Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Age and cohort differences in the association between polygenic scores and psychological distress: evidence from the 1958 and 1970 British Birth Cohorts

Esme Elsden

eelsden@ed.ac.uk

University of Edinburgh https://orcid.org/0000-0003-3867-3588

Alex Kwong

https://orcid.org/0000-0003-1953-2771

Andrew McIntosh

University of Edinburgh https://orcid.org/0000-0002-0198-4588

Ana Garfield

Praveetha Patalay

Article

Keywords:

Posted Date: March 3rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8958839/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors have declared there is **NO** conflict of interest to disclose

Title: Age and cohort differences in the association between polygenic scores and psychological distress: evidence from the 1958 and 1970 British Birth Cohorts

Authors:

Esme Elsden[1,2*], Alex S. F. Kwong[1], Andrew M. McIntosh[1], Victoria Garfield[3] & Praveetha Patalay[2,4]

Affiliations:

[1] Institute for Neuroscience and Cardiovascular Research, University of Edinburgh, Edinburgh, UK

[2] Centre for Longitudinal Studies, Social Research Institute, University College London, London, UK

[3] Department of Pharmacology and Therapeutics, University of Liverpool, Liverpool, UK

[4] Unit for Lifelong Health and Ageing, University College London, London, UK

*Corresponding Author:

Dr Esme Elsden

Division of Psychiatry
Edinburgh Royal Infirmary
49 Little France Cres
Edinburgh
EH16 4SB
eelsden@ed.ac.uk

Abstract

Psychological distress is influenced by genetic, environmental, and age-related factors. Although cohort differences in psychological distress have been evidenced, it is unclear whether these reflect genetic effects. We estimated the association between polygenic scores (PGS) for psychological distress and observed distress across adulthood in the 1958 National Child Development Study (1958c) and 1970 British Cohort Study (1970c). Psychological distress was measured using the Malaise Inventory between ages 23 and 50 across nine occasions. PGS were derived from a genome-wide association study of depressive symptoms. Age-specific associations were examined using linear regression, and cohort-stratified multilevel mixed-effects models tested PGS interactions with age and sex. Distress scores were higher in the 1970c and in females. In both cohorts, PGS were positively associated with psychological distress, with effect sizes increasing with age; incremental R^2 rose from 1.83% to 2.86% in the 1958c and from 1.44% to 2.16% in the 1970c. The 1958c displayed a steeper age-related increase in genetic effects, while the 1970c showed larger genetic effects in early adulthood (ages 26–30) that converged at older ages. A PGS-by-sex interaction was identified in the 1958c only, with females showing stronger genetic effects than males. Genetic liability to psychological distress exerts a modest but increasing influence across adulthood. Cohort differences appear driven more by environmental than genetic factors. The sex-specific genetic effect in the 1958c but not the 1970c suggests the interplay between genetic liability and sex varies across historical contexts, highlighting the need for sex-specific and cohort-specific approaches in future research.

Introduction

Psychological distress, encompassing the symptoms of anxiety and depression, poses a significant global public health concern [1]. The prevalence of psychological distress has been increasing, with anxiety and depression are leading causes of disease burden [2]. Psychological distress can vary in severity and symptom profiles across different stages of adulthood [3, 4]. This trajectory of distress is thought to be influenced by both genetic and environmental exposures, underscoring the importance of considering both factors when examining psychological distress. Prior studies have highlighted the complex interplay between genetic and environmental factors (GxE) in shaping psychological distress phenotypes and severity, where genetic contributions to distress can be amplified or attenuated by environmental contexts [5], such as the economic and social conditions experienced [6].

Genetic factors, such as polygenic scores (PGS), have been consistently shown to influence psychological distress, yet their utility in predicting distress is not uniform across all ages [7]. Previous research suggests genetic contributions to psychological traits, including distress, can vary with age, reflecting a dynamic interplay between genes and age-related environmental exposures [8]. For example, studies have observed that the heritability of depressive symptoms may increase in later adulthood. This may result from the expression of genetic effects becoming more pronounced as individuals age, experiencing fewer changes in their environments, and, therefore, have more stable life circumstances [9]. Conversely, other studies propose that the expression of genetic liability might be higher during periods of heightened life transitions, such as early adulthood or middle age, when individuals are exposed to significant life stressors (for example, employment changes, entry to parenthood, chronic illness onset, and care responsibilities) [8]. Existing research is limited in the conclusions it can draw due to analyses of different individuals at different ages, potentially conflating age and cohort effects.

Age-related manifestation of genetic risk for psychological distress is complex and may vary depending on the population studied, the underlying genetic architecture and environmental context throughout the lifespan. Environmental context moderates the genetic influence on psychological distress [10]. Research has shown that the same genetic predispositions can lead to different outcomes depending on the environmental context [11]. The environmental context can include socioeconomic status or life events at the personal level, and cultural norms at the societal level [12].

Cohort effects, where individuals born in the same period may experience shared environmental exposures that shape psychological outcomes differently compared to other cohorts [13]. Cohort effects differ from age effects, as they capture unique influences that affect a specific generation due to the historical and social circumstances they encounter [14]. The 1958 National Child Development Study (1958c) and the 1970 British Cohort Study (1970c) provide a valuable opportunity to explore these cohort effects [15, 16], the latter cohort experiencing higher levels of adulthood distress and an earlier mid-life peak in distress symptoms [3]. These cohorts, representing different socio-historical contexts in the UK, allow for a comparison of how genetic liability for psychological distress interacts with environmental conditions across generations [17, 18]. These cohorts' distinct socioeconomic landscapes and varying life experiences offer a framework to examine whether the genetic influence on distress differs by cohort.

Sex differences may moderate the strength of the association between polygenic scores and psychological distress by birth cohort [19]. However, Conley et al., tested whether the association between genetic liability to depression and psychological distress differed by sex across cohorts [9]. They found no significant sex differences in this association, nor did it vary significantly across cohorts. However, the three-way interaction used to test this may have been underpowered. Some evidence suggesting a clear gene-environment interaction effect for distress and sex comes from the candidate gene literature, though it is important to note that candidate gene findings have largely failed to replicate and much of this literature is now considered unreliable. The serotonin transporter gene (5-HTTLPR) showed gene-environment interaction effects for depressive symptoms in females compared to males [20]. However, one study found the 5-HTTLPR genotype interacted with county-level environmental factors (e.g. deprivation) to influence depressive symptoms in adolescent males but not females [21]. Given the well-documented replication failures in the candidate gene literature, these findings should be interpreted with caution; nevertheless, they illustrate the plausibility of sex-differentiated gene-environment interactions for distress and underscore the need for more methodologically rigorous investigation using polygenic approaches. In summary, while there are some contradictory findings, some studies suggest sex-specific gene-environment interactions contribute to psychological distress. Therefore, this study will further explore whether there are sex differences in genetic liability to psychological distress.

The current study aims to explore how the association between polygenic scores for psychological distress and observed psychological distress varies across adulthood in two British birth cohorts—the 1958c and the 1970c. By analysing data from ages 23 to 50 across nine-time points, this study investigates whether the relationship between polygenic score of psychological distress and observed psychological distress manifests differently across ages and cohorts with distinct environmental exposures. In addition, given well-established sex differences in the prevalence of psychological distress across ages, we also investigate an interaction by sex across ages and cohorts.

