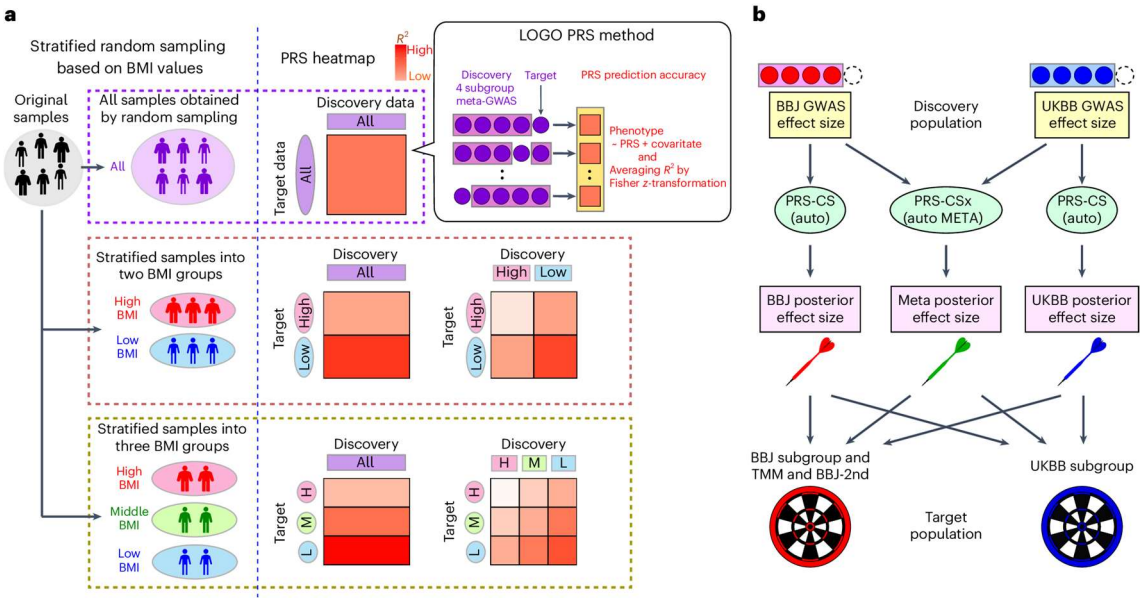
Body mass index stratification optimizes polygenic prediction of type 2 diabetes in cross-biobank analyses

***Nature Gnetics* 56**, 1100–1109 (2024).
https://doi.org/10.1038/s41588-024-01782-y

Takafumi Ojima, Shinichi Namba, Ken Suzuki, Kenichi Yamamoto, Kyuto Sonehara, Akira Narita, the Tohoku Medical Megabank Project Study Group, the Biobank Japan Project, Yoichiro Kamatani, Gen Tamiya, Masayuki Yamamoto, Toshimasa Yamauchi, Takashi Kadowaki, Yukinori Okada

Type 2 diabetes (T2D) shows heterogeneous body mass index (BMI) sensitivity. Here, we performed stratification based on BMI to optimize predictions for BMI-related diseases. We obtained BMI-stratified datasets using data from more than 195,000 individuals (nT2D = 55,284) from BioBank Japan (BBJ) and UK Biobank. T2D heritability in the low-BMI group was greater than that in the high-BMI group. Polygenic predictions of T2D toward low-BMI targets had pseudo-R2 values that were more than 22% higher than BMI-unstratified targets. Polygenic risk scores (PRSs) from low-BMI discovery outperformed PRSs from high BMI, while PRSs from BMI-unstratified discovery performed best. Pathway-specific PRSs demonstrated the biological contributions of pathogenic pathways. Low-BMI T2D cases showed higher rates of neuropathy and retinopathy. Combining BMI stratification and a method integrating cross-population effects, T2D predictions showed greater than 37% improvements over unstratified-matched-population prediction. We replicated findings in the Tohoku Medical Megabank (n = 26,000) and the second BBJ cohort (n = 33,096). Our findings suggest that target stratification based on existing traits can improve the polygenic prediction of heterogeneous diseases.


a
Stratified random sampling based on BMI values
PRS heatmap
High
Low
LOGO PRS method
Original samples
All samples obtained by random sampling
All
Discovery data
All
Target data
All
Discovery 4 subgroup meta-GWAS
Target
PRS prediction accuracy
Phenotype ~ PRS + covariate and Averaging R^2 by Fisher z-transformation
Stratified samples into two BMI groups
High BMI
Low BMI
Discovery
All
Target
High
Low
Discovery
High
Low
Target
High
Low
Stratified samples into three BMI groups
High BMI
Middle BMI
Low BMI
Discovery
All
Target
H
M
L
Discovery
H
M
L
Target
H
M
L
b
BBJ GWAS effect size
Discovery population
UKBB GWAS effect size
PRS-CS (auto)
PRS-CSx (auto META)
PRS-CS (auto)
BBJ posterior effect size
Meta posterior effect size
UKBB posterior effect size
BBJ subgroup and TMM and BBJ-2nd
Target population
UKBB subgroup