Predictive Performance of Exposome Score for Schizophrenia in the General Population

Lotta-Katrin Pries[1,8], Gamze Erzin[2,8], Jim van Os[1,3,4], Margreet ten Have[5], Ron de Graaf[5], Saskia van Dorsselaer[5], Maarten Bak[1,6], Bart P. F. Rutten[1], and Sinan Guloksuz*[,1,7]

[1]Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience, Maastricht University Medical Center, Maastricht, The Netherlands; [2]Department of Psychiatry, Ankara Dışkapı Yıldırım Beyazıt Training and Research Hospital, Ankara, Turkey; [3]Department of Psychiatry, UMC Utrecht Brain Center, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands; [4]Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK; [5]Department of Epidemiology, Netherlands Institute of Mental Health and Addiction, Utrecht, The Netherlands; [6]FACT, Mondriaan Mental Health, Maastricht, The Netherlands; [7]Department of Psychiatry, Yale University School of Medicine, New Haven, CT

[8]These authors contributed equally to the article.

*To whom correspondence should be addressed; tel: 31-433-88-4071, fax: 31433-88-4122, e-mail: sinan.guloksuz@maastrichtuniversity.nl

Previously, we established an estimated exposome score for schizophrenia (ES-SCZ) as a cumulative measure of environmental liability for schizophrenia to use in gene–environment interaction studies and for risk stratification in population cohorts. Hereby, we examined the discriminative function of ES-SCZ for identifying individuals diagnosed with schizophrenia spectrum disorder in the general population by measuring the area under the receiver operating characteristic curve (AUC). Furthermore, we compared this ES-SCZ method to an environmental sum score (Esum-SCZ) and an aggregate environmental score weighted by the meta-analytical estimates (Emet-SCZ). We also estimated ORs and Nagelkerke's R^2 for ES-SCZ in association with psychiatric diagnoses and other medical outcomes. ES-SCZ showed a good discriminative function (AUC = 0.84) and statistically significantly performed better than both Esum-SCZ (AUC = 0.80) and Emet-SCZ (AUC = 0.80). At optimal cut point, ES-SCZ showed similar performance in ruling out (LR− = 0.20) and ruling in (LR+ = 3.86) schizophrenia. ES-SCZ at optimal cut point showed also a progressively greater magnitude of association with increasing psychosis risk strata. Among all clinical outcomes, ES-SCZ was associated with schizophrenia diagnosis with the highest OR (2.76, $P < .001$) and greatest explained variance ($R^2 = 14.03\%$), followed by bipolar disorder (OR = 2.61, $P < .001$, $R^2 = 13.01\%$) and suicide plan (OR = 2.44, $P < .001$, $R^2 = 12.44\%$). Our findings from an epidemiologically representative general population cohort demonstrate that an aggregate environmental exposure score for schizophrenia constructed using a predictive modeling approach—ES-SCZ—has the potential to improve risk prediction and stratification for research purposes and may help gain insight into the multicausal etiology of psychopathology.

Key words: exposome/risk score/schizophrenia/environment/psychosis/prediction

Introduction

Schizophrenia spectrum disorder is a heterogeneous phenotype with a complex pathoetiology that involves a multitude of genetic and environmental risk factors and their interaction.[1] The progress of genome-wide association studies (GWAS) has paved the way for polygenic risk estimation of schizophrenia (polygenic risk score [PRS]: a weighted sum of trait-associated alleles) to measure molecular genetic liability as a single metric.[2] The recent release of the Psychiatric Genomics Consortium (PGC3) shows that PRS for schizophrenia (PRS-SCZ) explains up to 7.7% of variation on the liability scale to schizophrenia.[3] Recent studies investigating the electronic health records in the United States showed that PRS-SCZ was associated with schizophrenia diagnosis in the population.[4,5] Of PRS for mental disorder phenotypes that have been approximated thus far, PRS-SCZ seems

to be the most outstanding for testing PRS-based prediction for population risk stratification and clinical applications. However, PRS-SCZ is not distinctly associated with schizophrenia but is also associated with several psychiatric and other medical conditions[4,6,7] and subclinical multidimensional phenotypes,[7–10] as well as general mental and physical health outcomes.[11,12]

