Validating Orchid's Hypertension Genetic Risk Score

ORCHID

Written by Orchid Bioinformatics Staff

Introduction

Hypertension is defined by persistently high blood pressure, measured using systolic pressure (during a heartbeat) and diastolic pressure (between heartbeats). It is a major risk factor for cardiovascular disease, with risk increasing continuously as blood pressure rises. Key hypertension risk factors include obesity, physical inactivity, excessive sodium consumption, insufficient dietary potassium intake, and alcohol intake.[1] Hypertension is the single largest risk factor contributing to global all-cause mortality, leading to approximately 9.4 million deaths per year.[1] The CDC estimates that nearly half of U.S. adults have hypertension, with only about a quarter of those affected having it controlled by treatment.[2] Treatment typically involves a combination of lifestyle modification and blood pressure lowering medication. Common lifestyle interventions include dietary change and increased physical activity.[1]

Genetic Risk Score

Hypertension is influenced by both environmental and genetic factors. Rare variants in genes such as PDE3A can cause the disease,[1] but most cases arise from the combined effects of many genetic variants and environmental exposures. Genetic risk scores (GRS), which combine the small effects of many variants into a single score, can estimate genetic risk. Although not diagnostic, a GRS can indicate how likely an individual is to develop the disease.

Orchid's hypertension GRS was trained following current industry standards.[3,4] The GRS was constructed using the SBayesRC algorithm trained on publicly available FinnGen and Million Veterans Program summary statistics.[5,6] The summary statistics include 606,331 cases and 493,942 controls.[7] The resulting GRS contains over a million variants.

Risk predictions are adjusted to each individual's ancestry, with predictive power decaying as genetic distance from the predominately European training data increases.[8] Orchid considers a GRS meaningfully predictive if individuals at roughly the 97.7th percentile have an odds ratio (OR) of at least 2. The hypertension GRS meets this criterion for the European, Central South Asian, and East Asian ancestry groups and is available to individuals of these groups. Availability for an individual may vary due to admixture.

Evaluation on UK Biobank Data

We evaluated the predictive accuracy of Orchid's hypertension GRS using the UK Biobank (UKB), a research database of roughly 500,000 genotyped individuals from the United Kingdom.[9] We restricted the analysis to participants of British ancestry and defined hypertension as any diagnoses under ICD-10 codes I10.x, yielding 119,352 cases and 289,168 controls (29.2% 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.

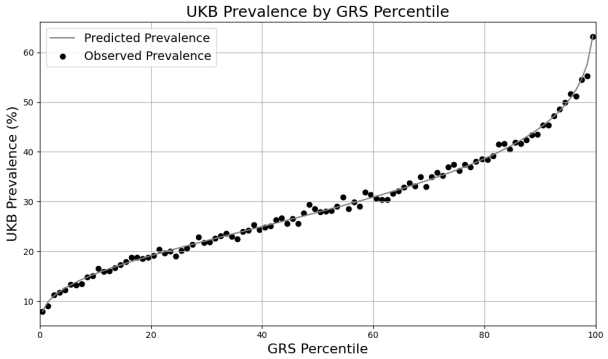

Figure 1. Risk Stratification. Predicted and observed prevalence in the UKB for individuals grouped by GRS percentile.

Table 1 shows the hypertension 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 hypertension, supporting the predictive accuracy of the GRS to identify individuals with elevated risk.

GRS Group	Observed UKB Prevalence	Odds Ratio
Baseline (50th percentile)	28.21%	1.00
Top 10%	50.27%	2.57
Top 5%	54.34%	3.03
Top 1%	63.18%	4.37

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

Estimating Lifetime Risk

The average observed prevalence of hypertension in the UKB was 29.2%. This is considerably lower than the lifetime

prevalence in the US general population, which has been estimated to be approximately 48.1%.[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.[10] 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).[11] 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).

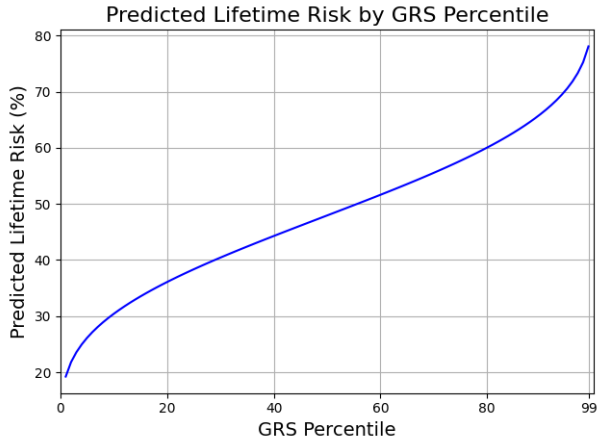

Figure 2. Adjusted Risk Stratification. Predicted risk estimates adjusted so that overall prevalence matches the 48.1% estimate.[2]

GRS Percentile	Predicted Lifetime Risk	Relative Risk
50th (baseline)	47.95%	1.00x
95th	70.58%	1.47x
97th	73.35%	1.53x
99th	78.10%	1.63x

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

Conclusion

In this study, we evaluated our hypertension 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 hypertension GRS meets Orchid's performance criteria for the European, Central South Asian, and East Asian ancestry groups and is available to individuals of these groups. Availability for an individual may vary due to admixture.

