Robust use of phenotypic heterogeneity at drug target genes for mechanistic insights: application of *cis*-multivariable Mendelian randomization to *GLP1R* gene region

Ashish Patel, Dipender Gill, Jack Bowden, & Stephen Burgess *et al*.

May 2023

Part I: Colocalization analysis of BMI and T2D

We start with a colocalization analysis of BMI and T2D at the *GLP1R* locus. Such an analysis is useful for providing insight on whether the joint genetically-predicted effects of BMI and T2D can be reliably estimated by multivariable Mendelian randomization.

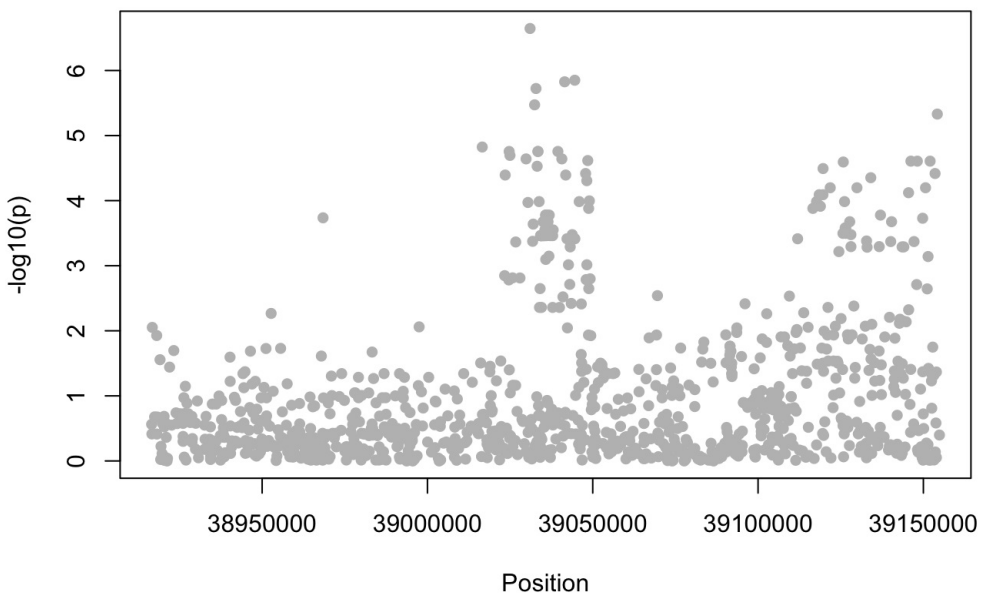
Load the coloc R package and harmonized genetic association data on BMI, T2D, and CAD. Here we also temporarily suppress warnings to keep things tidy. Lets plot the BMI and T2D associations through the relevant coloc functions.

```
default.warning <- getOption("warn"); options(warn=-1)
library(coloc)
```

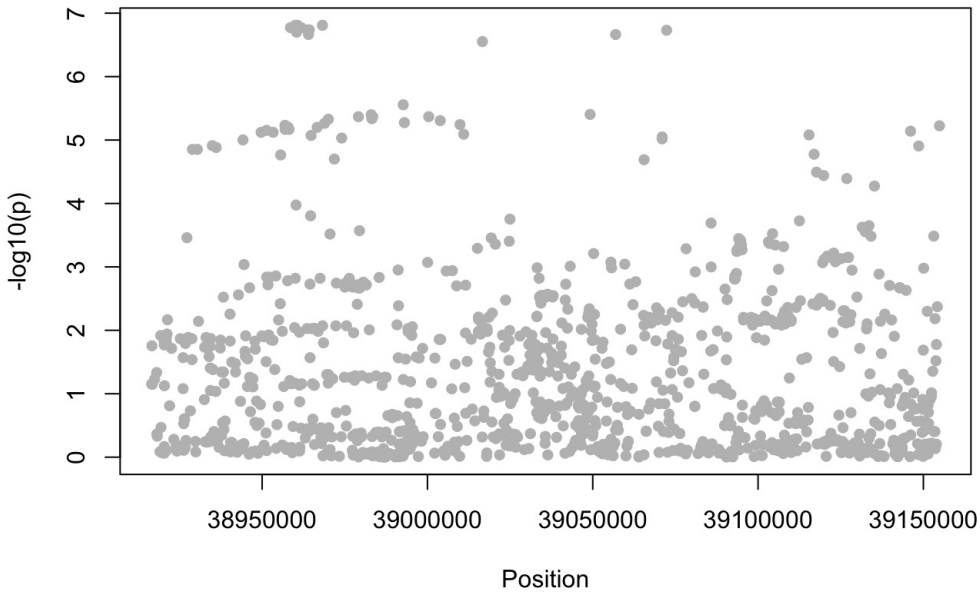
```
## This is a new update to coloc.
```

```
load("bmi_t2d_cad.RData")

# variant associations with BMI
bmi_coloc_dat=data.frame(bmi$Effect,(bmi$StdErr)^2,bmi$SNP,bmi$POS); colnames(bmi_coloc_dat) <- c("beta","varbeta
","snp","position")
bmi_coloc_dat$sdY <- 4.77152 #see https://biobank.ndph.ox.ac.uk/ukb/field.cgi?id=21001
bmi_coloc_dat$type <- "quant"
plot_dataset(bmi_coloc_dat)
```




```
# variant associations with T2D
t2dm_coloc_dat=data.frame(t2dm$Effect,(t2dm$StdErr)^2,t2dm$SNP,t2dm$position); colnames(t2dm_coloc_dat) <- c("bet
a","varbeta","snp","position")
t2dm_coloc_dat$type <- "cc"
plot_dataset(t2dm_coloc_dat)
```



