



Polygenic risk scores differentiate schizophrenia patients with toxoplasma gondii compared to toxoplasma seronegative patients

[Adriana Lori](#), *Emory University*
[Dimitrios Avramopoulos](#), *Johns Hopkins University*
[Alex W. Wang](#), *Emory University*
[Jennifer Mulle](#), *Emory University*
[Nicholas Massa](#), *Emory University*
[Erica Duncan](#), *Emory University*
[Abigail Powers](#), *Emory University*
[Karen Conneely](#), *Emory University*
[Charles Gillespie](#), *Emory University*
[Tanja Jovanovic](#), *Emory University*

Only first 10 authors above; see publication for full author list.

Journal Title: COMPREHENSIVE PSYCHIATRY

Volume: Volume 107

Publisher: W B SAUNDERS CO-ELSEVIER INC | 2021-05-01, Pages 152236-152236

Type of Work: Article | Final Publisher PDF

Publisher DOI: 10.1016/j.comppsy.2021.152236

Permanent URL: <https://pid.emory.edu/ark:/25593/vmn99>

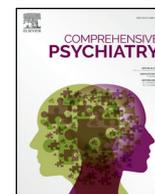
Final published version: <http://dx.doi.org/10.1016/j.comppsy.2021.152236>

Copyright information:

© 2021 Published by Elsevier Inc.

This is an Open Access work distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

Accessed December 8, 2021 12:00 PM EST



Polygenic risk scores differentiate schizophrenia patients with *Toxoplasma gondii* compared to toxoplasma seronegative patients



Adriana Lori ^a, Dimitrios Avramopoulos ^b, Alex W. Wang ^{c,d}, Jennifer Mulle ^e, Nicholas Massa ^{c,f}, Erica J. Duncan ^{a,f}, Abigail Powers ^a, Karen Conneely ^e, Charles F. Gillespie ^a, Tanja Jovanovic ^g, Kerry J. Ressler ^h, Brad D. Pearce ^{c,*}

^a Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, 101 Woodruff Circle, GA 30322, USA

^b Institute of Genetic Medicine, Johns Hopkins University, Baltimore, MD 21205, USA

^c Dept. of Epidemiology, Rollins School of Public Health; 1518 Clifton Rd., Atlanta, GA 30322, USA

^d Philadelphia College of Osteopathic Medicine, Suwanee, GA 30024, USA

^e Department of Human Genetics, Emory University, Atlanta, GA 30322, USA

^f Atlanta Veterans Affairs Health Care System, Decatur, GA 30033, USA

^g University of Texas at Austin Dell Medical School, Austin, TX, USA

^h Department of Psychiatry, Harvard Medical School, Boston, MA 02115, USA

ARTICLE INFO

ABSTRACT

Schizophrenia (SCZ) is an etiologically heterogeneous disease with genetic and environmental risk factors (e.g., *Toxoplasma gondii* infection) differing among affected individuals. Distinguishing such risk factors may point to differences in pathophysiological pathways and facilitate the discovery of individualized treatments. *Toxoplasma gondii* (TOXO) has been implicated in increasing the risk of schizophrenia. To determine whether TOXO-positive individuals with SCZ have a different polygenic risk burden than uninfected people, we applied the SCZ polygenic risk score (SCZ-PRS) derived from the Psychiatric GWAS Consortium separately to the TOXO-positive and TOXO-negative subjects with the diagnosis of SCZ as the outcome variable. The SCZ-PRS does not include variants in the major histocompatibility complex.

Of 790 subjects assessed for TOXO, the 662 TOXO-negative subjects (50.8% with SCZ) reached a Bonferroni corrected significant association ($p = 0.00017$, $R^2 = 0.023$). In contrast, the 128 TOXO-positive individuals (53.1% with SCZ) showed no significant association ($p = 0.354$) for SCZ-PRS and had a much lower R^2 ($R^2 = 0.007$). To account for Type-2 error in the TOXO-positive dataset, we performed a random sampling of the TOXO-negative subpopulation ($n = 130$, repeated 100 times) to simulate equivalent power between groups: the p -value was <0.05 for SCZ-PRS 55% of the time but was rarely (6% of the time) comparable to the high p -value of the seropositive group at $p > 0.354$.

We found intriguing evidence that the SCZ-PRS predicts SCZ in TOXO-negative subjects, as expected, but not in the TOXO-positive individuals. This result highlights the importance of considering environmental risk factors to distinguish a subgroup with independent or different genetic components involved in the development of SCZ.

