinct from the feelings that may potentially motivate or stem from such behavior, is yet to be established. Consequently, there is a gap in our knowledge about the extent to which SIB may represent a *causal and independent* risk for poor

<sup>1</sup>Department of Psychiatry, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Pl., New York, NY 10029, USA. <sup>2</sup>Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Pl., New York, NY 10029, USA. <sup>3</sup>Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Pl., New York, NY 10029, USA. <sup>4</sup>Department of Psychiatry, Perelman School of Medicine and the Lifespan Brain Institute, Penn Medicine and Children's Hospital of Philadelphia, University of Pennsylvania, 3400 Spruce, Philadelphia, PA 19104, USA. <sup>5</sup>MRC Integrative Epidemiology Unit, University of Bristol, Bristol BS8 2PR, UK. <sup>6</sup>Department of Population Health Sciences, Bristol Medical School, University of Bristol, Bristol BS8 2PR, UK. <sup>7</sup>Centre for Academic Mental Health, Bristol Medical School, University of Bristol, Bristol BS8 2PR, UK. <sup>8</sup>Division of Psychological Medicine and Clinical Neurosciences, MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, Cardiff CF24 4HQ, UK. <sup>9</sup>Department of Research, Mental Health Organization "GGZ Noord-Holland-Noord", Heerhugowaard, The Netherlands. <sup>10</sup>Present address: Regeneron Genetics Centre, Tarrytown, NY, USA. ✉email: adam.socrates@mssm.edu

Received: 22 February 2023 Revised: 2 May 2024 Accepted: 16 May 2024

Published online: 30 May 2024



**Fig. 1** Manhattan plot for social isolation behavior after conditioning on psychiatric disorders, based on the meta-analysis of 4 traits. The red horizontal line denotes genome-wide significance ( $P < 5 \times 10^{-8}$ ). See Supplementary Table 5 for details on genome-wide significant SNPs.

mental and physical health instead of being merely a direct consequence of other (clinical) symptomatology, for example due to stress or feelings of paranoia.

Twin studies have demonstrated that there is a similar genetic influence on both social isolation (as measured by access to social support; 40%) and loneliness (38%), but that they are only moderately genetically correlated, reflecting partially distinct constructs [18]. However, to our knowledge no prior study has carried out a genome-wide association study (GWAS) to elucidate the polygenic component of the purely behavioral aspects of SIB, as separate and distinct from feelings such as loneliness or relationship satisfaction. This information is pertinent, as the behavior itself could be partially driven by genetic factors, and could be detected and modified, providing early intervention targets if found to be on the causal pathway between inherited genetic variation and psychiatric disorders [19].

In order to better understand the genetic factors that influence SIB, the present study (1) conducted a novel GWAS for SIB in the UK Biobank cohort by meta-analyzing 4 behavioral social isolation traits; (2) derived Polygenic Risk Scores (PRS) from this GWAS for individuals in the Avon Longitudinal Study of Parents and Children (ALSPAC, UK) and used to examine associations with social traits for GWAS validation; (3) examined the genetic correlation between SIB and psychiatric disorders using GWAS results from the Psychiatric Genomics Consortium (PGC), as well as genetic correlations with personality and psychological traits such as neuroticism and loneliness; (4) Mendelian Randomization (MR) was applied to estimate causal effects between SIB and these traits/disorders; and (5) Genomic structural equation modeling (Genomic SEM) was used to model the shared genetic architecture of SIB and these traits.

## RESULTS

### GWAS

To investigate genetic propensity towards social isolation behavior (SIB), a GWAS was performed in the UK Biobank, based on a composite of 4 self-reported behavioral traits, constructed so higher scores represent increased SIB. A GWAS was performed on each of the 4 traits, after which Linkage disequilibrium (LD) score correlation revealed that the individual traits were genetically

correlated (see Supplementary Table 4). These were meta-analyzed with Multi-Trait Analysis of GWAS (MTAG) to produce a single GWAS before being conditioned on schizophrenia, major depressive disorder (MDD), and autism spectrum disorder (ASD), using multi-trait conditional and joint analysis (mtCOJO) to remove the effect of these psychiatric conditions. The final GWAS identified 19 loci, post-conditioning 17 loci remained at genome-wide significance ( $P < 5 \times 10^{-8}$ ; see Fig. 1).

The majority of the single nucleotide polymorphisms (SNPs) found to be associated with SIB were not previously associated with psychiatric or neurodevelopmental disorders. However, there are several exceptions. For example, the top lead SNP (rs6777906;  $P = 1.80 \times 10^{-15}$ ) is situated in the ARFGF2 gene, implicated in both bipolar disorder (BD) and schizophrenia but showing opposite directions of effect and hence may be a marker that could be used to differentiate between the two [20], as well as being linked to post-traumatic stress disorder (PTSD) [21, 22]. The second top SNP in chromosome 8, and the fourth top hit overall (rs2721942;  $1.47 \times 10^{-10}$ ), has also been associated with PTSD [23]. In chromosome 19, the lead SNP (rs28567442;  $P = 6.31 \times 10^{-10}$ ) is embedded in ZNF536, implicated in the development of the forebrain, and associated with schizophrenia [23]. Other genome-wide significant SNPs are in genes associated with schizophrenia (rs6125539;  $4.72 \times 10^{-9}$ , CSE1L) [24] and impulsivity (rs1248860;  $9.51 \times 10^{-9}$ , CADM2) [25]. In chromosome 13, rs17057528 ( $P = 8.82 \times 10^{-9}$ ) is in DIAPH3, identified as an autism risk gene [26], and is also implicated in hearing loss and impairment of speech perception [27].

