A comprehensive evaluation of polygenic score methods across cohorts in psychiatric disorders

Guiyan Ni*, Jian Zeng, Joana A Revez, Ying Wang, Tian Ge, Restaudi Restaudi, Jacqueline Kiewa, Dale R Nyholt, Jonathan R I Coleman, Jordan W Smoller, Schizophrenia Working Group of the Psychiatric Genomics Consortium, Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, Jian Yang, Peter M Visscher, & Naomi R Wray *

* Institute for Molecular Bioscience, University of Queensland, Brisbane, Queensland, 4072, Australia

Background

- There are now many methods to calculate polygenic scores (PGSs)
- Evaluation of new PGS methods are made using simulated data or single target cohort. In real data sets, however, there can be heterogeneity between target sample cohorts, such as phenotype definition, ascertainment of cases and controls, cohort-specific ancestry or technical artefacts in genotype generation.

➔ **What is the optimal method for psychiatric disorders ?**

Method

- **Prediction methods:** P+T[2] (benchmark), SBLUP[3], LDpred-Inf[4], LDpred-funct[5], LDpred[4], PRS-CS[6], PRS-CS-auto[6], SBayesR[7]

  P+T, LDpred, and PRS-CS need a tuning cohort to determine a critical parameter
- Out-of-sample prediction using Leave-one-cohort-out strategy
  - target sample: each cohorts with genotype data
- **Evaluation of prediction:** The area under curve, the proportion of variance explained by PGS on the scale of liability

Data

Schizophrenia:

- GWAS summary statistics:
  - From PGC SCZ Working group, 34 European ancestry cohorts, denoted as SCZ34; And another three cohorts from Pardiñas et al[1];
  - Comprises more than 8M imputed SNPs in 31K SCZ cases and 41K controls.
- Individual level genotype data: 25K cases and 30K controls of SCZ34

Major depression:

- GWAS summary statistics:
  - From seven studies: UK Biobank, 23andMe, GERA, iPSYCH, deCODE, GenScotland, and the PGC Major Depressive Disorder (MDD) Working group (denoted as MDD29).
  - Comprise almost 13M imputed SNPs from 248K cases and 563K controls
- Individual level genotype data:
  - 15K cases and 24K controls from 26 cohorts of MDD29

Method

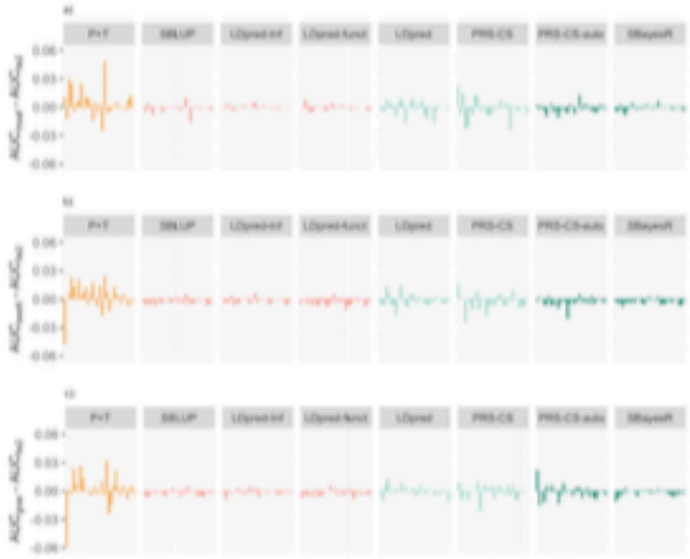

Figure 1. Differences in AUC of a PGS method when using different tuning cohorts. The different bars in each method (x-axis) refer to different validation cohorts ordered by sample size. The y-axis is the AUC difference when using alternative tuning cohort

The tuning cohort that generates higher PGS is method dependent and differs between cohorts.

Result - 2

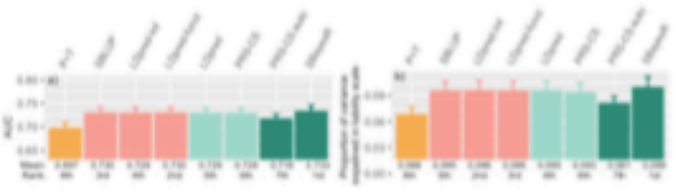

Figure 2. Results from prediction of SCZ case/control status using different PGS methods.

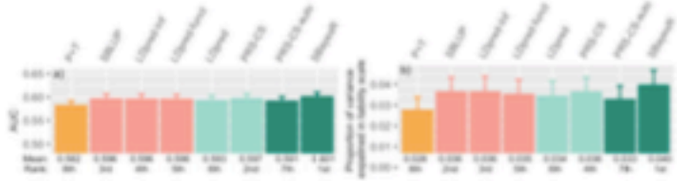

Figure 3. Results from prediction of MDD case/control status using different PGS methods.

- For SCZ across 30 target cohorts, the SBayesR PGS achieved a mean area under the receiver operator characteristic curve (AUC) of 0.733, and explained 9.9% of variance on the liability scale.
- For MDD across 26 target cohorts, the AUC and variance explained were 0.601 and 4.0%, respectively.
- The variance explained by the SBayesR PGS was 46% and 43% higher for SCZ and MDD, respectively, compared to the basic p-value thresholding P+T method.

Conclusions

- Except P+T, all methods have similar performance.
- SBayesR saw the highest prediction accuracy in most of the comparisons.
- The tuning cohort can have considerable impact.

Contact:
Naomi Wray: Naomi.wray@uq.edu.au
Guiyan Ni: uqgni@uq.edu.au

Reference

[1]Pardiñas, A.F., et al. (2018). Nat. Genet. 50, 381-389.
[2]Wray, N.R., et al. (2007). Genome Res. 17, 1520-1528.
[3]Robinson, M.R., et al. (2017). Nat. Hum. Behav. 1, 0016.
[4]Vilhjálmsson, B.J., et al. (2015). Am. J. Hum. Genet. 97, 576-592.
[5]Marquez-Luna, C., et al. (2018). bioRxiv, 375337.
[6]Ge, T., et al. (2019). Nat. Commun. 10, 1776.
[7]Lloyd-Jones, L.R., et al. (2019). Nat. Commun. 10.