© 2021 Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Schizophrenia (SCZ) is a highly debilitating and complex brain disorder affecting approximately 1% of the worldwide population. It is characterized by periods of psychosis, disorganized thought patterns and behaviors, and deficits in neurocognitive function. It is crucial to identify the risk factors that contribute to the increased risk of developing SCZ. Among those, the high heritability of this disorder has been recognized

as a contributing risk factor, with studies showing increased risk from 1% in the general population to over 40% in monozygotic twins of affected individuals. The high heritability of this disorder underscores the importance of genetic risk factors (1). Despite this high heritability, a single gene or variant rarely heightens the risk of developing this disorder. Instead, recent GWAS studies have highlighted its polygenic architecture by identifying over 108 loci that contribute to the overall risk for SCZ (2–4). However, this genetic predisposition alone cannot fully explain the incidence of SCZ since there may be interactions with environmental risk factors that also contribute to the risk of developing SCZ (5). One environmental factor known to heighten the risk for SCZ is chronic infection with *Toxoplasma gondii* (TOXO), a protozoan

* Corresponding author at: Rollins School of Public Health, Emory University, Office: 1518 Clifton Rd. CNR 4049 (1518-002-AA), Atlanta, GA 30322, USA.
E-mail address: bpearce@emory.edu (B.D. Pearce).

parasite (6). TOXO is a common neuro-invasive parasite that remains relatively quiescent in immune-competent mammalian hosts but can affect neurocircuitry and cause subtle behavioral changes in rodents and humans (7,8). Our study aims to examine whether people infected by TOXO have different genetic risk factors that predispose them to SCZ compared to non-infected subjects.

Polygenic risk scores (PRS) have been used to capture genetic effects of genome-wide markers that may not individually achieve GWAS significance in large-scale association studies. Markers are selected using an initial training sample, at different *p*-value threshold levels, and used to construct a score in an independent replication sample by forming the weighted sum of associated alleles within each subject. A significant association between the polygenic score and the trait in the replication sample implies that a genetic signal is present among the selected markers, establishing a common polygenic basis for the two datasets. Given the polygenic nature of SCZ and the large number of samples enrolled in the Schizophrenia Psychiatric Genomics Consortium, summary statistics derived by these datasets are particularly useful for PRS analyses.

We hypothesized that people with this known environmental risk for SCZ, chronic TOXO infection, may have less conventional (SNP-based) genetic risk for SCZ than those who had not been infected with TOXO. To test this hypothesis, we examined whether patients infected by TOXO have less polygenic risk burden as ascertained by a polygenic risk score (PRS) compared to uninfected patients. Therefore, we used the large-scale summary statistics from the Schizophrenia Psychiatric Genomics Consortium (training dataset) (2) to derive a PRS in a distinct cohort (target dataset) of individuals with a history of SCZ who were identified either as TOXO-positive or TOXO-negative.

If confirmed, this would support the hypothesis that TOXO exposure is a risk factor for SCZ that confers risk separately from the risk conferred by vulnerability genes included in the SCZ-PRS.

2. Materials and methods

2.1. Population

Inclusion criteria for participants of known Ashkenazi-Jewish (AJ) descent, SCZ diagnostic criteria, immunoassay measurement for antibodies of TOXO, and genotyping, were previously described by Avramopoulos et al. (9). The Johns Hopkins institutional review board approved the recruitment methods, protocols, and informed consent documents. All participating human subjects provided written, informed consent to participate in this study. A total of 519 SCZ patients and 271 healthy controls (790 total) were selected for genotyping and TOXO evaluation. The TOXO seropositivity status was determined by the presence of TOXO-IgG antibodies, as previously described by (9). Reagents for the TOXO IgG ELISA were obtained from IBL Laboratories, Hamburg, Germany. Cut points for seropositivity were clearly defined from density plots as described previously (9). Genotyping was performed with Affymetrix Human Genome-Wide SNP Array 6.0 at Emory University. Genotype quality control (QC) was performed using the software package PLINK (ver. 1.9) following the steps outlined in the Psychiatric GWAS consortium (2,10). As previously described (11), we first verified ancestry by merging the AJ autosomal pruned variants with the HapMap populations' common variants: the AJ cohort appeared as a harmonized highly ancestral homogenous population. Then, we calculated the principal component eigenvectors of the genetic relationship matrix (R; ver.3.2.2) in the AJ cohort alone, retaining those individuals who fell within three standard deviations of the medians of the first and second principal components. The Tracy-Widom test (12) was used to assess the number of principal components (PCs) that need to be included as covariates into the association analyses (2 ancestry PCs). However, to ensure that our results were not biased by ancestry, we repeated all the analyses using 5 PCs.

2.2. Polygenic risk scores (PRS)

The calculation of a Polygenic Risk Score (PRS) evaluates the genetic liability of several variants into a single per-individual score. In the current analysis, PRSice (v1.25) software (13) was used to calculate PRS. This PRSice software analysis used two datasets to calculate PRS: a training dataset (summary statistics from SCZ-PGC) to derive the sets of SNPs and their effect size based on association *p*-values; and a base dataset (genome-wide data from TOXO-positive, TOXO-negative, or the entire dataset), from which the PRS is calculated based on SNP individual genotypes additively coded. In the training dataset (summary statistics SCZ-PGC), SNPs surviving the clumping and thresholding steps were selected based on *p*-value thresholds (PT-score). For each set of SNPs, the PRS score was then calculated as their weighted sum of per loci risk allele in the base dataset: $PRS_j = \sum (S_i \times G_{ij})/M_j$, where S_i is the summary statistic for the effective allele for the genotype i , G_{ij} is the genotype i for the j individual (0,1,2), and M_j is the number of alleles included in the PRS of the j th individual.