### Polygenic risk scores

**ALSPAC.** To validate the SIB GWAS and PRS in an independent cohort, as well as explore its generalizability to a developmental cohort, PRS were generated in ALSPAC using the 13 significance thresholds for SNP inclusion ( $P_T < 5 \times 10^{-8}$ ,  $P_T < 5 \times 10^{-7}$ ,  $P_T < 5 \times 10^{-6}$ ,  $P_T < 5 \times 10^{-5}$ ,  $P_T < 5 \times 10^{-4}$ ,  $P_T < 0.001$ ,  $P_T < 0.01$ ,  $P_T < 0.05$ ,  $P_T < 0.1$ ,  $P_T < 0.2$ ,  $P_T < 0.3$ ,  $P_T < 0.4$ ,  $P_T < 0.5$ ; increasing in number of SNPs as the thresholds increase). The PRS were used to examine associations with friendship scores, comprising the 5 items relating to peer contact in  $n = 4934$  (at age 12) and  $n = 2909$  (at age 18) participants of the ALSPAC cohort (see Supplementary Tables 2 and 3 for details on friendship score items). As the

**Table 1.** Mendelian randomization results for MR-Egger analyses.

Exposure	Outcome	N SNPs	B	SE	P
SIB	ASD	65	-69.360	52.57	0.192
ASD	SIB	23	0.020	0.01	0.025
SIB	SCZ	65	1.008	54.72	0.985
SCZ	SIB	313	0.004	0.00	0.225
SIB	MDD	65	-70.943	52.86	0.184
MDD	SIB	34	0.000	0.01	0.974
SIB	Edu	65	0.361	0.28	0.196
Edu	SIB	440	0.085	0.05	0.120
SIB	Extra	65	-0.029	0.41	0.945
Extra	SIB	8	0.133	0.14	0.386
SIB	Loneliness	63	0.789	0.32	0.016
Loneliness	SIB	53	0.429	0.16	0.010

N SNPs number of SNPs in exposure, B Beta, SE standard errors, P p-value, SIB social isolation behavior, ASD autism spectrum disorder, SCZ schizophrenia, MDD major depressive disorder, Edu educational attainment, Extra extraversion.

genetic variants that contribute to SIB were ascertained in an adult sample, we expected the effect of the same genetic variants in a younger sample to be less pronounced, as genetically influenced behaviors may yet to be manifest.

The SIB PRS were not associated with friendship scores at age 12. At age 18, friendship score was significantly associated with the SIB PRS at the  $P_T = 0.05$  and  $P_T = 0.1$  threshold, with the latter being the most strongly associated ( $r^2 = 0.006$ ,  $P = 0.001$ ; see Supplementary Tables 6 and 7 for full results). The fewer SNPs included, the less predictive the model in terms of p-value, with the genome-wide significant only SNPs not associated with the friendships scores. This demonstrates that there are SNPs that do not reach genome-wide significance ( $5 \times 10^{-8}$ ) in the GWAS that still have a signal to detect SIB and contribute to the predictive power of the SIB PRS.

**LD score correlation.** LD score correlation was performed to investigate genetic correlations between SIB in the UK Biobank and schizophrenia, MDD, ASD, anorexia nervosa (AN), and BD from the Psychiatric Genomics Consortium (PGC). Three of these psychiatric disorders were correlated with SIB, with ASD having the strongest genetic correlation ( $rg = 0.23$ ,  $SE = 0.048$ ,  $P = 2.25 \times 10^{-06}$ ), followed by schizophrenia ( $rg = 0.102$ ,  $SE = 0.028$ ,  $P = 0.0002$ ) and MDD ( $rg = 0.093$ ,  $SE = 0.035$ ,  $P = 0.009$ ). This is in line with previous research which found negative genetic correlations with these outcomes and sociability [17]. AN and BD were not significantly genetically correlated with SIB ( $rg = -0.073$ ,  $SE = 0.041$ ,  $P = 0.073$ ;  $rg = -0.018$ ,  $SE = 0.035$ ,  $P = 0.61$ ). The results indicate that SIB genetics are associated with the genetics of certain psychiatric disorders and may form part of the genetic basis for them. This could occur if the genetics of SIB have downstream effects on behavior that could increase risk of symptoms and eventual diagnosis, or if the diagnosis itself leads to increased SIB.

LD score correlation was performed on anxiety, extraversion, neuroticism, loneliness, and educational attainment. Loneliness ( $rg = 0.29$ ,  $SE = 0.031$ ,  $P = 2.38 \times 10^{-20}$ ) was genetically correlated with SIB, suggesting the behavioral and perceptual aspects of social isolation are influenced by partially similar genetic factors. Educational attainment and extraversion were also significantly correlated with SIB ( $rg = 0.13$ ,  $SE = 0.025$ ,  $P = 2.11 \times 10^{-07}$ ;  $rg = -0.44$ ,  $SE = 0.07$ ,  $P = 1.29 \times 10^{-11}$ ). The latter is negatively correlated as might be expected, but the former's positive correlation is more surprising, and may be due to the way the phenotype is measured (i.e. years of education). Genetic

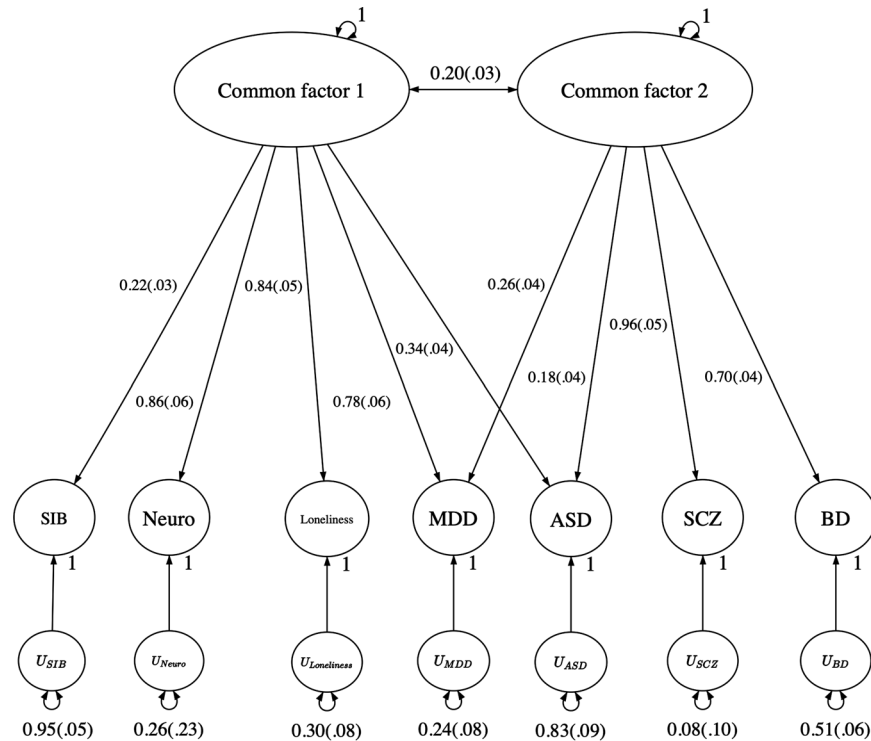
predisposition to more advanced education in the general population could potentially be linked to less social behavior due to prioritization of study and career over socializing (see Supplementary Table 8 for full results).

