**LD score regression**

**Case-control phenotypes**

for dx in afib cad celiac gallstone polyp t2d varicose vte; do /project\_root/ldsc/munge\_sumstats.py \

--sumstats /project\_root/ML\_GWAS/Step2\_Anc\_Comb\_LDSC/cases\_eur\_${dx}.txt \

--out /project\_root/ML\_GWAS/Step2\_Anc\_Comb\_LDSC/cases\_eur\_${dx}\_munged \

--merge-alleles /project\_root/ML\_GWAS/eur\_w\_ld\_chr/w\_hm3.snplist \

--chunksize 500000; done

for dx in afib cad celiac gallstone polyp t2d varicose vte; do /project\_root/ldsc/ldsc.py \

--h2 /project\_root/ML\_GWAS/Step2\_Anc\_Comb\_LDSC/cases\_eur\_${dx}\_munged.sumstats.gz \

--ref-ld-chr /project\_root/ML\_GWAS/eur\_w\_ld\_chr/ \

--w-ld-chr /project\_root/ML\_GWAS/eur\_w\_ld\_chr/ \

--out /project\_root/ML\_GWAS/Step2\_Anc\_Comb\_LDSC/cases\_eur\_${dx}\_ldsc; done

**Predicted phenotypes**

for dx in cad celiac gallstone polyp t2d varicose vte; do /project\_root/ldsc/munge\_sumstats.py \

--sumstats /project\_root/ML\_GWAS/Step2\_Anc\_Comb\_LDSC/preds\_eur\_${dx}.txt \

--out /project\_root/ML\_GWAS/Step2\_Anc\_Comb\_LDSC/preds\_eur\_${dx}\_munged \

--merge-alleles /project\_root/ML\_GWAS/eur\_w\_ld\_chr/w\_hm3.snplist \

--chunksize 500000; done

for dx in afib cad celiac gallstone polyp t2d varicose vte; do /project\_root/ldsc/ldsc.py \

--h2 /project\_root/ML\_GWAS/Step2\_Anc\_Comb\_LDSC/preds\_eur\_${dx}\_munged.sumstats.gz \

--ref-ld-chr /project\_root/ML\_GWAS/eur\_w\_ld\_chr/ \

--w-ld-chr /project\_root/ML\_GWAS/eur\_w\_ld\_chr/ \

--out /project\_root/ML\_GWAS/Step2\_Anc\_Comb\_LDSC/preds\_eur\_${dx}\_ldsc; done

**Calculate genetic correlations**

for dx in afib cad celiac gallstone polyp t2d varicose vte; do /project\_root/ldsc/ldsc.py \

--rg /project\_root/ML\_GWAS/Step2\_Anc\_Comb\_LDSC/preds\_eur\_${dx}\_munged.sumstats.gz,/project\_root/ML\_GWAS/Step2\_Anc\_Comb\_LDSC/cases\_eur\_${dx}\_munged.sumstats.gz \

--ref-ld-chr /project\_root/ML\_GWAS/eur\_w\_ld\_chr/ \

--w-ld-chr /project\_root/ML\_GWAS/eur\_w\_ld\_chr/ \

--out /project\_root/ML\_GWAS/Step2\_Anc\_Comb\_LDSC/both\_eur\_${dx}\_ldsc; done

**MTAG**

**Convert Regenie outputs into MTAG-compatible summary statistics**

for dx in ['afib','cad','celiac','gallstone','polyp','t2d','varicose','vte']:

for anc in ['eur']:

df = pd.read\_csv(f'./ML\_GWAS/Step2\_Anc\_Comb/preds\_{anc}\_{dx}.regenie', sep=' ')

df['P'] = 10\*\*(-df['LOG10P'])

df = df[['ID','CHROM','GENPOS','ALLELE1','ALLELE0','A1FREQ','N','BETA','SE','P']]

df = df.set\_axis(['snpid','chr','bpos','a1','a2','freq','n','beta','se','p'],axis=1)

df = df.loc[df['freq'] >= 0.009]

df['z'] = df['beta']/df['se']

df.to\_csv(f'./ML\_GWAS/MTAG/Input/preds\_{anc}\_{dx}.txt', sep='\t', index=False)

case\_proportions = {

'afib': 33909/(33909+391201),

'cad': 47012/(47012+371441),

'celiac': 3810/(3810+421300),

'gallstone': 29929/(29929+395181),

'polyp': 7640/(7640+417470),

't2d': 32691/(32691+386022),

'varicose': 34135/(34135+390975),

'vte': 21887/(21887+398220)

}

for dx in ['afib','cad','celiac','gallstone','polyp','t2d','varicose','vte']:

for anc in ['eur']:

df = pd.read\_csv(f'./ML\_GWAS/Step2\_Anc\_Comb/cases\_{anc}\_{dx}.regenie', sep=' ')

proportion\_cases = case\_proportions[dx]

