# Validating Orchid's Class III Obesity Genetic Risk Score

ORCHID

**Written by Orchid Bioinformatics Staff**

## Introduction

Class III obesity is defined as a body mass index (BMI) of at least 40 kg/m² and is the most severe category of obesity based on standard BMI classification. It is associated with markedly increased morbidity and mortality, including substantially elevated risks of type 2 diabetes, cardiovascular disease, sleep apnea, certain cancers, impaired physical function, and reduced quality of life. Environmental risk factors include living in settings with abundant energy-dense foods and limited opportunities for physical activity. Excessive weight gain by either the mother during pregnancy or by the child after birth is also a risk factor.[1]

Obesity is highly prevalent worldwide, with the WHO estimating that >2.1 billion adults were overweight or obese in 2014.[1] In the United States, the CDC estimates that 9.2% of adults suffered from class III obesity between 2017 and 2020.[2] Treatment of obesity focuses first on lifestyle changes such as improving diet and increasing physical activity. If these lower-risk approaches are not enough, medications can be used as add-ons to diet and exercise for selected patients. Bariatric surgery is an option for severe obesity that can produce substantial, longer-term weight loss but requires careful risk–benefit consideration.[1]

## Genetic Risk Score

Class III obesity is shaped by both environmental and genetic factors. 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.

Orchid's class III obesity 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 277,884 cases and 799,863 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 predominantly 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 class III obesity GRS meets this criterion for all common ancestry groups.

## Evaluation on UK Biobank Data

We evaluated the predictive accuracy of Orchid's class III obesity 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 class III obesity as a BMI of at least 40, yielding 7,931 cases and 400,589 controls (1.9% 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.

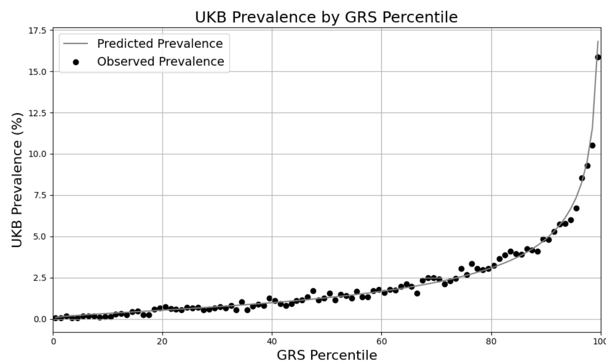


**Figure 1. Risk Stratification.** Predicted and observed prevalence in the UKB for individuals grouped by GRS percentile.

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

| GRS Group | Observed UKB Prevalence | Odds Ratio |
|---|---:|---:|
| Baseline (50th percentile) | 1.40% | 1.00 |
| Top 10% | 7.19% | 5.44 |
| Top 5% | 9.55% | 7.42 |
| Top 1% | 15.76% | 13.14 |

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

## Estimating Lifetime Risk

The average observed prevalence of class III obesity in the UKB was 1.9%. This is considerably lower than the prevalence in the US general population, which has been estimated to be approximately 9.2%.[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).

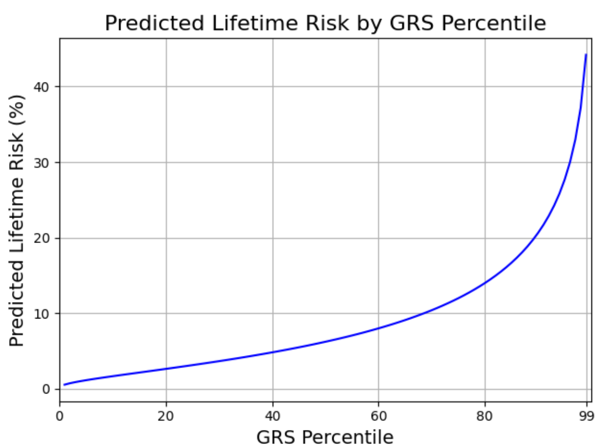


**Figure 2. Adjusted Risk Stratification.** Predicted risk estimates adjusted so that overall prevalence matches the 9.2% estimate.[2]

| GRS Percentile | Predicted Lifetime Risk | Relative Risk |
|---|---|---|
| 50th (baseline) | 6.23% | 1.00x |
| 95th | 27.70% | 4.45x |
| 97th | 33.00% | 5.30x |
| 99th | 44.19% | 7.10x |

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

## Conclusion

In this study, we evaluated our class III obesity 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 class III obesity GRS model is available to individuals of all ancestry groups.

