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

Erik Reidar Sund,1,2,3 Frank J van Lenthe,4,5 Mauricio Avendano,6,7 Parminder Raina,8,9,10 Steinar Krokstad1,3

ABSTRACT

Background Research suggests that genetic predisposition for common mental disorders may be moderated by the environment. This study examines whether a polygenic risk score (PRS) for depression is moderated by the level of residential area urbanicity using five symptoms of poor mental health as outcomes.

Methods The study sample consisted of 41,198 participants from the 2006–2008 wave of the Norwegian HUNT study. We created a weighted PRS for depression based on 99 variants identified in a recent genome-wide association study. Participants were classified into urban or rural place of residence based on wards that correspond to neighbourhoods. Mixed effects logistic regression models with participants nested in 477 neighbourhoods were specified.

Results A SD increase in PRS for depression was associated with a small but statistically significant increase in the odds of anxiety, comorbid anxiety and depression and mental distress. Associations for depression were weaker and not statistically significant. Compared with urban residents, rural resident had higher odds for reporting poor mental health. Genetic propensity for depression was higher for residents of urban than rural areas, suggesting gene–environment correlation. There was no sign of effect modification between genetic propensity and urbanicity for depression, anxiety, comorbid anxiety and depression, or mental distress.

Conclusion The PRS predicted small but significant odds of anxiety, comorbid anxiety and depression and mental distress, but we found no support for a differential effect of genetic propensity in urban and rural neighbourhoods for any of the outcomes.

INTRODUCTION

Common mental disorders are a leading contributor to morbidity and disability and represent a substantial public health problem worldwide. Both depressive disorders, characterised by sustained symptoms of sadness, low energy and sleep disturbances, as well as anxiety disorders, defined by excess worry, hyperarousal and fear, are highly prevalent and they show a high degree of comorbidity. The risk of common mental disorders varies by age, sex, socioeconomic status and has also been found to vary geographically.

The aetiology of both depression and anxiety is complex, but likely determined by genetic, social and environmental factors in a complex interplay. Discoveries from genome-wide association studies (GWAS) suggest that mental health disorders are highly polygenic, that is, they are influenced by hundreds or thousands of genetic variants each having a small effect, but overall determining an individual’s genetic predisposition. On their own, however, genetic factors are unlikely to explain a large share of variation in mental health disorders, which are also strongly influenced by the environment. One important environmental factor is captured by urbanicity, which refers to the impact of living in urban areas at a given point in time, and the presence of conditions that are more prevalent than in non-urban areas. This may confer both an urban penalty, for example, by increasing exposure to air pollution or violence, or an urban advantage, conferred by higher access to services, cultural activities or social networks. Individuals living in rural areas will generally experience a different environment, typically less stressful, less noise and with much less air pollution. A recent review found conflicting evidence for urban–rural variation prevalent for common mental disorders.

The recognition that both genes (‘nature’) and environments (‘nurture’) contribute to the aetiology of psychiatric disorders has motivated the study of gene–environment interactions (GxE). GxE studies examine to what extent genetic propensity modifies the association between environmental factors and mental health, or conversely, how environmental factors modify associations between genes and mental health. Conceptually, this line of inquiry builds on the diathesis–stress model that posits that genetic propensity (diathesis) interacts, for example, with stressful life events (SLE) to give rise to adverse mental health outcomes. According to this model, genes may exacerbate or buffer the effects of stressful environments. Previous studies on depression rooted in the diathesis–stress model and using polygenic risk scores (PRS) have shown inconsistent results.

A recent test of the diathesis–stress model on depression using PRS and SLE found a significant diathesis–stress interaction, but these results are yet to be reproduced. The majority of GxE studies adhere to the diathesis–stress model, but alternatives like the differential susceptibility model exist. According to this model, individuals vary in their susceptibility to both positive and negative environmental influences...
rather than claiming that specific genotypes are good or bad.

In this study, we aim to assess the hypothesis that the urban environment modifies the relationship between genes and mental health disorders. The majority of GxE studies within the domain of mental health have used the term ‘environment’ to refer to individual-level factors such as behaviour or major life events, while no studies have examined the interaction between genes and the wider physical and social environment. Our study is based on the Nord-Trøndelag Health study (HUNT), a large general population-based study with substantial variation in level of urbanicity and with detailed genetic data, that enables assessing differential effects of genetic propensity on five mental health outcomes by level of urbanicity.

