Twin Research and Human Genetics (2025), 1–9
doi:10.1017/thg.2025.10011

CAMBRIDGE UNIVERSITY PRESS

Article

Exploring the Relationship between Polygenic Scores, Community-Shared Socioeconomic Indicators and Major Depressive Disorder Outcome

Dante T. Sepulveda[1,2], Jackson G. Thorp[2,3], Penelope A. Lind[2,3], Nicholas G. Martin[2], Sarah E. Medland[2] and Brittany L. Mitchell[2,3,4]

[1]National Autonomous University of Mexico, Mexico, [2]Brain and Mental Health Program, QIMR Berghofer, Brisbane, Queensland, Australia, [3]School of Biomedical Sciences, Faculty of Medicine, University of Queensland, Brisbane, Queensland, Australia and [4]School of Biomedical Sciences, Queensland University of Technology, Brisbane, Queensland, Australia

Abstract

Depression, a leading cause of global disability, arises from a multifaceted combination of genetic and environmental components. This study explores the relationship between major depressive disorder (MDD) polygenic scores (PGS), characteristics and symptoms of depression, and community-shared socioeconomic factors derived from postal code data in a cohort of 12,646 individuals from the Australian Genetics of Depression Study (AGDS). Our findings reveal that people living in areas with relatively higher socioeconomic advantages and education/occupation scores are more likely to report experiencing fewer depressive symptoms during their worst depressive period, as well as fewer number of lifetime episodes. Additionally, participants who reported depression onset later in life tend to currently reside in wealthier areas. Interestingly, no significant interaction between genetic and socioeconomic factors was observed, suggesting their independent contribution to depression outcomes. This research underscores the importance of integrating socioeconomic factors into psychiatric evaluation and care, and points to the critical role of public policy in addressing mental health disparities driven by socioeconomic factors. Future research should aim to further elucidate the causal relationships within these associations and explore the potential for integrated genetic and socioeconomic approaches in mental health interventions.

Keywords: Genetics; Polygenic scores; Depression; Socioeconomic

(Received 2 June 2025; revise received 2 June 2025; accepted 4 June 2025)

Major depressive disorder (MDD) is a common mental disorder affecting millions of people worldwide, and it is considered the leading contributor to global disability and suicide (World Health Organization, 2017). Its prevalence differs across various demographic groups, influenced by factors such as sex, age, geographical location, and socioeconomic status (World Health Organization, 2017). The etiology of MDD is considered to be multifaceted, arising from a complex interplay between genetic, environmental, and sociodemographic elements, thereby highlighting the multidimensional nature of the disorder (Ferrari et al., 2013; Kendler & Karkowski-Shuman, 1997).

Advancements in genetic research, particularly through genomewide association studies (GWAS), have shed light on the genetic architecture of MDD. These studies identify MDD as a highly polygenic disorder, with genetics accounting for about 30% to 40% of the variation in susceptibility to this condition (Kendler & Karkowski-Shuman, 1997). In this context, the development and employment of polygenic scores (PGS) have been crucial. PGS aggregate the effects of numerous risk-associated single nucleotide polymorphisms (SNPs), offering a quantified insight into the cumulative genetic vulnerability to depression (Sullivan et al., 2000; Wray et al., 2018). Previous studies have confirmed the MDD PGS predictive capacity of depression diagnosis, symptomatology, and other relevant features such as the number of depressive episodes, the age at onset, and co-occurrence with other psychiatric disorders (Hyde et al., 2016).

While previous research has extensively investigated the relationship between MDD and various socioeconomic indicators (Guan et al., 2022; Lloyd-Jones et al., 2019), the integration of genetic risk factors within these analyses has not been as closely explored. The environment in which individuals grow up and reside, encompassing factors like early life stress, parental care, social interactions, and exposure to trauma, has been recognized as a significant moderator of genetic predisposition to depression (Jaffee & Price, 2007; Mitchell et al., 2021; Uher & McGuffin, 2010). Research that incorporates both genetic and environmental aspects acknowledges that depression PGS, combined with variables such as chronic stress, inadequate parenting, limited social interactions, or trauma, can aggravate depressive symptomatology (Brugha

Corresponding author: Brittany L. Mitchell; Email: Brittany.Mitchell@qimrb.edu.au.
Cite this article: Sepulveda DT, Thorp JG, Lind PA, Martin NG, Medland SE, Mitchell BL. Exploring the Relationship between Polygenic Scores, Community-Shared Socioeconomic Indicators and Major Depressive Disorder Outcome. *Twin Research and Human Genetics* https://doi.org/10.1017/thg.2025.10011

et al., 1985; Mullins et al., 2016; Peen et al., 2010). However, these studies often overlook the importance of socioeconomic disparities as a key element of the environmental impact. Some studies that have focused on socioeconomic factors as major environmental variables found significant associations between genetic risk for MDD and indicators of lower socioeconomic status, including unemployment, marital discord, and financial strain (Agerbo et al., 2021; Machlitt-Northen et al., 2022; Musliner et al., 2015). A recent study showed that educational attainment, lower wealth, and limited access to participation in cultural activities correlates with individual levels of depressive symptoms independently from their polygenic predisposition to depression (Zhang et al., 2023).

The influence of community-level socioeconomic factors, such as impaired accessibility to public resources and lack of social connectivity with neighbours, on moderating genetic predisposition to depression remains mostly unstudied. Disadvantaged communities experience elevated rates of residential mobility, unemployment, crime, and financial stress, among numerous other factors, which collectively magnify the detrimental impact on mental health due to their persistent daily exposure (Kosciuszko et al., 2023). Recent studies indicate that factors shared within a community, such as urbanicity, neighbourhood disadvantage, or poverty-related stress, are associated with the severity of psychiatric disorders (Arias-de la Torre et al., 2018; Kosciuszko et al., 2023; Oenning et al., 2018; Santiago et al., 2011). One study examining how urban life can act as an environmental risk factor for schizophrenia suggests a causal association between schizophrenia risk and the choice to live in more densely populated areas (Almeida et al., 2012). Conversely, a similar study examining the impact of urbanicity on depression outcomes found no significant relationship between depression PGS and urbanicity when estimating symptoms of poor mental health (Colodro-Conde et al., 2018). This study aims to delve deeper into the relationship between genetics and environment, particularly focusing on how community-level socioeconomic factors moderate the genetic risk for MDD outcomes. Our study is methodically structured to analyze the interaction between PRS and an array of socioeconomic and demographic factors, categorized by postal code, among individuals from the Australian Genetics of Depression Study. Through this examination, we aim to deepen our understanding of how collective socioeconomic circumstances at the community level influence the development and expression of depression. Ultimately, this investigation seeks to contribute to the creation of targeted strategies that effectively mitigate mental health disparities arising from socioeconomic inequalities.