Using LD score correlation, the SNP-heritability of SIB after conditioning on the genetically correlated psychiatric disorders was estimated to be  $h^2 = 0.04$  ( $SE = 0.0022$ ,  $P = 8.95 \times 10^{-77}$ ), suggesting a small but significant SNP-based heritable component.

**Mendelian randomization.** Using the MR-Egger method to account for horizontal pleiotropy, and performing bi-directional MR, there was no evidence of causal relationships between SIB and any of the genetically correlated psychiatric disorders or traits in either direction (see Table 1 below). MR-Egger is particularly robust for ascertaining causality due to allowing of pleiotropic effects i.e., variants in the model can contribute to multiple traits as is likely the case with the genetics of complex behavioral traits. However, this flexibility means weak-instrument bias, a larger estimate variance, and a loss of power. The inverse-variance weighting (IVW) method has more power, but the likely presence of horizontal pleiotropy in behavioral trait genetics means causal effects are likely to be biased. See Supplementary Table 9 for full results, including inverse variance weighted (IVW) MR.

**Genomic SEM.** To ascertain the genetic architecture, we initially tested a model in which all 11 traits (SIB, plus schizophrenia, MDD, ASD, AN, BD, anxiety, extraversion, neuroticism, loneliness, and educational attainment) loaded onto a single common factor. Model fit was fairly poor ( $\chi^2 = 1756.37$ ,  $AIC = 1800.37$ ,  $CFI = 0.53$ ,  $SRMR = 0.134$ ), with MDD, Neuroticism, and loneliness most strongly loading onto the common factor, and educational attainment, AN and extraversion loading negatively onto this latent factor. We then removed the negatively loading traits, as well as anxiety due to it being underpowered in the model and produced a model in which the remaining 7 traits loaded onto a single common factor. This model showed an improved fit, but was still not optimal ( $\chi^2 = 770.11$ ,  $AIC = 798.11$ ,  $CFI = 0.65$ ,  $SRMR = 0.141$ ), despite all the traits loading moderately to strongly onto the common factor, aside from SIB.

Finally, we ran an exploratory factor analysis which suggested a 2-factor model, with personality loading onto one factor and psychotic traits (schizophrenia and BD) loading onto another, with MDD and ASD loading onto both (see Fig. 2). Confirmatory factor analysis demonstrated that this model is a good fit for the genetic



**Fig. 2 Genomic structural equation model of 11 traits loading onto 2 common latent genetic factors.** The path diagram is displayed with the loadings of each trait with the standard errors in parentheses. *U* represents the residual variances after removing variance explained by the common factors.

data ( $\chi^2 = 87.06$ ,  $AIC = 121.06$ ,  $CFI = 0.97$ ,  $SRMR = 0.067$ ; see Supplementary Table 10 for full model fit statistics).

This model indicates 2 moderately correlated common genetic factors that distinctly influence personality and psychosis, with SIB loading less strongly than neuroticism and loneliness onto the personality factor. This finding may be due to SIB being a behavioral trait with the others being internalizing, despite being associated. Both personality and psychosis loading factors appear to influence MDD and ASD jointly. Overall, to some extent the genetic factors that increase SIB also increase neuroticism, loneliness, and depression, implicating behavior amongst the latter 3 traits previously found to be genetically correlated with each other and with general well-being [28, 29].

## DISCUSSION

In this first study of the genetic factors that contribute to the behavior of social isolation (SIB), a meta-analysis of individual 4 behavioral trait GWAS of the discovery sample identified 17 genetic loci which predispose towards social isolation behavior. Some of these were in genes previously associated with psychiatric and neurological disorders, as well as neurotransmitter and brain function. However, most were not previously associated with other mental health, neurodevelopmental, or personality traits. Polygenic risk scores (PRS) derived from the GWAS were associated, in an independent developmental sample (ALSPAC), with the friendship scores at age 18 and there was strong evidence supporting shared genetic etiology between SIB and major psychiatric disorders, personality traits, and educational attainment, based on genetic correlations.

The PRS generated in ALSPAC were associated with friendship scores at age 18 but not at age 12. These results suggest that the SIB GWAS is a valid indicator of social-related traits, with higher PRS associated with lower friendship scores and outcomes. The PRS association with scores at age 18, as opposed to age 12,

might indicate that genetically influenced personal social behavior does not necessarily manifest until later in adolescence. This finding could be due to confounding by gene-environment correlation [30]. At younger ages, children may have less control over their own social environments and interactions than at age 18, as their parents would likely select their environments for them, in which case behavior would be less strongly influenced by their own genetic predispositions. A similar effect is observed in intelligence genetics, in which heritability increases over time [31]. It is considered that genetic predisposition leads to active and passive correlations with school selection or teacher attention for example, creating a “snowball” effect in which those genetic influences are amplified over time. It is possible that similar effects are at play with behavioral genetics, in which SIB genetic predisposition leads to development, or lack thereof, of social skills and sociability, modulating real life social isolation over time.

SIB was found to be genetically correlated with schizophrenia, as well as with ASD and MDD. This pattern of results suggest that SIB is a feature that cuts across multiple psychiatric disorders and mental health generally. It is well known that social isolation is linked to poorer mental health [32], but here it is shown that there is a genetic association which indicates that SIB may form part of the etiological basis of these disorders. MR did not reveal evidence of causality of SIB on psychiatric disorders, but the specific method required for complex behavioral traits has limited power to detect such causality effects. Genomic SEM revealed a shared latent genetic component for SIB, MDD, ASD, loneliness, and neuroticism, and this component was also found to be moderately correlated with a shared genetic component which contributes to psychosis. Further studies with psychiatric populations will be required to test this hypothesis, but considering that social engagement is a modifiable intervention target [33], identifying those with a genetic predisposition towards SIB may be a useful strategy in mitigating mental health issues.

