

FEV1

Description

UK Biobank Population

The data release contains genotypes of 488,377 UK Biobank participants, and 825,927 markers are genotyped.

Quality control

The following QC steps are conducted

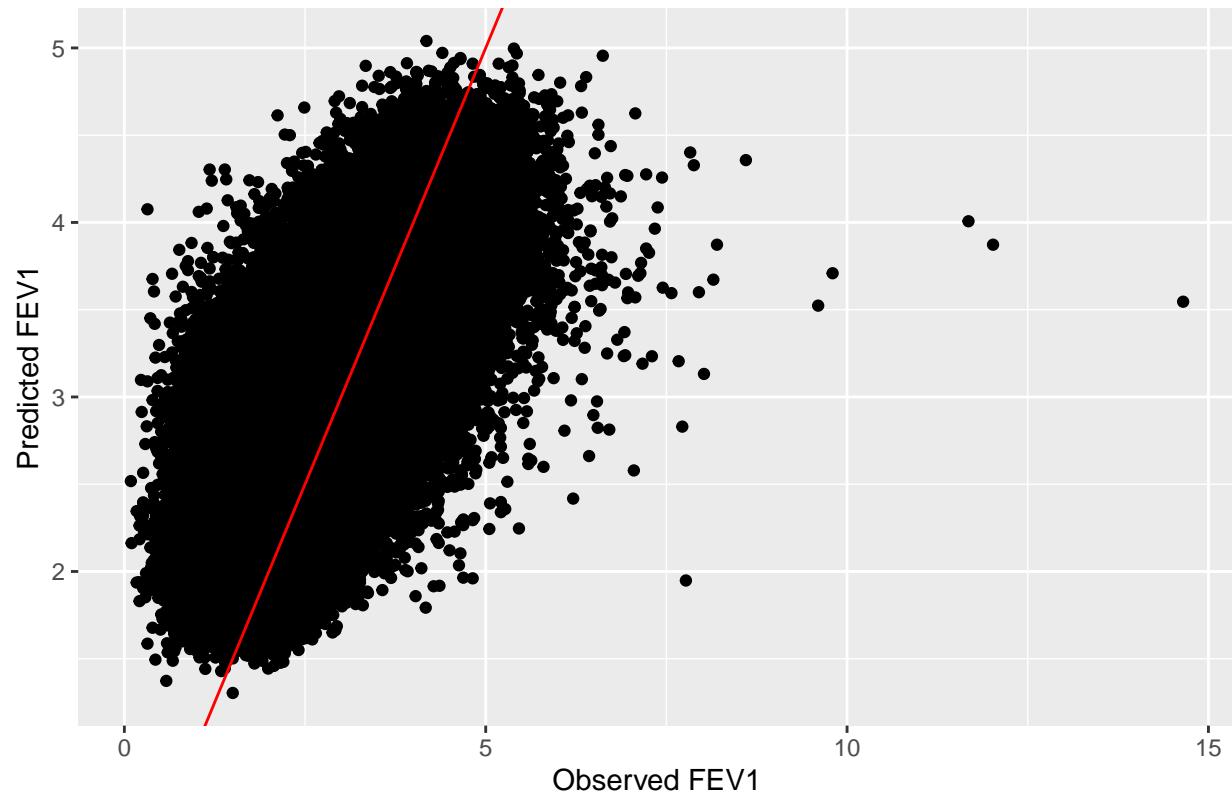
- Exclude individuals with > 10% missing genotypes.
- Exclude SNPs with maf < 0.05
- Include only SNPs with a 90% genotyping rate
- Exclude markers that fail the Hardy-Weinberg test at 0.001
- Include only White-British population (Code 1001 in Data field 21000, self reported)
- LD pruning. Window size = 50 SNPs, pairwise LD > 0.5 is removed.
- Exclude one member of each pair of samples with kinship coefficient greater than 0.117. (First degree relatives are removed)

After QC steps and removing individuals with missing age, sex, or phenotype height, 408207 samples and 211,587 genetic markers remained.