Materials and Methods

Participants

The 1958 National Child Development Study is a nationally representative longitudinal birth cohort study following the lives of over 17,415 children born in a single week of March 1958, with 9,100 individuals included in the latest wave (2013-14). Data was collected at birth and from ages 7, 11, 16, 23, 33, 42, 44, 46, 50, 55 [15]. The 1970s British Birth cohort is a nationally representative longitudinal birth cohort study of children born in a single week in 1970 in England, Scotland, and Wales. Data collection sweeps were conducted at birth, and then at ages 5, 10, 16, 26, 30, 34, 38, 42, 46, and 50 between 1970 and 2022 [16]

Genetic data was collected in both cohorts. For the 1958c, a biomedical sweep between 2002 and 2004 at ages 44-46 collected blood samples from 9,377 individuals. The 1970c also collected blood samples at age 46 biomedical sweep from 8,581

individuals. Full details on imputation and quality control of genetic variants are provided in the Supplementary Information.

Sample

To be included in the analytic sample, participants must have contributed data at least once for the malaise inventory score and provided genetic data. The present study included 6,312 participants from 1958c and 5,423 from 1970c.

Measures

Outcome: Psychological distress was measured in both cohorts. The 24-item version was used at ages 23, 33, and 42. The nine-item version of the Malaise Inventory was collected at ages 23, 33, 42, and 50 in the 1958c and at ages 26, 30, 34, and 46 in the 1970c [22, 23]. To increase occasions, the nine-item was used to enable harmonisation across cohorts. The Malaise Inventory ranges from 0 to 9, with higher scores indicating greater psychological distress. Psychological distress captures depression and anxiety symptoms [24]. Participants were asked items such as “Do you often feel depressed?” with a yes/no response for reporting how they felt generally. In both cohorts, the Malaise items were assessed via written self-completion. It has been used in general population studies and in investigations of high-risk groups [25]. The Malaise Inventory was previously found to be reliable, with fair internal consistency (Cronbach's alpha = 0.77 at age 23 and 0.80 at age 33 [26]; 0.70 to 0.80 is deemed good). The Malaise Inventory score has both configural and scalar measurement invariance across ages within each cohort and between the two cohorts [27]. This means the relationships between the latent construct of psychological distress and the observed items remained stable across different cohorts and time points. Therefore, the pattern in responses to items remained the same over time, enabling mean score comparisons.

Exposure: A polygenic score (PGS) for psychological distress was constructed using summary statistics from a large genome-wide association study (GWAS) of depressive symptom measurement (Baselmans et al., 2019; N = 1,067,913) [28]. This GWAS meta-analysis was conducted across cohorts that used depressive symptom measurement scales such as the Patient Health Questionnaire, the General Health Questionnaire, and the Centre for Epidemiologic Scale – Depression [29–31]. The 1958 National Child Development Study and the 1970 British Cohort Study were not included in the discovery GWAS. Polygenic scores were generated using PRSice-2. The primary exposure was a standardised PGS constructed at a p-value threshold of 0.005, consistent with prior work on related depressive symptom phenotypes [32]. To ensure comparability across cohorts, the PGS was jointly standardised (mean = 0, SD = 1), with higher scores indicating a greater polygenic propensity toward psychological distress. More information on genetic data imputation, quality control and polygenic score methods is in the Supplementary Materials.

Covariates: sex and 10 principal components (PCs) were adjusted for in all the non-sex-stratified analyses. The PCs summarise the distribution of genetic variation across geographical locations and ancestral backgrounds [33]. PCs are derived from genetic data and capture patterns of genetic diversity in the population. These PCs are included to control for residual genetic variation across geographical locations and

ancestral backgrounds, helping maintain a homogeneous population in the analysis (81).

Analytical Strategy

All analyses were performed in Stata version 17 and R v 4.3.2. Descriptive statistics were summarised to give the mean and 95% confidence intervals of the Malaise Inventory.

Cross-sectional age-specific associations

Cross-sectional linear regression analyses were performed to investigate the relationship between the polygenic score for psychological distress and the Malaise Inventory score at ages 23, 26, 30, 33, 34, 42, 46, and 50. Each model controlled for 10 principal components and sex as covariates. Effect sizes were compared visually by plotting the coefficient and incremental R^2 estimates. Beta coefficients were used to assess the strength of the association between polygenic scores and observed psychological distress. Incremental R^2 were calculated by deducting the null model without the PGS from the full model with the PGS at each age. This was used to assess the proportion of variance in psychological distress attributable to PGS, thereby representing PGS heritability. To account for multiple testing across time points (15 thresholds × 9 occasions = 135 tests), p-values were adjusted using the Benjamini–Hochberg false discovery rate (FDR) procedure, with an FDR threshold of 0.05 [34]. Confidence intervals were calculated using bootstrapping with 1,000 replications.

Multiple imputation was employed to handle missing data in the Malaise Inventory score for the cross-sectional analyses. It was used under the assumption that data are missing at random to improve estimation accuracy [35]. Multiple imputations utilised auxiliary variables to address missingness. Auxiliary variables included sex, father's social class, the polygenic score for psychological distress, crowding in childhood, whether breastfed, whether parents divorced during participants' childhood, family difficulties due to mental illness, poor maternal mental health, low birth weight, well-being alongside internalising and externalising behaviours during childhood [36–38].

Longitudinal age-patterns, cohort and sex differences

To examine age-related and cohort-specific patterns in PGS effects, a stepwise analytical strategy was employed. Separate mixed-effects models were fitted for each cohort to examine the interaction between the polygenic score and age. Age was modelled as a linear term. This approach enables age-specific estimation of PGS effects within each cohort over time, maximising the benefits of mixed-effects modelling, including accounting for correlation in repeated measures and utilising all available data points. Confidence intervals were bootstrapped. For each cohort, marginal effects of the PGS were calculated at each age, with age being mean-centred, within the follow-up range (e.g., ages 23 to 50). This step provides insights into how the PGS effect varies by age within each cohort. For overlapping ages across cohorts, the differences in marginal PGS effects at each age were calculated.

Following this age-cohort analysis, we tested for sex differences with mixed-effects models in the separate cohorts and tested interactions between the polygenic score

and sex while controlling for covariates. For each cohort, the marginal effects of PGS and sex were calculated at each age, with age mean-centred within the follow-up range (e.g., ages 23 to 50). This step provides insights into how the PGS effect varies by sex across age within each cohort.

Sensitivity Analysis

Specification curve plots were used to assess the robustness of findings across 15 p-value thresholds, ranging from $p<5\times10^{-8}$ to $p<0.5$. As the optimal threshold for capturing genetic liability to psychological distress is unknown a priori and may vary across populations and developmental periods, examining consistency of findings across thresholds strengthens confidence that results are not artefacts of arbitrary threshold selection and enhances comparability with studies employing different thresholding approaches.

Results

Descriptive Statistics

Mean psychological distress scores were consistently higher in the 1970 cohort than the 1958 cohort across comparable ages, and females showed higher mean distress than males in both cohorts at every time point (Figure 1; Supplementary Table 1). Both cohorts displayed a dip in distress in early-to-mid adulthood, followed by an increase in midlife, though this pattern was more pronounced in the 1958 cohort (Figure 1). The 1970 cohort showed greater variability in distress scores at each time point compared to the 1958 cohort, and females showed greater variability than males consistently across both cohorts and all ages (Supplementary Table 1)

Cross-sectional age-specific associations

Polygenic scores for psychological distress and observed distress levels (measured by the Malaise Inventory) showed consistent, increasing associations in both the 1958 and 1970 cohorts (Figure 2). Beta coefficients ranged from 0.21 to 0.33, and effect sizes increased with age in both cohorts (Table 1).