The current study demonstrated a heritable genetic component to SIB by utilizing a large sample size and detailed phenotype information in the discovery sample, allowing the development comprehensive and valid SIB trait. This was confirmed by the PRS generated from the discovery GWAS being validated in an independent sample, and several genome-wide significant SNPs were found associated with SIB. However, LD score correlation only estimated 4% heritability for SIB and PRS were only able to explain 0.6% of the variance in friendship scores in ALSPAC replication sample. The SNP-heritability is likely to be a lower bound estimate, as this only takes into account the common SNPs genotyped and not rare variants or de novo mutations [34]. Further, despite having up to 450,000 individuals available for the discovery GWAS, the most powerful GWAS such as educational attainment are becoming increasingly predictive with approximately 3 million participants [35]. Thus, increasing sample size will allow the detection of more SNPs that contribute to SIB behavior and increase both heritability estimates and the predictive power of PRS. In ALSPAC the target sample also had relatively few participants at age 18 ( $n = 2909$ ) compared to age 12 ( $n = 4934$ ), which likely contributed to lower bound variance explained.

Despite the likelihood that our results represent a lower bound estimate, the low heritability of SIB indicates that environmental factors also play a substantial role. Twin studies have shown that social isolation (as measured by proxy as low social support) has a 60% contribution of non-shared environmental factors [18]. This suggests that although genetics may drive social behavior and self-selection of isolating environments to some extent, socio-demographic factors, life events and peer influence may have a larger overall effect on the motivators of SIB. The heritable component of perceived social isolation or loneliness has been suggested as an adaptive mechanism that motivates connection and incentivizes group integration in the wake of the environmental risk of social isolation [36]. Our study suggests that genetic, as well as environmental factors play a role in the act of isolating the self, perhaps indicating adaptive benefits e.g., mitigating disease risk, competition, and risk of injury [37].

In order to further investigate how the genetic component of SIB manifests in behavior and the development of psychiatric disorders, further studies will be required establishing whether or not SIB PRS can predict case control status for disorders such as schizophrenia, MDD and ASD. If so, it will be necessary to consider which specific behaviors are influenced by genetics, and how they manifest in the development and diagnosis of psychiatric disorders. By targeting behavior, our present study has laid the foundation for identifying a possible target for intervention that can be addressed in real world scenarios. However, the relatively small effect sizes of individual SNPs and the resulting low predictive power of PRS indicate further investigation.

## METHODS

### Study cohorts

**Discovery sample.** The UK Biobank (UKB) is a detailed prospective study with over 502,650 participants aged 40–69 years when recruited in 2006–2010 and includes both genetic and phenotypic data on complex traits [38]. The recruitment process was coordinated around 22 centers in the UK (between 2007 and 2010). Individuals within traveling distance of these centers were identified using NHS patient registers (response rate = 5.47%). Invitations were sent using a stratified approach to ensure demographic parameters were in concordance with the general population. All participants provided written informed consent and the current study was ethically approved by the UK Biobank Ethics and Governance Council (REC reference 11/NW/0382; UK Biobank application reference 18177).

**Genetic data.** Blood samples from 488,366 UK Biobank participants were genotyped using the UK BiLEVE array or the UK Biobank axiom array. Further details on the genotyping and quality control (QC) can be found on

the UK Biobank website (<http://www.ukbiobank.ac.uk/scientists-3/genetic-data/>). In the current study, SNPs were removed if they had missingness <0.02 and a minor allele frequency (MAF) <0.01. Exclusions based on heterozygosity and missingness were implemented according to UK Biobank recommendations (<http://biobank.ctsu.ox.ac.uk/showcase/label.cgi?id=100314>). Samples were removed if they were discordant for sex. SNPs deviating from Hardy-Weinberg equilibrium (HWE) were removed at a threshold of  $P < 1 \times 10^{-8}$ . Genotype data was imputed according to standard UK Biobank procedure, on 487,442 samples [39], excluding variants with an MAF <0.01 and an imputation quality score <0.3. After basic QC procedures and exclusions, 487,409 samples with phenotype data remained for genetic analysis. As genetic variants and polygenic risk scores (PRS) have low generalizability between ancestries, 4-means clustering on the first two principal components was performed, retaining 449,609 individuals from the largest cluster corresponding to European ancestry.

### Phenotype data

**Social isolation:** To derive a comprehensive measure of social isolation behavior (SIB), we ran a data-driven principal component analyses (using Promax rotation) on all available self-reported answers to questions that (1) directly probed for the quantity or quality of social engagement, (2) were available for at least 90% of study participants, and (3) did not include internalizing, perceptions, or feelings related to sociability e.g. “Friendship satisfaction”. Based on these criteria, we included data on the following 3 items, that all loaded on a single factor: “Frequency of family/friend visits”, “Being able to confide in others”, and “Number of social activities a week”. The items “Frequency of family/friend visits” and “Being able to confide with others” were both rated on a seven-point Likert scale (i.e. ‘Almost daily’, ‘2–4 times a week’, ‘about once a week’, ‘about once a month’, ‘once every few months’, ‘never or almost never’, and ‘no friends/ family outside of household’). The items “Frequency of family/friend visits” and “Being able to confide with others” were considered continuously and recoded so that higher values corresponded to greater social isolation. Answer options for the item “Number of a/social activities a week” included attending a sports club, pub, social club, religious group, adult educational classes, or other group activities and were summed to represent the ‘total number of social activities a week’, also considered continuously.

To complement the answers to the self-report, sociodemographic information about the number of people in the household was added as an additional proxy of social contact, as representing a behavioral decision to isolate. This “Number in household” item was dichotomized as a binary trait representing living alone, with 0 others in household coded as ‘1’ for SIB and any greater number in household as ‘0’. As such, higher scores for each of the 4 traits corresponded to higher SIB. See supplementary material for full phenotype and coding details, and <https://biobank.ndph.ox.ac.uk/showcase/> for further information. For all items, individuals with missing data, or who preferred not to answer were excluded (maximum  $N = 18949$ ). Participants who were wheelchair users ( $N = 426$ ) and/or morbidly obese ( $BMI > 40$ ;  $N = 9689$ ) were also excluded from the analysis, as these factors may arguably hamper the level of social activity but are unrelated to genetic or psychiatric vulnerability to SIB [40, 41].