We used the meta-analyses results of the SCZ dataset from the PGC [scz2.prs.txt.gz (2)] as a training dataset. SNPs in linkage disequilibrium ($r^2 < 0.2$ within 500 kb window) were excluded (clumping). The major histocompatibility complex locus (termed HLA in humans) was also excluded. The PRSs were calculated using: a) *p*-value thresholds between 0 and 0.5 in the training dataset, with a 0.01 increment for a total of 50 different *p*-values (Bonferroni corrected *p*-value = 0.001); and b) as target datasets, either the TOXO-positive individuals, the TOXO-negative individuals or the entire cohort (TOXO-positive plus TOXO-negative). Polygenic risk scores were then tested for association with SCZ-case versus SCZ-control status in the target group using a logistic regression model, covarying for age, sex, and the first 2 ancestry PCs. In the entire cohort, a similar logistical regression model was conducted but TOXO status was added as an additional covariate. All the analyses were repeated using 5 PCs. Nagelkerke pseudo R-squared measures (R^2) were used to compare the two main models: PRS from the TOXO-negative vs. TOXO-positive individuals.

2.3. PRS sample size randomization for Type 2 error assessment

Since the number of TOXO-negative subjects ($n = 662$) was substantially larger than the TOXO-positive group ($n = 128$), we performed additional analysis to rule out Type 2 error. A SCZ-PRS was generated for 100 subsets of 130 randomly selected TOXO-negatives. Of note, the TOXO-negative subsets ($n = 130$) were all chosen to demographically represent the TOXO-negative cohort, and therefore were composed of 67.7% SCZ subjects and 41.5% female subjects (Table 1).

3. Results

Table 1 shows the AJ population's diagnostic and demographic characteristics divided by TOXO and SCZ status. In the entire cohort,

Table 1
Demographic of AJ Cohort by Toxoplasma seropositivity.

	Toxoplasma positive		Toxoplasma negative		Total Sample	
	No.	%	No.	%	No.	%
Patient Status						
Schizophrenia	68	53.1	451	68.1	519	65.7
Control	60	46.9	211	31.9	271	34.3
Gender ^a						
Male	63	49.2	387	58.5	450	57
Female	65	50.8	275	41.5	340	43
Age in years (SD) ^b	56.7	13.7	46.6	12.7	48.2	13.4

^a $p = 0.06$, with fisher *t*-test.

^b $p < 0.0001$, with two-tailed *t*.

we did not observe an association between TOXO and SCZ ($X^2 = 0.252$, $p = 0.616$).

A polygenic genetic burden associated with TOXO status was used in a SCZ-PGC training dataset to predict SCZ (Fig. 1). Specifically, a SCZ-PRS was calculated at a set of p -value threshold (PT) bins to achieve marginal evidence for association for many causal variants. In the entire cohort, a PT bin of 0.07 reached significance p -values ($p = 0.001$, $R^2 = 0.015$), with all other PT reaching nominal ($p = 0.05$) significance (Supplementary Fig. 1). In the TOXO-negative subjects, several PT bins reached Bonferroni corrected significance as predictors of SCZ (29/50 with p -values ≤ 0.001 ; Supplementary Fig. 2A) with a maximum at a p -value threshold bin of 0.13 ($p = 0.00017$, $R^2 = 0.023$; Fig. 1A). The polygenic scores from the best P -value threshold ($p = 0.00017$, $PT = 0.13$) for the TOXO-negative (Supplementary Fig. 3A and 3D) showed a clear difference in the distribution between SCZ cases and controls, with SCZ cases having higher PRS scores compared to controls without SCZ (Supplementary Fig. 3G). In contrast, in the TOXO-positive subgroup, none of the p -values reached a nominal significance level, with the lowest significance at $PT = 0.02$ ($p = 0.354$, $R^2 = 0.007$; Fig. 1B, Supplementary Fig. 2B). The PRS for either the same PT value ($p = 0.97$ $PT = 0.13$; Supplementary Fig. 3B and 3E), or the best PT threshold in the TOXO-positive subgroup ($p = 0.354$ $PT = 0.02$, Supplementary Fig. 3C and 3F) had overlapping distributions in SCZ cases and controls, with no differences between the two groups (Supplementary Fig. 3H-J). Further, there was also a clear difference in the distributions of R^2 , with the R^2 for the TOXO-negative higher than the TOXO-positive (Fig. 2A, Supplementary Fig. 2).