At the youngest ages (23 and 26) , the association was $\beta = 0.21$ [95% CI: 0.17, 0.25], $p^{FDR}=1.12 \times 10^{-23}$ in the 1958c, and $\beta = 0.23$ [95% CI: 0.17, 0.28], $p^{FDR}=1.31 \times 10^{-15}$ in the 1970c. At age 30 in the 1970c, the association was $\beta = 0.21$ [95% CI: 0.16, 0.26], $p^{FDR}=1.31 \times 10^{-15}$. At age 33 in the 1958c, the association was $\beta = 0.21$ [95% CI: 0.17, 0.24], $p^{FDR}=5.15 \times 10^{-26}$, while at age 34 in the 1970c it was $\beta = 0.26$ [95% CI: 0.21, 0.32], $p^{FDR}=2.20 \times 10^{-19}$. At the overlapping age of 42, effect sizes were similar across cohorts (1958c: $\beta = 0.27$ [95% CI: 0.23, 0.32], $p^{FDR}=6.97 \times 10^{-34}$; 1970c: $\beta = 0.26$ [95% CI: 0.21, 0.33], $p^{FDR}=2.20 \times 10^{-19}$). At the oldest ages, the association strengthened further (1958c at age 50: $\beta = 0.33$ [95% CI: 0.27, 0.37], $p^{FDR}=1.40 \times 10^{-35}$; 1970c at age 46: $\beta = 0.31$ [95% CI: 0.24, 0.36], $p^{FDR}=3.23 \times 10^{-21}$).

The proportion of variance in Malaise Inventory scores explained by polygenic scores showed a modest contribution of genetic factors to psychological distress throughout adulthood (Figure 2b). Incremental R-squared values ranged from 1.83% [95%CI: 1.11%-2.52%] at age 23 to 2.86% [95%CI: 1.94%-3.68%] at age 50 in the 1958c, and from 1.80% [95%CI: 1.09%-2.69%] at age 26 to 2.16% [95%CI: 1.39%-3.19%] at age 46 in the 1970c, with higher values generally observed at older ages. There was a

cohort difference at the comparable age of 42 where 1970c had lower variance than 1958c: 1.85% [95%CI: 1.08%-2.75%] versus 2.46% [95%CI: 1.64%-3.22%].

Specification curve plots indicated that the choice of p-value threshold for the polygenic score moderately influenced the explained variance in distress, with higher thresholds including more SNPs and explaining more variance, albeit with the risk of including variants related to other traits and only marginally associated with psychological distress (Supplementary Figure 1).

Longitudinal age-patterns, cohort and sex differences

In both cohorts, the polygenic score was independently associated with psychological distress (β_{1958c}= 0.24, [95%CI: 0.23, 0.26], p=5.26 x 10^{-50}; β_{1970c} = 0.25, [95%CI: 0.23, 0.27], p=7.95 x 10^{-32}) (Table 2). There was evidence of an interaction between polygenic score and age in both cohorts (β_{1958c} = 0.004, [95% CI: 0.003, 0.005], p=4.49 x 10^{-7}; β_{1970c} = 0.004, [95%CI: 0.001, 0.006], p=4.46 x 10^{-3}) (Table 2). The predictive margins plot (Figure 3) showed that the 1970 cohort had a higher marginal effect at younger ages, but both cohorts converged at older age points. The 1958 cohort showed a steeper age-related increase in genetic effects.

The multilevel model revealed sex differences in psychological distress. Females showed higher levels of observed psychological distress compared to males in both cohorts (β_{1958c} = 0.66, [95%CI: 0.60, 0.62], p=1.04 x 10^{-91}; β_{1970c} = 0.53, [95%CI: 0.45, 0.62], p=1.42 x 10^{-35}) (Table 2). The PGS was predictive of observed distress outcomes in both cohorts, with the 1970c having a marginally higher beta coefficient of 0.24 (95% CI [0.18, 0.30], p=1.82 x 10^{-15}) compared to a beta of 0.20 ([95% CI: 0.16, 0.25], p=8.06 x 10^{-18}) in the 1958c. There was some evidence of an interaction between the polygenic score and sex in the 1958c (β_{1958c} = 0.09, [95% CI: 0.01, 0.11], p=5.01 x 10^{-3}), indicating that the effect of genetic liability for psychological distress was greater among females than males. However, there was no evidence of an interaction in the 1970c (β_{1970c} = 0.02, [95%CI: -0.06, 0.11], p=0.584). Marginal effect estimates derived from the age-centred model confirmed these results and illustrate the 1958c interaction more clearly (Figure 4a & 4b; Supplementary Figure 3).

Missingness

To examine potential sample bias in the 1958 and 1970 cohorts, we conducted multiple logistic regressions to evaluate patterns of missingness.

First, we assessed whether missing data at the oldest time points (age 50 in the 1958 cohort and age 46 in the 1970 cohort) was associated with higher psychological distress scores at the youngest time points (age 23 in 1958c and age 26 in 1970c). Logistic regression analyses revealed that in both cohorts, higher distress at younger ages significantly increased the odds of missingness at later ages, with a 57% increase in the 1970 cohort (OR=1.57, [95%CI: 1.35-1.82], p=0.001) and a 67% increase in the 1958 cohort (OR=1.67, [95%CI: 1.42-1.95], p=0.001).

Second, we evaluated whether higher psychological distress scores were linked to non-participation in genetic data collection in both cohorts. Higher distress scores were associated with a 20% increase in the odds of not contributing genetic data in the 1958 cohort (OR=1.20, [95%CI: 1.03-1.41], p=0.024) and a 52% increase in the 1970 cohort

(OR=1.52, [95%CI: 1.30-1.77], p=0.001). As PGS cannot be computed for non-participants, the inverse question — whether genetic liability to distress predicts non-participation — cannot be directly tested; however, these findings suggest the genotyped samples may be selected toward lower distress and, by extension, lower genetic liability, which should be considered when interpreting PGS associations.

In the 1958 cohort, 34% of participants (N=3,748) had dropped out by age 50. Those missing at this age showed higher psychological distress at age 23 (Mean = 1.40, SD = 1.72) compared to those who completed follow-up at age 50 (Mean = 1.24, SD = 1.58). Similarly, in the 1970 cohort, 41% of participants (N=2,312) had dropped out by age 46, with higher distress at age 26 for those who were missing at age 46 (Mean = 1.92, SD = 2.27) compared to those who completed follow-up (Mean = 1.71, SD = 2.07).

Finally, we tested whether there was a sample bias due to sex. The polygenic scores differed significantly by sex (p=0.001), whereby females had a higher mean PGS (Mean=0.004 SD=1.00) compared to males (Mean=-0.01, SD=0.99).

Discussion

This study found that the association between polygenic scores for psychological distress and observed distress levels increased in strength across adulthood. However, the proportion of variance explained by polygenic scores remained modest, rising from approximately 1.8% in early adulthood to 2.9% by age 50 in the 1958 cohort and 2.2% by age 46 in the 1970 cohort. Both cohorts showed evidence of an interaction between polygenic scores and age, with the 1970 cohort experiencing larger effects in early adulthood (26-30y), which converged with 1958 at older ages. The 1958 showed a steeper age-related increase in genetic effects on psychological distress. Sex differences emerged, with females showing both higher levels of observed distress and stronger genetic effects in the 1958 cohort, where a polygenic score-by-sex interaction was identified. The relatively small genetic contribution points to the potential role of changing societal conditions and environmental contexts in shaping female psychological distress in younger cohorts

Age Differences

Our study found that the direction of association between polygenic scores for psychological distress and observed distress levels is consistent across adulthood. This finding aligns with previous research by Nivard et al. (2015), who reported that genetic influences on internalising problems are present and relatively stable across adulthood compared to childhood [39]. However, our findings are inconsistent with Nivard's decreasing additive genetic effects at older ages. We observed a trend of increasing effect sizes with age in both cohorts. This observation is consistent with the findings of Kendler et al. (2008), using data from the Swedish Twin Study of Child and Adolescent Development, who reported that genetic factors became increasingly important in explaining variance in depressive symptoms as individuals age [40].