### Validation cohort

**ALSPAC cohort:** The Avon Longitudinal Study of Parents and Children (ALSPAC) is a prospective birth cohort which recruited pregnant women with expected delivery dates between April 1991 and December 1992 from Bristol UK. 14,541 pregnant women were initially enrolled with 14,062 children born and 13,988 alive at 1 year of age. Detailed information on health and development of children and their parents were collected from regular clinic visits and completion of questionnaires. Please note that the study website contains details of all the data that is available through a fully searchable data dictionary and variable search tool” and reference the following webpage: <http://www.bristol.ac.uk/alspac/researchers/our-data/>. A detailed description of the cohort has been previously published [42, 43]. Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees.

**Genotype data.** 11,343 participants in ALSPAC have genotype data available, genotyped using the Illumina HumanHap550 quad chip genotyping platforms, and standard quality control (QC) procedures applied. Individuals with non-European ancestry were removed to minimize bias due to ancestral population stratification. SNPs with a

MAF of  $<0.01$ , a call rate of  $<0.95$  or evidence for violations of Hardy–Weinberg equilibrium ( $P < 5 \times 10^{-07}$ ) were removed. Data was imputed using standard ALSPAC procedure using the HapMap 2 reference panel, keeping SNPs with MAF  $>0.02$  and an INFO score  $>0.9$ . After these quality control measures, 9115 individuals and 4,731,235 SNPs remained in the analysis. Full quality control procedures can be found at: [https://alspac.github.io/omics\\_documentation/alspac\\_omics\\_data\\_catalogue.html](https://alspac.github.io/omics_documentation/alspac_omics_data_catalogue.html)

**Phenotype data.** To test the validity of the SIB construct, 2 friendship scores were derived from 5 questions from clinical questionnaires based on questions from the Cambridge Hormones and Moods Project Friendship Questionnaire [44], completed by the parents of offspring at ages 12 and 18 respectively e.g. “*Teenager is happy with number of friends*”. Each question consisted of 4–6 categorical responses, corresponding to a 4–6 point scale e.g. “1 = Very happy, 2 = Quite happy, 3 = Quite unhappy, 4 = Unhappy, 5 = No friends”. Responses were summed to create a continuous scale, with higher scores corresponding to lower friendship quality and greater SIB. 4, 934 of the cohort had the phenotype information at age 12, and 2909 at age 18. See supplementary Table 2 for full details on questions.

**GWAS summary statistics:** To test for genetic correlations between SIB and associated psychiatric disorders using LD score correlation, the SIB GWAS based on UK Biobank data was used along 10 base genome-wide association summary statistics for schizophrenia, depression (MDD), autism spectrum disorder (ASD), anorexia nervosa (AN), bipolar disorder (BD), anxiety, extraversion, neuroticism, loneliness, and educational attainment. These were the Psychiatric Genomics Consortium Wave 3 (PGC3) schizophrenia GWAS [45], the 2019 PGC MDD Working Group GWAS [46], and the 2017 PGC ASD Working Group GWAS [47], the 2019 Eating Disorders Working Group of the Psychiatric Genomics Consortium anorexia GWAS [48], the BD Working group of the Psychiatric Genomics Consortium GWAS [49], the 2019 iPSYCH anxiety GWAS [50] the Genetics of Personality-2 (GPC-2) extraversion and neuroticism GWAS [51, 52], a 2019 GWAS meta-analysis of loneliness [53], and the 2018 Social Science Genetic Association Consortium (SSGAC) GWAS of educational attainment [54].

### Statistical analyses

**GWAS analysis.** Association testing of autosomal SNPs was carried out on each of the 4 SIB traits (“*Frequency of family/friend visits*”, “*Being able to confide in others*”, “*Number of social activities a week*”, and “*Number in household*”) using BOLT Bayesian linear mixed models (BOLT-LMM) [55] to account for relatedness and cryptic population stratification, while increasing power and controlling for false positives. Age, sex, batch, and center were included as covariates, as well as education, income, and Townsend deprivation index (TDI) to account for socio-economic status (SES). The top 15 principal components (PCs) were also included to control for main population stratification. Multi-Trait Analysis of GWAS (MTAG) [56] was used to meta-analyze the individual “*Frequency of family/friend visits*”, “*Being able to confide in others*”, “*Number of social activities a week*”, and “*Number in household*” outcomes to form a single, composite SIB GWAS. This score is achieved by leveraging power across correlated GWAS estimates in overlapping samples. Finally, multitrait-based conditional and joint analysis (mtCOJO) [57] was used to adjust the SIB GWAS summary statistics for the effects of psychiatric disorders, specifically schizophrenia, major depressive disorder (MDD), and autism spectrum disorder (ASD), using European ancestry GWAS summary statistics for each. These are the psychiatric disorders which are commonly considered to lead to increased risk of social withdrawal and isolation [2–4, 7, 8] and were conditioned on to remove potential downstream effects of psychiatric disorders. SNPs were selected as instruments at  $5 \times 10^{-05}$ , clumped 1MB apart or with LD  $r^2 < 0.2$  based on the 1000 Genomes Project Phase 3 reference panel for independence. mtCOJO uses these SNPs Generalized Summary-data-based Mendelian Randomization (GSMR) to estimate the effect of the exposures (psychiatric disorders) on the outcome (SIB), producing conditioned effect sizes and  $p$ -values. Statistically significant independent signals were identified using 1MB clumping and a genome-wide significance threshold of  $P < 5 \times 10^{-08}$ .