We conducted the same analyses using 5 PCs which lead to comparable results. The TOXO-negative subset reached the lowest p -values with Bonferroni correction significance and the highest R^2 ($PT = 0.13$, $p = 0.0001$ $R^2 = 0.023$) (Supplementary Fig. 4).

To verify that the results obtained were not due simply to Type-2 error, we randomly generated 100 sub-sets of the TOXO-negative

cohort, with the correct proportion of SCZ-affected patients and distribution of gender. We compared the R^2 of the 100 subsets with the lowest p -values to the R^2 of the TOXO-negative and TOXO-positive cohorts (Fig. 2A). The R^2 of the 100 subsets mainly overlapped with the TOXO-negative cohort, which had predominately higher R^2 values compared to the TOXO-positive cohort. We observed 5 subsets with the lowest PT bins reaching Bonferroni correction significance, 55 subsets with the lowest p -value less than 0.05, and 6 subsets with $p > 0.354$ (Fig. 2B-C). This indicates that there was a stronger association with the TOXO-negative subsets.

4. Discussion

This study used the large-scale Schizophrenia Psychiatric Genomics Consortium derived schizophrenia polygenic risk score and GWAS data from a genetically homogeneous AJ population to compare the genetic burden for SCZ in relation to TOXO infection. We report that patients with SCZ and a history of TOXO infection have a different conventional polygenic risk score (SCZ-PRS) than those with SCZ who were not TOXO-infected.

Schizophrenia is a highly heterogeneous disorder with both genetic and environmental components (14). Recent studies of schizophrenia genetics have emphasized the importance of multiple genetic variants, each with a small effect size. However, few of these studies have considered the relationship between polygenic risk and environmental factors (2,15). Remarkably, SCZ has been associated with environmental insults, such as infectious disease (e.g., TOXO infection), which may also be sensitive to the genetic background. Indeed, substantial indirect evidence has implicated TOXO infection in SCZ (6,8). Studies in humans and animal models suggest that at least some of the epidemiological associations of TOXO with neuropsychiatric disorders can be attributed to specific neurobiological alterations (7,16–18). Infection with TOXO is one possible environmental factor that has both epidemiological and