## Acknowledgments

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

## References

1. P. González-Muniesa, M. A. Martínez-González, F. Hu, et al. Obesity. *Nat Rev Dis Primers*, 3:17034, 2017. doi:10.1038/nrdp.2017.34.
2. Centers for Disease Control and Prevention. Adult Obesity Facts. https://www.cdc.gov/obesity/adult-obesity-facts/. Accessed 2025-12-17.
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. Accessed 2025-12-05.
7. FinnGen. FinnGen+MVP+UKBB Phenotypes. https://mvp-ukbb.finngen.fi. Accessed 2025-12-15.
8. Florian 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 |
|---|---|---|
| 6.23% | 2.90 | 8.42 |

**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 $R^2$ |
|---|---|---|
| 1.9% | 9.2% | 19.97% |

**Table 4. Liability $R^2$.** The estimated liability $R^2$ using a population prevalence of 9.2%.

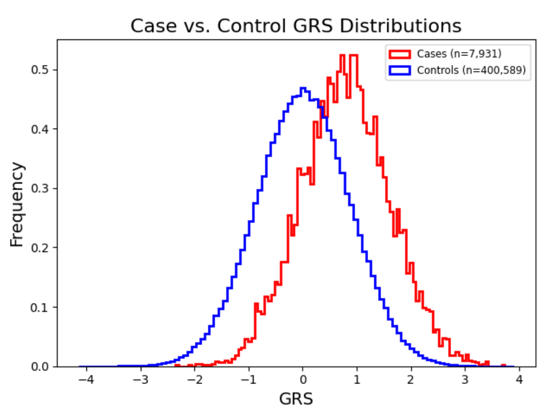


**Figure 3. GRS histograms.** GRS distributions for cases and controls. Both are approximately normal, with the case distribution shifted noticeably higher compared to the controls.

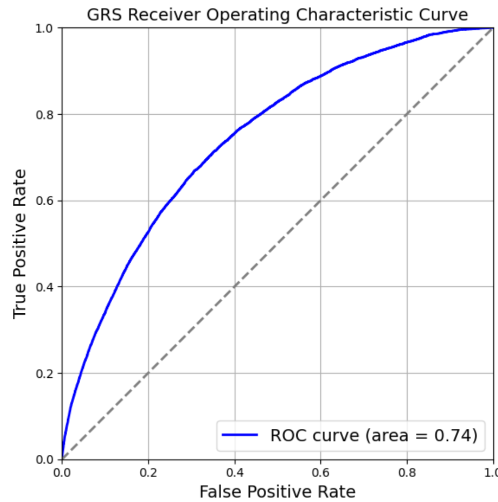


**Figure 4. The receiver operating characteristic (ROC) used to compute the ROC area under the curve (AUC).** The ROC curve is a graphical representation of a binary classifier's performance, plotting the True Positive Rate (TPR) against the False Positive Rate (FPR) across different decision thresholds.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

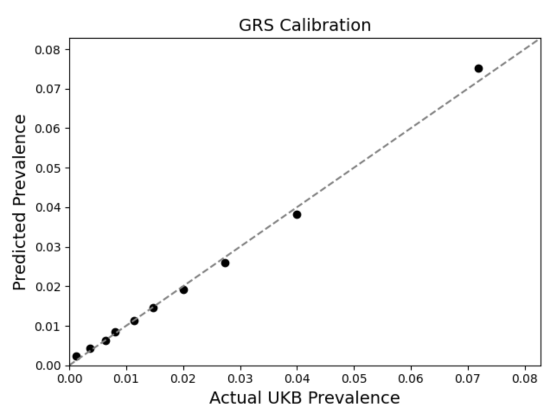


**Figure 5. Calibration Curve.** Calibration plot showing observed disease prevalence versus predicted risk across GRS deciles.