Validating Orchid's Bipolar Disorder Genetic Risk Score

ORCHID

Written by Orchid Bioinformatics Staff

Introduction

Bipolar disorder is a chronic psychiatric condition characterized by recurrent episodes of abnormally elevated mood and energy (mania or hypomania) and episodes of depression, with periods of relative mood stability in between.[1] The lifetime prevalence of bipolar disorder among adults in the US general population has been estimated to be approximately 4.4%.[2] Individuals with bipolar disorder have a reduced life expectancy compared with the general population, with increased risk of suicide and higher rates of comorbid medical conditions contributing to this difference.[3,4] Treatment of bipolar disorder usually includes long-term use of mood-stabilizing medications, such as lithium or valproate, often combined with psychotherapy, although many individuals continue to experience recurrent mood episodes or residual symptoms despite treatment.[5]

Genetic Risk Score

Risk for bipolar disorder is shaped partly by genetic factors.[6] Monogenic testing is not available because no single gene causes the condition. Genetic risk scores (GRS), which combine the small effects of many variants into a single score, are currently the only way to estimate genetic risk. Although not diagnostic, a GRS can indicate how likely an individual is to develop the disease compared to the population baseline risk.

The bipolar disorder GRS included in Orchid's reports was sourced from the PGS Catalog, an open database of published GRS models, and contains more than 900,000 variants.[7,8] Risk predictions are adjusted to each individual's ancestry, with predictive power decaying as genetic distance from the predominantly European training data increases.[9] Orchid considers a GRS meaningfully predictive if individuals at approximately the 97.7th percentile have an odds ratio (OR) of at least 2. The bipolar disorder GRS meets this criterion for all common ancestry groups.

Evaluation on UK Biobank Data

We evaluated the predictive accuracy of Orchid's bipolar disorder GRS using the UK Biobank (UKB), a research database of roughly 500,000 genotyped individuals from the United Kingdom.[10] We restricted the analysis to individuals of British ancestry and defined bipolar disorder using the F31 ICD-10 code, yielding 1,412 cases and 407,108 controls (0.35% prevalence). We then grouped individuals by GRS percentile and compared the observed disease prevalence within each group to our model's predictions (Figure 1). For additional technical details, see the Supplementary Information.

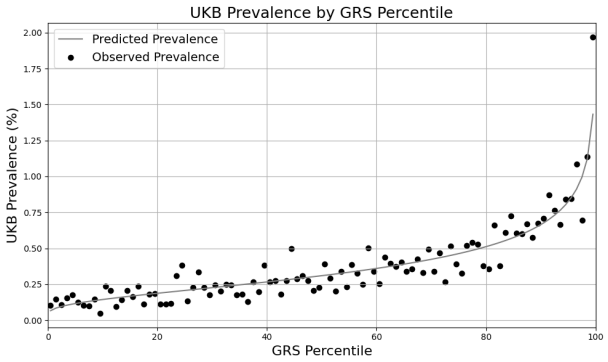

Figure 1. Risk Stratification. Predicted and observed prevalence in the UKB for individuals grouped by GRS percentile.

Table 1 shows the bipolar disorder observed prevalence for individuals in the UKB grouped by GRS percentile range (top 10%, 5%, and 1%), as well as how their risk compares to the baseline risk at the 50th GRS percentile. Those with higher GRS relative to the population baseline also had substantially higher observed prevalence of bipolar disorder, supporting the predictive accuracy of the GRS to identify individuals with elevated risk.

GRS Group	Observed UKB Prevalence	Odds Ratio
Baseline (50th percentile)	0.31%	1.00
Top 10%	0.91%	3.00
Top 5%	1.09%	3.59
Top 1%	1.97%	6.54

Table 1. Observed prevalence of bipolar disorder in the UKB by GRS percentile range. Those with higher GRS relative to the population baseline also had substantially higher observed prevalence of bipolar disorder.

Estimating Lifetime Risk

The average observed prevalence of bipolar disorder in the UKB was 0.35%. This is considerably lower than the lifetime prevalence, which has been estimated to be approximately 4.4%.[2] This is likely due in part to the fact that UKB participants tend to be healthier than the general population, which leads to lower observed disease prevalence.[11] Additionally, the observed prevalence in the UKB includes people still living who could develop the disease when they are older, and so does not capture the full lifetime risk of the disease.