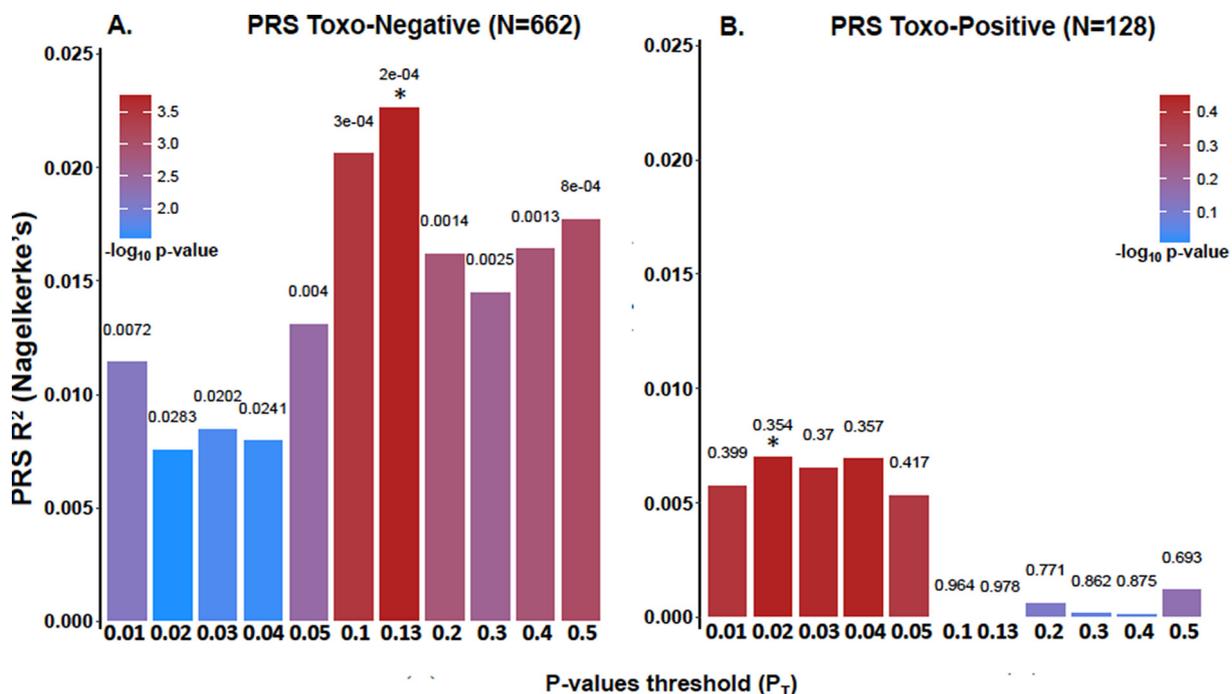


Fig. 1. PGC-SCZ PRS p -value bins in TOXO-Negative and TOXO-Positive participants. In this classical representation of the PRS results we show on the x-axis 0.01 increments for p -value bins up to 0.05, and 0.1 to 0.5 with 0.1 increments for subsets of PRS SNPs predicting SCZ in the PGC data. Broad p -value thresholds for schizophrenia PRS from the PGC predicting SCZ status in TOXO-negative ($n = 662$, panel A), and TOXO-positive participants ($n = 128$, Panel B). The x-axis also shows the bin for the lowest p -value association ($PT = 0.13$ for A, and $PT = 0.02$ for B, asterisk); the entire PT values are shown in supplementary fig. 2. The height of each bar on the y-axis shows the portion of the variance in our current data explained for predicting SCZ by the corresponding set of SNPs (Nagelkerke R^2). The numbers above each bar are the p -values for that set of SNPs in our current dataset, and the bars are color-coded by $-\log_{10}$ p -values. Note that the range used for the color-coded scale differs for panels A and B to allow internal comparison. TOXO-negative samples show an overall lower p -value and higher portion of the variance explained by the PRS for the SCZ outcome compared to the TOXO-positive group.

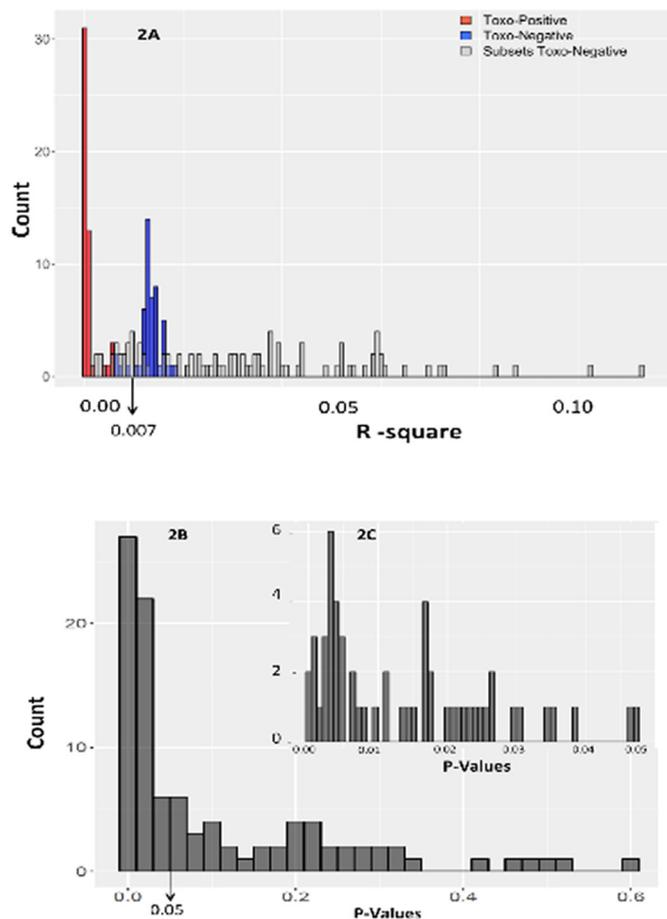


Fig. 2. Sample Size Randomization distributions of R^2 and p-values comparing TOXO-negative and TOXO-positive subsets. Histograms of R^2 (2A) and lowest p-values (2B,C). Fig. 2A, the x-axis shows the portion of the variance explained, and the y-axis is the count for the corresponding bar for TOXO-negative (blue) and TOXO-positive (red) subjects plus the 100 subsets of TOXO-negative samples (gray, $n = 130$). Note that the distribution of R^2 in which little of the variance is explained for the TOXO-positive subjects ($N = 128$) but much more of variance is explained for the 100 subsets of TOXO-negative samples ($n = 130$, in gray), which overlaps mostly with the R^2 of the TOXO-negative sample ($N = 662$, in blue). The arrow outlines the separation between the TOXO-positives (in red) and TOXO-negatives R^2 (in blue) without any overlapping values between red and blue. Fig. 2B shows p values of the 130 TOXO-negative samples. The arrow indicates the lowest p-values reaching nominal ($p = 0.05$) significance; an enlarged portion of these p-values is shown in Fig. 2C. Only 6% of the samples in this $N = 130$ set of TOXO-negative had p-values > 0.354 , which is the most significant p-value reached by the TOXO-positive samples.

neurobiological plausibility, although the specific mechanistic link between a positive TOXO serology and SCZ remains unclear (7,17,19).

