University of Colorado
Anschutz Medical Campus

Polygenic Risk Scores of COVID-19

Christopher H. Arehart on behalf of CCPM COVID-19 working group
University of Colorado Anschutz Medical Campus, Aurora, CO

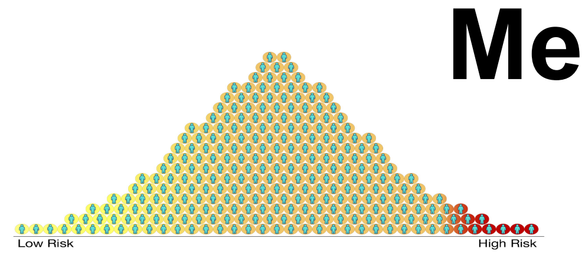
@iconicomics

SCHOOL OF MEDICINE
Biomedical Informatics
& Personalized Medicine
UNIVERSITY OF COLORADO ANSCHUTZ MEDICAL CAMPUS

Background

- COVID-19 symptoms are highly variable
- 3p21.31 gene cluster has been identified as a key susceptibility locus in patients with COVID-19.
- UK Biobank PRS AUC = 0.51 (Toh et al., Hum Genomics 2020).
- Individuals of some ethnic groups (Black, Asian and Hispanic) are more likely to be infected with SARS-CoV-2, compared to those of White ethnicity (Sze et al., EClinicalMedicine 2020).
- PRSs might provide insight into genetic risk differences within ancestry groups.

Methods



$$PRS_i = \sum_{j}^{M} \hat{\beta}_j \times dosage_{ij}$$

Training Dataset

- GWAS: Covid vs. population
- $N_{CASES} = 36{,}590$
- $N_{CONTROLS} = 1{,}668{,}938$

Genotyped Validation Dataset

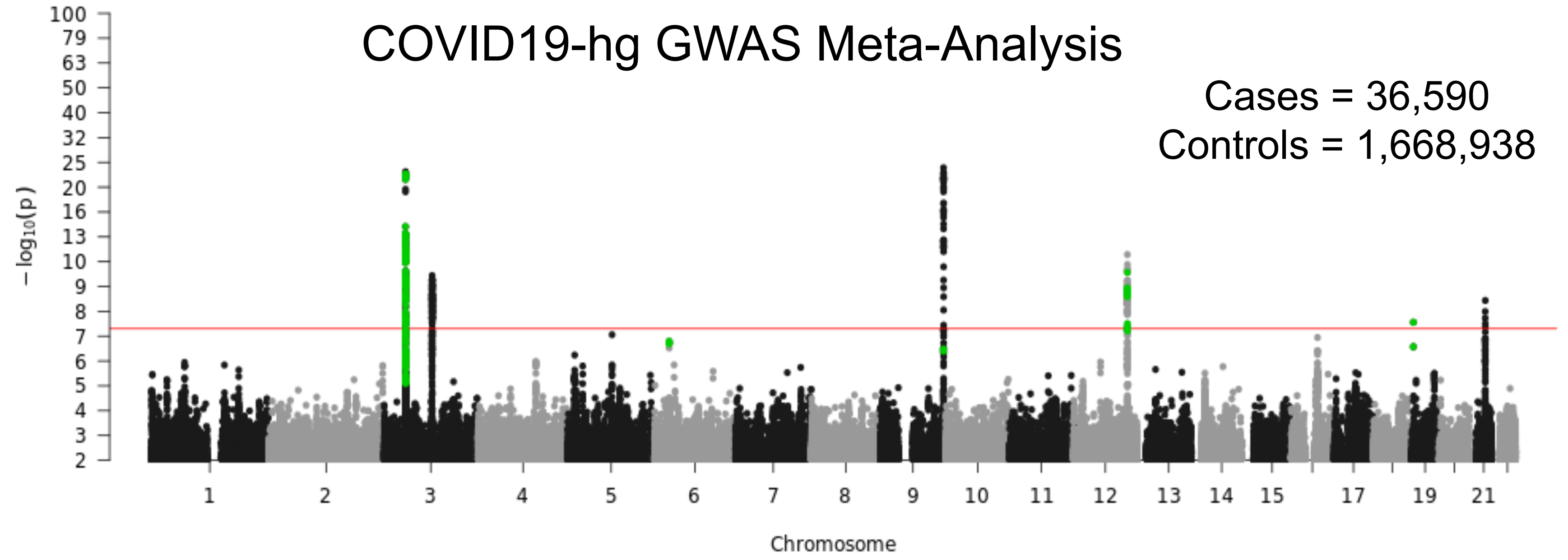
- CCPM Biobank Freeze 2: N ≅ 34,000
- MEGA Chip Imputed with TOPMED

