

Polygenic Risk Score (PRS) Introduction 001

Overview and topic-stratified PRS reference list

Drs. Lei Sun, Wei Deng, Yanyan Zhao

Department of Statistical Sciences, FAS

Division of Biostatistics, DLSPH

University of Toronto

10 October, 2021

McLaughlin Centre PRS Team

- Critical mass of methodologists motivated by PRS-relevant data and implementation challenges
- Built on CIHR STAGE: over a decade of productive interdisciplinary research collaborations and capacity building

PRS Team Major Goal

Develop methodologies, platform and expertise for robust applications of PRS for precision health with a population impact



**McLAUGHLIN
CENTRE**

camh
Centre for Addiction and Mental Health



Statistical Sciences
UNIVERSITY OF TORONTO



Sinai
Health | Lunenfeld-Tanenbaum
Research Institute

Dalla Lana
School of Public Health

SickKids

Outline of the whole PRS workshop

1. Statistical fundamentals of PRS
2. Real-life users FAQ & solutions
3. Hands-on exercises (restricted)
4. Advanced topics
5. Clinical implications
6. Discussion

Outline of 1. Statistical Fundamentals of PRS

Lei Sun, Wei Deng and Yanyan Zhao

- ▶ slides-PRS-001-overview: a topic-stratified PRS reference list
- ▶ slides-PRS-101-foundation: GWAS, h^2 and prediction
- ▶ **slides-PRS-201-basic:** PRS calculation and performance evaluation
- ▶ **slides-PRS-301-basic-plus:** some obvious or not so obvious follow-up Qs
- ▶ **slides-PRS-401-heterogeneity:** heterogeneity and transportability
- ▶ slides-PRS-501-LD: linkage disequilibrium

Misc notes

- ▶ **Deeper** understanding of the **basic** PRS
- ▶ Step-by-step **R-assisted ‘hand calculation’**
- ▶ (almost) **without black box**
- ▶ **Reproducible** examples and results

All .pdf and source .Rmd files will be made open resource soon after the IGES workshop at <https://github.com/LeiSunUofT>

- ▶ Workshop teaching: \approx 3 hours (with breaks), **covering only some of the notes. Please go over the 001 and 101 slides distributed prior to the workshop, so that more workshop time can be dedicated to other materials. Thank you!**
- ▶ If interested: ∞ hours of self-study
- ▶ If interested: quizzes throughout the lecture notes

Goal of this lecture

- ▶ Overview of the topics to be covered in 1. statistical fundamentals of PRS of this workshop.
- ▶ PRS is intuitive with a simple expression:

$$PRS_i = \sum_{j=1}^J \hat{\beta}_j G_{ij}$$

- ▶ But, there are many (answered and open) complex (methodological and practical) questions.
- ▶ Provide a topic-stratified reference list for self-study.

If interested: PRS by NIH-National Human Genome Research Institute

A "polygenic risk score" is one way by which people can learn about their risk of developing a disease, based on the total number of changes related to the disease.

(environmental factors, populations)

If interested: PRS by Wiki (March 17, 2021)

*In genetics, a polygenic score, also called a polygenic risk score (PRS), genetic risk score, or genome-wide score, is a number that summarises the estimated effect of many genetic variants on an individual's phenotype, typically calculated as a **weighted sum of trait-associated alleles**.^{[1][2][3]}*

It reflects an individual's estimated genetic predisposition for a given trait and can be used as a predictor for that trait.^{[4][5][6][7][8]}

Polygenic scores are widely used in animal breeding and plant breeding (usually termed genomic prediction or genomic selection) due to their efficacy in improving livestock breeding and crops.^[9]

*They are also increasingly being used for **risk prediction** in humans for **complex diseases** [10] which are typically affected by many genetic variants that each confer a small effect on overall risk.^{[11][12]}*