Using GWAS of TOXO infection in the same AJ population, we recently identified several putative TOXO susceptibility genes with the most significant association within the chitinase locus (11). Prior studies in animal models have shown that the production of chitinase by macrophages in the brain responding to TOXO infection is crucial for controlling the TOXO infection (20). This finding highlights the ability to potentially reveal new genes associated with TOXO infection given the unique homogeneity of the AJ population. However, this finding does not address the question of whether TOXO-mediated SCZ has different genetic risk than TOXO-negative SCZ. Hence, in the current study, we addressed whether TOXO-associated schizophrenia differs from typical schizophrenia, in which there is a major contribution of common SNPs.

The current literature on PRS-environmental interactions has mainly focused on psychosocial variables, or overt obstetric complications, rather than biological measures such as TOXO (15,21,22). Our study shows that in the TOXO-negative subset, the SCZ-PRS very robustly predicted SCZ reaching a Bonferroni corrected p-value. In contrast, in the TOXO-positive group, the SCZ-PRS did not predict SCZ, consistent with the increased non-genetic burden (e.g., TOXO). Interestingly, when we analyzed the entire cohort, we still reached a significant p-value for one PT bin (albeit with a lower R^2), indicating that in a general setting, TOXO positive individuals might be incorporated in the overall p-value and may be indistinguishable unless analyzed separately. Because SNPs in the major histocompatibility complex (HLA) locus were not in the SCZ-PRS, we cannot rule out that genes at this locus played a role in SCZ risk in the TOXO-positive sample-sets. Indeed, SCZ may be associated with differences in HLA-directed immune responses to TOXO as suggested by the finding that an HLA-allele (HLA-C*04:01) can be protective against TOXO in people with SCZ but have the opposite effect in healthy controls (23). It is possible that there are polygenic associations in the MHC region that could explain the apparent differences in the PRS results between TOXO-positive schizophrenia and TOXO-negative schizophrenia. If this were the case, it would nevertheless support our argument that TOXO positive schizophrenia is linked to a different set of genes than TOXO-negative schizophrenia. Specifically, since the PGC consortium PRS excludes MHC, then the polygenic risk in conventional (TOXO-negative) schizophrenia that has been reported by the consortium (among others) demonstrates that genes outside MHC contribute to schizophrenia risk. Our study shows that this is not the case for TOXO-positive SCZ.

To exclude the possibility that the negative results observed in the TOXO-positive patients were due to a small sample size (Type 2 error) of TOXO-positive participants, we performed a PRS Sample Size Randomization. The PRS sample size randomization indicated that the lack of a SCZ-PRS association in the TOXO-positive cohort was not due to the small sample size (type 2-error), as evidenced by the observed p-values and R^2 distributions. Analysis of the distribution of the R^2 values calculated from the random TOXO-negative subsets with a small sample size (low power) predicted the SCZ phenotype 91% of the time. Although the results support our hypothesis, we cannot rule out a Type 2 error in the discovery sample and, therefore, larger sample sizes should be examined in the future.

In this cohort, TOXO seropositivity was not statistically significantly higher among schizophrenia patients compared to controls (9). Thus, there may be an alternative explanation for why our schizophrenia PRS findings were able to differentiate the TOXO-positive SCZ group from the TOXO-negative SCZ group. One possibility is that there are a sufficient number of false negatives for TOXO serology that are “contaminating” the seronegative group, such that a statistically significant difference in serology result was not obtained (24,25). This possibility derives from evidence that some individuals who are infected long before the sample was taken (prenatally or in childhood) may have titers of antibodies that have dropped below the level of detection (24). If this were the case, it implies that our sample of putative TOXO seronegative samples contains some number of actual TOXO-seropositive samples. Still, it would not explain why the PRS scores performed well in predicting schizophrenia in the seronegative group but not in the dependably seropositive group. Another possibility is that TOXO is not actually contributing to schizophrenia risk at any point in a patient’s trajectory (i.e., that TOXO infection does not play any causal role in schizophrenia). In this case, the TOXO serology may be acting as a proxy variable identifying alternative cellular and molecular pathways leading to schizophrenia that do not involve conventional SNP-based PRS.

The major limitation of the current study is the relatively modest sample size of the AJ cohort, particularly the TOXO-positive subset. However, for predicting individual trait values, it is more critical for the training set to be large (26), a requirement that has been satisfied

in the current study by using the SCZ-PGC. Another limitation of this study is the use of an AJ population. Given the homogeneity within the AJ population further studies should expand the study population characteristics; e.g. by including individuals with different ancestry. While the genetic homogeneity of the AJ population in some ways is limiting, in this case it is ideal for our purposes since its homogeneity may enhance the power to identify differences in PRS in a subset of the population exposed to different environmental insults (TOXO exposure). Furthermore, the use of Schizophrenia Psychiatric Genomics Consortium in the training dataset, and the AJ in the testing dataset, may better assure the independence of study subjects between those two cohorts, decreasing the possibility of inflation in the *p*-values.