Schizophrenia spectrum disorder, similar to its polygenic composition, has been associated with several environmental exposures, including cannabis use, childhood adversities (eg, sexual abuse, peer-bullying, and emotional neglect), obstetric and pregnancy complications, proxies of social exclusion (eg, ethnic minority and hearing impairment immigration), and season of birth (winter birth).[13] These environmental factors are often correlated to a degree and comprise a network at population level: the exposome.[14,15] To supplement PRS-SCZ in gene–environment interaction studies and risk stratification in population cohorts, we have estimated an exposome score for schizophrenia (ES-SCZ) as a cumulative measure of environmental liability for schizophrenia.[16] Recently, we have demonstrated that the ES-SCZ is associated with psychosis risk states,[17] as well as mental and physical health in the general population.[12]

In the present study, we aimed to examine the discriminative function of ES-SCZ for identifying individuals diagnosed with schizophrenia spectrum disorder in the general population by measuring the area under the receiver operating characteristic (AUC), sensitivity, specificity, and positive and negative likelihood ratios. We also estimated ORs and Nagelkerke's R^2 for ES-SCZ in association with psychiatric diagnoses and other medical outcomes.

Methods

Study Population

The Netherlands Mental Health Survey and Incidence Study-2 (NEMESIS-2) was conducted to study the prevalence, incidence, course, and consequences of mental disorders in the Dutch general population. The study was approved by the Medical Ethics Review Committee for Institutions on Mental Health Care and written informed consent was collected from participants at each wave. A multistage random sampling procedure was applied to ensure the representativeness of the sample in terms of age (between the ages of 18 and 65 at baseline), region, and population density. Dutch illiteracy was an exclusion criterion. Details of NEMESIS-2 were provided elsewhere.[18,19] From 2007 to 2009, the first wave (T0) enrolled 6646 participants (response rate 65.1%; average interview duration: 95 min) who were followed up at 3 visits within 9 years: successive response rates at year 3 (T1), year 6 (T2), and year 9 (T3) were 80.4% (n = 5303; excluding those who deceased; interview duration: 84 min), 87.8% (n = 4618; interview duration: 83 min), and 86.8%

(n = 4007; interview duration: 102 min), respectively. Rates at baseline reflect lifetime occurrence; rates at T1–T3 reflect 3-year interval (T0–T1, T1–T2, and T2–T3) occurrence. Attrition between T0 and T3 was not significantly associated with any of the individual 12-month mental disorders at T0 after controlling for sociodemographic characteristics.[20] For this cross-sectional analysis, data from baseline (n = 6646) were utilized.

Exposome Score

The exposome score in the current analyses was calculated based on our previously validated estimates[16] for constructing cumulative environmental load. Using the log odds from our previous report, we generated the ES-SCZ by summing log-odds-weighted environmental exposures (cannabis use, winter birth, hearing impairment, and childhood adversities [emotional neglect, psychological abuse, physical abuse, sexual abuse, and peer victimization]) at baseline (supplementary methods and supplementary table 1). For comparison, an environmental sum score (hereafter, Esum-SCZ) by adding each binary exposure per individual as 0 = absent and 1 = present (ranging from 0 to 8) and an aggregate environmental score weighted by the meta-analytical estimates for each exposure (hereafter, Emet-SCZ) conforming to a previous study were generated.[21]

Outcomes

Nonclinician, trained interviewers applied the Composite International Diagnostic Interview version 3.0.[22,23] To examine the discriminative function of ES-SCZ for identifying individuals diagnosed with schizophrenia spectrum disorder, lifetime *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition (DSM-IV) diagnosis of schizophrenia at baseline was used. To examine the association of ES-SCZ with clinical risk strata, we used psychosis risk strata that were previously defined based on the degree of positive psychotic symptomatology, help-seeking attempt, antipsychotic treatment, and service use and admission for psychotic symptomatology. Psychosis risk strata consisted of 5 distinct categories (no-risk, low-risk, moderate-risk, high-risk, and clinical psychosis). For further details, see elsewhere.[17] For the outcome-wide association of ES-SCZ, we used lifetime diagnoses of DSM-IV disorders, lifetime suicide thoughts, plans, and attempts, self-reported chronic somatic health problems in the last 12 months, and general traits of neuroticism and extraversion as listed (supplementary method and supplementary table 1).