Methods

The Australian Genetics of Depression Study

The Australian Genetics of Depression Study (AGDS) provides the fundamental basis for this research, with its primary aim being to explore the genetic and environmental risk factors of depression, alongside understanding treatment responses in individuals diagnosed with depression at some point in their lives (Sund et al., 2021). Comprehensive details regarding the recruitment process and sample characteristics can be found in previous publications (Byrne et al., 2020; Sund et al., 2021). Participants completed a series of questionnaires to assess diagnostic criteria for depression and self-reported psychiatric history. They also provided additional information about their family, general health, and other clinical features relevant to MDD.

An essential aspect of the AGDS was the collection of saliva samples from a group of over 16,000 participants for genotyping analysis. These participants granted their informed consent for the study via an online platform, adhering to the ethical standards set forth. The QIMR Berghofer Medical Research Institute's Human Research Ethics Committee (QIMR) granted full approval for all study protocols under the ethics code P2118.

To ensure the quality and accuracy of the dataset, individuals who did not align with the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5; American Psychiatric Association, 2013) criteria for MDD or presented significant gaps in their data were excluded from analysis. Following these exclusions, the final cohort comprised 12,646 participants, of whom 76% were women (mean age = 48, *SD* = 13, min = 18, max = 86) and 24% male (mean age = 42, *SD* = 15, min= 18, max = 90).

Depression Polygenic Scores

In this study we used polygenic scores for MD that were constructed using the summary statistics provided by Howard et al. (2019), including 23andMe. These PGS were created as part of the AGDS study, using leaving-one-out summary statistics, which specifically excluded QIMR samples to eliminate bias stemming from potential sample overlap. The process of generating these PGS followed rigorous quality control measures. This included the removal of SNPs with alleles not matching the genotyping data and the exclusion of indels. The construction of PGS employed the SBayesR method, a Bayesian approach that models SNP effects as originating from a mixture of four zero-mean normal distributions, each with distinct variances (Rojas-Garcia et al., 2015). For the linkage disequilibrium (LD) reference, the methodology aligned with that used by Lloyd-Jones et al. (2019). In their approach, the LD matrix was constructed based on the HapMap3 SNPs from a carefully selected sample of 50,000 unrelated UK Biobank individuals. Using the posterior SNP effects, the polygenic scores for each individual in the study were calculated, employing the –score function in PLINK (Patria, 2022).

Socio-Economic and Demographic Measures

The Socioeconomic Indexes for Areas (SEIFA), derived from the 2016 Australian Census of Population and Housing, offers a comprehensive view of the socioeconomic landscape across Australian regions (Gomez et al., 2023). SEIFA encompasses four distinct indexes, each one evaluating a different dimension of socioeconomic status. The Index of Relative Socioeconomic Disadvantage (IRSD), evaluates both the presence and absence of disadvantage in an area, incorporating factors like unemployment rates, internet access in households, and the proportion of laborers. The Index of Relative Socioeconomic Advantage and Disadvantage (IRSAD), expands upon this by considering aspects of socioeconomic advantage, such as higher household incomes (above AUD 78,000), professional or managerial occupations, and households with multiple bedrooms. The Index of Education and Occupation (IEO) focuses on educational and occupational variances, considering both the advantages and disadvantages associated with educational qualifications and vocational skills. Lastly, the Index of Economic Resources (IER) delves into financial aspects like car ownership, single-parent households, and high mortgage payments.

Each geographic area in Australia is assigned a SEIFA score for every index, reflecting its relative socioeconomic position. Table 1 illustrates the SEIFA score characteristics derived from the

Table 1. Socioeconomic indexes for areas in the AGDS cohort ($N = 12,646$)

Score	Mean	*SD*	Min–Max
Disadvantage (IRSD)	1018	61	565–1143
Disadvantage-Advantage (IRSAD)	1021	73	690–1181
Education-Occupation (IEO)	1028	85	759–1234
Economic Resources (IER)	1003	60	584–1194

Note: AGDS, Australian Genetics of Depression Study; IRSD, Index of Relative Socioeconomic Disadvantage; IRSAD, Index of Relative Socioeconomic Advantage and Disadvantage; IEO, Index of Education and Occupation; IER, Index of Economic Resources.

Table 2. Distribution of depression outcomes

Number of episodes		Number of symptoms	
13 +	36.0%	9	50.44%
12	0.75%	8	29.60%
11	0.20%	7	13.17%
10	7.21%	6	4.95%
9	0.69%	5	1.85%
8	3.14%	Anxiety comorbidity	
7	1.90%	Present	56.30%
6	7.03%	Absent	43.70%
5	9.31%	Age of onset	
4	10.83%	Mean	21 years
3	11.62%	*SD*	11 years
2	7.40%	Max	79 years
1	3.94%	Min	1 year

participants' residential areas in our study. For a more in-depth description of how these indexes are developed, their foundational concepts, and the data sources used, the SEIFA Technical Paper (Gomez et al., 2023) provides thorough explanations.

It is important to note that index scores are allocated to Statistical Area Level 1 (SA1) units, not to individuals. Our research examines the link between the average socioeconomic environment of a participant's residential area and their depression outcomes. Therefore, the scores applied in our analysis represent the mean of SA1 level scores within specific postal areas. Despite the strong correlation between IRSD and IRSAD, the existence of extremely disadvantaged or advantaged pockets within an area might skew perceptions of the area's overall socioeconomic condition. Consequently, index scores may not precisely capture the level of disadvantage (or advantage) experienced by individuals residing in areas marked by such variability. This is further addressed in the discussion as one of the study's limitations. Additionally, a detailed overview is presented in the supplemental material from the SEIFA Technical Paper (Australian Bureau of Statistics, 2021), and Supplementary Figure 1 includes a detailed breakdown of the variables constituting each socioeconomic index.

The classification of urbanicity for each postal area was determined using the Modified Monash Model (MMM), a framework that classifies areas in Australia into metropolitan, regional, rural, and remote regions based on geographic remoteness and population size. The Modified Monash (MM) scale ranges from 1 to 7. The distribution of participants across the MM categories is as follows: (1) Metropolitan areas (70.9%), (2) regional centers (9.9%), (3) large rural towns (4.6%), (4) medium rural towns (1.9%), (5) small rural towns (11.4%), (6) remote communities (0.7%), and (7) very remote communities (0.5%). Further information about this measure can be found in the MMM fact sheet on the Australian Department of Health and Aged Care website (Wise & Mathews, 2011).

Study Groups

Each participant's PGS was linked to the corresponding SEIFA scores and MM category based on their postal code. The primary sample ($N = 12,646$) included participants from 1642 different postal codes. The concentration of participants per postal code varied significantly, with the highest being 77 individuals in a single code, and as many as 342 codes having only one participant.