Orchid's clinical reports include predicted lifetime disease risk, which we calculate by first estimating how disease risk varies across GRS in the UKB and then rescaling that pattern so the average matches the known lifetime population risk (Figure 2).[12] People at the high end of the GRS distribution are predicted to have an elevated lifetime risk of the disease relative to the population (Table 2).

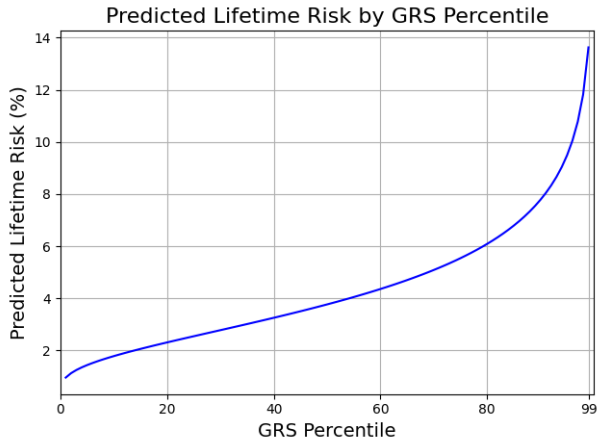

Figure 2. Adjusted Risk Stratification. Predicted risk estimates adjusted so that overall prevalence matches the approximately 4.4% estimate.[2]

GRS Percentile	Predicted Lifetime Risk	Relative Risk
50th (baseline)	3.77%	1.00x
95th	9.50%	2.52x
97th	10.78%	2.86x
99th	13.63%	3.61x

Table 2. Predicted lifetime prevalence of bipolar disorder at different GRS percentiles. Individuals with the highest GRS percentiles are predicted to have an increased risk of bipolar disorder relative to those at the 50th percentile.

Conclusion

In this study, we evaluated the bipolar disorder GRS on data from the UKB. We found that it performed well, particularly for identifying individuals with elevated risk of the disease relative to the population. In our embryo and couple reports, we adjust the model to predict lifetime risk, which is generally higher than observed prevalence in the UKB. The bipolar disorder GRS model is available to individuals of all ancestry groups.

Acknowledgements

This research was conducted using the UK Biobank Resource under Application Number 80545.

References

1. World Health Organization. Bipolar disorder, 2025. URL: https://www.who.int/news-room/fact-sheets/detail/bipolar-disorder.
2. National Institute of Mental Health. Bipolar Disorder — Statistics, 2025. URL: https://www.nimh.nih.gov/health/statistics/bipolar-disorder.
3. R. Voelker. What Is Bipolar Disorder? *JAMA*, 2024. URL: https://jamanetwork.com/journals/jama/fullarticle/2815378.
4. Kathleen R. Merikangas, Robert Jin, Jian-Ping He, Ronald C. Kessler, et al. Prevalence and correlates of bipolar spectrum disorder in the World Mental Health Survey Initiative. *Archives of General Psychiatry*, 68(3):241–251, 2011. doi:10.1001/archgenpsychiatry.2011.12.
5. National Institute of Mental Health. Bipolar Disorder, 2025. URL: https://www.nimh.nih.gov/health/publications/bipolar-disorder.
6. Jennifer H. Barnett and Jordan W. Smoller. The genetics of bipolar disorder. *Neuroscience*, 164(1):331–343, 2009. doi:10.1016/j.neuroscience.2009.03.080.
7. PGS Catalog. PGS002786: BD_SDPR (Bipolar disorder Polygenic Score), 2022. URL: https://www.pgscatalog.org/score/PGS002786/.
8. Yuanyuan Gui, Xiaocheng Zhou, Zixin Wang, Yiliang Zhang, et al. Sex-specific genetic association between psychiatric disorders and cognition, behavior and brain imaging in children and adults. *Translational Psychiatry*, 12(1):347, 2022. doi:10.1038/s41398-022-02041-6.
9. Florian Privé, Hugues Aschard, Shai Carmi, Lasse Folkersen, et al. Portability of 245 polygenic scores when derived from the UK Biobank and applied to 9 ancestry groups from the same cohort. *American Journal of Human Genetics*, 109(1):12–23, 2022. doi:10.1016/j.ajhg.2021.11.008.
10. C. Sudlow, J. Gallacher, N. Allen, et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Medicine*, 12(3):e1001779, 2015. doi:10.1371/journal.pmed.1001779.
11. A. Fry, T.J. Littlejohns, C. Sudlow, et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *American Journal of Epidemiology*, 186:1026–1034, 2017. doi:10.1093/aje/kwx246.
12. N. Chatterjee, J. Shi, M. García-Closas, et al. Developing and evaluating polygenic risk prediction models for stratified disease prevention. *Nature Reviews Genetics*, 17:392–406, 2016. doi:10.1038/nrg.2016.27.