PRS Model

- Genetic Variant and individual QC
- QC between datasets
- Account for LD and compute h^2
- LDpred2 auto
- LDpred2 infinitesimal

Results

COVID19-hg GWAS Meta-Analysis

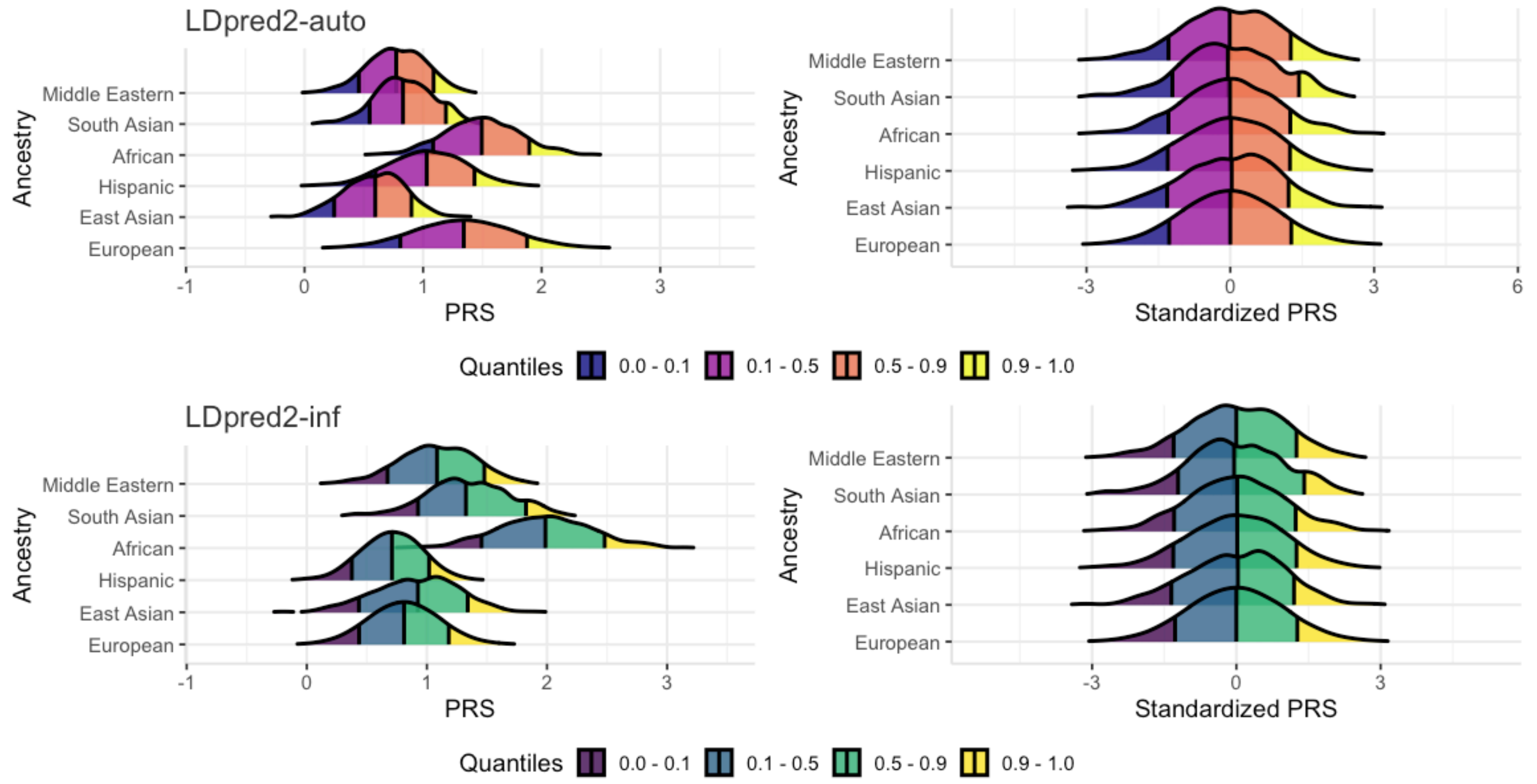


Cases = 36,590
Controls = 1,668,938

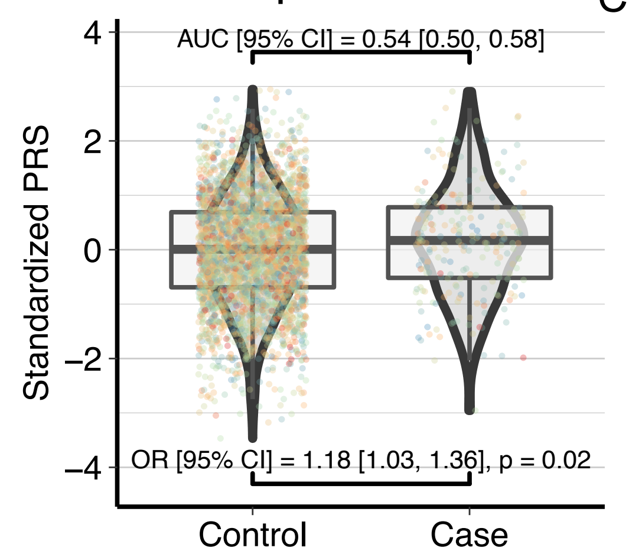
CCPM Biobank Validation Dataset

Ancestry	≅ N
MLE (Middle Eastern)	1,000
SAS (South Asian)	400
AFR (African)	1,400
AMR (Admixed American)	3,500
EAS (East Asian)	850
EUR (European)	26,500