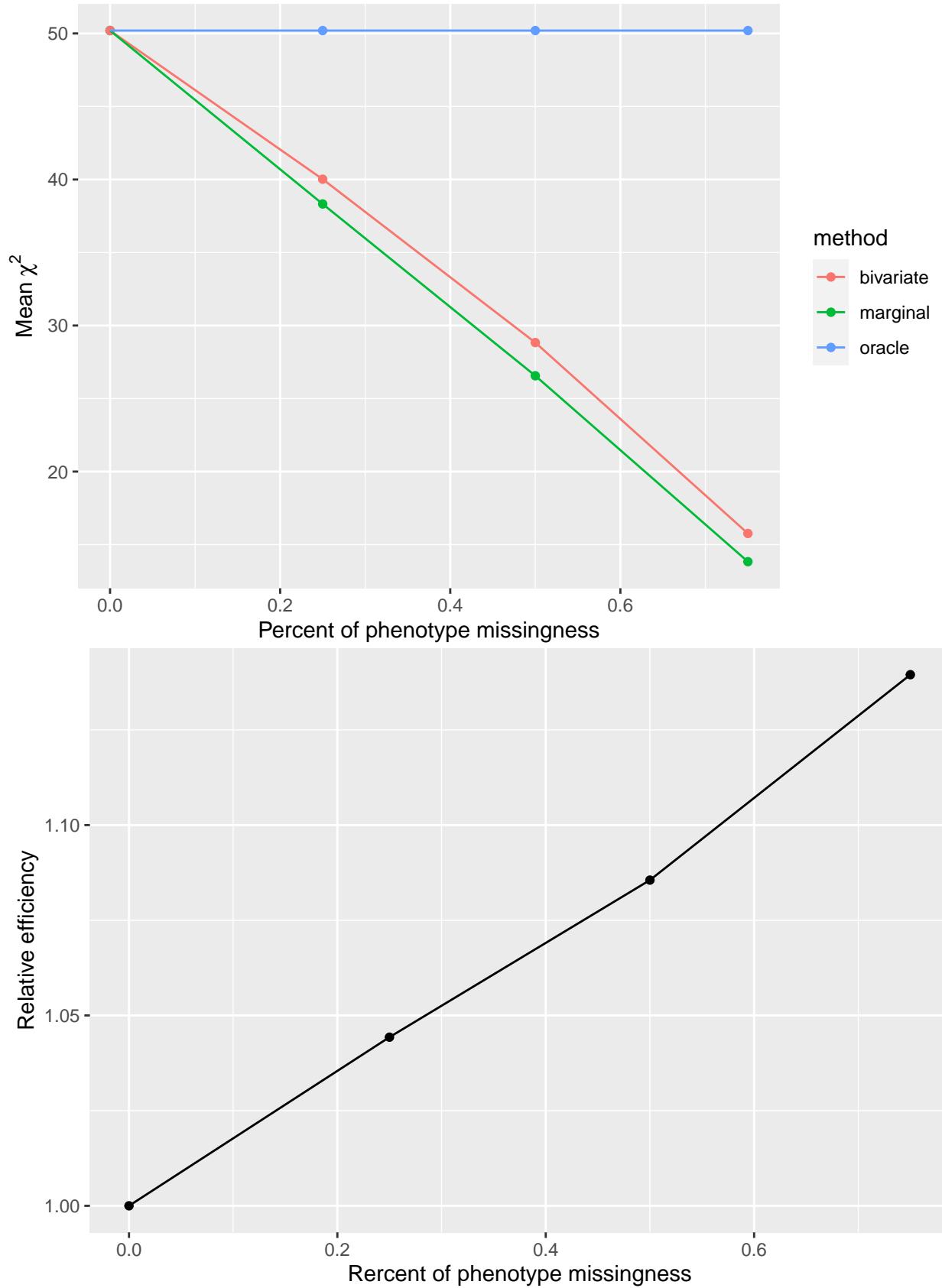
FEV1 Prediction

In total, 25697 individuals were removed in the association analysis due to being first-degree relatives with other people. After excluding people with missing weight, 23030 remaining individuals served as training set. The covariates are age, sex, waist circumference, hip circumference, height, smoking status, BMI, basal metabolic rate and impedance of whole body.

