**Sample-level filtering**

Initial sample

n = 487'409

Sample after call-rate filtering

n = 429'874

Sample after filtering based on UK Biobank data-fields

n = 429'874

Ambiguous sex

n = 85

Data-fields filtering

n = 57’450

Call-rate < 95%

n = 0

Sample after ambiguous sex filtering

n = 487'324

**SNP-level filtering**

Initial SNPs

p = 730’150

Duplicates

p = 2'145

Call-rate < 95%

p = 0

MAF < 1%

r2 < 0.04

HWE > 10-6

p = 665’268

SNPs after call-rate filtering

p = 728'005

SNPs after removing duplicates

p = 728'005

SNPs after LD pruning, MAF and HWE filtering

p = 62’737