One explanation for the increasing impact of genetic liability with age may be explained by gene-environment interplay, with individuals increasingly being selected into or choosing environments that align with their genetic predispositions as they age [41].

For instance, individuals with a genetic predisposition to psychological distress may be more likely to avoid stressful situations, thereby reinforcing their anxious tendencies over time [18]. Another explanation could be specific genetic factors may only become relevant or active later in life, contributing to the observed increase in genetic influence with age [42]. A methodological explanation of the study's finding might also reflect using a GWAS predominantly derived from middle-aged samples [43]. This could mean the SNPs identified are more relevant to mid-life psychological distress, as different genes may be expressed at different life stages [44]. A further methodological consideration is individuals with higher genetic liability who experience severe distress early in life may be less likely to participate in studies at older ages. While this selective dropout might reduce the variation, it does not necessarily explain why genetic effects would become greater in magnitude on the absolute scale. Additionally, this could affect the proportion of variance explained (R^2) by genetic factors, as a more homogeneous sample might make genetic influences appear more pronounced (37).

Our results indicate a modest contribution of genetic factors to psychological distress throughout adulthood. The proportion of variance in Malaise Inventory scores explained by polygenic scores (incremental R-squared values) ranged from 1.83% to 2.86% in the 1958 cohort, and from 1.44% to 2.16% in the 1970 cohort, with higher values generally observed at older ages. These findings are consistent with previous studies that have reported modest effect sizes for polygenic scores in predicting psychological traits [45]. The increase in the variance explained by polygenic scores over time may also reflect cumulative gene expression, overshadowing short-term environmental influences or the activation of new genetic pathways later in life [46]. The modest explained variance underscores the complex aetiology of psychological distress, suggesting that while genetic factors play a role, environmental and other non-genetic factors still account for a substantial portion of the variability in distress levels. Twin studies have decomposed these factors into additive genetic effects, the non-shared environment and the shared environment [39, 47]. This highlights the importance of considering both genetic and environmental factors in understanding and addressing psychological distress.

Cohort Differences

Individuals born in the 1958 cohort showed lower psychological distress scores compared to those born in the 1970 cohort. This aligns with previous research by Gondek et al. (2021, 2022), who reported higher distress levels in more recent cohorts [3, 4]. The relationship between genetic liability and psychological distress showed similar patterns across cohorts. The 1970 cohort demonstrated marginally higher genetic effects in early adulthood compared to the 1958 cohort, but these differences converged at older ages. Both cohorts showed evidence of gene-age interactions, with the 1958 cohort displaying a steeper increase in genetic effects across age. While the 1970 cohort experienced higher overall distress levels, particularly in early adulthood (ages 26-30) the genetic contribution to these differences was modest. The convergence of genetic effects at older ages, despite maintained cohort differences in distress levels, suggests environmental factors may play a more substantial role in cohort differences than genetic liability [48]. Even with increases in the predictive ability of polygenic scores with larger sample sizes, a recent family-based GWAS found shrinkage in direct genetic effect estimates on depression phenotype from ~20% to ~2.5% [49]. These results align with research across multiple countries that have

found increasing levels of psychopathology in more recent birth cohorts, attributing these changes to societal shifts precipitating a depressogenic environment, rather than genetic factors [1, 50–53]. Genes and the environment are contingently entangled, so it may be through gene-environment correlations that the effect of genes is through indirect pathways [54].

Sex Differences

Our study identified an interaction between the polygenic score and sex only in the 1958c. Previous research has shown sex differences in the genetic architecture of depression and anxiety disorders, which corroborates the 1958c as the data used in these studies was collected at a similar time point, meaning the sample may share the same cohort [19]. However, measurement errors may undermine PGS accuracy, as most European-ancestry GWAS studies have historically had a female bias in sampling, and the current study may also exhibit a potential sample bias towards females. However, there is a lack of evidence of sexual interaction in the 1970c females, even though they have higher mean levels of distress and more variability in distress scores. Therefore, another explanation could be certain environmental stressors may have more of an effect on females, which may reflect broader societal or environmental influences that exacerbate distress in females more than in males for the 1970c [55]. The lack of evidence for an interaction in the 1970c suggests no consistent trend in sex differences in genetic effects.

Strengths and Limitations

Strengths include the use of data from two nationally representative longitudinal cohorts with consistent measures of genetic exposure and psychological distress. The study employed a polygenic score derived from a genome-wide association study (GWAS) of depressive symptoms. This score served as an effective proxy for psychological distress, as the GWAS phenotype closely aligned with the Malaise Inventory score through comparable depressive symptom measurement scales. The phenotypic similarity between the two measures strengthened the validity of using the polygenic score in this context. This design allows for a direct comparison of age and cohort effects within the same framework, addressing a gap in previous research that often relied on cross-sectional data or single-cohort longitudinal studies. A further strength is the use of specification curve analysis across 15 p-value thresholds. Psychological traits are highly polygenic, with heritability potentially residing in common variants with small individual effects that fall below genome-wide significance thresholds. Stricter thresholds capture only variants with robust GWAS associations, while more liberal thresholds incorporate variants with small effects that may collectively better predict phenotypic variance. Demonstrating consistency of findings across this range supports the robustness of the observed PGS-distress associations and reduces concern that results reflect arbitrary analytical decisions.

Although we used consistent measures across cohorts, there may be subtle differences in how distress is experienced or reported across generations that are not captured by our instruments. However, the malaise inventory score showed configural and scalar measurement invariance in both cohorts, enabling mean-difference comparisons across ages [27]. The polygenic scores used in this study assume that

genetic liability factors for depressive symptoms are stable over time, based on the premise that SNPs do not vary significantly across cohorts. Therefore, this GWAS approach may not isolate cohort-specific genetic signals due to the inclusion of multiple generations in the sample, which may prevent the detection of cohort differences. The present study uses birth year as an exogenous variable, allowing for the estimation of gene-cohort interactions and addressing concerns of environmental endogeneity. Several sources of sample bias warrant consideration when interpreting these findings. Genetic data were obtained from a subset of the total study participants. This sampling approach may introduce bias into the estimates, potentially leading to underestimation of the true effects. The reason for this is the 'healthy survivor' phenomenon: individuals experiencing higher levels of distress are more likely to drop out in later waves of cohort studies. This attrition issue is a persistent challenge in longitudinal research [36, 38]. Participants with higher psychological distress were more likely to be missing at follow-up and less likely to contribute genetic data, suggesting the analytic samples are selected toward lower distress and potentially lower genetic liability. Additionally, the observed sex differences in polygenic score distributions are unexpected, given that the PGS was derived from autosomal chromosomes and should theoretically be balanced across sexes in a representative sample, suggesting further potential selection bias in who contributed genetic data. Together, these patterns suggest that effect sizes for PGS associations with distress may be underestimated, and that sex-stratified findings should be interpreted with caution. Lastly, the study focused on a predominantly white British population, as these cohorts preceded the increase in migration in the latter half of the 20th century [56]. Future research should investigate whether these genetic associations hold across different ancestral backgrounds and ages, as genetic influences vary across populations [57].

Conclusion

This study confirms that polygenic scores for psychological distress have an increasing influence across adulthood psychological distress, but similar across generations. Genetic liability does not fully explain the higher distress levels observed in younger cohorts, highlighting the importance of considering both genetic and environmental factors in understanding psychological distress across generations. Future research should explore specific environmental stressors or mechanisms and include more diverse cohorts for comprehensive analysis. The observed sex differences in genetic influence underscore the need for sex-specific approaches when addressing psychological distress across the lifespan.