Given the nature of the SEIFA scores being calculated for SA1 regions rather than individual or broader geographic units such as postal codes, our study had to account for variations within these areas. Since a single postal code can encompass numerous SA1 areas, and considering that the scores in our study represent an average of all SA1 scores within a specific postal code, we constructed additional subgroups to investigate the influence of score diversity in postal codes containing multiple SA1 areas.

Two distinct subgroups were formulated based on the variability of SEIFA scores within the postal codes. The first subgroup encapsulated participants residing in postal codes where the disparity between the highest and lowest SA1 scores for each of the four SEIFA indexes was confined to less than one standard deviation from the mean. This subgroup, representing areas with low score diversity, included 1542 participants. Conversely, the second subgroup was defined with a more lenient criterion, including postal codes where the score difference reached up to two standard deviations. This subgroup, indicative of moderate score diversity, encompassed 6134 participants.

The formation of these subgroups allowed for a nuanced exploration of how socioeconomic score diversity within a postal code might influence the association between PGS and depression outcomes, providing a more granular understanding of the interplay between genetics and environment.

Depression Outcomes

Our study conducted an analysis to explore the relationship between MDD PGS, SEIFA scores, and MMM classifications, and their collective impact on various depression outcomes. These outcomes were specifically gauged across four dimensions for each participant: (1) age at onset of depression, (2) lifetime number of reported depressive episodes, (3) number of symptoms during the worst period of depression, and (4) comorbidity with self-reported anxiety disorder. Table 2 shows the distribution of depression outcomes from the primary cohort. A notable observation from this table is the high prevalence of severe and recurrent depression cases within the cohort. Specifically, a significant majority (80.0%), reported experiencing 8 or 9 symptoms, which is indicative of severe depressive episodes. Furthermore, a substantial proportion of the cohort (36%), disclosed having encountered 13 or more depressive episodes over their lifetime, underlining the recurrent nature of the disorder in our study population.

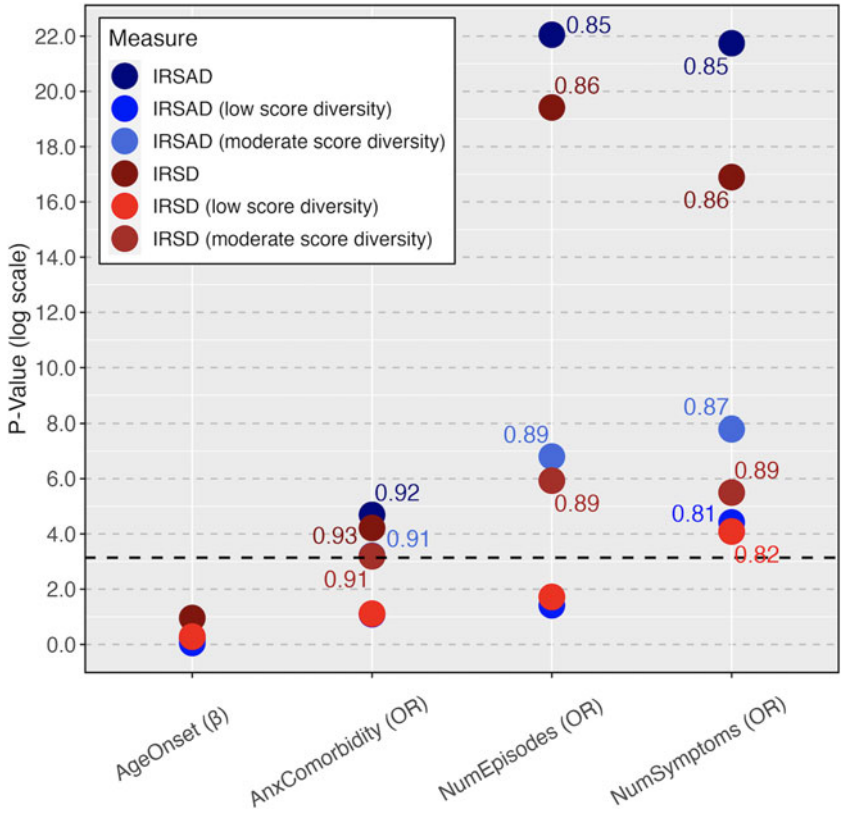

Figure 1. Comparison between the association of the IRSAD and the IRSD on all depression outcomes. A logarithmic scale was used to display *p* values in order to improve visualization. Colors represent the socioeconomic variables and the group where the regression analysis was performed. Associations with significant standardized beta coefficients/odds ratios are shown next to the corresponding point. The dashed line indicates the corrected significance threshold ($p = 7.2e\text{-}04$).

Statistical Analysis

The statistical analyses for this research were carried out using R software (version 4.2.2); specifically, the MASS, ggplot2, patchwork, and ggrepel packages. Our approach involved a three-model strategy, each designed to evaluate specific types of associations. Given the diverse scales of our response variables — age of onset (continuous), number of symptoms and episodes (ordinal), and comorbidity with anxiety (binary) — we applied different regression techniques: linear regression for continuous variables, proportional odds logistic regression for ordinal variables, and logistic regression for the binary variable.

For clarity and ease of interpretation, all independent variables, including SEIFA scores, age, and PGS, were standardized to a mean of zero and a standard deviation of 1. The standardized variables cover the following ranges: PGS (−4.15 to 3.79), IRSAD (−4.50 to 2.16), IEO (−3.15 to 2.40), IER (−6.9 to 3.13), and MM Class (−0.54 to 3.67). All models accounted for the covariates sex, age, and the first 10 ancestry principal components. When exploring interactions between socioeconomic factors and depression PGS, terms for PGS-by-covariate and SEIFA-score-by-covariate were included in the models. Additionally, a Bonferroni correction was applied to all statistical analyses for multiple testing, adjusting the significance threshold to $p = .05/69 = 7.2e\text{-}04$.

Model 1 evaluated the predictive capability of the depression PGS on depression outcomes without accounting for demographic or socioeconomic factors (Eq. 1). Model 2 introduced socioeconomic measures (IRSD, IRSAD, IEO, IER, MM) as independent variables to assess their additional impact on the associations found in Model 1 (Eq. 2). Additionally, this model examined the regression effects on subgroups with low and moderate SEIFA scores diversity and determined whether IRSAD consistently outperformed IRSD in its predictive power.