- [1] Dudbridge (2013). *PLOS Genetics*. Power and predictive accuracy of polygenic risk scores.
- [2] Torkamani et al. (2018). *Nature Reviews Genetics*. The personal and clinical utility of polygenic risk scores.
- [3] Lambert et al. (2019). *Human Molecular Genetics*. Towards clinical utility of polygenic risk scores.
- [4] de Vlaming and Groenen (2015). *BioMed Research International*. The Current and Future Use of Ridge Regression for Prediction in Quantitative Genetics.
- [5] Lewis and Vassos E (2017). *Genome Medicine*. Prospects for using risk scores in polygenic medicine.
- [6] Khera et al. (2018). *Nature Genetics*. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations.
- [7] Yanes et al. (2020). *Clinical Genetics*. Uptake of polygenic risk information among women at increased risk of breast cancer.
- [8] Vilhjalmsson et al.(2015). *American Journal of Human Genetics*. Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. (LDpred)
- [9] Spindel and McCouch (2016). *The New Phytologist*. When more is better: how data sharing would accelerate genomic selection of crop plants. (Caution: more is only better if you know how to deal with potential heterogeneity.)
- [10] Regalado (2019). *MIT Technology Review*. 23andMe thinks polygenic risk scores are ready for the masses, but experts aren't so sure.
- [11] Visscher et al. (2017) *American Journal of Human Genetics*. 10 Years of GWAS Discovery: Biology, Function, and Translation. (*The Utility of GWAS-Derived Genetic Predictors*)
- [12] Spiliopoulou et al. (2015). *Human Molecular Genetics*. Genomic prediction of complex human traits: relatedness, trait architecture and predictive meta-models.

PRS = a weighted sum of trait-associated alleles

a simple expression with many complex questions

$$PRS_i = \sum_{j=1}^J \hat{\beta}_j G_{ij}$$

- ▶ PRS_i : PRS for individual i
which population? environmental exposure?
- ▶ J : the total number of '**relevant**' bi-allelic (common) SNPs
rare variants? other types of genetic markers?
- ▶ $\hat{\beta}_j$: **estimated effect size** of SNP j
portability: which population? phenotype/pleiotropy? direction of effects?
- ▶ G_{ij} : the number of copies of the risk allele
counting a or A? the minor or major allele?
- ▶ Σ : how to deal with **dependency/linkage disequilibrium (LD)**?

Other Complications/Considerations (= Opportunities)

Autosomes → **the X-chromosome** (and Y-chromosome)

Dr. Wei Deng, McMaster University

Main → **Interaction effect**

Variance → Covariance, **genetic overlap**

Population → **Family data**

Cross-sectional → **Longitudinal data** Dr. Laurent Briollais,
Luenfeld-Tanenbaum

Association and prediction → **Causal inference**

Dr. Linbo Wang, UofT Stat

(dry) lab-environment method development and evaluation →
real-life implementation and clinical implication Dr. Delnaz

Roshandel, Sickkids, Dr. Jennifer Brooks, UofT Epi and Dr. Andrew Paterson,
Sickkids

PRS for risk prediction for complex diseases

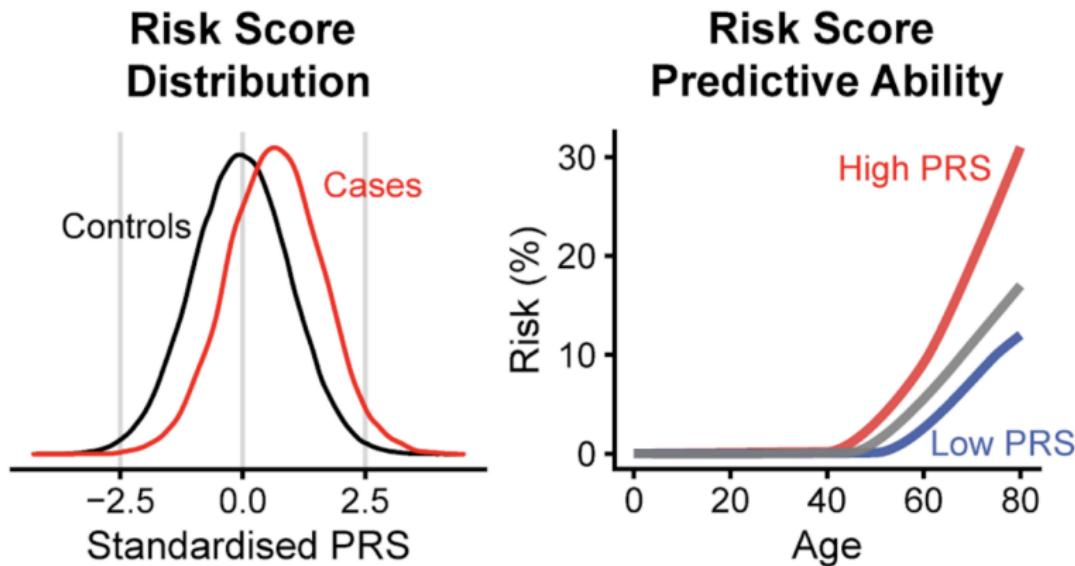


Figure 1 of Wand et al. (2021). *Nature*. Improving reporting standards for polygenic scores in risk prediction studies.