In conclusion, we found evidence that the SCZ-PRS predicts SCZ status in TOXO-negative subjects but not TOXO-infected patients, suggesting that SCZ patients affected by TOXO may have less established genetic components involved in the development of SCZ. Distinguishing populations based on environmental and genetic risks may elucidate disease mechanisms, improve diagnosis of SCZ subtypes, and discover individualized treatment targets.

Author contributions

AL and BDP conceived the design of the experiment and contributed to the interpretation of the data, drafted the primary manuscript, as well as editing the final version. AL also performed the statistical analysis. DA performed the acquisition of the data, and determined subject characteristics, implementation of the GWAS analysis, and edited the manuscript. AWW performed early steps in the acquisition and cleaning of the data. JM performed the acquisition and cleaning of the dataset as well as contributing the original GWAS analysis and statistical plan, and editing the manuscript. NM performed assays, helped clean the dataset, and edited the manuscript. EJD helped with demographic and diagnostic category information, helped clean the dataset, and with editing the manuscript. AP provided diagnostic category input and edited the manuscript. KC provided input on statistical analyses, contributed to the original GWAS, and edited the manuscript. CFG provided input on data acquisition and edited the manuscript. TJ contributed to acquisition of the data and editing the manuscript. KJR contributed to the design the experiment, interpretation of the data, and writing the manuscript.

Support

1R01MH092512 (BDP, DA), NIH R01MH094757, R01MH096764.

Declaration of Competing Interest

Dr. Pearce reported no biomedical financial interests or potential conflicts of interest. He has received grant support from the NIH and the United States Defense Advanced Research Projects Agency and Congressionally Directed Medical Research Programs.

Dr. Duncan has received research support for work unrelated to this project from Posit Science, Auspex Pharmaceuticals, Inc., and Teva Pharmaceuticals, Inc. E.D. is a full-time attending psychiatrist in the Mental Health Service Line at the Atlanta Veterans Affairs Health Care System, Decatur, GA. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Department of Veterans Affairs.

Dr. Avramopoulos reported no biomedical financial interests or potential conflicts of interest.

Dr. Powers Lott reported no biomedical financial interests or potential conflicts of interest.

Dr. Jovanovic reported no biomedical financial interests or potential conflicts of interest.

Mr. Massa reported no biomedical financial interests or potential conflicts of interest. He is a full-time employee in the Research and Development Service at the Atlanta Veterans Affairs Medical Center, Decatur, GA.

Dr. Conneely reported no biomedical financial interests or potential conflicts of interest.

Dr. Lori reported no biomedical financial interests or potential conflicts of interest.

Dr. Mulle reported no biomedical financial interests or potential conflicts of interest.

Dr. Ressler serves on advisory boards for Takeda, Janssen, and Verily, and he has received sponsored research support from Alkermes and Brainsway. He receives funding from NIH and the Brain and Behavior Research Fund.

Dr. Gillespie reported no biomedical financial interests or potential conflicts of interest. He serves as a paid consultant to Cohen Veterans Bioscience.

Mr. Wang reported no biomedical financial interests or potential conflicts of interest.

Acknowledgments

This work was made possible due the work of Ann E. Pulver PhD, who is the director of the Epidemiology/Genetics Research Program in Psychiatry at Johns Hopkins University. Dr. Pulver and her team recruited the participants, collected samples, and performed key evaluations through her epidemiology genetics program. We are also grateful to Robert Yolken, MD (Theodore and Vada Stanley Distinguished Professor of Neurovirology in Pediatrics at Johns Hopkins) whose laboratory performed the serology assays.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.comppsy.2021.152236>.

References

- [1] Gejman PV, Sanders AR, Kendler KSJ. Genetics of schizophrenia: new findings and challenges, 12; 2011; 121–44.
- [2] Schizophrenia Working Group of the Psychiatric Genomics C, et al. Nature. 2014;511(7510):421–7. <https://doi.org/10.1038/nature13595> (PubMed PMID: 25056061; PMID: PMC4112379).
- [3] Sullivan PF. Puzzling over schizophrenia: schizophrenia as a pathway disease. Nat Med. 2012;18(2):210–1. <https://doi.org/10.1038/nm.2670> (PubMed PMID: 22310687; PMID: PMC4188922).
- [4] Pardin AF, Holmans P, Pocklington AJ, Escott-Price V, Ripke S, Carrera N, et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. Nat Genet. 2018;50(3):381–9 Epub 2018/02/28 <https://doi.org/10.1038/s41588-018-0059-2>. (PubMed PMID: 29483656; PMID: PMC5918692).
- [5] Nimgaonkar VL. Solutions for a perplexing conundrum: Immunologic abnormalities in schizophrenia. US: Oxford University Press; 2018.
- [6] Torrey EF, Bartko JJ, Yolken RH. *Toxoplasma gondii* and other risk factors for schizophrenia: an update. Schizophr Bull. 2012;38(3):642–7.
- [7] Vyas A, Kim SK, Giacomini N, Boothroyd JC, Sapolsky RM. Behavioral changes induced by toxoplasma infection of rodents are highly specific to aversion of cat odors. Proc Natl Acad Sci U S A. 2007;104(15):6442–7.
- [8] Flegel J. Influence of latent toxoplasma infection on human personality, physiology and morphology: pros and cons of the toxoplasma–human model in studying the manipulation hypothesis. J Exp Biol. 2013;216(1):127–33.
- [9] Avramopoulos D, Pearce BD, McGrath J, Wolyniec P, Wang R, Eckart N, et al. Infection and inflammation in schizophrenia and bipolar disorder: a genome wide study for interactions with genetic variation. PLoS One. 2015;10(3):e0116696 Epub 2015/03/18 <https://doi.org/10.1371/journal.pone.0116696>. (PubMed PMID: 25781172; PMID: PMC4363491).
- [10] Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559–75 S0002-9297(07)61352-4 [pii] <https://doi.org/10.1086/519795>. (PubMed PMID: 17701901).
- [11] Wang AW, Avramopoulos D, Lori A, Mulle J, Conneely K, Powers A, et al. Genome-wide association study in two populations to determine genetic variants associated with *Toxoplasma gondii* infection and relationship to schizophrenia risk. Progress in