METHODS

Data material

Data from the third wave of the Nord-Trøndelag Health study (HUNT3) was used. The total population above 19 years in the Nord-Trøndelag county were invited (N=93 860) of which 50 802 participated, yielding a response of 54%. The data include questionnaire information on health, lifestyle, drug treatment and relational issues like family situation. Clinical measurement data and blood samples were collected at screening stations established on several locations (N=23) in the county. Due to the administration of the two main questionnaires (the first sent by mail and brought to the screening station and the second received at the screening station and mailed afterwards), a lower number of respondents had answered the second questionnaire that contained questions on mental health (N=41 198). A study among non-respondents conducted after HUNT3 found that non-participants were more likely to have lower socio-economic status, higher mortality and a higher prevalence of chronic diseases. The regional committee for medical research ethics approved the study and all participants provided written consent.

Outcome measures

Two different measurement instruments for mental health were used in HUNT3: The Hospital Anxiety and Depression Scale (HADS) measures symptoms of anxiety and depression and consists of 14 questions (seven each for anxiety and depression) to depression (HADS-D). Each subscale ranges from 0 to 21 and a score of ≥8 has been found to be the optimal cut-off with a sensitivity and specificity of ca. 0.8. Comorbid anxiety and depression were also constructed based on these cut-offs. For the depression subscale, we additionally chose a cut-off of 11 (≥11) to indicate a more severe symptom load.

The Mental Health Index (MHI) consists of seven items with values from 1 to 4. The 8 has been found to be the optimal cut-off which were given values from 1 to 4. The scale (as ratios). While interactions on the multiplicative scale in non-linear models are readily available, additive interactions require some extra calculations and here we followed recommendations from recent methodological literature. Specifically, from model 2 we calculated the marginal effects of the PRS on the urban and rural categories and less than 20% were classified as ‘semi-urban’. The remainder living in wards where more than 20% were living close were classified as ‘urban’.

Covariates

All models controlled for age (entered as a restricted cubic spline (RCS) with 4 knots), sex and five ancestry-informative principal components (PCs), which account for population stratification.

Statistical analysis

Mixed effect logistic regression models were used to account for the data structure with individuals nested in 477 wards. First, we regressed each outcome on the PRS adjusting for age (RCS), sex and the first five ancestry-informative PCs (model 1). Second, we added urbanicity (model 2), and third, we expanded the models by adding an interaction term between the PRS and urbanicity (model 3). Fixed effects are reported as ORs with 95% CIs and random effects as variances on the log-odds scale. Effects from interaction terms in non-linear models are scale-dependent and the current advice is to report interactions on both the additive (as differences) and multiplicative scale (as ratios). While interactions on the multiplicative scale in non-linear models are readily available, additive interactions require some extra calculations and here we followed recommendations from recent methodological literature. Specifically, from model 2 we calculated the marginal effects of the PRS on the urban and rural categories and less than 20% were classified as ‘semi-urban’. The remainder living in wards where more than 20% were living close were classified as ‘urban’.
marginal effects were different between urban and rural individuals using p<0.05 as the threshold for statistical significance. In an additional test for additive interactions, we also specified linear probability models. Given that interactions can be hard to interpret, we visualised the predictions according to the urban–rural place of residence and the PRS for one of the outcomes (HADS-D8).

We also specified a model to investigate gene-environment correlations (rGE) by regressing urbanicity on the PRS adjusting for age, sex and ancestry. Checking for rGE is important because what appears as interactions may in fact be correlations, that is, the level of genetic propensities may be different in urban and rural wards. We performed a complete case analysis excluding participants with missing values. Data management and statistical modelling were performed in Stata v.15.26

RESULTS

Table 1 shows the descriptive characteristics of the sample. Their mean age was 54.4 years, there were more women (56%) than men, and most participants lived in urban neighbourhoods (70%). There were between 4% and 7.4% missing on the outcomes. Symptoms of anxiety were the most prevalent condition (70%), while symptoms of severe depression (HADS-D cut-off 11) were the least prevalent condition (2.2%).

Model 1 in table 2 shows the main effects of the PRS on the five mental health outcomes adjusted for age, sex and ancestry. A SD increase in PRS was associated with a significant 1.08 (95% CI 1.05 to 1.12) increased odds of moderate-to-severe anxiety (HADS-D & A&D), a 1.05 (95% CI 1.00 to 1.10) increased odds of comorbid A&D and a 1.08 (95% CI 1.04 to 1.12) increased odds of mental distress. By contrast, associations were not significant for moderate-to-severe symptoms of depression (HADS-D8) (1.03, 95% CI 0.99 to 1.06) and severe depression (HADS-D11) (1.05, 95% CI 0.98 to 1.12).