These manhattan plots suggest there are independent genetic predictors of both BMI and T2D. Lets formally test this with coloc.

```
coloc.res <- coloc.abf(dataset1=bmi_coloc_dat,dataset2=t2dm_coloc_dat)
```

```
## PP.H0.abf PP.H1.abf PP.H2.abf PP.H3.abf PP.H4.abf
##   0.01670   0.00557   0.73300   0.24400   0.00080
## [1] "PP abf for shared variant: 0.08%"
```

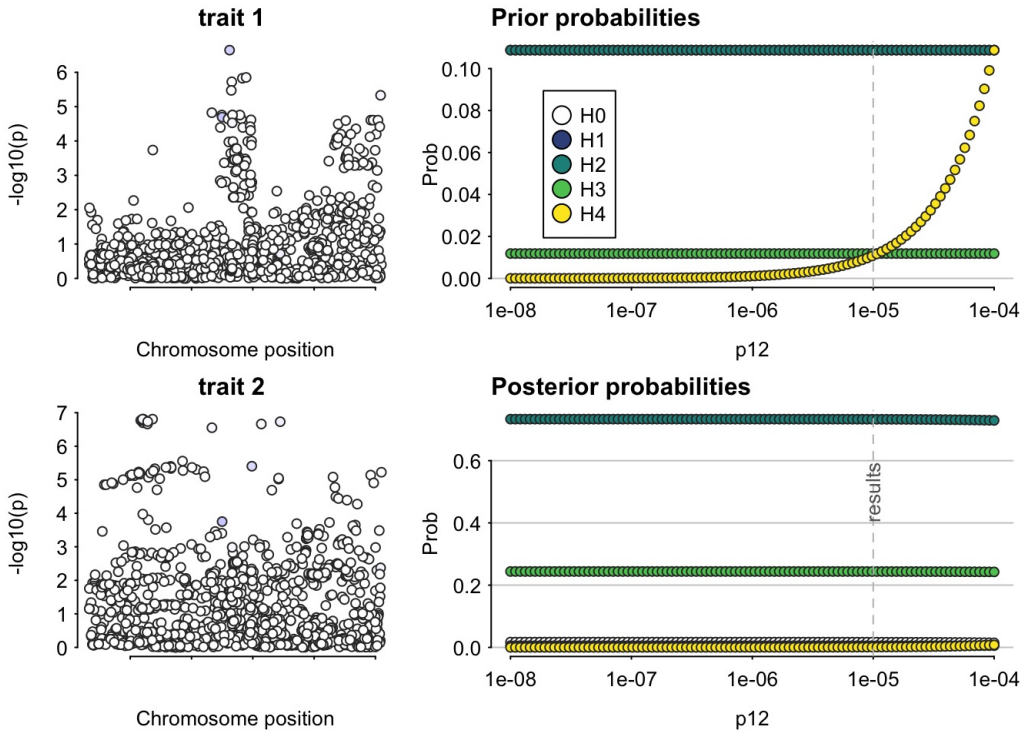
We see that coloc provides no evidence for a shared causal variant for both BMI and T2D at the *GLP1R* locus. At the default prior values, the posterior probability of the shared causal variant hypothesis (H_4) is 0.08%.