Statistical Analyses

All analyses were performed using Stata, version 16.[24] P < .05 (2 tailed) was considered nominally statistically significant. To determine the discriminative function

of ES-SCZ for identifying individuals diagnosed with schizophrenia, receiver operating characteristic (ROC) analysis was performed using the ROCREG command[25] that applies a nonparametric estimator of the 95% CIs around the AUC using a bootstrap method (n = 1000 repetitions). The ROCCOMP command was used to compare the ROC areas of ES-SCZ, Esum-SCZ, and Emet-SCZ.[26] By applying the ROCTAB command, the optimal cut point for ES-SCZ was estimated using the Liu method that maximizes the product of sensitivity and specificity.[27] The sensitivity, specificity, correct classification rate, and positive (LR+) and negative likelihood ratios (LR−) were reported. Given that the covariates may influence the discriminative function of ES-SCZ, we performed ROC analysis controlled for the covariates (age, sex, and education) using the CTRLCOV option.[28] We performed multinomial logistic regression models using the MLOGIT command to analyze the association of ES-SCZ at the optimal cut point with psychosis risk strata ("no-risk" group as the reference). Finally, we applied logistic regression models to test the association of ES-SCZ

with psychiatric diagnoses and other medical outcomes. Unadjusted OR with corresponding 95% CI and Nagelkerke's R^2 of ES-SCZ were reported for each model. Each model was also controlled for covariates (age, sex, and education).

Results

Baseline frequencies of demographic variables, exposures, outcome variables, and missing values of the total sample are shown in the supplementary table 1. At baseline, 43 individuals (0.7%) were diagnosed with a lifetime schizophrenia spectrum disorder.

Predictive Performance of the ES-SCZ

Distinguishing individuals diagnosed with schizophrenia spectrum disorder from controls, the unadjusted ROC analyses indicated the highest AUC of 0.84 (95% CI: 0.77; 0.91) for the ES-SCZ, whereas AUCs were similar for Emet-SCZ with 0.80 (95% CI: 0.73; 0.87) and for Esum-SCZ with 0.80 (95% CI: 0.72; 0.87). AUCs comparisons demonstrated a significant difference between ES-SCZ and Emet-SCZ (x^2 = 7.29, P = .007), as well as between ES-SCZ and Esum-SCZ (x^2 = 6.66, P = .010). No significant difference was found between Emet-SCZ and Esum-SCZ (x^2 = 0.14, P = .711). For visualization, covariate-adjusted ROC of ES-SCZ against a reference line is shown in figure 1.

Table 1 reports the sensitivity, specificity, likelihood ratios, and classification accuracy at the >50%, >75%, and the optimal cut point for the ES-SCZ. The models indicated higher sensitivity (84.62%–89.74%) relative to specificity (43.73%–78.06%), with the classification accuracy varying between 44.04% and 78.10%. At the optimal cut point (3.22), ES-SCZ showed high sensitivity (84.62%) and specificity (78.06), with 78.10% correct classification. The positive and negative likelihood ratios were 3.86 and 0.20, respectively.

Association Between ES-SCZ and Psychosis Risk Strata

ES-SCZ at the optimal cut point was significantly associated with the low-risk, moderate-risk, high-risk, and clinical psychosis strata (table 2). Additional post hoc group comparisons across strata showed significant differences
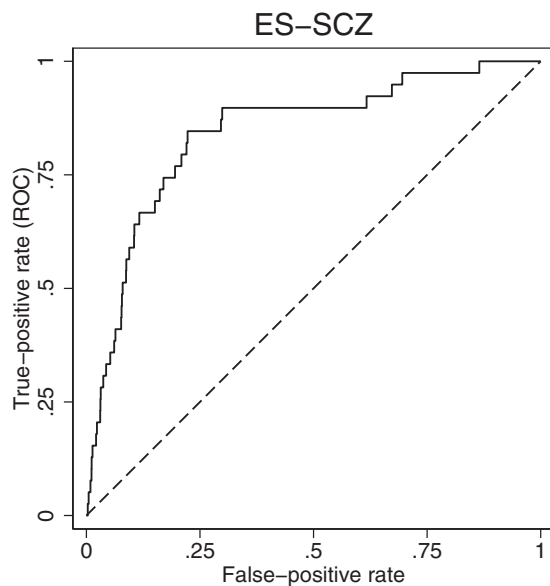
Figure 1. Receiver operating characteristic of ES-SCZ on schizophrenia spectrum disorder. Adjusted for sex, age (continuous), and 4-level education (1—primary school, 2—lower secondary education, 3—higher secondary education, and 4—higher professional education). Dashed line indicates the reference line.