In model 2, the indicator for urban–rural place of residence was added together with variables from model 1. Compared with urban residents, rural resident had an increased odds for reporting poor mental health on all outcomes except for mental distress. Figure 1 depicts ORs and 95% CIs from model 2.

Model 3 (table 2) expands model 2 by including an interaction term between the PRS and urban–rural living. In model 3, the main effect of the PRS for urban participants was 1.04 (95% CI 1.00 to 1.09) for HADS-D8 and 1.09 (95% CI 1.00 to 1.18) for HADS-D11, whereas the other main effects for urban participants were similar to the effects in model 1 for all participants. The interaction terms suggest a decreased risk for rural participants compared with urban participants associated with 1 SD increase in polygenic scores for moderate-to-severe depression (OR 0.96, 95% CI 0.89 to 1.03) and severe depression (OR 0.91, 95% CI 0.79 to 1.05), but these associations were not statistically significant. We found no evidence of interactions on the additive scale (online supplementary table 1). No interactions were found in models stratified either by sex or age (over/under 50 years).

Figure 2 shows the predicted probability (95% CI) for moderate-to-severe symptoms of depression according to PRS and urbanicity and shows a different effect of the PRS for urban participants compared to rural participants. A test for simple slope for urban participants was not statistically significant (p=0.06).

Analyses with a three-group classification of urbanicity showed that there was a dose–response relationship with urbanicity, where the odds of reporting poor mental health increased with decreasing level of urbanicity (online supplementary table 2). No interactions were found between the PRS and urbanicity.

DISCUSSION

Our results confirm prior findings suggesting that a PRS for depression has a small but significant association with the risk of mental health outcomes. However, we found no evidence that the effect of genetic propensity differs between urban and rural areas for any of the mental health outcomes examined.

Comparison with previous research

Few previous studies have used a truly environmental spatial construct to investigate moderated effects of genetic propensity for mental health phenotypes. One study from the USA found that the genetic propensity for smoking predicted higher mean number of cigarettes smoked per day in neighbourhoods with a low level of social cohesion than in neighbourhoods with high social cohesion.27 A more recent study from the Netherlands...
tested interactions between a PRS for substance abuse and a number of neighborhood characteristics and found that only 1 of 14 tested interactions was statistically significantly related to substance abuse.28 Another recent study suggests that a PRS for schizophrenia was more strongly related to treatment-resistant schizophrenia in rural and semiurban areas (HR: 1.20) compared with the capital area.29 Our study adds to the evidence of inconsistent findings in the GxE literature looking at higher-order environmental features. There may be methodological issues causing these inconsistencies or more fundamental flaws in the underlying theoretical models. Most studies have been rooted in the diathesis–stress framework, but the differential susceptibility model may also be important. However, variants from GWAS might not capture differential susceptibility and thus not constitute the best measure for GxE studies.30

### Interpretation of findings
The PRS we tested on five different symptoms of poor mental health was significantly associated with several of the mental health outcomes examined, but associations were relatively small. As a consequence, our ability to find GxE was small. While the GWAS found the reported genetic variants to be robust across three studies, they replicated poorly for the phenotypes in our sample (details available from the corresponding author).