Acknowledgements

The authors would like to thank Professor David Bann and Dr Liam Wright for their valuable comments on earlier drafts of this manuscript and for their contributions to the conceptualisation of the work. We also thank Dr Gemma Shireby for her support and guidance with the polygenic scoring procedures.

Conflict of Interest

The authors have no conflicts of interest to declare.

References

1. James SL, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet. 2018;392:1789–1858.
2. World Health Organization. Global burden of mental disorders and the need for a comprehensive, coordinated response from health and social sectors at the country level: Report by the Secretariat. World Health Organisation; 2011.
3. Gondek D, Bann D, Patalay P, Goodman A, McElroy E, Richards M, et al. Psychological distress from early adulthood to early old age: Evidence from the 1946, 1958 and 1970 British birth cohorts. Psychological Medicine. 2021. 2021. https://doi.org/10.1017/S003329172000327X.
4. Gondek D, Lacey RE, Blanchflower DG, Patalay P. How is the distribution of psychological distress changing over time? Who is driving these changes? Analysis of the 1958 and 1970 British birth cohorts. 2021. 1 June 2021. https://doi.org/10.1101/2021.05.31.21257562.
5. Akingbuwa WA, Hammerschlag AR, Bartels M, Middeldorp CM. Systematic Review: Molecular Studies of Common Genetic Variation in Child and Adolescent Psychiatric Disorders. Journal of the American Academy of Child & Adolescent Psychiatry. 2022;61:227–242.
6. Ye J, Wen Y, Sun X, Chu X, Li P, Cheng B, et al. Socioeconomic Deprivation Index Is Associated With Psychiatric Disorders: An Observational and Genome-wide Gene-by-Environment Interaction Analysis in the UK Biobank Cohort. Biological Psychiatry. 2021;89:888–895.
7. Mitchell BL, Campos AI, Whiteman DC, Olsen CM, Gordon SD, Walker AJ, et al. The Australian Genetics of Depression Study: New Risk Loci and Dissecting Heterogeneity Between Subtypes. Biological Psychiatry. 2022;92:227–235.
8. Athanasiadis G, Meijsen JJ, Helenius D, Schork AJ, Ingason A, Thompson WK, et al. A comprehensive map of genetic relationships among diagnostic categories based on 48.6 million relative pairs from the Danish genealogy. Proc Natl Acad Sci USA. 2022;119:e2118688119.
9. Conley D, Laidley TM, Boardman JD, Domingue BW. Changing Polygenic Penetrance on Phenotypes in the 20th Century Among Adults in the US Population. Sci Rep. 2016;6:30348.
10. Belsky J, Pluess M. Beyond diathesis stress: Differential susceptibility to environmental influences. Psychological Bulletin. 2009;135:885–908.
11. Keers R, Pluess M. Childhood quality influences genetic sensitivity to environmental influences across adulthood: A life-course Gene × Environment interaction study. Dev Psychopathol. 2017;29:1921–1933.
12. Bronfenbrenner U. Environments in developmental perspective: Theoretical and operational models. In: Friedman SL, Wachs TD, editors. Measuring environment across the life span: Emerging methods and concepts., Washington: American Psychological Association; 1999. p. 3–28.
13. Kuh D, Ben-Shlomo Y, Lynch J, Hallqvist J, Power C. Life Course Epidemiology. Journal of Epidemiology and Community Health. 2003;57:778–783.

14. Bell A. Age period cohort analysis: a review of what we should and shouldn't do. Annals of Human Biology. 2020;47:208–217.
15. Power C, Elliott J. Cohort profile: 1958 British birth cohort (National Child Development Study). International Journal of Epidemiology. 2006;35:34–41.
16. Sullivan A, Brown M, Hamer M, Ploubidis GB. Cohort Profile Update: The 1970 British Cohort Study (BCS70). International Journal of Epidemiology. 2023;52:e179–e186.
17. Mills MC, Tropf FC. Sociology, Genetics, and the Coming of Age of Sociogenomics. Annu Rev Sociol. 2020;46:553–581.
18. Boyce WT, Ellis BJ. Biological sensitivity to context: I. An evolutionary–developmental theory of the origins and functions of stress reactivity. Develop Psychopathol. 2005;17.
19. Silveira PP, Pokhvisneva I, Howard DM, Meaney MJ. A Sex-Specific Genome-Wide Association Study of Depression Phenotypes in UK Biobank. 2022. 31 March 2022. https://doi.org/10.1101/2022.03.30.22273201.
20. Zhao L, Han G, Zhao Y, Jin Y, Ge T, Yang W, et al. Gender Differences in Depression: Evidence From Genetics. Front Genet. 2020;11:562316.
21. Uddin M, Koenen KC, De Los Santos R, Bakshis E, Aiello AE, Galea S. Gender differences in the genetic and environmental determinants of adolescent depression. Depress Anxiety. 2010;27:658–666.
22. Rutter M, Tizard J, Whitmore K. Education, Health And Behaviour. London: Longman; 1970. London: Longman; 1970.
23. Mcelroy E, Villadsen A, Patalay P, Goodman A, Richards M, Northstone K, et al. Harmonisation and measurement properties of mental health measures in six British cohorts. 2021.
24. Drapeau A, Marchand A, Beaulieu-Prevost D. Epidemiology of Psychological Distress. In: LAbate L, editor. Mental Illnesses - Understanding, Prediction and Control, InTech; 2012.
25. Furnham A, Cheng H. The stability and change of malaise scores over 27years: Findings from a nationally representative sample. Personality and Individual Differences. 2015;79:30–34.
26. Rodgers B, Pickles A, Power C, Collishaw S, Maughan B. Validity of the Malaise Inventory in general population samples. Social Psychiatry and Psychiatric Epidemiology. 1999;34:333–341.
27. Ploubidis GB, McElroy E, Moreira HC. A longitudinal examination of the measurement equivalence of mental health assessments in two british birth cohorts. Longitudinal and Life Course Studies. 2019;10:471–489.
28. Baselmans BML, Jansen R, Ip HF, van Dongen J, Abdellaoui A, van de Weijer MP, et al. Multivariate genome-wide analyses of the well-being spectrum. Nature Genetics. 2019;51:445–451.
29. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population.
30. Goldberg D. Manual of the general health questionnaire. Nfer Nelson; 1978.
31. Kroenke K, Spitzer RL, Williams JBW. Patient Health Questionnaire-9. 2011.
32. Kwong ASF, Morris TT, Pearson RM, Timpson NJ, Rice F, Stergiakouli E, et al. Polygenic risk for depression, anxiety and neuroticism are associated with the severity and rate of change in depressive symptoms across adolescence. Journal of Child Psychology and Psychiatry and Allied Disciplines. 2021. 2021. https://doi.org/10.1111/jcpp.13422.