Table 1. Predictive ability of exposome score for schizophrenia on schizophrenia spectrum disorder at different cut points

	Cut points	Sensitivity %	Specificity %	Positive likelihood ratio	Negative likelihood ratio	Correctly classified %
<50%	2.03	89.74	43.73	1.59	0.23	44.04
<75%	2.86	84.62	73.67	3.21	0.21	73.74
Optimal cut point	3.22	84.62	78.06	3.86	0.20	78.10

Table 2. Associations between exposome score for schizophrenia and psychosis risk strata

	Reference group ("no-risk")			Psychosis low-risk state		Psychosis moderate-risk state		Psychosis high-risk state	
	RRR	95% CI	P	Wald χ^2	P	Wald χ^2	P	Wald χ^2	P
Psychosis low-risk state	1.53	1.23–1.90	<.001	—	—	—	—	—	—
Psychosis moderate-risk state	2.79	2.17–3.89	<.001	13.64	<.001	—	—	—	—
Psychosis high-risk sate	4.06	3.15–5.23	<.001	35.18	<.001	4.52	.033	—	—
Clinical psychosis	7.27	3.58–14.73	<.001	17.34	<.001	6.35	.012	2.35	.125

Note: Reference group = 84.1%; psychosis low-risk state = 6.9%; psychosis moderate-risk state = 4.3%; psychosis high-risk state = 4.1%; clinical psychosis = 0.5%.
RRR, relative risk ratio.
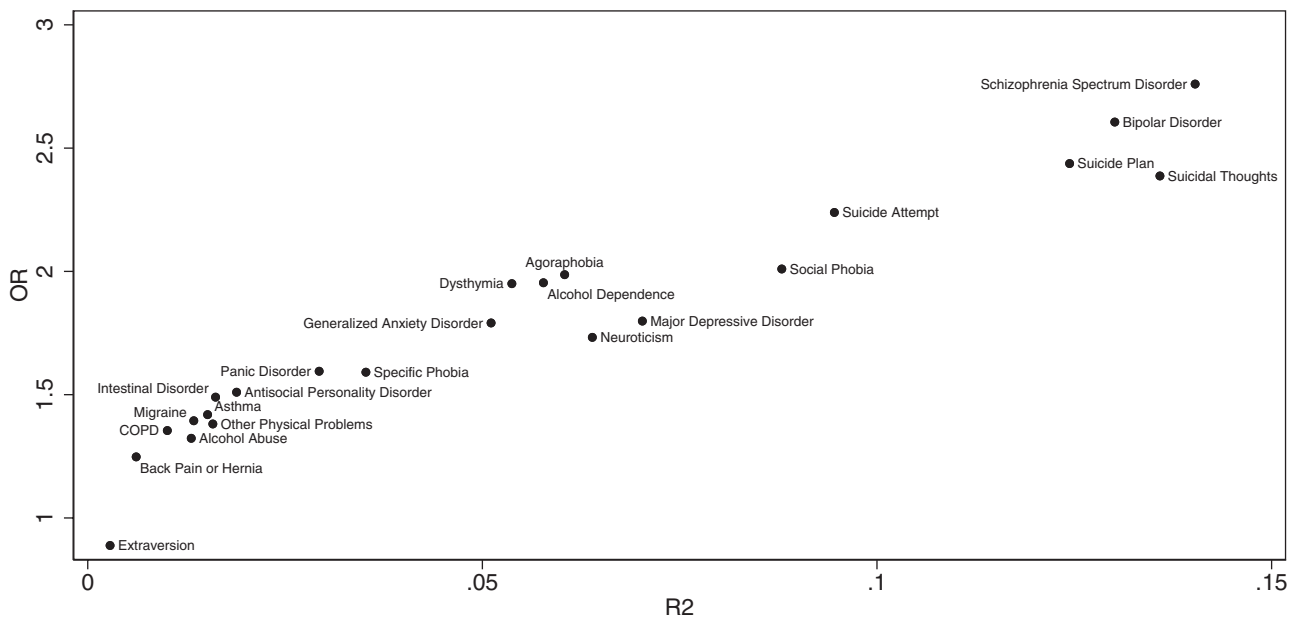
Figure 2. Unadjusted variances and ORs of the association between ES-SCZ and multiple mental and physical health outcomes. The figure shows 23 significant associations after Bonferroni correction ($P < .05/33$); COPD, Chronic obstructive pulmonary disease; ES-SCZ: Exposome score for schizophrenia; R^2, Nagelkerke's R^2.

for the low-risk vs moderate-risk, low-risk vs high-risk, low-risk vs clinical psychosis, moderate-risk vs high-risk, and moderate vs clinical psychosis. The comparison between high-risk vs clinical psychosis was not significantly different. The results were similar in the covariate-adjusted analyses (supplementary table 2).