#### Table 2 Associations§ between a polygenic risk score for depression and five mental health outcomes.

<table>
<thead>
<tr>
<th></th>
<th>HADS-D (≥8)</th>
<th>HADS-D (≥11)</th>
<th>HADS-A (≥8)</th>
<th>Comorbid A&amp;D</th>
<th>Mental health score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1 Fixed effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRS (z-score)</td>
<td>1.03 (0.99 to 1.06)</td>
<td>1.05 (0.98 to 1.12)</td>
<td>1.08 (1.05 to 1.12)</td>
<td>1.05 (1.00 to 1.10)</td>
<td>1.08 (1.04 to 1.12)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>1 ref.</td>
<td>1 ref.</td>
<td>1 ref.</td>
<td>1 ref.</td>
<td>1 ref.</td>
</tr>
<tr>
<td>Men</td>
<td>1.20 (1.12 to 1.28)</td>
<td>1.14 (0.99 to 1.30)</td>
<td>0.55 (0.52 to 0.59)</td>
<td>0.78 (0.70 to 0.86)</td>
<td>0.81 (0.75 to 0.88)</td>
</tr>
<tr>
<td>Random effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance (SE)</td>
<td>0.070 (0.015)</td>
<td>0.097 (0.039)</td>
<td>0.036 (0.009)</td>
<td>0.068 (0.022)</td>
<td>0.081 (0.019)</td>
</tr>
<tr>
<td>Model 2 Fixed effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRS (z-score)</td>
<td>1.03 (0.99 to 1.06)</td>
<td>1.05 (0.98 to 1.12)</td>
<td>1.08 (1.05 to 1.12)</td>
<td>1.05 (1.00 to 1.10)</td>
<td>1.08 (1.04 to 1.12)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>1 ref.</td>
<td>1 ref.</td>
<td>1 ref.</td>
<td>1 ref.</td>
<td>1 ref.</td>
</tr>
<tr>
<td>Men</td>
<td>1.19 (1.11 to 1.28)</td>
<td>1.13 (0.98 to 1.30)</td>
<td>0.55 (0.52 to 0.59)</td>
<td>0.77 (0.70 to 0.85)</td>
<td>0.81 (0.75 to 0.88)</td>
</tr>
<tr>
<td>Urbanicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>1 ref.</td>
<td>1 ref.</td>
<td>1 ref.</td>
<td>1 ref.</td>
<td>1 ref.</td>
</tr>
<tr>
<td>Rural</td>
<td>1.18 (1.08 to 1.29)</td>
<td>1.33 (1.14 to 1.56)</td>
<td>1.08 (1.00 to 1.16)</td>
<td>1.16 (1.03 to 1.30)</td>
<td>1.01 (0.91 to 1.12)</td>
</tr>
<tr>
<td>Random effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance (SE)</td>
<td>0.060 (0.014)</td>
<td>0.080 (0.035)</td>
<td>0.034 (0.009)</td>
<td>0.060 (0.022)</td>
<td>0.081 (0.019)</td>
</tr>
<tr>
<td>Model 3 Fixed effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRS (z-score)</td>
<td>1.04 (1.00 to 1.09)</td>
<td>1.09 (1.00 to 1.18)</td>
<td>1.08 (1.04 to 1.12)</td>
<td>1.05 (1.00 to 1.12)</td>
<td>1.08 (1.03 to 1.14)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>1 ref.</td>
<td>1 ref.</td>
<td>1 ref.</td>
<td>1 ref.</td>
<td>1 ref.</td>
</tr>
<tr>
<td>Men</td>
<td>1.19 (1.11 to 1.28)</td>
<td>1.13 (0.99 to 1.30)</td>
<td>0.55 (0.52 to 0.59)</td>
<td>0.77 (0.70 to 0.85)</td>
<td>0.81 (0.75 to 0.88)</td>
</tr>
<tr>
<td>Urbanicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>1 ref.</td>
<td>1 ref.</td>
<td>1 ref.</td>
<td>1 ref.</td>
<td>1 ref.</td>
</tr>
<tr>
<td>Rural</td>
<td>1.18 (1.08 to 1.29)</td>
<td>1.34 (1.15 to 1.56)</td>
<td>1.08 (1.00 to 1.16)</td>
<td>1.16 (1.03 to 1.30)</td>
<td>1.01 (0.91 to 1.12)</td>
</tr>
<tr>
<td>PRS*Rural</td>
<td>0.96 (0.89 to 1.03)</td>
<td>0.91 (0.79 to 1.05)</td>
<td>1.00 (0.94 to 1.07)</td>
<td>0.98 (0.89 to 1.09)</td>
<td>0.99 (0.90 to 1.08)</td>
</tr>
<tr>
<td>Random effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance (SE)</td>
<td>0.060 (0.014)</td>
<td>0.080 (0.035)</td>
<td>0.034 (0.009)</td>
<td>0.060 (0.022)</td>
<td>0.081 (0.019)</td>
</tr>
</tbody>
</table>

§Adjusted for age (restricted cubic splines with 4 knots) and first five principal components.
HADS, Hospital Anxiety and Depression Scale; PRS, polygenic risk score. ref.=reference category

![Figure 1](Figure 1 OR and 95% CI (95% CI) for poor mental health in rural areas (ref=urban areas).)
A possible explanation for this discrepancy is that the genetic variants used to calculate the PRS came from a GWAS on major depression, while the phenotypes we studied were symptoms of poor mental health.

Urbanicity may constitute a very heterogeneous environmental construct encompassing both risk factors and protective factors, for example, urban environments may be more stressful, but at the same time, access to health services or social networks may reduce stress and depression. Previous studies have largely studied environmental conditions that operate at the individual level, such as childhood trauma, SLE and social support. By contrast, a characteristic of the area where individuals reside capture higher-order effects that are more difficult to capture when using individual-level data, making it also more challenging to identify GxE interactions.