An illustration of the distribution of a polygenic risk score in both cases (individuals with a disease/trait) and controls (those without) [left], and

the predictive ability where individuals with a high polygenic risk score have an increased risk of disease/trait over time [right].

PRS for risk prediction, cont'd

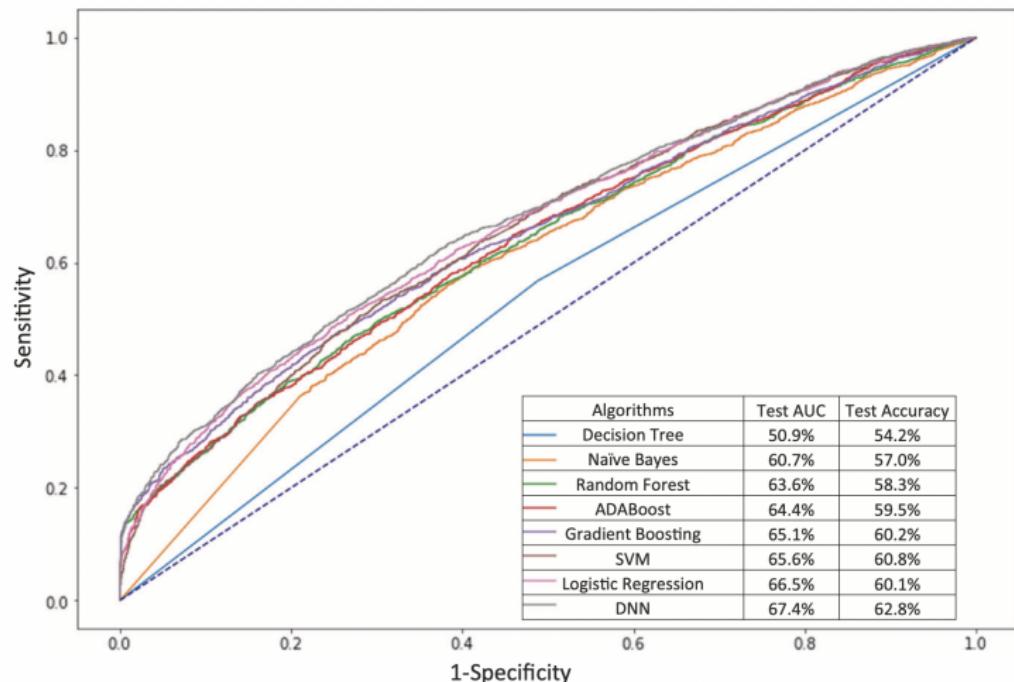


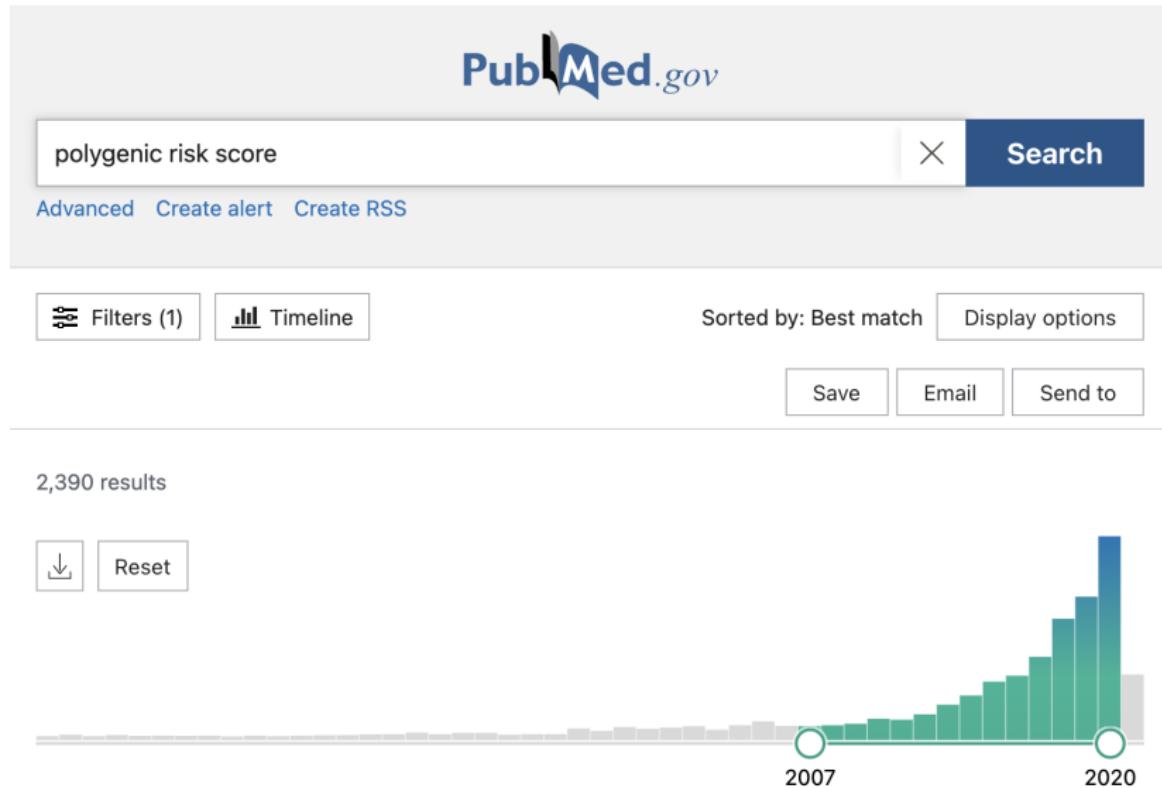
Figure 3 of Badre et al. (2021). *Journal of Human Genetics*. Deep neural network improves the estimation of polygenic risk scores for breast cancer.