Association Between ES-SCZ and Multiple Outcomes

In the univariate analyses, ES-SCZ was significantly (Bonferroni corrected $P < .05/33$) associated with 23 out of 33 outcomes (figure 2), while the multivariate analyses indicated that ES-SCZ was significantly associated with 25 outcomes (supplementary figure 1). Table 3 shows the outcome-wide association of ES-SCZ. Unadjusted significant ORs varied between 0.89 and 2.76, with the explained variance varying between 0.28% and 14.03%. The association between

ES-SCZ and schizophrenia spectrum disorder indicated the highest OR (2.76 [95% CI: 2.20; 3.46], $P < .001$) with an explained variance of 14.03%. This was followed by bipolar disorder (OR = 2.61 [95% CI: 2.19; 3.10], $P < .001$, $R^2 = 13.01\%$), suicide plan (OR = 2.44 [95% CI: 2.16; 2.75], $P < .001$, $R^2 = 12.44\%$), suicidal thoughts (OR = 2.39 [95% CI: 2.19; 2.60], $P < .001$, $R^2 = 13.58\%$), and suicide attempt (OR = 2.24 [95% CI: 1.95; 2.57], $P < .001$, $R^2 = 9.46\%$). The analyses adjusted for covariates showed similar results for the top 5 associations (supplementary table 3): schizophrenia spectrum disorder (OR = 2.71 [95% CI: 2.16; 3.40], $P < .001$, $R^2 = 15.87\%$), bipolar disorder (OR = 2.59 [95% CI: 2.17; 3.09], $P < .001$, $R^2 = 14.69\%$), suicide plan (OR = 2.46 [95% CI: 2.17; 2.78], $P < .001$, $R^2 = 12.88\%$), suicidal thoughts (OR = 2.41 [95% CI: 2.21; 2.64], $P < .001$, $R^2 = 14.08\%$), and suicide attempt (OR = 2.24 [95% CI: 1.95; 2.57], $P < .001$, $R^2 = 10.56\%$).

Table 3. Unadjusted outcome-wide association of exposome score for schizophrenia

Outcome variable	OR	95% CI	P	R^2
Extraversion	0.89	0.83–0.95	<.001	0.28
Poor eyesight	0.90	0.57–1.44	.666	0.05
High blood pressure	0.98	0.88–1.08	.652	0.01
Cancer	1.12	0.85–1.48	.429	0.09
Diabetes	1.13	0.97–1.32	.108	0.15
Thyroid abnormality	1.15	0.98–1.35	.097	0.18
Minor depressive disorder	1.15	0.98–1.35	.081	0.19
Heart disease	1.20	0.96–1.49	.105	0.27
Joint wear	1.20	1.06–1.36	.004	0.40
Back pain or hernia	1.25	1.11–1.40	<.001	0.61
Joint inflammation	1.26	1.09–1.46	.002	0.57
Alcohol abuse	1.32	1.22–1.43	<.001	1.31
COPD	1.35	1.17–1.57	<.001	1.01
Other physical problems	1.38	1.26–1.51	<.001	1.58
Migraine	1.39	1.23–1.58	<.001	1.34
Asthma	1.42	1.25–1.61	<.001	1.52
Intestinal disorder	1.49	1.23–1.80	<.001	1.62
Antisocial personality disorder	1.51	1.29–1.77	<.001	1.89
Ulcers	1.52	1.14–2.02	.005	1.51
Specific phobia	1.59	1.46–1.74	<.001	3.52
Panic disorder	1.59	1.42–1.79	<.001	2.93
Neuroticism	1.73	1.62–1.85	<.001	6.39
Generalized anxiety disorder	1.79	1.61–1.99	<.001	5.11
Major depressive disorder	1.80	1.68–1.93	<.001	7.03
Dysthymia	1.95	1.62–2.35	<.001	5.37
Alcohol dependence	1.95	1.67–2.29	<.001	5.77
Agoraphobia	1.99	1.69–2.34	<.001	6.04
Social phobia	2.01	1.85–2.18	<.001	8.79
Suicide attempt	2.24	1.95–2.57	<.001	9.46
Suicidal thoughts	2.39	2.19–2.60	<.001	13.58
Suicide plan	2.44	2.16–2.75	<.001	12.44
Bipolar disorder	2.61	2.19–3.10	<.001	13.01
Schizophrenia spectrum disorder	2.76	2.20–3.46	<.001	14.03

23 significant associations after Bonferroni correction (P < .05/33); COPD, chronic obstructive pulmonary disease; R^2, Nagelkerke's R^2.