Model 3 investigated interactions between MDD PGS and socioeconomic/demographic factors on both additive and multiplicative scales (Eq. 3). For multiplicative interaction analysis, 'age of onset' was converted to a binary variable, classifying participants into 'early' or 'late' onset groups based on a cut-off of 23 years. The other dependent variables remained unchanged. This conversion enabled inclusion in logistic regression or proportional odds logistic regression analyses. Additive interactions were explored through linear regressions for all outcomes; including interaction terms in the model allowed us to estimate the effects of PGS and socioeconomic factors beyond their individual contributions. All the regressions conducted at this stage incorporated the PGS-by-covariate and socioeconomic factor-by-covariate terms.

$$Depression\ Outcomes \sim PGS + Age + Sex + PCs \quad \text{Eq. 1}$$

$$Depression\ Outcomes \sim PGS + SEIFA + Age + Sex + PCs \quad \text{Eq. 2}$$

$$Depression\ Outcomes \sim PGS : SEIFA + PGS : Covariate + SEIFA : Covariate \quad \text{Eq. 3}$$

$$PGS SEIFA + Age + Sex + PCs \quad \text{Eq. 4}$$

As a sensitivity analysis, we replicated our three-model approach using the most recent SEIFA scores, derived from the 2021 Census data. Despite minor adjustments from the 2016 to 2021 SEIFA versions, such as the revision of cut-off values for variables like 'high income' or 'occupation skill level', the 2021 edition preserves the core methodologies and consistency with its 2016 predecessor. Considering that the AGDS questionnaire was completed closer in time to 2016, leveraging more contemporary

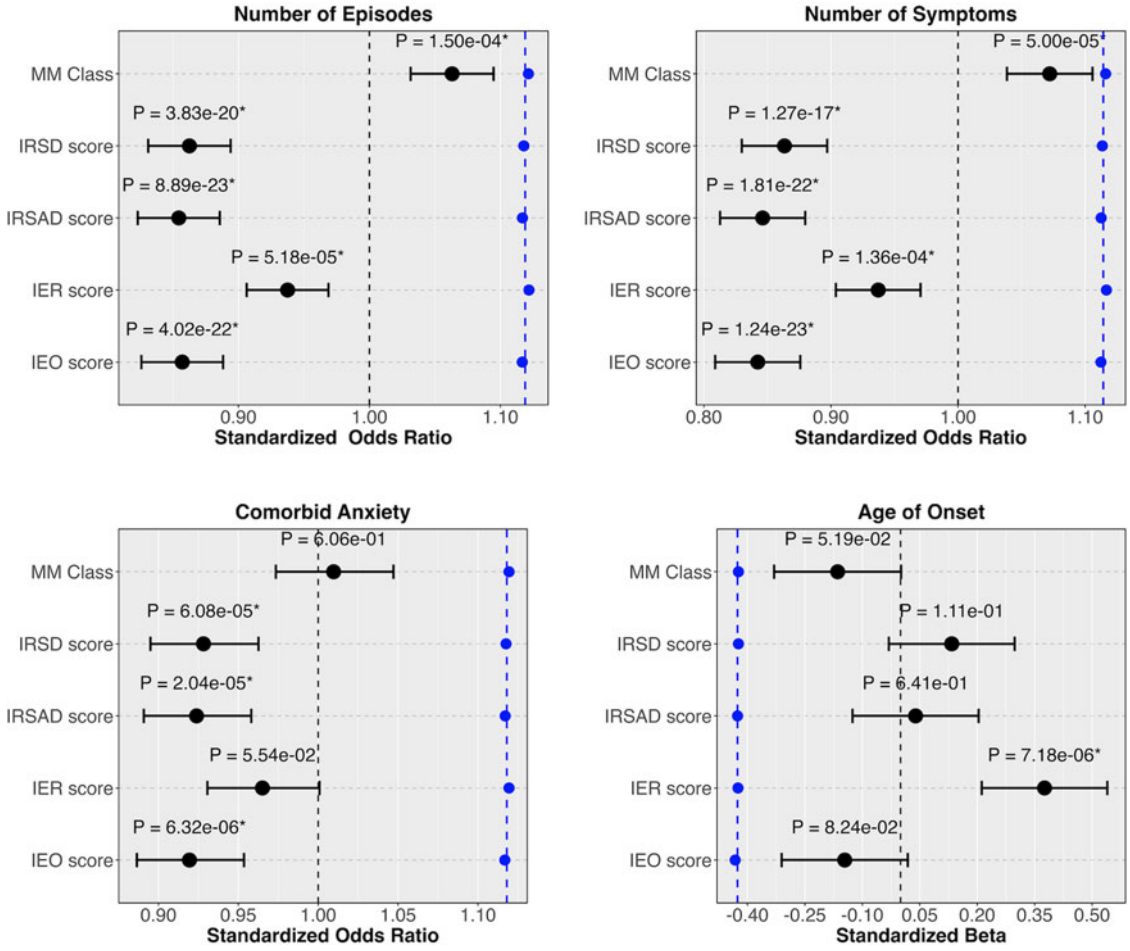

Figure 2. Relationships between socioeconomic and demographic indicators (IRSAD, IEO, IER, and MM Class), and depression outcomes.
Note: IRSAD, Index of Relative Socioeconomic Advantage and Disadvantage; IEO, Index of Education and Occupation; IER, Index of Economic Resources; MM, Modified Monash scale.

data could potentially enhance the accuracy of our findings. This additional analysis phase enabled us to evaluate the impact of employing different SEIFA versions on the performance of our models.

Finally, a series of regression analyses were carried out to explore the relationship between SEIFA scores and depression PGS (Eq. 4). This exploration aimed to understand how the socioeconomic context of an individual's place of residence might reflect their genetic vulnerability to depression. These analyses, while separate from the initial three-model strategy, adhered to the same protocols for standardization and included the same covariates, ensuring a consistent and rigorous examination of the relationship between socioeconomic factors and genetic predisposition to depression.

Results

Main Effects of Depression PGS

The results from the first model revealed significant associations between MDD PGS and all the evaluated depression outcomes. We observed that a higher MDD PGS correlates positively with an increase in the number of depressive episodes ($OR = 1.12$, 95% CI [1.09, 1.16], $p = 3.9e\text{-}13$) and the number of symptoms ($OR = 1.12$; 95% CI [1.08, 1.15], $p = 6.1e\text{-}11$). Additionally, a higher PGS corresponded to increased odds of comorbid anxiety ($OR = 1.12$; 95% CI [1.08, 1.16], $p = 1.1e\text{-}09$). Conversely, there was a negative relationship between MDD PGS and the age at the onset of depression ($\beta = -0.44$; $p = 3.5e\text{-}07$). These findings indicate that individuals with higher MDD PGS are more likely to experience their first depressive episode at an earlier age, in addition to having more episodes and symptoms, and a higher likelihood of anxiety comorbidity.

Main Effects SEIFA Scores and MM Class

The second model highlighted significant negative correlations between both the number of depressive symptoms and episodes and the likelihood of comorbid anxiety with IRSAD and IRSD scores. Figure 1 shows that IRSAD, which accounts for socioeconomic advantages and disadvantages, exhibited stronger and more statistically significant effects in comparison to IRSD, focusing solely on socioeconomic disadvantage. Notably, in the single instance where IRSAD did not surpass IRSD, no statistically significant results were observed.