**Polygenic risk score analysis.** Polygenic risk scores (PRS) were generated in ALSPAC using PRSice-2 [58], using the discovery SIB GWAS to sum and weight risk alleles for individuals in each cohort. SIB GWAS results were pruned for linkage disequilibrium (LD) using the  $p$ -value informed clumping method in PLINK (-clump-p1 1 - clump-p2 1 -clump-r2 0.1

-clump-kb 250). This method preferentially retains SNPs with the strongest evidence of association and removes SNPs in LD ( $r^2 > 0.1$ ) that show weaker evidence of association within 250Kb windows, based on LD structure from the HRC reference panel. Subsets of SNPs were selected from the results at 13 increasingly liberal  $P$  value thresholds (ranging from  $p < 5 \times 10^{-08}$ , to  $p < 0.5$ ). Risk alleles were included and tested to predict outcomes at 13 different significance thresholds, allowing the utilization of the most predictive PRS and threshold. These PRS were tested for associations with the friendship scores in ALSPAC, using linear regression models and including age, sex and 10 PCs as covariates. To account for the multiple testing of 13 PRS thresholds and 2 friendship scores, a Bonferroni correct significance threshold of  $P < 0.002$  was used.

**LD score correlation.** Genetic correlations and heritability estimates were conducted using LD score correlation [59], to investigate associations between SIB and schizophrenia, MDD, ASD, AN, BD, anxiety, extraversion, neuroticism, loneliness, and educational attainment, using GWAS summary statistics from the SIB GWAS conducted in the UK Biobank and each psychiatric disorder or trait from the PGC, GPC, iPSYCH, or SSGAC. LD score correlation provides an accurate estimation of the genetic correlation between two traits by separating genuine polygenic effects from confounders by regressing test statistics against LD scores (correlation between genomic sites, and SNPs tagging causal variants), as SNPs with test statistics describing true associations are positively correlated with LD scores, and the intercept estimates confounding accurately.

**Mendelian randomization.** To test for causality between SIB and psychiatric outcomes, bi-directional Mendelian Randomization was conducted using the *r*-package TwoSampleMR (<https://github.com/MRCIEU/TwoSampleMR>) [60]. Instrumental variables for the exposures (both SIB and the significant genetically correlated schizophrenia, MDD, ASD, extraversion, loneliness, and educational attainment) were extracted at genome-wide significance and at  $p < 5 \times 10^{-06}$  after strict LD clumping at 10,000 kb windows and LD  $r^2 < 0.001$  to ensure instruments were independent. Exposure and outcomes were harmonized and MR-Egger was used in the primary analyses to account for horizontal pleiotropy. The inverse variance weighted (IVW) was also used as a less conservative, more powerful approach as its weighting combines instrumental variable estimates more efficiently, but it may not be valid for these traits as it assumes no horizontal pleiotropy is present. To account for multiple testing, a Bonferroni corrected  $p$ -value threshold of  $P < 0.004$  was used to ascertain significance.

**Genomic SEM.** To investigate the genetic architecture of the traits investigated along with SIB, we performed exploratory and confirmatory factor analysis using the *r*-package Genomic Structural Equation Modelling (Genomic SEM; <https://github.com/GenomicSEM/GenomicSEM>) [61]. Genomic SEM uses multivariable LD score correlation to fit models using genetic and sampling covariance matrices. A single common genomic factor is derived with loadings for each trait which represent shared genetic contribution. Exploratory factor analysis was then performed of the genetic covariance matrix to determine loadings across multiple factors. Confirmatory factor analysis was then utilized to compare and test model fits.

### DATA AVAILABILITY

UK Biobank data are available through a procedure described at <http://www.ukbiobank.ac.uk/using-the-resource/>. ALSPAC data access is through a system of managed open access. The steps below highlight how to apply for access to the data included in this paper and all other ALSPAC data. 1. Please read the ALSPAC access policy ([http://www.bristol.ac.uk/media-library/sites/alspac/documents/researchers/data-access/ALSPAC\\_Access\\_Policy.pdf](http://www.bristol.ac.uk/media-library/sites/alspac/documents/researchers/data-access/ALSPAC_Access_Policy.pdf)) which describes the process of accessing the data and biological samples in detail, and outlines the costs associated with doing so. 2. You may also find it useful to browse our fully searchable research proposals database (<https://proposals.epi.bristol.ac.uk/>), which lists all research projects that have been approved since April 2011. 3. Please submit your research proposal (<https://proposals.epi.bristol.ac.uk/>) for consideration by the ALSPAC Executive Committee using the online process. You will receive a response within 10 working days to advise you whether your proposal has been approved. If you have any questions about accessing data, please email: [alspac-data@bristol.ac.uk](mailto:alspac-data@bristol.ac.uk). Schizophrenia, Autism spectrum disorder, major depressive disorder, bipolar disorder, and anorexia nervosa GWAS summary statistics are publicly available from the PGC (<https://www.med.unc.edu/pgc/download-results/>). Anxiety GWAS summary statistics are available from iPSYCH (<https://ipsych.dk/en/research/downloads>). Extraversion

and Neuroticism GWAS summary statistics are publicly available from the GPC (<https://tweelingenregister.vu.nl/gpc/>). Loneliness GWAS summary statistics are available at the following link: <https://t.co/ARgS84uwKl>. Educational attainment GWAS summary statistics are available from the SSGAC data portal (<https://thessgac.com/>).

## CODE AVAILABILITY

Software code for PRSice-2 is available at <https://www.prsice.info/>. All other code used is available upon request.