33. Privé F, Luu K, Blum MGB, McGrath JJ, Vilhjálmsson BJ. Efficient toolkit implementing best practices for principal component analysis of population genetic data. Bioinformatics. 2020;36:4449–4457.
34. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B: Statistical Methodology. 1995;57:289–300.
35. Li P, Stuart EA, Allison DB. Multiple Imputation: A Flexible Tool for Handling Missing Data. JAMA. 2015;314:1966.
36. Centre for Longitudinal Studies. NCDS response and handling missing data. 2024. https://cls.ucl.ac.uk/cls-studies/1958-national-child-development-study/ncds-response-and-handling-missing-data/.
37. Mostafa T, Wiggins RD. Handling attrition and non-response in the 1970 British Cohort Study. Centre for Longitudinal Studies; 2014.
38. Katsoulis M, Narayanan M, Dodgeon B, Ploubidis G, Silverwood R. A data driven approach to address missing data in the 1970 British birth cohort. 2024.
39. Nivard MG, Dolan CV, Kendler KS, Kan KJ, Willemsen G, Van Beijsterveldt CEM, et al. Stability in symptoms of anxiety and depression as a function of genotype and environment: A longitudinal twin study from ages 3 to 63 years. Psychological Medicine. 2015;45:1039–1049.
40. Kendler KS, Gardner CO, Lichtenstein P. A developmental twin study of symptoms of anxiety and depression: evidence for genetic innovation and attenuation. Psychol Med. 2008;38:1567–1575.
41. Scarr S, McCartney K. How People Make Their Own Environments: A Theory of Genotype --> Environment Effects. Child Development. 1983;54:424.
42. Jiang X, Holmes C, McVean G. The impact of age on genetic risk for common diseases. PLoS Genet. 2021;17:e1009723.
43. Hewitt J, Walters M, Padmanabhan S, Dawson J. Cohort profile of the UK Biobank: diagnosis and characteristics of cerebrovascular disease. BMJ Open. 2016;6:e009161.
44. Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu X, Li M, et al. Spatio-temporal transcriptome of the human brain. Nature. 2011;478:483–489.
45. Mitchell BL, Thorp JG, Wu Y, Campos AI, Nyholt DR, Gordon SD, et al. Polygenic Risk Scores Derived From Varying Definitions of Depression and Risk of Depression. JAMA Psychiatry. 2021:1–9.
46. Wolf S, Melo D, Garske KM, Pallares LF, Lea AJ, Ayroles JF. Characterizing the landscape of gene expression variance in humans. PLoS Genet. 2023;19:e1010833.
47. Polderman TJC, Benyamin B, De Leeuw CA, Sullivan PF, Van Bochoven A, Visscher PM, et al. Meta-analysis of the heritability of human traits based on fifty years of twin studies. Nature Genetics. 2015;47:702–709.
48. Ferri E, Bynner J, Wadsworth M, editors. Changing Britain, changing lives: three generations at the turn of the century. London: Inst. of Education, Univ. of London; 2003.
49. Tan T, Jayashankar H, Guan J, Nehzati SM, Mir M, Bennett M, et al. Family-GWAS reveals effects of environment and mating on genetic associations. 2024.
50. Bell A. Life-course and cohort trajectories of mental health in the UK, 1991–2008 – A multilevel age–period–cohort analysis. Social Science & Medicine. 2014;120:21–30.

51. Beller J. Age-period-cohort analysis of depression trends: are depressive symptoms increasing across generations in Germany? Eur J Ageing. 2022;19:1493–1505.
52. Keyes KM, Gary D, O'Malley PM, Hamilton A, Schulenberg J. Recent increases in depressive symptoms among US adolescents: trends from 1991 to 2018. Soc Psychiatry Psychiatr Epidemiol. 2019;54:987–996.
53. Okui T. An age-period-cohort analysis for prevalence of common psychiatric disorders in Japan, 1999–2017. Soc Psychiatry Psychiatr Epidemiol. 2021;56:639–648.
54. Jaffee SR, Price TS. Gene–environment correlations: a review of the evidence and implications for prevention of mental illness. Mol Psychiatry. 2007;12:432–442.
55. Martin J, Khramtsova EA, Goleva SB, Blokland GAM, Traglia M, Walters RK, et al. Examining Sex-Differentiated Genetic Effects Across Neuropsychiatric and Behavioral Traits. Biological Psychiatry. 2021;89:1127–1137.
56. Office for National Statistics. Explore 50 years of international migration to and from the UK. 2024.
57. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. Nature Genetics. 2019;51:584–591.

692

Tables

Table 1: Linear regression associations between polygenic score for psychological distress with overlapping SNPs and 9-item Malaise Inventory Score at ages 23-50 in 1958c and 1970c (PGS Threshold=0.005, 1958c N=6,312; 1970c N=5,423)

Year	Age	Cohort	Mean (SD)	R^2	R^2 95% CI Bootstrapped	Beta	Beta 95%CI Bootstrapped	P-value	FDR p-value
1981	23	1958c	1.15 (1.50)	1.83%	(1.11%-2.52%)	0.21	(0.17-0.25)	1.12×10^{-23}	1.12×10^{-23}
1996	26	1970c	1.63 (1.65)	1.80%	(1.09%-2.69%)	0.23	(0.17-0.28)	1.05×10^{-15}	1.31×10^{-15}
2000	30	1970c	1.43 (1.65)	1.44%	(0.82%-2.27%)	0.21	(0.16-0.26)	1.16×10^{-14}	1.16×10^{-14}
1991	33	1958c	0.94 (1.46)	2.02%	(1.25%-2.70%)	0.21	(0.17-0.24)	3.86×10^{-26}	5.15×10^{-26}
2004	34	1970c	1.52 (1.77)	2.10%	(1.32%-3.12%)	0.26	(0.21-0.32)	8.78×10^{-20}	2.20×10^{-19}
2012	42	1970c	1.74 (1.89)	1.85%	(1.08%-2.75%)	0.26	(0.21-0.33)	1.78×10^{-17}	2.97×10^{-17}
2000	42	1958c	1.49 (1.74)	2.46%	(1.64%-3.22%)	0.27	(0.23-0.32)	3.49×10^{-34}	6.98×10^{-34}
2016	46	1970c	1.68 (2,06)	2.16%	(1.39%-3.19%)	0.31	(0.24-0.36)	6.46×10^{-22}	3.22×10^{-21}
2008	50	1958c	1.68 (2.06)	2.86%	(1.94%-3.68%)	0.33	(0.27-0.37)	3.49×10^{-36}	1.40×10^{-35}

R^2: amount of variance explained by the PGS. Analyses were adjusted for sex and the first 10 principal components of ancestry.
NB: blue indicates the 1958c, and red indicates the 1970c

Table 2: Results from the separate multilevel mixed effects model interacting polygenic score with age and then by sex for 1958c and 1970c (N_{1958C}=6,312; N_{1970c}=5,423)			
	Beta Coefficient	Bootstrapped 95% CI	p-value
Cohort			
1958 Model			
PGS	**0.24**	**0.23-0.26**	**5.26×10^{-50}**
Age	**0.01**	**0.01-0.02**	**2.03×10^{-67}**
PGS*Age	**0.004**	**0.003-0.005**	**4.49×10^{-7}**
1970 Model			
PGS	**0.25**	**0.23-0.27**	**7.95×10^{-32}**
Age	**0.007**	**0.005-0.01**	**1.10×10^{-08}**
PGS*Age	**0.004**	**0.001– 0.006**	**4.46×10^{-3}**
Sex			
1958 Model			
Male (ref)			
Female	**0.66**	**0.60-0.71**	**1.04×10^{-91}**
PGS	**0.20**	**0.16-0.25**	**8.06×10^{-18}**
PGS*Female	**0.09**	**0.01-0.11**	**5.01×10^{-3}**
1970 Model			
Male (ref)			
Female	**0.53**	**0.45-0.62**	**1.42×10^{-35}**
PGS	**0.24**	**0.18-0.30**	**1.82×10^{-15}**
PGS*Female	0.02	-0.06 - 0.11	0.584

