* Variant information necessary to apply the PGS to new samples (variant rsID and/or genomic position, weights/effect sizes, effect allele, genome build).
	+ genome build : GRCh37
* Information about how the PGS was developed (computational method, variant selection, relevant parameters).
	+ computational method : PRS-CSx
	+ variant selection : We used all HapMap3 variants which are remained in variant QC in discovery cohort (BBJ and UKBB) and test cohort (TMM or BBJ2nd).
	+ relevant parameters : phi='auto', --meta=TRUE
* Descriptions of the samples used to develop the score (e.g. discovery of the variant associations [*GWAS samples, which can usually be extracted directly from the GWAS Catalog using GCST IDs*], as well as other samples used to develop/train the PGS) and external evaluation.
	+ In discovery GWAS, we performed T2D-GWAS for BMI-stratified datasets from BioBank Japan and UK Biobank using plink2. For BMI stratification, we conducted stratified random sampling to keep case-control ratio in each biobank.
	+ Sample numbers of discovery GWAS are described below.

|  |  |
| --- | --- |
| **Number of****BMI-stratified groups** | **type 2 diabetes** |
| **Male****Case** | **Male****Control** | **Female****Case** | **Female****Control** | **Total** |
| **All samples** | 35676 | 90660 | 19608 | 49824 | 195768 |
| **Two BMI groups** | 17838 | 45330 | 9804 | 24912 | 97884 |
| **Three BMI groups** | 11892 | 30220 | 6536 | 16608 | 65256 |

* + For test dataset for polygenic prediction, we used two independent biobanks, Tohoku Medical Megabank (TMM, *N*T2D=6,000, *N*control=20,000) or the second cohort of BioBank Japan (BBJ-2nd, *N*T2D=11,236, *N*control=21,860).
* Establishment of the PGS' analytic validity, and a description of its predictive performance (e.g. effect sizes [beta, OR, HR, etc.], classification accuracy, proportion of the variance explained (R2), and any covariates evaluated in the PGS prediction).
	+ We evaluated PGS by AUROC and liability R2.
	+ We included age, age2, sex, sex × age, sex × age2 and the top 20 principal components as covariates.