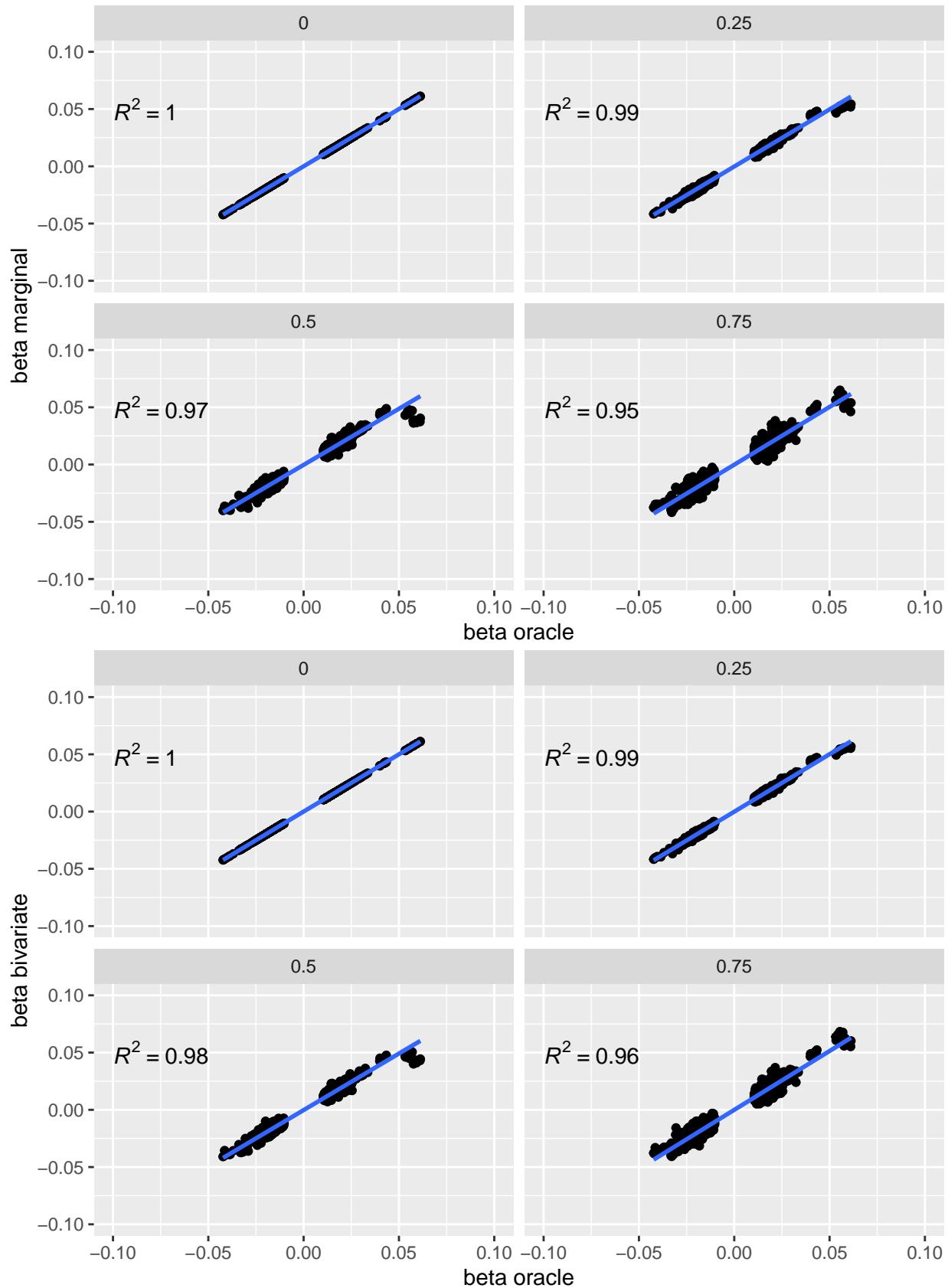
Random Forest Prediction



Relative efficiency



Estimated Effect Size



FN and FP

missing	FNR.marginal	FPR.marginal	FNR.binormal	FPR.binormal
0.00	0.0000000	0.00e+00	0.0000000	0.00e+00
0.25	0.4357213	7.60e-05	0.3846909	7.60e-05
0.50	0.7369971	2.85e-05	0.6947988	4.27e-05
0.75	0.9342493	4.70e-06	0.9224730	4.70e-06