Discussion

In this study, we investigated the discriminative capacity and risk stratification properties of ES-SCZ in the general population. Our findings were that ES-SCZ showed a good discriminative function (AUC: 84) for identifying schizophrenia in the general population. The AUC comparison showed that ES-SCZ significantly performed better than both the Esum-SCZ (AUC: 80) and Emet-SCZ (AUC: 80). At optimal cut point, ES-SCZ showed similar performance in ruling out (LR− = 0.20) and ruling in (LR+ = 3.86) schizophrenia. ES-SCZ at optimal cut point showed a progressively greater magnitude of association with increasing psychosis risk strata.

This is the first study investigating the discriminative function of an aggregate environment risk score for schizophrenia, the ES-SCZ, which is generated using a training sample to estimate the sum of the weighted effect sizes of environmental exposures. Unlike other previous approaches, the ES-SCZ uses estimates from the multivariate model that takes into account the interdependency of exposures. ES-SCZ showed significantly better discriminative function than the Esum-SCZ and Emet-SCZ. This finding provides further support that approaches that take into account the correlation between exposures prevent overestimation of the weights per exposure and achieve better predictive performance than those assuming independence (eg, simple summation of exposures or weighted estimates of individuals exposures from meta-analyses).[16] The finding should be anticipated given the fact that environmental risk factors for mental disorder phenotypes, such as schizophrenia, are often linked with each other through causal and noncausal paths observed in the general population.[14]

Our findings showed that although AUC results are generally considered very good for values between 0.8 and 0.9, ES-SCZ at the optimal cut point generated small to moderate changes in probability for predicting schizophrenia. There have been no comparable studies for ES-SCZ; according to a proposed guideline for a clinically relevant risk prediction, the LRs should be optimally over 10 for LR+ and under 0.1 for LR− for decisive shifts from pretest to posttest probability.[29] Therefore, ES-SCZ cannot provide the risk prediction utility that is required for predicting individual diagnosis in the general population. However, as an environmental liability index for schizophrenia, ES-SCZ may offer improved solutions in research settings. First, ES-SCZ can be particularly useful for risk stratification in large population data sets as evidenced by our present findings showing a progressively greater magnitude of the association between increasing psychosis risk strata and ES-SCZ at the optimal cut point. Second, ES-SCZ may be useful for risk enrichment to target selective smaller samples for expensive, experimental, or time-consuming methods. Finally, the quantification of environmental liability as a single metric enhances statistical power over multiple testing of each exposure.[12,30] Certainly, the integration of ES-SCZ with electronic health records and molecular genetic markers, such as PRS-SCZ, has the potential to boost prediction in the future.

ES-SCZ was associated with schizophrenia diagnosis with the highest OR (2.76) and greatest explained variance (Nagelkerke R^2 = 14%) among all outcomes. However, in line with converging evidence suggesting that environmental exposures are not distinctly associated with psychosis spectrum phenotype only but instead are more universally related to broad psychopathology,[12,31,32] ES-SCZ was also associated with several psychiatric diagnoses and other medical outcomes in the general population. It should be noted that results on physical and mental health outcomes may not be directly comparable as mental disorders reflect lifetime diagnoses, whereas physical health outcomes reflect the previous 12-month period.

Pleiotropy is a rule rather than an exception for psychiatric diagnoses and behavioral phenotypes as also

demonstrated in GWAS.[33] Similar to PRS-SCZ, ES-SCZ can provide only little to no benefit in discriminating schizophrenia from another adjacent diagnostic category, such as bipolar disorder. However, it may be used for risk stratification of broader mental ill health in the general population.

A major strength of our approach was constructing ES-SCZ in independent training and validation case-control samples and consequently testing ES-SCZ in a large population data set derived from the same country of origin with matching environmental assessment. However, several limitations should be noted. Although the random sampling procedure applied in this data set increases epidemiological representativeness, individuals who refrained from participating in this study (eg, because of trust issues) may be slightly underrepresented. Additionally, ES-SCZ is inherently limited to the exposure estimates derived from the original model using reliably measured and equally available exposures in the training and the validation case-control samples. Several other exposures associated with schizophrenia can be considered for addition to ES-SCZ. However, some of these exposures are largely unavailable in collected or ongoing cohort studies, such as obstetric and pregnancy complications, which are also extremely difficult to reliably assess without detailed birth records and maternal interviews and, therefore, impossible to collect in retrospect. Furthermore, some of these exposures, such as urbanicity, do not display a consistent association with psychosis phenotypes across countries and ethnic groups.[34] Also, given the lower predictive performance of genetic scores due to limited population diversity in GWAS,[35] the addition of some exposures, such as ethnic minority and migration, may decrease the utility of ES-SCZ when integrated with genetic data. Although generating a universal environmental loading score is extremely challenging, increasing efforts to the harmonization of available samples and determining a limited set of measures for core environmental assessment for epidemiological cohorts will pave the way for wider application.[15,36]

In conclusion, our findings from an epidemiologically representative general population cohort demonstrate that an aggregate environmental exposure score for schizophrenia constructed using a predictive modeling approach—ES-SCZ—has the potential to improve risk prediction and stratification for research purposes and may help gain insight into multicausal etiology of pluripotent psychopathology.