N.B. The classical logistic regression is competitive!

Reference – PRS Paper ‘0’

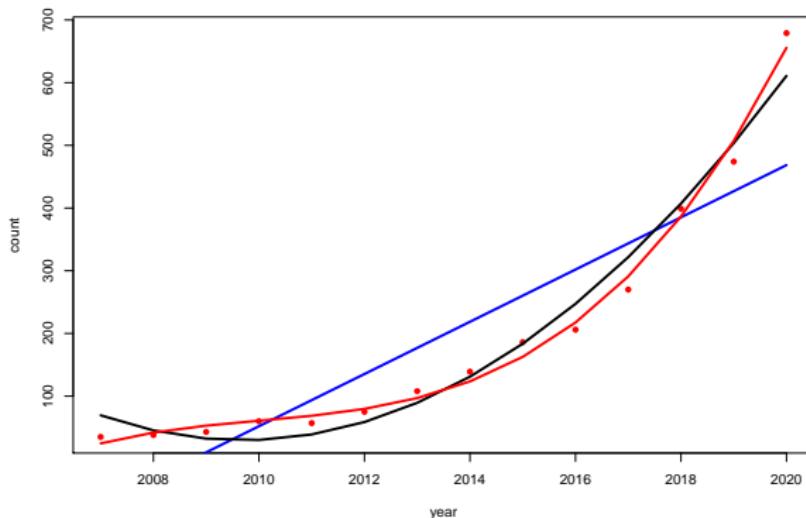
- ▶ Wray, Goddard, Visscher PM (2007). *Genome Research*. Prediction of individual genetic risk to disease from genome-wide association studies. (514 citation in PMC; March 17, 2021)
- ▶ International Schizophrenia Consortium; Purcell et al. (2009). *Nature*. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder (4123 citations in PMC; March 17, 2021)