Additionally, this analysis indicated that narrowing the focus to participants in postal codes with low or moderate SEIFA score diversity did not enhance the significance or strength of the associations. Consequently, further analyses included the entire sample of 12,646 participants. The high correlation between IRSD and IRSAD suggests that findings for IRSAD are generally applicable to IRSD, except where explicitly stated otherwise.

Figure 2 illustrates the associations between socioeconomic and demographic factors at the postal code level and depression outcomes. Results from Model 2 also revealed significant negative associations between IEO and IRSAD scores and the number of depressive symptoms, number of episodes, and odds of comorbid anxiety, indicating that higher scores in these indexes are linked to fewer reported symptoms and episodes, and lower odds of having comorbid anxiety. Interestingly, IER was the only indicator displaying a significant positive association with the age of onset, suggesting a propensity for later-onset depression in more affluent areas. Furthermore, significant links were identified between urbanicity (as indicated by MM Class) and both the number of symptoms and episodes. However, given the marginal significance of these associations and the uneven distribution of participants across the MM scale, these findings warrant cautious interpretation.

The IRSAD is significantly associated with the number of depressive episodes ($OR = 0.85$; 95% CI [0.82, 0.89], $p = 8.90e\text{-}23$), the number of symptoms ($OR = 0.84$; 95% CI [0.81, 0.88], $p = 1.81e\text{-}22$), and comorbid anxiety ($OR = 0.92$; 95% CI [0.89, 0.95], $p = 2.04e\text{-}05$). The IEO shows a significant coefficient of 0.86 (OR; 95% CI [0.83, 0.89], $p = 4.02e\text{-}22$) for the number of episodes, 0.83 (OR; 95% CI [0.80, 0.87], $p = 1.24e\text{-}23$) for the number of symptoms, and 0.91 (OR; 95% CI [0.87, 0.95], $p = 6.32e\text{-}06$) for comorbid anxiety. Additionally, the IER is associated with the age of onset ($\beta = 0.37$, $p = 7.18e\text{-}06$, R^2=0.278), the number of symptoms ($OR = 0.93$; 95% CI [0.90, 0.97], $p = 1.36e\text{-}04$) and the number of episodes ($OR = 0.94$; 95% CI [0.90, 0.96], $p = 5.18e\text{-}05$). The MM Class shows a small but significant coefficient with the number of episodes ($OR = 1.06$; 95% CI [1.03, 1.09], $p = 1.50e\text{-}04$) and the number of symptoms ($OR = 1.07$; 95% CI [1.04, 1.10], $p = 5.00e\text{-}05$). Statistically significant associations are denoted by asterisks, adhering to a Bonferroni-corrected threshold of $p < 7.2e\text{-}04$. The black dashed line indicates the threshold for a null effect, while blue points represent polygenic score coefficients/odds ratios, with their average depicted by the dashed blue line.

Interaction Effects

The third model focused on exploring interaction effects revealed no significant additive or multiplicative interactions between depression PGS and any of the socioeconomic or demographic factors. Table 3 presents results from the interaction analyses between depression PGS and IRSAD scores (results across all variables are available in the supplementary material). Additionally, the analysis considered potential interaction effects between the covariates 'Sex' and 'Age' and SEIFA scores or MM classification. Although a few instances displayed p values below .05, none reached statistical significance after adjusting for multiple testing.

This absence of significant interactions suggests that the combined influence of depression PGS and each socioeconomic measure on depression outcomes is merely additive or independent, without exceeding the sum or product of their individual effects. Consequently, this analysis provides no evidence to support the presence of interactions between depression PGS and the examined socioeconomic indicator.

SEIFA 2021 and 2016 Versions

Our comparative analysis between the 2021 and 2016 versions of the SEIFAs revealed no statistically significant differences in the outcomes derived from these datasets, as detailed in the supplementary materials. Although we observed consistency across the two SEIFA versions, we decided to use the 2016 SEIFA version since it provides a more accurate representation of the socioeconomic conditions and experiences relevant to our participants during the study period.

Table 3. Additive and multiplicative interactions effects (IRSAD x DepPGS)

Additive			Multiplicative		
Depression outcome	Beta	*p* value	Depression outcome	*OR*	*p* value
Age of onset	0.013	.872	Early age of onset	0.989	.619
Comorbid anxiety	−0.0006	.891	Comorbid anxiety	0.997	.884
Number symptoms	0.002	.733	Number symptoms	1.003	.840
Number episodes	0.031	.417	Number episodes	1.012	.421

Note: IRSAD, Index of Relative Socioeconomic Advantage and Disadvantage; DepPGS, depression polygenic scores.

Relationship Between SEIFA Scores and Depression PGS

The results obtained from the analysis between depression PGS and socioeconomic factors revealed significant correlations. Specifically, we found that depression PGS were negatively associated with both IRSAD ($\beta = -2.38$, $p = 2.7e\text{-}04$) and IEO ($\beta = -2.8$, $p = 1.5e\text{-}04$), suggesting that higher depression PGS are prevalent in areas with more pronounced socioeconomic disadvantages and diminished educational opportunities. Interestingly, among all the socioeconomic variables we analyzed, the IER showed the weakest association with depression PGS ($\beta = -0.4$, $p = .45$). This relatively weak and nonsignificant correlation implies that the economic resources context, as measured by IER, might not be a strong determinant in the residential patterns of individuals with higher genetic predispositions to depression.

Discussion

The complexity of depression stems from the interplay between genetic predispositions and environmental influences. Previous studies have extensively investigated the correlation between MDD and socioeconomic factors, yet the cumulative impact of genetic risk alongside community-shared socioeconomic contexts on depression outcomes has been less explored. This study aimed to bridge this gap by examining the intersection of community-shared socioeconomic and demographic factors, genetic risk, and MDD outcomes.

We confirmed that higher MDD PGS are significantly associated with more severe depressive symptoms, an increased number of episodes, a higher likelihood of comorbid anxiety, and an earlier onset of depression in the AGDS sample. This aligns with existing literature underscoring the relevance of PGS in MDD diagnosis and its outcomes (Halldorsdottir et al., 2019; Musliner et al., 2021; Purcell et al., 2007).

Our findings indicate that residing in areas with higher socioeconomic advantage, particularly as measured by IRSAD and IEO, correlates with fewer reported depressive symptoms and episodes, and lower odds of having comorbid anxiety. This aligns with prior research suggesting that favorable socioeconomic conditions in neighborhoods serve as protective factors against the aggravation and occurrence of depressive symptoms (Beard

et al., 2009; Wray et al., 2018). Similarly, we observed a notable correlation between later depression onset and currently residing in wealthier areas (IER). In contrast, the association between urbanicity and depression outcomes was less pronounced, indicating the need for further studies that include a more diverse representation of rural and remote communities to fully understand the role of urbanicity in depression outcomes.