Supplementary Material

Supplementary material is available at *Schizophrenia Bulletin*.

Acknowledgments

NEMESIS-2 is conducted by the Netherlands Institute of Mental Health and Addiction (Trimbos Institute) in Utrecht. The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Data Availability

Data available on reasonable request from the authors.

References

1. Guloksuz S, Pries LK, Delespaul P, et al; Genetic Risk and Outcome of Psychosis (GROUP) investigators. Examining the independent and joint effects of molecular genetic liability and environmental exposures in schizophrenia: results from the EUGEI study. *World Psychiatry*. 2019;18(2):173–182.
2. Choi SW, Mak TS-H, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. *Nat Protocols*. 2020;15:2759–2772.
3. Ripke S, Walters JT, O'Donovan MC. 'Mapping genomic loci prioritises genes and implicates synaptic biology in schizophrenia', medRxiv, doi:10.1101/2020.09.12.20192922, 13 September 2020, preprint: not peer reviewed.
4. Zheutlin AB, Dennis J, Karlsson Linnér R, et al. Penetrance and pleiotropy of polygenic risk scores for schizophrenia in 106,160 patients across four health care systems. *Am J Psychiatry*. 2019;176(10):846–855.
5. Kember RL, Merikangas AK, Verma SS, et al. Polygenic risk of psychiatric disorders exhibits cross-trait associations in electronic health record data from european ancestry individuals. *Biol Psychiatry*. 2020:1–10. doi:10.1016/j.biopsych.2020.06.026.
6. Richardson TG, Harrison S, Hemani G, Davey Smith G. An atlas of polygenic risk score associations to highlight putative causal relationships across the human phenome. *Elife*. 2019;8:e43657.
7. Mistry S, Harrison JR, Smith DJ, Escott-Price V, Zammit S. The use of polygenic risk scores to identify phenotypes associated with genetic risk of schizophrenia: systematic review. *Schizophr Res*. 2018;197:2–8.
8. Jones HJ, Stergiakouli E, Tansey KE, et al. Phenotypic manifestation of genetic risk for schizophrenia during adolescence in the general population. *JAMA Psychiatry*. 2016;73(3):221–228.
9. van Os J, van der Steen Y, Islam MA, Gülöksüz S, Rutten BP, Simons CJ; GROUP Investigators. Evidence that polygenic

risk for psychotic disorder is expressed in the domain of neurodevelopment, emotion regulation and attribution of salience. *Psychol Med.* 2017;47(14):2421–2437.

10. van Os J, Pries LK, Delespaul P, et al. Replicated evidence that endophenotypic expression of schizophrenia polygenic risk is greater in healthy siblings of patients compared to controls, suggesting gene-environment interaction. The EUGEI study. *Psychol Med.* 2019;113:1–14.

11. Marsman A, Pries L-K, ten Have M, et al. Do current measures of polygenic risk for mental disorders contribute to population variance in mental health? *Schizophr Bull.* 2020;46(6):1353–1362.

12. Pries LK, van Os J, Ten Have M, et al. Association of recent stressful life events with mental and physical health in the context of genomic and exposomic liability for schizophrenia. *JAMA Psychiatry*. 2020:1–9. doi:10.1001/jamapsychiatry.2020.2304.

13. Belbasis L, Köhler CA, Stefanis N, et al. Risk factors and peripheral biomarkers for schizophrenia spectrum disorders: an umbrella review of meta-analyses. *Acta Psychiatr Scand.* 2018;137(2):88–97.

14. Guloksuz S, Rutten BPF, Pries LK, et al; European Network of National Schizophrenia Networks Studying Gene-Environment Interactions Work Package 6 (EU-GEI WP6) Group. The complexities of evaluating the exposome in psychiatry: a data-driven illustration of challenges and some propositions for amendments. *Schizophr Bull.* 2018;44(6):1175–1179.

15. Guloksuz S, van Os J, Rutten BPF. The exposome paradigm and the complexities of environmental research in psychiatry. *JAMA Psychiatry.* 2018;75(10):985–986.