If interested: Growth of PRS Research



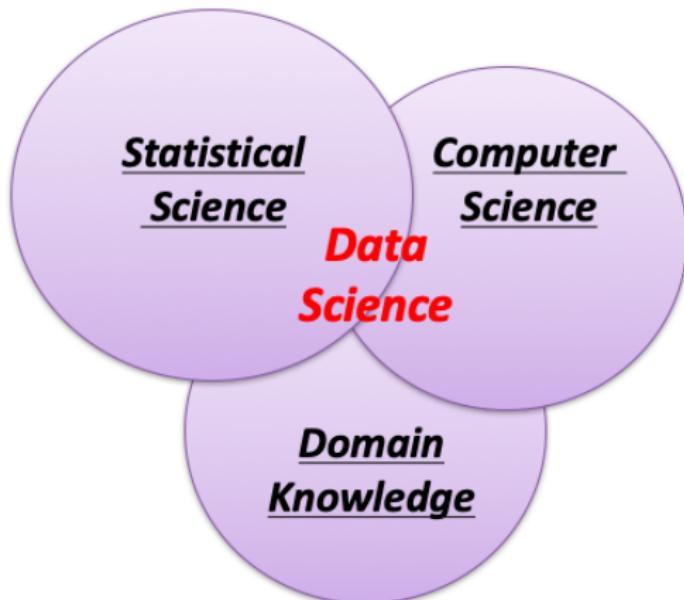
If interested: Not a Linear Growth!

```
# Search query: polygenic risk score;
# Search date: March 17, 2021
year=c(2007,2008,2009,2010,2011,2012,2013,2014,2015,2016,2017,2018,2019,2020)
count=c(35,38,43,60,57,75,108,139,186,206,270,399,474,679)
plot(year,count,col="red",pch=16)
lines(year,fitted(lm(count~year)),col='blue',lwd=3)
lines(year,fitted(lm(count~poly(year,2))),col='black',lwd=3)
lines(year,fitted(lm(count~poly(year,3))),col='red',lwd=3)
```



N.B. prediction \neq model fitting \neq 'best' fit

If interested: Discussion, PRS research is data science =
interdisciplinary/multidisciplinary research



If interested: Discussion cont'd

Harvard Data Science Review (HDSR)

Data Science Education

Reproducibility and Replicability

AI and Responsible Data Science

Genetic Diversity and Ancestry

Akinyemi Oni-Orisan, Yusuph Mavura, Yambazi Banda, Timothy A Thornton, Ronnie Sebro (2021). *New England Journal of Medicine*
Embracing Genetic Diversity to Improve Black Health

Topic-Stratified References

Not an Exhaustive List

Reference – PRS tutorial and ‘famous’ software

- ▶ Choi et al. (2019). *Nature Protocols*. Tutorial: a guide to performing polygenic risk score analyses.
- ▶ Purcell et al. (2007). *American Journal of Human Genetics*. PLINK: a toolset for whole-genome association and population-based linkage analysis. (PLINK)
- ▶ Vilhjalmsson et al.(2015). *American Journal of Human Genetics*. Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. (LDpred)
- ▶ Euesden et al. (2015). *Bioinformatics*. PRSice: Polygenic Risk Score software. (PRSice)
- ▶ Mak et al. (2017). *Genetic Epidemiology*. Polygenic scores via penalized regression on summary statistics. (lassosum; R package).

Reference – PRS review and perspectives

- ▶ Dudbridge (2013). *PLOS Genetics*. Power and predictive accuracy of polygenic risk scores.
- ▶ Wray et al. (2014). *Journal of Child Psychology and Psychiatry*. Research review: Polygenic methods and their application to psychiatric traits.
- ▶ Chatterjee et al. (2016). *Nature Review Genetics*. Developing and evaluating polygenic risk prediction models for stratified disease prevention.
- ▶ Visscher et al. (2017). *American Journal of Human Genetics*. 10 Years of GWAS Discovery: Biology, Function, and Translation. (*The Utility of GWAS-Derived Genetic Predictors*)
- ▶ Dudbridge et al. (2018) *Genetic Epidemiology*. Predictive accuracy of combined genetic and environmental risk scores.
- ▶ See a full [video](#) of “Panel Discussion: The Personal and Clinical Utility of Genetic Risk Scores: Are they ready for PRIME TIME?” from the CANSSI-Ontario Research Day on March 5, 2020.

Reference – PRS utility and clinical implication

- ▶ Lewis and Vassos (2017). *Genome Medicine*. Prospects for using risk scores in polygenic medicine.
- ▶ Hasin et al. (2017). *Genome Biology*. Multi-omics approaches to disease.
- ▶ Torkamani et al. (2018). *Nature Reviews Genetics*. The personal and clinical utility of polygenic risk scores.
- ▶ Lambert et al. (2019). *Human Molecular Genetics*. Towards clinical utility of polygenic risk scores.
- ▶ Janssens (2019). *Human Molecular Genetics*. Validity of polygenic risk scores: are we measuring what we think we are?.
- ▶ Lewis and Vassos (2020). *Genome Medicine*. Polygenic risk scores: from research tools to clinical instruments.
- ▶ Li et al. (2020). *Nature Review Genetics*. Electronic health records and polygenic risk scores for predicting disease risk.