Supplementary Information

Baseline Risk	OR per SD	OR per 2 SD
3.77%	1.82	3.31

Table 3. OR per SD. The baseline risk for an individual with a median GRS, and the predicted OR at one and two SDs, respectively. A GRS must have a predicted OR >2 at 2 SD to be included in Orchid's clinical reports.

UKB Prevalence	Population Prevalence	Liability R^2
0.35%	4.4%	6.44%

Table 4. Liability R^2. The estimated liability R^2 using a population prevalence of 4.4%.

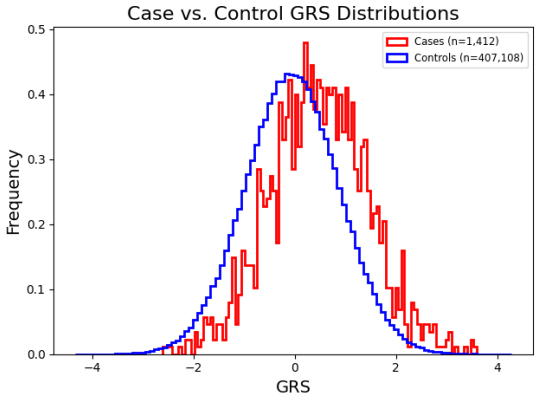

Figure 3. GRS histograms. GRS distributions for cases and controls. Both are approximately normal, with the case distribution shifted noticeably higher compared to the controls.

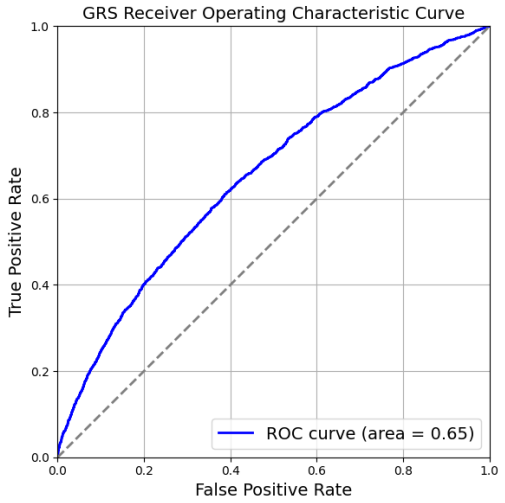

Figure 4. The receiver operating characteristic (ROC) used to compute the ROC area under the curve (AUC). The ROC curve is a graphical representation of a binary classifier's performance, plotting the True Positive Rate (TPR) against the False Positive Rate (FPR) across different decision thresholds. A curve closer to the top-left indicates a better model, while a diagonal line (AUC = 0.5) represents random guessing.

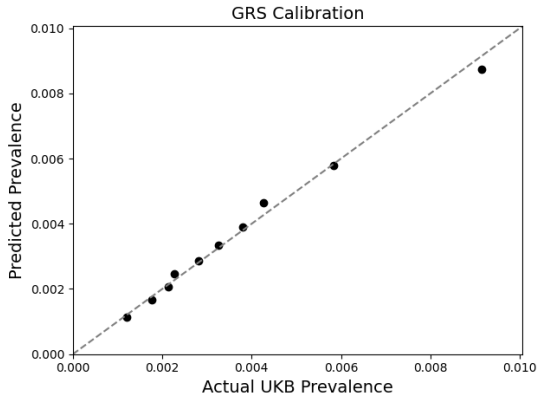

Figure 5. Calibration Curve. Calibration plot showing observed disease prevalence versus predicted risk across GRS deciles.