Debate exists between social causation and social selection theories: the first one proposes that social variables, including urbanicity and socioeconomic status, may lead to mental health issues, while the second one suggests that those with mental health disorders may end up living in specific socioeconomic environments through mechanisms related to their disorder. Therefore, caution must be taken when attributing a direction of causation between genetic risk, phenotype, and socioeconomic or environmental factors. These results may support the role of community-shared socioeconomic factors in modulating mental disorder prevalence, although the causal direction with mental health remains ambiguous. Notably, the overlap among SEIFA scores due to shared variables complicates their interpretation as entirely independent indicators. This is exemplified in Figure 2, where IRSD, IRSAD, and IEO indices demonstrate similar patterns and statistical significance, likely attributable to shared variables. For instance, the IEO index includes variables such as the proportion of nontertiary educated, unemployed, or schooling individuals, which are also prevalent in IRSAD. However, IRSAD and IER share distinct variables like income levels, mortgage status, and homeownership, absent in IEO. Crucially, these overlaps underline the pivotal influence of education and employment variables on the association between SEIFA scores and lifetime depression severity, as indicated by symptom count and episode frequency. The association between depression outcomes and lower educational-occupational scores in the area in which participants reside at the time they enrol in the AGDS may suggest that experiencing depression, and its corresponding severity, might act as an obstacle to educational attainment or occupational trajectory. On the other hand, previous literature has shown that higher education can provide individuals with better coping mechanisms, a broader social network, and increased self-efficacy (Wang et al., 2010), all of which can act as buffers against mental health challenges.

Our analysis did not uncover significant interaction effects between depression PGS and socioeconomic variables, suggesting that the influence of these variables on depression outcomes might operate independently, without amplifying or mitigating each other's impact. Nonetheless, since interaction effects are known to be harder to detect, the absence of interaction effects might also be attributed to the limitations in our study such as dataset size or composition.

An interesting result is the significant association between SEIFA scores and MDD PGS, which highlights the relationship between genetic susceptibility to depression and shared socioeconomic factors within communities. The observed link between higher depression PGS and residing in socioeconomically disadvantaged areas could indicate that individuals genetically predisposed to depression may be more affected by the stressors and challenges of economic adversity. This vulnerability may arise from limited access to resources, exposure to daily stress, or reduced chances for social and economic mobility. On the other hand, individuals at greater risk of depression might face difficulties in securing professional roles or achieving optimal performance in their careers. This can lead to financial struggles and stressful situations, further confining them to economically deprived areas.

These findings underscore the importance of considering socioeconomic contexts in psychiatric care and the formulation of public policies aimed at mitigating the impact of socioeconomic factors on mental health in disadvantaged communities. The lack of significant gene-environment interaction effects in our findings suggests an independent effect of genetic and socioeconomic factors on depression outcomes. However, this interpretation requires further investigation for greater clarity. Further research, potentially using longitudinal data or methodologies like Mendelian randomization, is needed to clarify these associations and determine the causal directions of our findings.

Our study, while insightful, has several limitations that must be acknowledged. First, the use of postal code-level data to approximate community-shared socioeconomic factors offers a broad overview but may not accurately reflect the socioeconomic experiences of participants. This generalized approach could potentially overlook nuanced, individual-level socioeconomic variations. Second, our reliance on self-reported data for depression outcomes introduces the possibility of recall bias and inconsistencies in reporting, either as under-reporting or over-reporting. Additionally, our socioeconomic indicators, though comprehensive, may not encompass all relevant factors that could impact depression characteristics, suggesting the need for a more exhaustive set of variables in future research. A further limitation is the temporal mismatch between the assessment of depression outcomes — based on an individual's lifetime or their worst episode — and the collection of postcode data, which reflects their current residence. This discrepancy means we cannot ascertain the socioeconomic context during the actual period of depression experience. Lastly, the composition of our study sample, predominantly wealthier and more educated individuals, restricts the diversity and hence the generalizability of our findings. A more heterogeneous sample would be beneficial in future studies to ensure broader applicability of the results.

Conclusion

This study has demonstrated the presence of significant, yet independent, associations between genetic predispositions to depression (reflected through MDD PRS) and shared socioeconomic factors within communities. Our analysis indicates that individuals displaying fewer depressive symptoms and episodes, along with a reduced likelihood of comorbid anxiety, tend to reside in areas characterized by higher socioeconomic status and better education-occupation outcomes. Additionally, a trend towards later onset of depression was observed in participants who reported living in wealthier areas. Importantly, we found no significant interaction between genetic predispositions and socioeconomic factors, implying that these factors independently contribute to depression outcomes. This observation calls for further research to substantiate these findings and deepen our understanding of the complex interplay between genetics and environment in mental health. The findings of this research underscore the critical importance of considering socioeconomic factors in the delivery of psychiatric care and the formulation of public policies aimed at reducing mental health disparities. By integrating insights on both genetic susceptibilities and socioeconomic environments, future interventions can be better tailored to address the multifaceted nature of depression and improve mental health outcomes across diverse communities.

Supplementary material. To view supplementary material for this article, please visit https://doi.org/10.1017/thg.2025.10011.

Data availability statement. Code used in this study are available from request to the corresponding author.

Acknowledgments. We thank the participants for their time and continuous support for this project. We wish to thank all the people who helped in the conception, implementation, beta testing, media campaign and data cleaning. We would specifically like to acknowledge Dale Nyholt for advice on using the PBS for research; Ken Kendler, Patrick Sullivan, Andrew McIntosh and Cathryn Lewis for input on the questionnaire; Lorelle Nunn, Mary Ferguson, Lucy Winkler and Natalie Garden for data and sample collection; Natalia Zmicerevska, Alissa Nichles and Candace Brennan for participant recruitment support; Jonathan Davies, Luke Lowrey and Valeriano Antonini for support with IT aspects; Vera Morgan and Ken Kirkby for help with the media campaign. We would like to thank VIVA! Communications for their effort in promoting the study. We also acknowledge David Whiteman and Catherine Olsen from QSkin.

Author contribution statement. D.T.S performed the analyses, wrote the manuscript and integrated input and feedback from all co-authors. B.L.M designed and supervised the study. Support and input was provided by S.E.M, who along with P.A.L and N.G.M were responsible for the cohort design, genotyping, and data acquisition, processing and quality control. J.G.T provided support with analyses. All authors read and approved the final manuscript.