References - PRS in action

- ▶ Khera et al. (2018). *Nature Genetics*. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations.
- ▶ Mavaddat et al. (2019). *American Journal of Human Genetics*. Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes.
- ▶ Lee et al. (2019). *Genetics in Medicine* BOADICEA: A comprehensive breast cancer risk prediction model incorporating genetic and nongenetic risk factors.
- ▶ Fritsche et al. (2020). *The American Journal of Human Genetics*. Cancer PRSweb: An Online Repository with Polygenic Risk Scores for Major Cancer Traits and Their Evaluation in Two Independent Biobanks.
- ▶ Mars et al. (2020). *Nature Medicine*. Polygenic and clinical risk scores and their impact on age at onset and prediction of cardiometabolic diseases and common cancers.
- ▶ Richardson et al. (2020). *Elife*. An atlas of polygenic risk score associations to highlight putative causal relationships across the human genome.
- ▶ Yanes et al. (2020). *Clinical Genetics*. Uptake of polygenic risk information among women at increased risk of breast cancer.
- ▶ Meyers et al. (2021). *Translational Psychiatry*. The association of polygenic risk for schizophrenia, bipolar disorder, and depression with neural connectivity in adolescents and young adults: examining developmental and sex differences.

Reference – PRS methods with ‘famous’ acronyms

- ▶ Vilhjalmsson et al.(2015). *American Journal of Human Genetics* Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. (LDpred)
- ▶ Prive et al.(2020). *Bioinformatics*. LDpred2: better, faster, stronger. (LDpred2)
- ▶ Euesden et al. (2015). *Bioinformatics* PRSice: Polygenic Risk Score software. (PRSice)
- ▶ Choi and O'Reilly (2019). *Giga Science*. PRSice-2: Polygenic Risk Score software for biobank-scale data. (PRSice-2)
- ▶ Mak et al. (2017). *Genetic Epidemiology*. Polygenic scores via penalized regression on summary statistics. (lassosum).
- ▶ Turley et al. (2018). *Nature Genetics*. Multi-trait analysis of genome-wide association summary statistics using MTAG. (MTAG)

Reference – PRS methods on J and Σ (LD)

- ▶ International Schizophrenia Consortium; Purcell et al. (2009). *Nature*. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. (Pruning and Thresholding; P+T)
- ▶ Goldstein et al. (2015). *Genetic Epidemiology*. Contemporary Considerations for Constructing a Genetic Risk Score: An Empirical Approach.
- ▶ Prive (2019). *American Journal of Human Genetics*. Making the Most of Clumping and Thresholding for Polygenic Scores. (Stacked Clumping + Thresholding)

Reference – PRS methods on J and Σ (LD) and $\hat{\beta}_j$

- ▶ de Vlaming and Groenen (2015). *BioMed Research International*. The Current and Future Use of Ridge Regression for Prediction in Quantitative Genetics.
- ▶ Shi et al. (2016). *PLoS Genetics*. Winner's Curse Correction and Variable Thresholding Improve Performance of Polygenic Risk Modeling Based on Genome-Wide Association Study Summary-Level Data.
- ▶ Mak et al. (2016). *Behavior Genetics*. Local True Discovery Rate Weighted Polygenic Scores Using GWAS Summary Data.
- ▶ Mak et al. (2017). *Genetic Epidemiology*. Polygenic scores via penalized regression on summary statistics. (lassosum).
- ▶ Zhang et al. (2018). *Nature Genetics*. Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies across 32 complex traits.

Reference – PRS Bayesian methods

- ▶ Vilhjalmsson et al.(2015). *American Journal of Human Genetics*. Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. (LDpred)
- ▶ So and Sham (2017). *Scientific Reports*. Improving polygenic risk prediction from summary statistics by an empirical Bayes approach.
- ▶ Zeng and Zhou (2017). *Nature Communications*. Non-parametric genetic prediction of complex traits with latent Dirichlet process regression models.
- ▶ Zhu and Stephens (2017). *Annals of Applied Statistics*. Bayesian large-scale multiple regression with summary statistics from genome-wide association studies. (RSS)
- ▶ Ge et al. (2019). *Nature Communications*. Polygenic prediction via Bayesian regression and continuous shrinkage priors. (PRS-CS)