## REFERENCES

- The cooperative human (editorial). *Nat Hum Behav.* 2018;2:427–8. <https://doi.org/10.1038/s41562-018-0389-1>.
- Green MF, Horan WP, Lee J, McCleery A, Reddy LF, Wynn JK. Social disconnection in schizophrenia and the general community. *Schizophrenia Bull.* 2018;44:242–9.
- Velthorst E, Froudust-Walsh S, Stahl E, Ruderfer D, Ivanov I, Buxbaum J, et al. Genetic risk for schizophrenia and autism, social impairment and developmental pathways to psychosis. *Transl Psychiatry.* 2018;8:204.
- Santini ZI, Jose PE, Cornwell EY, Koyanagi A, Nielsen L, Hinrichsen C, et al. Social disconnectedness, perceived isolation, and symptoms of depression and anxiety among older Americans (NSHAP): a longitudinal mediation analysis. *Lancet Public Health.* 2020;5:e62–e70.
- Sharma T, Padala PR, Mehta JL. Loneliness and social isolation: determinants of cardiovascular outcomes. *Curr Cardiol Rev.* 2021;17:e051121190873.
- Brinkhues S, Dukers-Muijers NH, Hoebe CJ, Van Der Kallen CJ, Dagnelie PC, Koster A, et al. Socially isolated individuals are more prone to have newly diagnosed and prevalent type 2 diabetes mellitus—the Maastricht study. *BMC Public Health.* 2017;17:1–12.
- Noguchi T, Saito M, Aida J, Cable N, Tsuji T, Koyama S, et al. Association between social isolation and depression onset among older adults: a cross-national longitudinal study in England and Japan. *BMJ Open.* 2021;11:e045834.
- Velthorst E, Fett AKJ, Reichenberg A, Perlman G, van Os J, Bromet EJ, et al. The 20-year longitudinal trajectories of social functioning in individuals with psychotic disorders. *Am J Psychiatry.* 2017;174:1075–85.
- Pancani L, Marinucci M, Aureli N, Riva P. Forced social isolation and mental health: a study on 1,006 Italians under COVID-19 lockdown. *Front Psychol.* 2021;12:1540.
- Velthorst E, Nieman DH, Becker HE, van de Fliert R, Dingemans PM, Klaassen R, et al. Baseline differences in clinical symptomatology between ultra high risk subjects with and without a transition to psychosis. *Schizophrenia Res.* 2009;109:60–65.
- Velthorst E, Meijer C, GROUP Investigators. The association between social anhedonia, withdrawal and psychotic experiences in general and high-risk populations. *Schizophrenia Res.* 2012;138:290–4.
- Kamenov K, Cabello M, Caballero FF, Cieza A, Sabariego C, Raggi A, et al. Factors related to social support in neurological and mental disorders. *PLoS One.* 2016;11:e0149356.
- Kennedy N, Foy K, Sherazi R, McDonough M, McKeon P. Long-term social functioning after depression treated by psychiatrists: a review. *Bipolar Disord.* 2007;9:25–37.
- Simons CJ, Bartels-Velthuis AA, Pijnenborg GH, Genetic Risk and Outcome of Psychosis (GROUP) Investigators. Cognitive performance and long-term social functioning in psychotic disorder: a three-year follow-up study. *PLoS One.* 2016;11:e0151299.
- Bellido-Zanin G, Pérez-San-Gregorio MÁ, Martín-Rodríguez A, Vázquez-Morejón AJ. Social functioning as a predictor of the use of mental health resources in patients with severe mental disorder. *Psychiatry Res.* 2015;230:189–93.
- Day FR, Ong KK, Perry JR. Elucidating the genetic basis of social interaction and isolation. *Nat Commun.* 2018;9:1–6.
- Bralten J, Mota NR, Klemann CJ, De Witte W, Laing E, Collier DA, et al. Genetic underpinnings of sociability in the general population. *Neuropsychopharmacology.* 2021;46:1627–34.
- Matthews T, Danese A, Wertz J, Odgers CL, Ambler A, Moffitt TE, et al. Social isolation, loneliness and depression in young adulthood: a behavioural genetic analysis. *Soc psychiatry Psychiatr Epidemiol.* 2016;51:339–48.
- Lewis CM, Vassos E. Polygenic risk scores: from research tools to clinical instruments. *Genome Med.* 2020;12:1–11.
- Ruderfer DM, Ripke S, McQuillin A, Boocock J, Stahl EA, Pavlides JMW, et al. Genomic dissection of bipolar disorder and schizophrenia, including 28 sub-phenotypes. *Cell.* 2018;173:1705–15.
- Huckins LM, Chatzinakos C, Breen MS, Hartmann J, Klengel T, da Silva Almeida AC, et al. Analysis of genetically regulated gene expression identifies a prefrontal PTSD gene, SNRNP35, specific to military cohorts. *Cell Rep.* 2020;31:107716.
- Maihofer AX, Choi KW, Coleman JR, Daskalakis NP, Denckla CA, Ketema E, et al. Enhancing discovery of genetic variants for posttraumatic stress disorder through integration of quantitative phenotypes and trauma exposure information. *Biol Psychiatry.* 2022;91:626–36.
- Xie P, Kranzler HR, Yang C, Zhao H, Farrer LA, Gelernter J. Genome-wide association study identifies new susceptibility loci for posttraumatic stress disorder. *Biol Psychiatry.* 2013;74:656–63.
- Thyme SB, Pieper LM, Li EH, Pandey S, Wang Y, Morris NS, et al. Phenotypic landscape of schizophrenia-associated genes defines candidates and their shared functions. *Cell.* 2019;177:478–91.
- Sanchez-Roige S, Fontanillas P, Elson SL, Gray JC, de Wit H, MacKillop J, et al. Genome-wide association studies of impulsive personality traits (BIS-11 and UPPS-P) and drug experimentation in up to 22,861 adult research participants identify loci in the CACNA11 and CADM2 genes. *J Neurosci.* 2019;39:2562–72.
- Vorstman JA, van Daalen E, Jalali GR, Schmidt ER, Pasterkamp RJ, de Jonge M, et al. A double hit implicates DIAPH3 as an autism risk gene. *Mol Psychiatry.* 2011;16:442–51.
- Santarelli R. Information from cochlear potentials and genetic mutations helps localize the lesion site in auditory neuropathy. *Genome Med.* 2010;2:1–10.
- Weiss A, Baselmans BM, Hofer E, Yang J, Okbay A, Lind PA, et al. Personality polygenes, positive affect, and life satisfaction. *Twin Res Hum Genet.* 2016;19:407–17.
- Baselmans BML, Van de Weijer MP, Abdellaoui A, Vink JM, Hottenga JJ, Willemsen G, et al. A genetic investigation of the well-being spectrum. *Behav Genet.* 2019;49:286–97.
- Plomin R. Commentary: Missing heritability, polygenic scores, and gene–environment correlation. *J Child Psychol Psychiatry.* 2013;54:1147–9.
- Plomin R, Deary IJ. Genetics and intelligence differences: five special findings. *Mol Psychiatry.* 2015;20:98–108.
- George M, Maheshwari S, Chandran S, Manohar JS, Rao TS. Understanding the schizophrenia prodrome. *Indian J Psychiatry.* 2017;59:505.
- Cornblatt BA, Auther AM, Niendam T, Smith CW, Zinberg J, Bearden CE, et al. Preliminary findings for two new measures of social and role functioning in the prodromal phase of schizophrenia. *Schizophrenia Bull.* 2007;33:688–702.
- Gibson G. Rare and common variants: twenty arguments. *Nat Rev Genet.* 2012;13:135–45.
- Okbay A, Wu Y, Wang N, Jayashankar H, Bennett M, Nehzati SM, et al. Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nat Genet.* 2022;54:437–49.
- Cacioppo JT, Hawkey LC, Ernst JM, Burleson M, Bertson GG, Nouriani B, et al. Loneliness within a nomological net: An evolutionary perspective. *J Res Personal.* 2006;40:1054–85.
- Brent LNJ, Ruiz-Lambides A, Platt ML. Persistent social isolation reflects identity and social context but not maternal effects or early environment. *Sci Rep.* 2017;7:17791.
- Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* 2015;12:e1001779.
- Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature.* 2018;562:203–9.
- Hajek A, Kretzler B, König HH. The association between obesity and social isolation as well as loneliness in the adult population: a systematic review. *Diab Metab Syndr Obes.* 2021;14:2765–73.
- Emerson E, Fortune N, Llewellyn G, Stancliffe R. Loneliness, social support, social isolation and wellbeing among working age adults with and without disability: Cross-sectional study. *Disabil Health J.* 2021;14:100965.
- Boyd A, Golding J, Macleod J, Lawlor DA, Fraser A, Henderson J, et al. Cohort profile: the ‘children of the 90s’—the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol.* 2013;42:111–27.
- Fraser A, Macdonald-Wallis C, Tilling K, Boyd A, Golding J, Davey Smith G, et al. Cohort profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol.* 2013;42:97–110.
- Goodyer IM, Wright C, Altham PME. Recent friendships in anxious and depressed school age children. *Psychological Med.* 1989;19:165–74.
- Ripke S, Walters JT, O’Donovan MC, Schizophrenia Working Group of the Psychiatric Genomics Consortium. Mapping genomic loci prioritises genes and implicates synaptic biology in schizophrenia. *MedRxiv.* 2020. <https://doi.org/10.1101/2020.09.12.20192922v1>.
- Howard DM, Adams MJ, Clarke TK, Hafferty JD, Gibson J, Shirali M, et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat Neurosci.* 2019;22:343–52.
- Anney RJ, Ripke S, Anttila V, Grove J, Holmans PA, Huang H, et al. 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.* 2017;8:21.