N\_total = df['N']

N\_cases = N\_total \* proportion\_cases

N\_controls = N\_total \* (1 - proportion\_cases)

effective\_n = 4 \* N\_cases \* N\_controls / (N\_cases + N\_controls)

df['P'] = 10\*\*(-df['LOG10P'])

df = df[['ID', 'CHROM', 'GENPOS', 'ALLELE1', 'ALLELE0', 'A1FREQ', 'N', 'BETA', 'SE', 'P']]

df = df.set\_axis(['snpid', 'chr', 'bpos', 'a1', 'a2', 'freq', 'n', 'beta', 'se', 'p'], axis=1)

df = df.loc[df['freq'] >= 0.009]

df['n'] = effective\_n

df['z'] = df['beta'] / df['se']

df.to\_csv(f'./ML\_GWAS/MTAG/Input/cases\_{anc}\_{dx}.txt', sep='\t', index=False)

**Run MTAG**

for dx in afib cad celiac gallstone polyp t2d varicose vte; do python /project\_root/Software/mtag/mtag.py \

--sumstats /project\_root/ML\_GWAS/MTAG/Input/preds\_eur\_\${dx}.txt,/project\_root/ML\_GWAS/MTAG/Input/cases\_eur\_\${dx}.txt \

--ld\_ref\_panel /project\_root/Software/ukb\_10k\_ld/ld\_eur\_@ \

--out /project\_root/ML\_GWAS/MTAG/Output/mtag\_ukbeur\_\${dx} \

--eaf\_name freq \

--n\_min 0.0 \

--fdr \

--stream\_stdout; done

**Polygenic risk scores**

**Run PRS-CSx for case-control phenotypes**

for pheno in cases; do for dx in afib cad gallstone polyp t2d varicose vte; do for chr in {1..22}; do python /project\_root/PRS/PRScsx/PRScsx.py --ref\_dir=/project\_root/PRS/1KG --bim\_prefix=/project\_root/MR/PRS/biome\_1kg --sst\_file=/project\_root/MR/PRS/Input/${pheno}\_afr\_${dx}.sumstat,/project\_root/MR/PRS/Input/${pheno}\_amr\_${dx}.sumstat,/project\_root/MR/PRS/Input/${pheno}\_eas\_${dx}.sumstat,/project\_root/MR/PRS/Input/${pheno}\_eur\_${dx}.sumstat,/project\_root/MR/PRS/Input/${pheno}\_sas\_${dx}.sumstat --n\_gwas=6802,995,2782,426844,9107 --pop=AFR,AMR,EAS,EUR,SAS --out\_dir=/project\_root/MR/PRS/Output --out\_name=${pheno}\_${dx}\_chr${chr} --chrom=${chr}; done; done; done

*For celiac disease, we exclude AMR and EAS as there are too few cases*

for pheno in cases; do for dx in celiac; do for chr in {1..22}; do python /project\_root/PRS/PRScsx/PRScsx.py --ref\_dir=/project\_root/PRS/1KG --bim\_prefix=/project\_root/MR/PRS/biome\_1kg --sst\_file=/project\_root/MR/PRS/Input/${pheno}\_afr\_${dx}.sumstat,/project\_root/MR/PRS/Input/${pheno}\_eur\_${dx}.sumstat,/project\_root/MR/PRS/Input/${pheno}\_sas\_${dx}.sumstat --n\_gwas=6802,426844,9107 --pop=AFR,EUR,SAS --out\_dir=/project\_root/MR/PRS/Output --out\_name=${pheno}\_${dx}\_chr${chr} --chrom=${chr}; done; done; done

**Run PRS-CSx for predicted phenotypes**

for pheno in preds; do for dx in afib cad celiac gallstone polyp t2d varicose vte; do for chr in {1..22}; do python /project\_root/PRS/PRScsx/PRScsx.py --ref\_dir=/project\_root/PRS/1KG --bim\_prefix=/project\_root/MR/PRS/biome\_1kg --sst\_file=/project\_root/MR/PRS/Input/${pheno}\_afr\_${dx}.sumstat,/project\_root/MR/PRS/Input/${pheno}\_amr\_${dx}.sumstat,/project\_root/MR/PRS/Input/${pheno}\_eas\_${dx}.sumstat,/project\_root/MR/PRS/Input/${pheno}\_eur\_${dx}.sumstat,/project\_root/MR/PRS/Input/${pheno}\_sas\_${dx}.sumstat --n\_gwas=6802,995,2782,426844,9107 --pop=AFR,AMR,EAS,EUR,SAS --out\_dir=/project\_root/MR/PRS/Output --out\_name=${pheno}\_${dx}\_chr${chr} --chrom=${chr}; done; done; done