When studying gene-environment interactions (GxE), it is important to simultaneously check for gene-environment correlations (rGE), because what appears as interactions may in fact reflect clustering according to genetic propensities. While rGE reflect genetic differences in exposure to particular environments, GxE refers to genetic differences in susceptibility to particular environments. When testing rGE, we found the PRS predicted urban residence, thus suggesting gene-environment correlations. When interpreting this finding, it is possible that our suggestive gene-environment interaction for depression is in fact gene-environment correlation, that is, genetic propensity for depression is more prevalent in urban areas. A higher prevalence may occur when individuals self-select environments guided by their genetic predispositions. This makes the interpretation of GxE cumbersome, as the interaction might arise as a result of genetic propensities for urban residential choice. A closely related interpretation of this finding is that polygenic scores influence the risk of depression and anxiety earlier in life and that the latter influence the probability of residing in urban areas, reflecting ‘reverse causality’. While we have treated rGE as a disturbing element in the pursuit of GxE, it is an interesting phenomenon largely ignored in the GxE literature, but it might be equally or even more important in the aetiology of mental health problems.

Our study has several strengths. It is conducted in a large general population sample and we used validated instruments as outcomes. Urbanicity, constructed from an external data source, was based on a detailed classification of place of residence in accordance with Statistics Norway’s definition of urban areas. Delineating urban–rural neighbourhoods based on wards is preferable, because this is the lowest spatial scale possible and corresponds closely with neighbourhoods, thus making them sociodemographic homogenous within and heterogeneous between. We developed a PRS based on the most recent GWAS reporting 102 genome-wide significant associations with major depression in populations of European ancestry. Thus, we had a very large and independent discovery sample that allowed us to derive the PRS.

Nevertheless, a number of limitations should be considered in this study. The response rate was 54% and a non-participation study has shown that non-participants had poorer health. Missing was in general low (<5%), but the MHI index with 7.4% missingness can be biased. The symptom scores used as outcomes were collected at one timepoint only. The genetic variants used to calculate the PRS were derived from a GWAS on major depression, and while the phenotypes we have studied are closely related to major depression, they are nevertheless symptoms and not clinically assessed diagnoses. Further, we lacked the possibility to adjust analyses for genotyping arrays. Finally, we performed an analysis on participants with valid information and made no attempt to impute missing data.

CONCLUSION

The PRS had a significant but small association with symptoms of anxiety, comorbid anxiety and depression and mental distress. We found no support for a differential effect of genetic propensity between urban and rural neighbourhoods. While our findings do not support the hypothesis of gene-environment interactions using PRS, other approaches such as genome-wide by environment interaction studies represents a potential alternative to understand how genetic variants interact with specific features of the urban environment. The value of doing GxE studies ultimately lies in their potential for advancing our understanding of causal pathways with respect to both genetic and environmental mechanisms in the origin of adverse mental health.

What is already known on this topic

- Studies suggest that genetic factors play an important role in both anxiety and depression and that genetic propensity may be contingent on environmental characteristics, that is, environment may modify the effect of genetic propensity.

What this study adds

- Genetic propensity for major depression, operationalised through a polygenic risk score, was associated with symptoms of anxiety, depression and mental distress, but there was no evidence of modification by residential urbanicity.

Author affiliations

1HUNT Research Centre, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Levanger, Norway
2Faculty of Nursing and Health Sciences, Nord University, Levanger, Norway
3Levanger Hospital, Nord-Trøndelag Hospital Trust, Levanger, Norway
4Department of Public Health, Erasmus MC, Rotterdam, Netherlands
The Nord-Trøndelag Health Study (HUNT) is a collaboration between the HUNT Research Centre (Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology), the Nord-Trøndelag County Council, and the Norwegian Institute of Public Health.

Contributors All authors contributed to the conceptualisation and design of the study. ERS performed the analysis and FJvL, MA, PR and SK contributed to the interpretation of the data. ERS wrote the first draft of the article and FJvL, MA, PR and SK revised it critically for important intellectual content. All authors have read and approved the final version of the manuscript. ERS is the guarantor for the overall content of the study.

Funding This work was supported by European Union Horizon 2020 award “Mindmap: Promoting mental wellbeing in the ageing urban population” (grant number 667661). Parminder Raina holds a Tier 1 Canada Research Chair in Geroscience and the Raymond and Margaret Labarge Chair in Research and Knowledge Application for Optimal Aging. The authors have not declared aspec aspecific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Data sharing statement Data are available upon reasonable request.

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: https://creativecommons.org/licenses/by/4.0/.

ORCID ids Erik Reidar Sund http://orcid.org/0000-0002-0736-5139
Mauricio Avendano http://orcid.org/0000-0002-7295-2911

REFERENCES
26. StataCorp. Stata statistical software. College Station, TX: StataCorp LP, 2015.