Figures

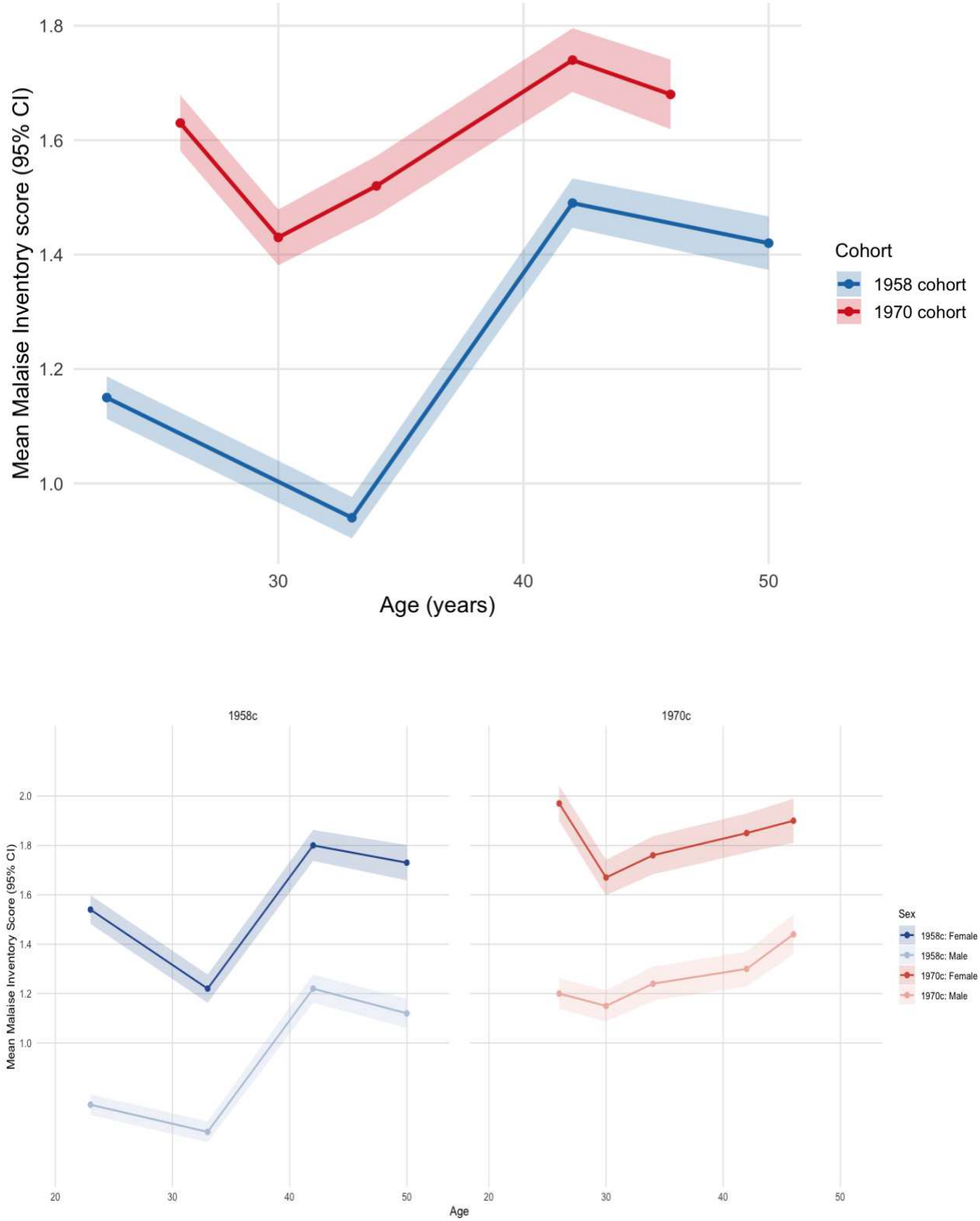


Figure 1: Mean and 95% confidence intervals of the malaise inventory score in a) both 1958c and 1970c, and b) by sex in both cohorts

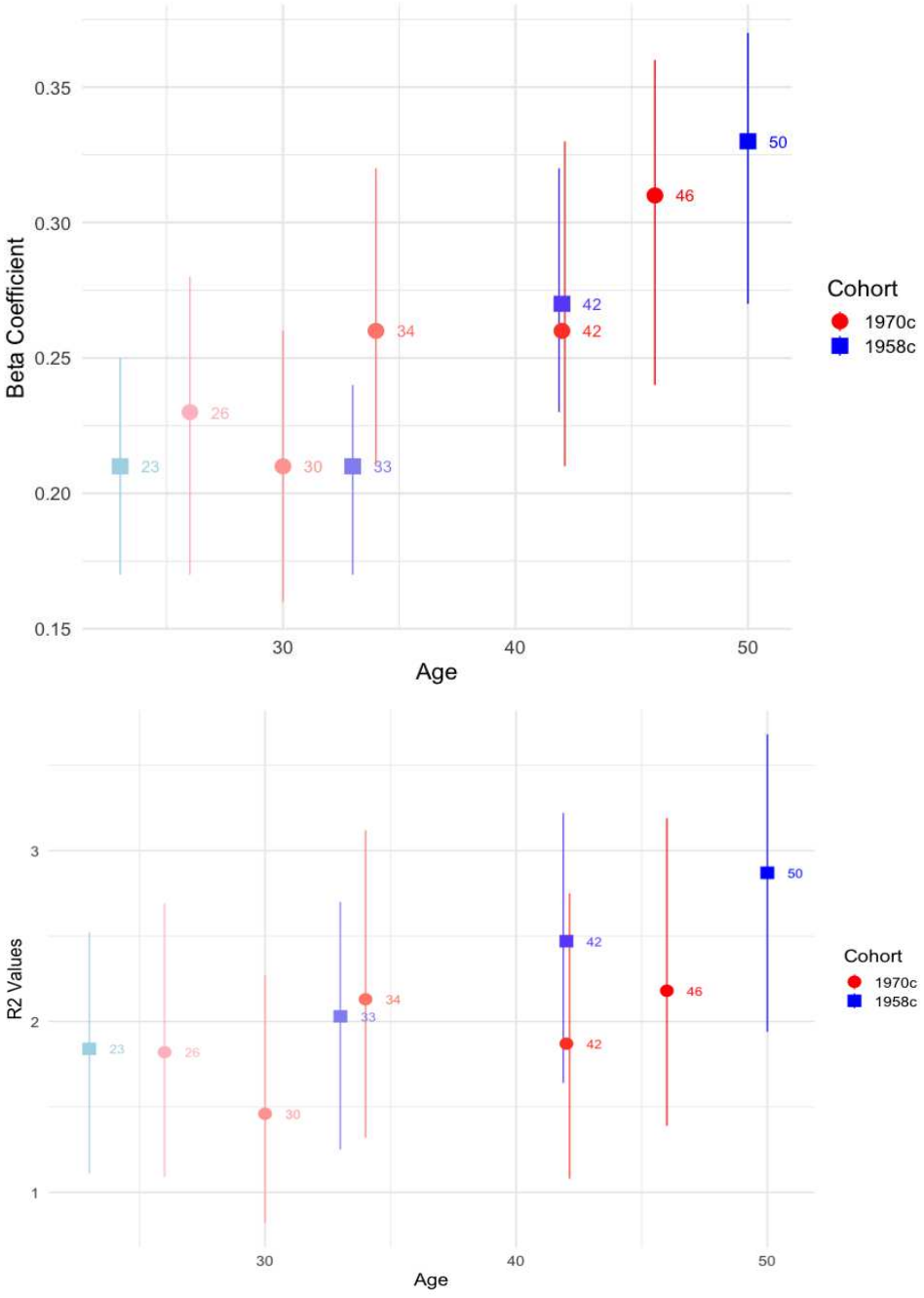


Figure 2: a) Beta coefficients and b) incremental adjusted r-squared from linear regression estimates of the association between the z-score polygenic score for psychological distress with overlapping SNPs at 0.005 p-value threshold and the 9-item Malaise Inventory score in 1958c and 1970c (N_{1958c}=6,312; N_{1970c}=5,423)