- Neuropsychopharmacology & Biological Psychiatry. 2019;92:133–47 Epub 2019/01/06 <https://doi.org/10.1016/j.pnpb.2018.12.019>. (PubMed PMID: 30610941).
- [12] Widom CATaH. Level-spacing distributions and the airy kernel. *Communications in Mathematical Physics*. 1994;159(1):151–74.
- [13] Euesden J, Lewis CM, O'Reilly PF. PRSice: polygenic risk score software. *Bioinformatics*. 2015;31(9):1466–8. <https://doi.org/10.1093/bioinformatics/btu848> (PubMed PMID: 25550326; PMCID: PMC4410663).
- [14] Tsuang MT, Stone WS, Faraone SV. Genes, environment and schizophrenia. *British Journal of Psychiatry - Supplementum*. 2001;40:s18–24.
- [15] Guloksuz S, Pries LK, Delespaul P, Kenis G, Luyck JJ, Lin BD, et al. 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):173–82.
- [16] Gaskell EA, Smith JE, Pinney JW, Westhead DR, McConkey GA. A unique dual activity amino acid hydroxylase in *Toxoplasma gondii*. *PLoS One*. 2009;4(3):e4801 Electronic Resource.
- [17] Webster JP, McConkey GA. *Toxoplasma gondii*-altered host behaviour: clues as to mechanism of action. *Folia Parasitol*. 2010;57(2):95–104.
- [18] Kamerkar S, Davis PH. *Toxoplasma* on the brain: understanding host-pathogen interactions in chronic CNS infection. *J Parasitol Res*. 2012;2012.
- [19] Xiao J, Prandovszky E, Kannan G, Pletnikov MV, Dickerson F, Severance EG, et al. *Toxoplasma gondii*: biological parameters of the connection to schizophrenia. *Schizophr Bull*. 2018;44(5):983–92.
- [20] Nance JP, Vannella KM, Worth D, David C, Carter D, Noor S, et al. Chitinase dependent control of protozoan cyst burden in the brain. *PLoS Pathog*. 2012;8(11):e1002990 Epub 2012/12/05 <https://doi.org/10.1371/journal.ppat.1002990>. (PubMed PMID: 23209401; PMCID: PMC3510238).
- [21] Trotta A, Iyegbe C, Di Forti M, Sham PC, Campbell DD, Cherny SS, et al. Interplay between schizophrenia polygenic risk score and childhood adversity in first-presentation psychotic disorder: a pilot study. *PLoS One*. 2016;11(9).
- [22] Ursini G, Punzi G, Chen Q, Marengo S, Robinson JF, Porcelli A, et al. Convergence of placenta biology and genetic risk for schizophrenia. *Nat Med*. 2018;24(6):792–801.
- [23] Parks S, Avramopoulos D, Mulle J, McGrath J, Wang R, Goes FS, et al. HLA typing using genome wide data reveals susceptibility types for infections in a psychiatric disease enriched sample. *Brain Behav Immun*. 2018;70:203–13. <https://doi.org/10.1016/j.bbi.2018.03.00129574260>.
- [24] Rougier S, Montoya JG, Peyron F. Lifelong persistence of toxoplasma cysts: a questionable dogma? *Trends Parasitol*. 2017;33(2):93–101.
- [25] Flegr J, Tureček P. New approach and new permutation tests with R programs for analyses of false-negative-contaminated data in medicine and biology. *Biology Open*. 2020;9(1).
- [26] Dudbridge F. Power and predictive accuracy of polygenic risk scores. *PLoS Genet*. 2013;9(3):e1003348. <https://doi.org/10.1371/journal.pgen.1003348> (PubMed PMID: 23555274; PMCID: PMC3605113).