Funding statement. The AGDS was primarily funded by National Health and Medical Research Council (NHMRC) of Australia grant 1086683. This work was further supported by NHMRC grants 1145645, 1078901 and 1087889 and by the Queensland Government through Queensland Health. SEM and BLM were supported by Australian NHMRC Investigator grants (APP1172917 and APP2017176 respectively).

Competing interests. The authors declare no competing interests.

Ethics statement. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

References

Agerbo, E., Trabjerg, B. B., Børglum, A. D., Schork, A. J., Vilhjálmsson, B. J., Pedersen, C. B., Hakulinen, C., Albiñana, C., Hougaard, D. M., Grove, J., McGrath, J. J., Bybjerg-Grauholm, J., Mors, O., Plana-Ripoll, O., Werge, T., Wray, N. R., Mortensen, P. B., & Musliner, K. L. (2021). Risk of early-onset depression associated with polygenic liability, parental psychiatric history, and socioeconomic status. *JAMA Psychiatry*, *78*, 387–397. https://doi.org/10.1001/jamapsychiatry.2020.4172

Almeida, O. P., Pirkis, J., Kerse, N., Sim, M., Flicker, L., Snowdon, J., Draper, B., Byrne, G., Lautenschlager, N. T., Stocks, N., Alfonso, H., & Pfaff, J. J. (2012). Socioeconomic disadvantage increases risk of prevalent and persistent depression in later life. *Journal of Affective Disorders*, *138*, 322–331. https://doi.org/10.1016/j.jad.2012.01.021

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). https://doi.org/10.1176/appi.books.9780890425596.

Arias-de la Torre, J., Vilagut, G., Martín, V., Molina, A. J., & Alonso, J. (2018). Prevalence of major depressive disorder and association with personal and socio-economic factors. Results for Spain of the European Health Interview Survey 2014–2015. *Journal of Affective Disorders*, *239*, 203–207. https://doi.org/10.1016/j.jad.2018.06.051

Australian Bureau of Statistics. (2021). *Socio-Economic Indexes for Areas (SEIFA): Technical paper*. https://www.abs.gov.au/statistics/detailed-methodology-information/concepts-sources-methods/socio-economic-indexes-areas-seifa-technical-paper/2021.

Beard, J. R., Cerdá, M., Blaney, S., Ahern, J., Vlahov, D., & Galea, S. (2009). Neighborhood characteristics and change in depressive symptoms among older residents of New York City. *American Journal of Public Health*, *99*, 1308–1314. https://doi.org/10.2105/AJPH.2007.125104

Brugha, T., Bebbington, P., Tennant, C., & Hurry, J. (1985). The List of Threatening Experiences: A subset of 12 life event categories with considerable long-term contextual threat. *Psychological Medicine*, *15*, 189–194. https://doi.org/10.1017/s003329170002105

Byrne, E. M., Kirk, K. M., Medland, S. E., McGrath, J. J., Colodro-Conde, L., Parker, R., Cross, S., Sullivan, L., Statham, D. J., Levinson, D. F., Licinio, J., Wray, N. R., Hickie, I. B., & Martin, N. G. (2020). Cohort profile: The Australian genetics of depression study. *BMJ Open*, *10*, e032580. https://doi.org/10.1136/bmjopen-2019-032580

Colodro-Conde, L., Couvy-Duchesne, B., Whitfield, J. B., Streit, F., Gordon, S., Kemper, K. E., Yengo, L., Zheng, Z., Trzaskowski, M., de Zeeuw, E. L., Nivard, M. G., Das, M., Neale, R. E., MacGregor, S., Olsen, C. M., Whiteman, D. C., Boomsma, D. I., Yang, J., Rietschel, M., ¼ Martin, N. G. (2018). Association between population density and genetic risk for schizophrenia. *JAMA Psychiatry*, *75*, 901–910. https://doi.org/10.1001/jamapsychiatry.2018.1581

Ferrari, A. J., Charlson, F. J., Norman, R. E., Patten, S. B., Freedman, G., Murray, C. J., Vos, T., & Whiteford, H. A. (2013). Burden of depressive disorders by country, sex, age, and year: Findings from the global burden of disease study 2010. *PLoS Medicine*, *10*, e1001547. https://doi.org/10.1371/journal.pmed.1001547

Gomez, L., Díaz-Torres, S., Colodro-Conde, L., Garcia-Marin, L. M., Yap, C. X., Byrne, E. M., Yengo, L., Lind, P. A., Wray, N. R., Medland, S. E., Hickie, I. B., Lupton,. M. K., Rentería, M. E., Martin, N. G., & Campos, A. I. (2023). Phenotypi
and genetic factors associated with donation of DNA and consent to record linkage for prescription history in the Australian Genetics of Depression Study. *European Archives of Psychiatry and Clinical Neuroscience*, *273*, 1359–1368. https://doi.org/10.1007/s00406-022-01527-0

Guan, N., Guariglia, A., Moore, P., Xu, F., & Al-Janabi, H. (2022). Financial stress and depression in adults: A systematic review. *PloS One*, *17*, e0264041. https://doi.org/10.1371/journal.pone.0264041

Halldorsdottir, T., Piechaczek, C., Soares de Matos, A. P., Czamara, D., Pehl, V., Wagenbuechler, P., Feldmann, L., Quickenstedt-Reinhardt, P., Allgaier, A.-K., Freisleder, F. J., Greimel, E., Kvist, T., Lahti, J., Räikkönen, K., Rex-Haffner, M., Arnarson, E. Ö., Craighead, W. E., Schulte-Körne, G., & Binder, E. B. (2019). Polygenic risk: Predicting depression outcomes in clinical and epidemiological cohorts of youths. *American Journal of Psychiatry*, *176*, 615–625. https://doi.org/10.1176/appi.ajp.2019.18091014

Howard, D. M., Adams, M. J., Clarke, T.-K., Hafferty, J. D., Gibson, J., Shirali, M., Coleman, J. R., Hagenaars, S. P., Ward, J., Wigmore, E. M, Alloza, C., Shen, X., Barbu, M. C., Xu, E. Y., Whalley, H. C., Marioni, R. E., Porteous, D. J., Davies, G., Deary, I. J., ¼ McIntosh, A. M. (2019). Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nature Neuroscience*, *22*, 343–352. https://doi.org/10.1038/s41593-018-0326-7

Hyde, C. L., Nagle, M. W., Tian, C., Chen, X., Paciga, S. A., Wendland, J. R., Tung, J. Y., Hinds, D. A., Perlis, R. H., & Winslow, A. R. (2016). Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nature Genetics*, *48*, 1031–1036. https://doi.org/10.1038/ng.362

Jaffee, S. R., & Price, T. S. (2007). Gene–environment correlations: A review of the evidence and implications for prevention of mental illness. *Molecular Psychiatry*, *12*, 432–442. https://doi.org/10.1038/sj.mp.4001950