Reference – PRS Bayesian methods, cont'd

- ▶ Lloyd-Jones et al. (2019). *Nature Communications*. Improved polygenic prediction by Bayesian multiple regression on summary statistics. (SBayesR)
- ▶ Newcombe (2019). *Genetic Epidemiology*. A flexible and parallelizable approach to genome-wide polygenic risk scores. (JAMPred).
- ▶ Song et al. (2020). *PLoS Computational Biology*. Leveraging effect size distributions to improve polygenic risk scores derived from summary statistics of genome-wide association studies. (EB-PRS;Empirical Bayes)
- ▶ Yang et al. (2020). *The American Journal of Human Genetics*. Accurate and Scalable Construction of Polygenic Scores in Large Biobank Data Sets. (DBSLMM)

Reference – PRS methods on data integration – diverse populations

- ▶ Marquez-Luna et al. (2017). *Genetic Epidemiology*. Multiethnic polygenic risk scores improve risk prediction in diverse populations.
- ▶ Duncan et al. (2019). *Nature Communications*. Analysis of polygenic risk score usage and performance in diverse human populations.
- ▶ Grinde et al. (2019). *Genetic Epidemiology*. Generalizing polygenic risk scores from Europeans to Hispanics/Latinos.
- ▶ Gola et al. (2020). *Circulation: Genomic and Precision Medicine*. Population Bias in Polygenic Risk Prediction Models for Coronary Artery Disease.
- ▶ Cavazos and Witte (2021). *Human Genetics and Genomics Advances*. Inclusion of variants discovered from diverse populations improves polygenic risk score transferability.
- ▶ Atkinson et al. (2021). *Nature Genetics*. Tractor uses local ancestry to enable the inclusion of admixed individuals in GWAS and to boost power (from global to local ancestry and admixture.)

Reference – PRS methods on data integration – pleiotropy

- ▶ Hu et al. (2017). *PLoS Genetics*. Joint modeling of genetically correlated diseases and functional annotations increases accuracy of polygenic risk prediction. (PleioPred)
- ▶ Fritzsche et al. (2018). *The American Journal of Human Genetics*. Association of Polygenic Risk Scores for Multiple Cancers in a Phenome-wide Study: Results from The Michigan Genomics Initiative.
- ▶ Krapohl (2018). *Molecular Psychiatry*. Multi-polygenic score approach to trait prediction. (MPS)
- ▶ Maier (2018). *Nature Communications*. Improving genetic prediction by leveraging genetic correlations among human diseases and traits. (SMTpred)
- ▶ Turley et al. (2018). *Nature Genetics*. Multi-trait analysis of genome-wide association summary statistics using MTAG. (MTAG)
- ▶ Chung et al. (2019). *Nature Communications*. Efficient cross-trait penalized regression increases prediction accuracy in large cohorts using secondary phenotypes. (CTPR)
- ▶ Chen et al. (2021). *Journal of the American Statistical Association*. A Penalized Regression Framework for Building Polygenic Risk Models Based on Summary Statistics From Genome-Wide Association Studies and Incorporating External Information. (PANPRS; also annotation)

Reference – PRS methods on data integration – others (e.g. gene-expression, functional annotation)

- ▶ Hasin et al. (2017). *Genome Biology*. Multi-omics approaches to disease.
- ▶ Hu et al. (2017). *PLoS Computational Biology*. Leveraging functional annotations in genetic risk prediction for human complex diseases. (AnnoPred)
- ▶ Marigorta et al. (2017). *Nature Genetics*. Transcriptional risk scores link GWAS to eQTLs and predict complications in Crohn's disease.
- ▶ Pare et al. (2017) *Scientific Reports*. A machine-learning heuristic to improve gene score prediction of polygenic traits. (GraBLD)
- ▶ Wu and Pan. (2018). *Genetic Epidemiology*. Integrating eQTL data with GWAS summary statistics in pathway-based analysis with application to schizophrenia.
- ▶ Gusev et al. (2019). *Nature Genetics*. Transcriptome-wide association study of schizophrenia and chromatin activity yields mechanistic disease insights.

Reference – PRS Other Topics

- ▶ Weiner et al. (2017). *Nature Genetics*. Polygenic transmission disequilibrium confirms that common and rare variation act additively to create risk for autism spectrum disorders. (pTDT)
- ▶ Meisner et al. (2019) *American Journal of Epidemiology*. Case-Only Analysis of Gene-Environment Interactions Using Polygenic Risk Scores. (Interaction)
- ▶ Deng et al. (2019). *Genetic Epidemiology*. Analytical strategies to include the X-chromosome in variance heterogeneity analyses: Evidence for trait-specific polygenic variance structure. (Xchr & Interaction)
- ▶ Yonova-Doing et al. (2021). *Nature Genetics*. An atlas of mitochondrial DNA genotype– phenotype associations in the UK Biobank (mitochondrial DNA)
- ▶ Andrews et al. (2020). *Neurobiology Aging*. Mitonuclear interactions influence Alzheimer's disease risk (mito-nuclear Interaction)