Acknowledgments

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

References

1. S. Oparil, M. C. Acelajado, G. L. Bakris, et al. Hypertension. *Nat Rev Dis Primers*, 4:18014, 2018. doi:10.1038/nrdp.2018.14.
2. Centers for Disease Control and Prevention. Hypertension prevalence. Million Hearts. https://millionhearts.hhs.gov/data-reports/hypertension-prevalence.html, 2025. Accessed 2025-12-15.
3. S. Moore, I. Davidson, J. Anomaly, et al. Development and validation of polygenic scores for within-family prediction of disease risks. *medRxiv*, 2025. doi:10.1101/2025.08.06.25333145.
4. S. Cordogan, D. B. Starr, N. R. Treff, et al. Within- and between-family validation of nine polygenic risk scores developed in 1.5 million individuals: implications for IVF, embryo selection, and reduction in lifetime disease risk. *medRxiv*, 2025. doi:10.1101/2025.10.24.25338613.
5. Z. Zheng, S. Liu, J. Sidorenko, et al. Leveraging functional genomic annotations and genome coverage to improve polygenic prediction of complex traits within and between ancestries. *Nat Genet*, 56:767–777, 2024. doi:10.1038/s41588-024-01704-y.
6. FinnGen. FinnGen+MVP+UKBB Summary Statistics. https://mvp-ukbb.finngen.fi/about, 2025. Accessed 2025-12-05.
7. FinnGen. FinnGen+MVP+UKBB Phenotypes. https://mvp-ukbb.finngen.fi, 2025. Accessed 2025-12-15.
8. F. Privé 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.
9. 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.

10. 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. *Am J Epidemiol*, 186:1026–1034, 2017. doi:10.1093/aje/kwx246.

11. N. Chatterjee, J. Shi, M. García-Closas, et al. Developing and evaluating polygenic risk prediction models for stratified disease prevention. *Nat Rev Genet*, 17:392–406, 2016. doi:10.1038/nrg.2016.27.

Supplementary Information

Baseline Risk	OR per SD	OR per 2 SD
47.95%	1.79	3.20

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
29.2%	48.1%	9.16%

Table 4. Liability R^2. The estimated liability R^2 using a population prevalence of 48.1%.

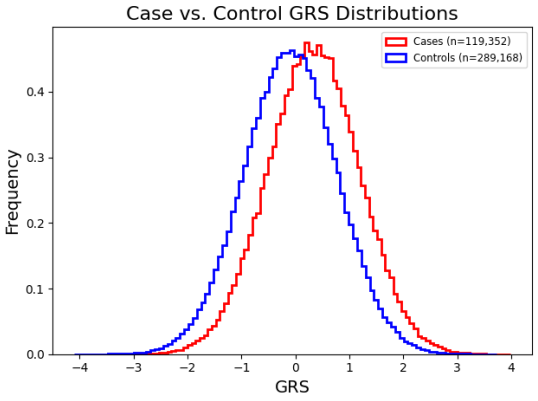

Figure 3. GRS histograms. GRS distributions for cases and controls. Both are approximately normal, with the case distribution shifted noticeably higher compared to the controls.

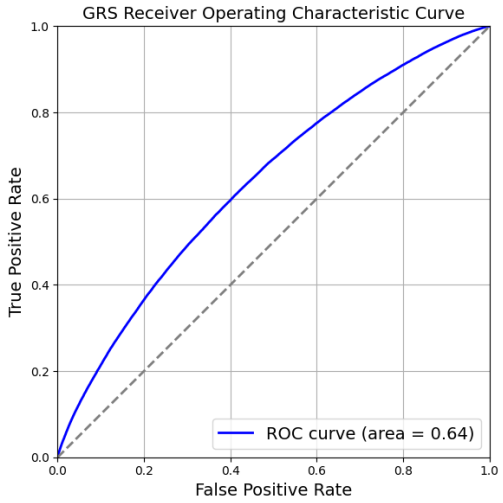

Figure 4. The receiver operating characteristic (ROC) used to compute the ROC area under the curve (AUC). The ROC curve is a graphical representation of a binary classifier's performance, plotting the True Positive Rate (TPR) against the False Positive Rate (FPR) across different decision thresholds.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

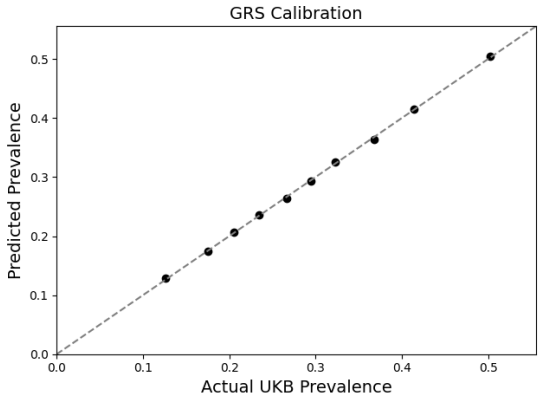

Figure 5. Calibration Curve. Calibration plot showing observed disease prevalence versus predicted risk across GRS deciles.