Kendler, K. S., & Karkowski-Shuman, L. (1997). Stressful life events and genetic liability to major depression: Genetic control of exposure to the environment? *Psychological Medicine*, *27*, 539–547. https://doi.org/10.1017/s0033291797004716

Kosciuszko, M., Steptoe, A., & Ajnakina, O. (2023). Genetic propensity, socioeconomic status, and trajectories of depression over a course of 14 years in older adults. *Translational Psychiatry*, *13*, 68. https://doi.org/10.1038/s41398-023-02367-9

Lloyd-Jones, L. R., Zeng, J., Sidorenko, J., Yengo, L., Moser, G., Kemper, K. E., Wang, H., Zheng, Z., Magi, R., Esko, T., Metspalu, A., Wray, N. R., Goddard, M. E., Yang, J., & Visscher, P. M. (2019). Improved polygenic prediction by Bayesian multiple regression on summary statistics. *Nature Communications*, *10*, 5086. https://doi.org/10.1038/s41467-019-12653-0

Machlitt-Northen, S., Keers, R., Munroe, P. B., Howard, D. M., & Pluess, M. (2022). Polygenic risk scores for schizophrenia and major depression are associated with socio-economic indicators of adversity in two British community samples. *Translational Psychiatry*, *12*, 477. https://doi.org/10.1038/s41398-022-02247-8

Mitchell, B. L., Thorp, J. G., Wu, Y., Campos, A. I., Nyholt, D. R., Gordon, S. D., Whiteman, D. C., Olsen, C. M., Hickie, I. B., Martin, N. G., Medland, S. E., Wray, N. B., & Byrne, E. M. (2021). Polygenic risk scores derived from varying definitions of depression and risk of depression. *JAMA Psychiatry*, *78*, 1152–1160. https://doi.org/10.1001/jamapsychiatry.2021.1988

Mullins, N., Power, R., Fisher, H., Hanscombe, K., Euesden, J., Iniesta, R., Levinson, D., Weissman, M. M., Potash, J., Shi, J., Uher, R., Cohen-Woods, S., Rivera, M., Jones, L., Jones, I., Craddock, N., Owen, M. J., Korszun, A., Craig, I. W., ¼ Lewis, C. M. (2016). Polygenic interactions with environmental adversity in the aetiology of major depressive disorder. *Psychological Medicine*, *46*, 759–770. https://doi.org/10.1017/S0033291715002172

Musliner, K. L., Agerbo, E., Vilhjálmsson, B. J., Albiñana, C., Als, T. D., Østergaard, S. D., & Mortensen, P. B. (2021). Polygenic liability and recurrence of depression in patients with first-onset depression treated in hospital-based settings. *JAMA Psychiatry*, *78*, 792–795. https://doi.org/10.1001/jamapsychiatry.2021.0701

Musliner, K. L., Seifuddin, F., Judy, J. A., Pirooznia, M., Goes, F. S., & Zandi, P. P. (2015). Polygenic risk, stressful life events and depressive symptoms in older adults: A polygenic score analysis. *Psychological Medicine*, *45*, 1709–1720. https://doi.org/10.1017/S0033291714002839

Oenning, N. S. X., Ziegelmann, P. K., De Goulart, B. N. G., & Niedhammer, I. (2018). Occupational factors associated with major depressive disorder: A Brazilian population-based study. *Journal of Affective Disorders*, *240*, 48–56. https://doi.org/10.1016/j.jad.2018.07.022

Patria, B. (2022). The longitudinal effects of education on depression: Finding from the Indonesian national survey. *Frontiers in Public Health*, *10*, 1017995. https://doi.org/10.3389/fpubh.2022.1017995

Peen, J., Schoevers, R. A., Beekman, A. T., & Dekker, J. (2010). The current status of urban-rural differences in psychiatric disorders. *Acta Psychiatrica Scandinavica*, *121*, 84–93. https://doi.org/10.1111/j.1600-0447.2009.01438.x

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., De Bakker, P. I., & Daly, M. J. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, *81*, 559–575. https://doi.org/10.1086/519795

Rojas-Garcia, A., Ruiz-Perez, I., Rodriguez-Barranco, M., Bradley, D. C. G., Pastor-Moreno, G., & Ricci-Cabello, I. (2015). Healthcare interventions for depression in low socioeconomic status populations: A systematic review and meta-analysis. *Clinical Psychology Review*, *38*, 65–78. https://doi.org/10.1016/j.cpr.2015.03.001

Santiago, C. D., Wadsworth, M. E., & Stump, J. (2011). Socioeconomic status, neighborhood disadvantage, and poverty-related stress: Prospective effects on psychological syndromes among diverse low-income families. *Journal of Economic Psychology*, *32*, 218–230. https://doi.org/10.1016/j.joep.2009.10.008

Sullivan, P. F., Neale, M. C., & Kendler, K. S. (2000). Genetic epidemiology of major depression: Review and meta-analysis. *American Journal of Psychiatry*, *157*, 1552–1562. https://doi.org/10.1176/appi.ajp.157.10.1552.

Sund, E. R., Van Lenthe, F. J., Avendano, M., Raina, P., & Krokstad, S. (2021). Does urbanicity modify the relationship between a polygenic risk score for depression and mental health symptoms? Cross-sectional evidence from the observational HUNT Study in Norway. *Journal of Epidemiology and Community Health*, *75*, 420–425. https://doi.org/10.1136/jech-2020-214256

Uher, R., & McGuffin, P. (2010). The moderation by the serotonin transporter gene of environmental adversity in the etiology of depression: 2009 update. *Molecular Psychiatry*, *15*, 18–22. https://doi.org/10.1038/mp.2009.123

Wang, J. L., Schmitz, N., & Dewa, C. (2010). Socioeconomic status and the risk of major depression: The Canadian National Population Health Survey. *Journal of Epidemiology & Community Health*, *64*, 447–452. https://doi.org/10.1136/jech.2009.090910

Wise, P., & Mathews, R. (2011). *Socio-economic indexes for areas: Getting a handle on individual diversity within* areas [Report]. Australian Bureau of Statistics.

World Health Organization. (2017). *Depression and other common mental disorders: Global health estimates*. https://www.who.int/publications/i/item/depression-global-health-estimates

Wray, N. R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., Abdellaoui, A., Adams, M. J., Agerbo, E., Air, T. M., & Andlauer, T. M. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics*, *50*, 668–681. https://doi.org/10.1038/s41588-018-0090-3.

Zhang, X., Qiao, Y., Wang, M., Liang, X., Zhang, M., Li, C., Cairang, J., Wang, J., Bi, H., & Gao, T. (2023). The influence of genetic and acquired factors on the vulnerability to develop depression: A review. *Bioscience Reports*, *43*, BSR20222644. https://doi.org/10.1042/BSR20222644