16. Pries LK, Lage-Castellanos A, Delespaul P, et al; Genetic Risk and Outcome of Psychosis (GROUP) investigators. Estimating exposome score for schizophrenia using predictive modeling approach in two independent samples: the results from the EUGEI Study. *Schizophr Bull.* 2019;45(5):960–965.

17. Guloksuz S, Pries LK, Ten Have M, et al. Association of preceding psychosis risk states and non-psychotic mental disorders with incidence of clinical psychosis in the general population: a prospective study in the NEMESIS-2 cohort. *World Psychiatry.* 2020;19(2):199–205.

18. de Graaf R, Ten Have M, van Dorsselaer S. The Netherlands Mental Health Survey and Incidence Study-2 (NEMESIS-2): design and methods. *Int J Methods Psychiatr Res.* 2010;19(3):125–141.

19. de Graaf R, ten Have M, van Gool C, van Dorsselaer S. Prevalence of mental disorders and trends from 1996 to 2009. Results from the Netherlands Mental Health Survey and Incidence Study-2. *Soc Psychiatry Psychiatr Epidemiol.* 2012;47(2):203–213.

20. de Graaf R, van Dorsselaer S, Tuithof M, ten Have M. Sociodemographic and psychiatric predictors of attrition in the third follow-up of the Netherlands Mental Health Survey and Incidence Study-2 (NEMESIS-2). Utrecht, Netherlands: Trimbos-instituut; 2018.

21. Padmanabhan JL, Shah JL, Tandon N, Keshavan MS. The "polyenviromic risk score": aggregating environmental risk factors predicts conversion to psychosis in familial high-risk subjects. *Schizophr Res.* 2017;181:17–22.

22. de Graaf R, Ormel J, ten Have M, Burger H, Buist-Bouwman M. *Mental disorders and Service Use in The Netherlands. Results from the European Study of the Epidemiology of Mental Disorders (ESEMeD)*. New York, NY: Cambridge University Press; 2008.

23. Alonso J, Angermeyer MC, Bernert S, et al; ESEMeD/MHEDEA 2000 Investigators. Sampling and methods of the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl.* 2004;109(420):8–20.

24. StataCorp. *Stata Statistical Software: Release 16*. College Station, TX: StataCorp LLC; 2019.

25. Lora D, Contador I, Pérez-Regadera JF, de la Cámara AG. Features of the area under the receiver operating characteristic (ROC) curve. A good practice. *Stata J.* 2016;16(1):185–196.

26. Cleves MA. From the help desk: Comparing areas under receiver operating characteristic curves from two or more probit or logit models. *Stata J.* 2002;2(3):301–313.

27. Liu X. Classification accuracy and cut point selection. *Stat Med.* 2012;31(23):2676–2686.

28. Janes H, Longton G, Pepe M. Accommodating covariates in ROC analysis. *Stata J.* 2009;9(1):17–39.

29. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *JAMA.* 1994;271(9):703–707.

30. Pries L-K, Ferro GAD, van Os J, et al. Examining the independent and joint effects of genomic and exposomic liabilities for schizophrenia across the psychosis spectrum. *Epidemiol Psychiatr Sci.* 2020. doi:10.1017/S2045796020000943.

31. Guloksuz S, van Nierop M, Lieb R, van Winkel R, Wittchen HU, van Os J. Evidence that the presence of psychosis in non-psychotic disorder is environment-dependent and mediated by severity of non-psychotic psychopathology. *Psychol Med.* 2015;45(11):2389–2401.

32. Pries LK, Guloksuz S, Ten Have M, et al. Evidence that environmental and familial risks for psychosis additively impact a multidimensional subthreshold psychosis syndrome. *Schizophr Bull.* 2018;44(4):710–719.

33. Lee PH, Anttila V, Won H, et al. Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. *Cell.* 2019;179(7):1469–1482. e1411.

34. Fett AJ, Lemmers-Jansen ILJ, Krabbendam L. Psychosis and urbanicity: a review of the recent literature from epidemiology to neurourbanism. *Curr Opin Psychiatry.* 2019;32(3):232–241.

35. Bigdeli TB, Genovese G, Georgakopoulos P, et al. Contributions of common genetic variants to risk of schizophrenia among individuals of African and Latino ancestry. *Mol Psychiatry.* 2019;25:2455–2467.

36. Hollander JA, Cory-Slechta DA, Jacka FN, et al. Beyond the looking glass: recent advances in understanding the impact of environmental exposures on neuropsychiatric disease. *Neuropsychopharmacology.* 2020;45(7):1086–1096.