48. Watson HJ, Yilmaz Z, Thornton LM, Hübel C, Coleman JR, Gaspar HA, et al. Genome-wide association study identifies eight risk loci and implicates metabolic-psychiatric origins for anorexia nervosa. *Nat Genet.* 2019;51:1207–14.
49. Stahl EA, Breen G, Forstner AJ, McQuillin A, Ripke S, Trubetskoy V, et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat Genet.* 2019;51:793–803.
50. Meier SM, Tronetti K, Purves KL, Als TD, Grove J, Laine M, et al. Genetic variants associated with anxiety and stress-related disorders: a genome-wide association study and mouse-model study. *JAMA Psychiatry.* 2019;76:924–32.
51. Van den Berg SM, de Moor MH, Verweij KJ, Krueger RF, Luciano M, Arias Vasquez A, et al. Meta-analysis of genome-wide association studies for extraversion: findings from the genetics of personality consortium. *Behav Genet.* 2016;46:170–82.
52. De Moor MH, Van Den Berg SM, Verweij KJ, Krueger RF, Luciano M, Vasquez AA, et al. Meta-analysis of genome-wide association studies for neuroticism, and the polygenic association with major depressive disorder. *JAMA Psychiatry.* 2015;72:642–50.
53. Abdellaoui A, Sanchez-Roige S, Sealock J, Treur JL, Dennis J, Fontanillas P, et al. Phenome-wide investigation of health outcomes associated with genetic predisposition to loneliness. *Hum Mol Genet.* 2019;28:3853–65.
54. Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M, et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat Genet.* 2018;50:1112–21.
55. Loh PR, Tucker G, Bulik-Sullivan BK, Vilhjalmsson BJ, Finucane HK, Salem RM, et al. Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat Genet.* 2015;47:284–90.
56. Turley P, Walters RK, Maghzian O, Okbay A, Lee JJ, Fontana MA, et al. Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nat Genet.* 2018;50:229–37.
57. Zhu Z, Zheng Z, Zhang F, Wu Y, Trzaskowski M, Maier R, et al. Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat Commun.* 2018;9:1–12.
58. Choi SW, O'Reilly PF. PRSice-2: Polygenic Risk Score software for biobank-scale data. *Gigascience.* 2019;8:giz082.
59. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, et al. An atlas of genetic correlations across human diseases and traits. *Nat Genet.* 2015;47:1236–41.
60. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenome. *elife.* 2018;7:e34408.
61. Grotzinger AD, Rhemtulla M, de Vlaming R, Ritchie SJ, Mallard TT, Hill WD, et al. Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nat Hum Behav.* 2019;3:513–25.

## ACKNOWLEDGEMENTS

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses. This work was supported in part by grant R01 MH128971 from the National Institute of Mental Health (AJS, EV, AR), and by the Beatrice and Samuel A Seaver Foundation (AJS, EV, AR). HJ & SZ are supported by the NIHR Biomedical Research Centre at University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol. RCG and REG are

supported by R01 MH119219. The UK Medical Research Council and Wellcome (Grant ref: 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors and Adam Socrates and Abraham Reichenberg will serve as guarantors for the contents of this paper.

## AUTHOR CONTRIBUTIONS

EV conceived the study. AJS, ES, and EV designed the analyses. AJS and EV interpreted the results and wrote the manuscript. AJS and HJ performed all the analyses and presented the results. AJS, NM, AR, PFO, SZ, RCG, and REG interpreted the results and revised the manuscript.

## COMPETING INTERESTS

ES is a current employee of Regeneron Pharmaceuticals, Inc.

## ETHICS APPROVAL AND CONSENT TO PARTICIPATE

All participants provided written informed consent and the current study was ethically approved by the UK Biobank Ethics and Governance Council (REC reference 11/NW/0382; UK Biobank application reference 18177). Ethical approval for the study was also obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41380-024-02617-2>.

**Correspondence** and requests for materials should be addressed to Adam J. Socrates.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© Crown 2024