Unstandardized PRS Distributions Differ by Ancestry

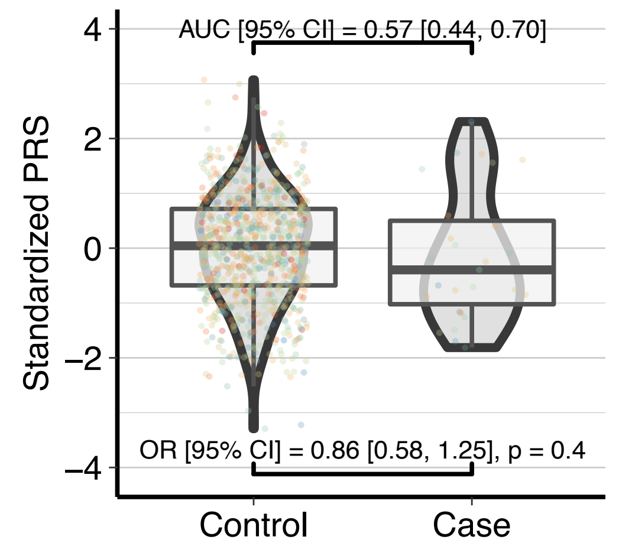


AMR LDpred2-inf



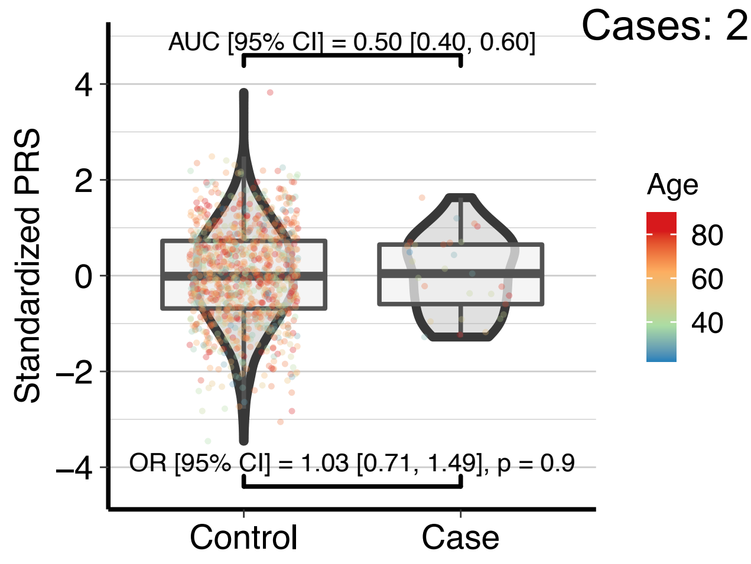
Controls: 3,175
Cases: 215

EAS LDpred2-auto



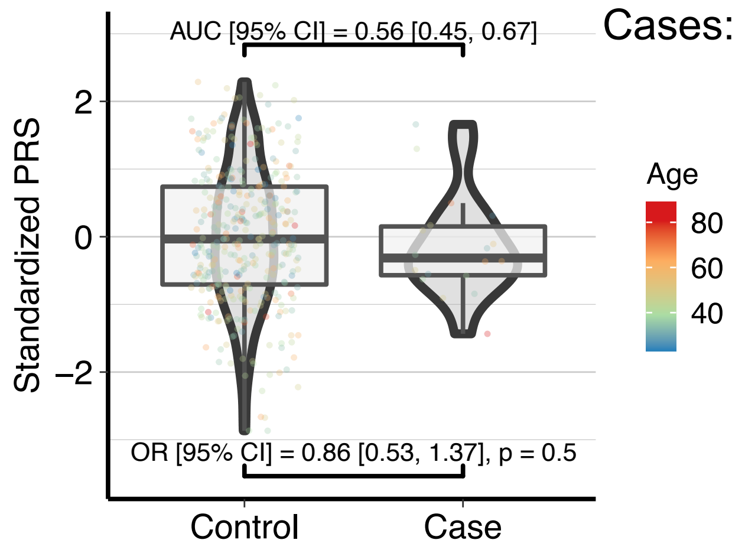
Controls: 817
Cases: 27

MLE LDpred2-auto



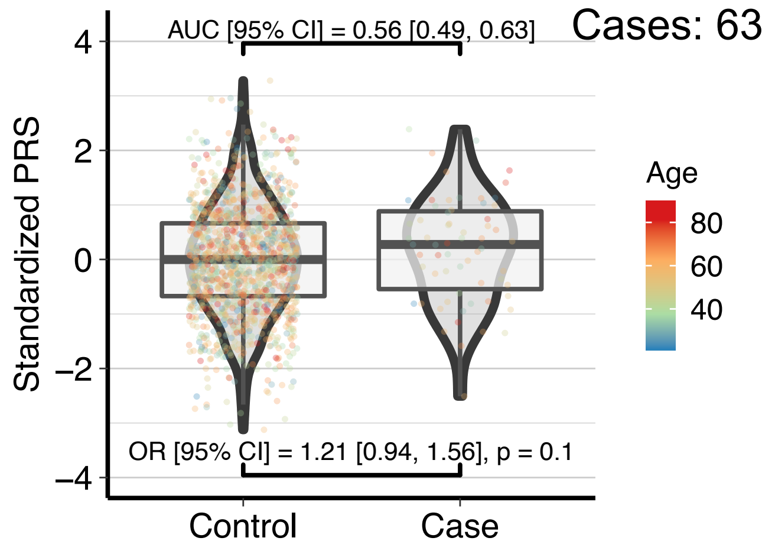
Controls: 973
Cases: 29

SAS LDpred2-auto



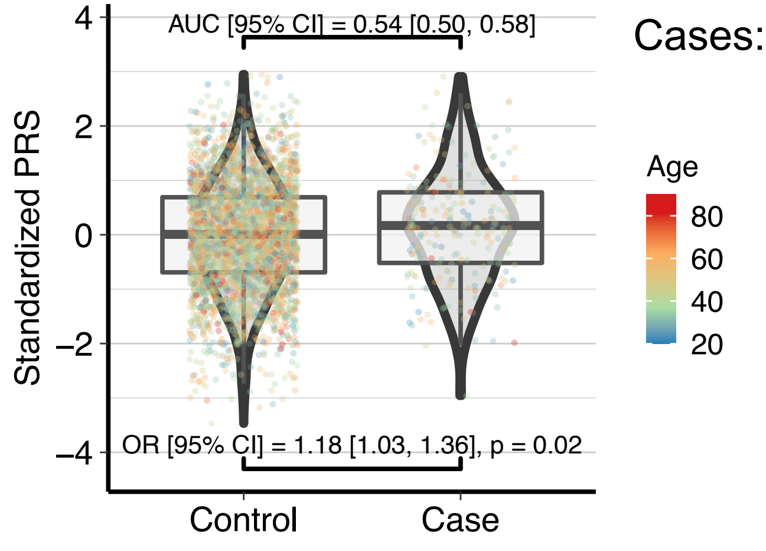
Controls: 354
Cases: 18

AFR LDpred2-inf



Controls: 1,274
Cases: 63

EUR LDpred2-inf



Controls: 25,388
Cases: 848

Conclusions

- PRSs for Admixed Americans (AMR) were predictive for COVID-19:
  - AUC = 0.54 [0.50, 0.58]
  - OR = 1.18 [1.03, 1.36], p = 0.02
- Many of the ancestry groups lacked power due to limited number of COVID-19 cases in the genotyped Freeze 2 cohort (AFR PRS p = 0.1).
- PRS models for MLE, SAS, AFR, EAS, and EUR did not have predictive utility.
- Results were generally consistent when using COVID-19 severity phenotypes.

Implications

- PRSs will likely improve across the board with increased sample sizes (more genotyped individuals) in the CCPM Biobank and increased representation of non-European individuals in the training GWAS.
- Understanding the underlying genetic architecture of COVID-19 susceptibility may improve our understanding of the wide variation in disease severity.
- Next Step: PRS PheWAS

Disclosures

- We would like to thank all study participants and have no relevant disclosures