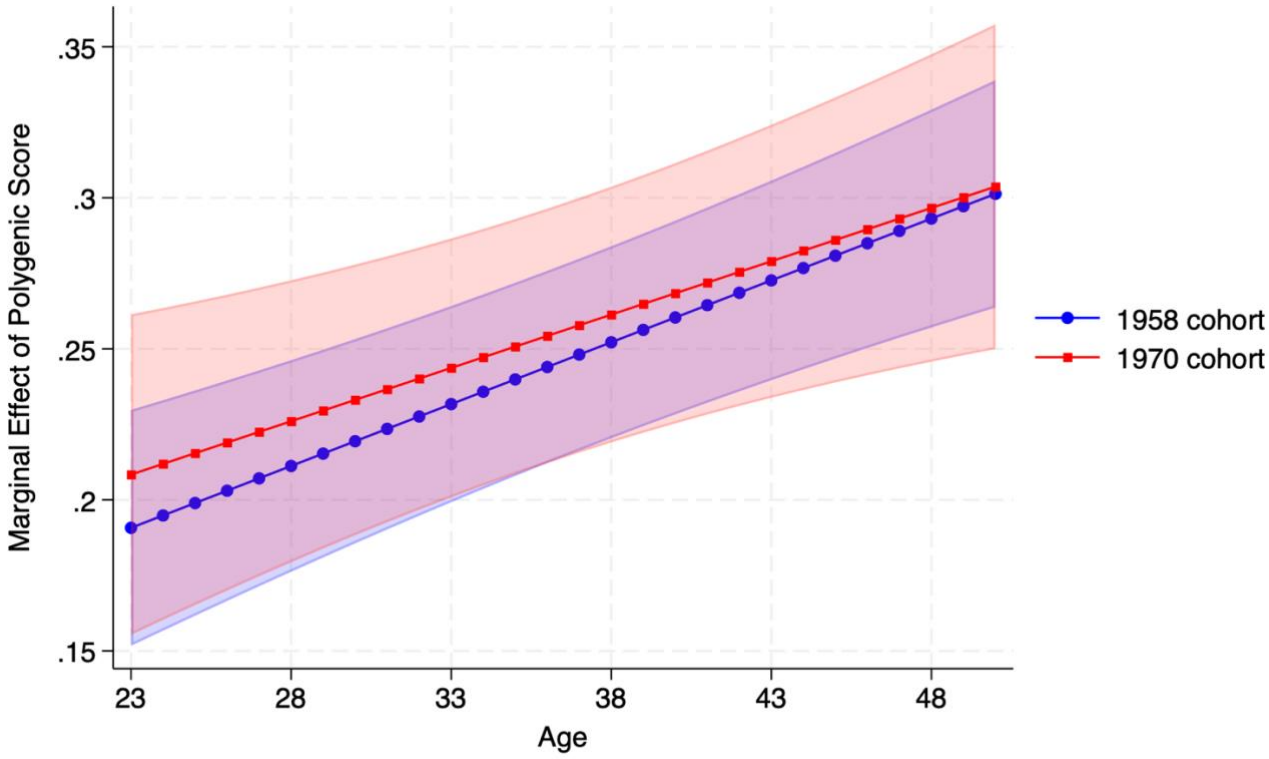


Figure 3: Predictive margins of psychological distress for 1958c versus 1970c across the polygenic score with confidence intervals

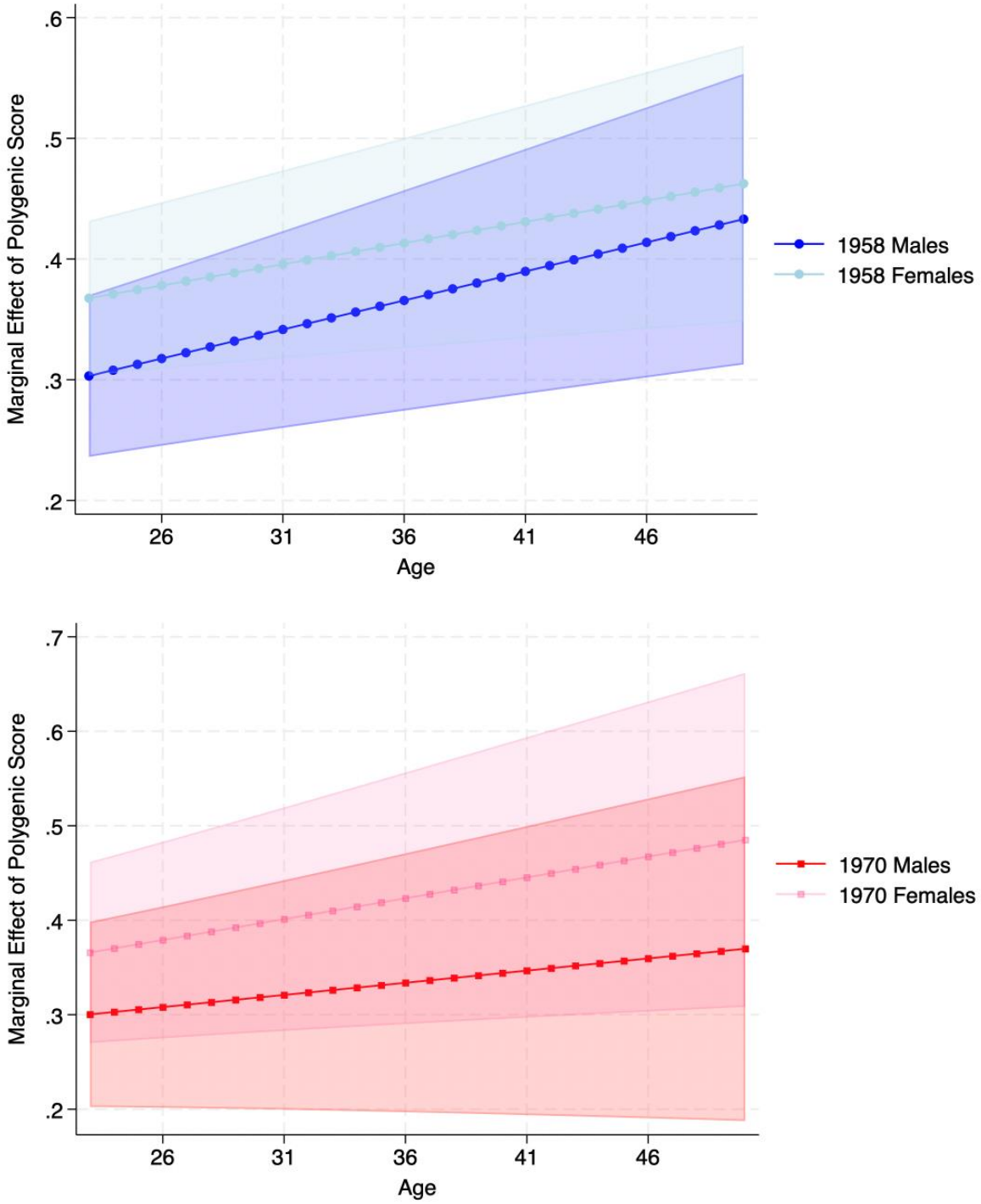


Figure 4: a) Predictive margins of psychological distress for males versus females across the polygenic score in 1958 and b) 1970 cohort

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- AgeCohortPGSCLSElsdenSupplementaryMaterials.pdf