A recent method comparison study **in the context of psychiatric disorders**

Ni et al. (2021). *Biological Psychiatry. A Comparison of Ten Polygenic Score Methods for Psychiatric Disorders Applied Across Multiple Cohorts.*

Methods: The Psychiatric Genomics Consortium Working Groups for schizophrenia and major depressive disorder bring together many independently collected case-control cohorts. We used these resources (31,328 schizophrenia cases, 41,191 controls; 248,750 major depressive disorder cases, 563,184 controls) in repeated application of leave-one-cohort-out meta-analyses, each used to calculate and evaluate PGS in the left-out (target) cohort.

Ten PGS methods (the baseline PC+T method and 9 methods that model genetic architecture more formally: SBLUP, LDpred2-Inf, LDpred-funct, LDpred2, Lassosum, PRS-CS, PRS-CS-auto, SBayesR, MegaPRS) were compared.

Conclusions: Although the methods that more formally model genetic architecture have similar performance, . . .

Personal opinion: no UMP tests/methods for complex problems.

Reference – PRS ‘bad news’ (= more opportunities)

- ▶ Martin et al. (2019). *Nature Genetics*. Clinical use of current polygenic risk scores may exacerbate health disparities
- ▶ Mostafavi (2020). *eLIFE*. Variable prediction accuracy of polygenic scores within an ancestry group
- ▶ Gola et al. (2020). *Circulation: Genomic and Precision Medicine*. Population Bias in Polygenic Risk Prediction Models for Coronary Artery Disease.
- ▶ Wang et al. (2020). *Nature Communication*. Theoretical and empirical quantification of the accuracy of polygenic scores in ancestry divergent populations.
- ▶ Wand et al. (2021). *Nature*. Improving reporting standards for polygenic scores in risk prediction studies

Recap the goal of this lecture

- ▶ PRS is intuitive with a simple expression:

$$PRS_i = \sum_{j=1}^J \hat{\beta}_j G_{ij}$$

- ▶ But, there are many (answered or open) complex (methodological and practical) questions.
- ▶ Provide a topic-stratified reference list for self-study.

What's next?

Overview of the topics to be covered and learning goal: a deeper understanding of

1. PRS foundation: GWAS, h² and prediction

- ▶ the multiple hypothesis testing issue inherent in GWAS
- ▶ the (high) variability inherent in the h² estimates
- ▶ h² as a function of both genetic effect beta and MAF
- ▶ the 'genetic effect size' of a SNP as a function of beta and MAF
- ▶ a conceptual PRS construction based on the ground truth, PRS.oracle
- ▶ DIY ROC plotting and AUC calculation for a PRS-based prediction

2. PRS basic: PRS calculation and performance evaluation

- ▶ the complexity of constructing a good PRS even under the simplest setting without LD or any heterogeneities; 10 out 5000 independent SNPs are truly associated with the same effect size of 0.3 but varying MAFs.
- ▶ the trouble introduced by false positives, due to multiple hypothesis testing and low power.
- ▶ 'the more is not always better' statement: PRS based on 6 'genome-wide' significant SNPs vs. 66 SNPs significant at 0.01.
- ▶ the various over-fitting or selection biases, and winner's curse in beta estimates for both false positives and true positives.

Learning goal cont'd, a **deeper** understanding of

3. PRS basic-plus: some obvious or not so obvious follow-up Qs

- ▶ Effects of ex.nsample and ex.beta.true on AUC: easy to answer.
- ▶ Answers to these Qs are less obvious: **If we decrease ex.beta.true from 0.3 to 0.1 but increase ex.nsnp.true from 10 to 90, h^2 and SNP h^2 ? AUC in general? AUC between PRS.gw and PRS.01?**

4. PRS heterogeneity and transportability

- ▶ First, why reference allele (genome build) matching is so consequential
- ▶ Then, population and locus heterogeneity including
 - my.maf \neq ex.maf
 - my.beta.true \neq ex.beta.true
 - my.nsnp.true \neq ex.nsnp.true

5. PRS LD consideration

- ▶ Some basic understanding of our **limited understanding of LD**.

Ready for the ride? (Hopefully a fun one!)