The posterior probability that there are distinct causal variants for BMI and T2D conditional on both variants having a causal variant $H_3 / (H_3 + H_4)$ is 99.7%.

Now lets consider a sensitivity analysis to the choice of colocalization prior parameter p_{12}.

```
sensitivity(coloc.res,rule="H4 > 0.5")
```

```
## Results fail decision rule H4 > 0.5
```



This sensitivity analysis indicates little support for the shared causal variant hypothesis at plausible values of the p_{12} parameter. We conclude that our colocalization analysis suggests the presence of phenotypic heterogeneity at the *GLP1R* locus.

Part II: Mendelian randomization analyses for BMI and T2D on CAD risk

Now we conduct robust PCA-GMM estimation of genetically-predicted BMI and T2D effects on CAD risk. Lets start by setting up our inputs. One of the required inputs is cor.x, which is the 2x2 correlation matrix of BMI and T2D. The off-diagonal elements of cor.x are the correlation of BMI and T2D, which is set to 0, but the resulting analysis is quite robust to this choice.

```
by <- cad$Effect; sy <- cad$StdErr; ny <- mean(cad$n); cor.x <- cbind(c(1,0),c(0,1))
bx <- cbind(bmi$Effect,t2dm$Effect); sx <- cbind(bmi$StdErr,t2dm$StdErr); nx <- c(nx.bmi,nx.t2dm)
```

The number of principal components used to instrument BMI and T2D are chosen to explain a certain percentage of variation in a weighted genetic correlation matrix of the *GLP1R* gene region. Lets calculate what numbers of principal components are needed to explain 99.9%, 99%, and 95% of variation.

```
pca.no <- function(k){
  Phi <- ((rowSums(abs(bx))/sy)%*%t(rowSums(abs(bx))/sy))*ld
  return(which(cumsum(prcomp(Phi,scale=FALSE)$sdev^2/sum((prcomp(Phi,scale=FALSE)$sdev^2)))>k)[1])}
pca.thres <- c(0.999,0.99,0.95)
pca.no <- sapply(pca.thres,pca.no)
cbind(pca.thres,pca.no)
```

```
##      pca.thres pca.no
## [1,]     0.999     40
## [2,]     0.990     17
## [3,]     0.950      8
```

We find that the first 40 principal components explain 99.9% of weighted genetic variation in *GLP1R*. Lets apply the robust PCA-GMM method using these 40 principal components as instruments.

```
source("PCA-GMM.R")
pca_res <- function(k){PCA_GMM(bx=bx,sx=sx,by=by,sy=sy,ld=ld,cor.x=cor.x,nx=nx,ny=ny,r=pca.no[k])}

# robust PCA-GMM results using principal components explaining 99.9% of weighted genetic variation
pca_res0 <- pca_res(1)

ci <- cbind(pca_res0$liml,pca_res0$liml - qnorm(1 - 0.05/2)*pca_res0$se.liml,pca_res0$liml + qnorm(1 - 0.05/2)*pc
a_res0$se.liml,2*(1-pnorm(abs(pca_res0$liml/pca_res0$se.liml))),rep(pca.thres[1],2),rep(pca_res0$factors,2),pca_r
es0$condF,(1-pchisq((pca_res0$condF*(pca_res0$factors-ncol(bx)+1)),pca_res0$factors-ncol(bx)+1)))

colnames(ci) <- c("estimate","lower","upper","p-value","prop. var. expl.","PCs", "cond. F-statistic","cond. F p-v
alue"); rownames(ci) <- c("bmi","t2dm")
round(ci,3)
```

```
##      estimate  lower upper p-value prop. var. expl. PCs cond. F-statistic
## bmi     1.470  0.621 2.319   0.001            0.999  40             2.371
## t2dm   -0.051 -0.239 0.136   0.591            0.999  40             3.028
##      cond. F p-value
## bmi                0
## t2dm               0
```

We find that genetically-predicted BMI was associated with CAD risk (log odds ratio estimate per 1 standard deviation increase in BMI 1.470, 95% confidence interval [CI] 0.621, 2.319), but genetically-predicted T2D liability was not (log odds ratio estimate per 1 unit increase in the log odds of T2D risk -0.051, 95% CI -0.239, 0.136).

The conditional F-statistics for BMI and T2D are 2.371 and 3.028. Our estimate of the overdispersion heterogeneity parameter is:

```
# overdispersion heterogeneity parameter estimate
pca_res0$kappa
```

```
## [1] 3.133687
```

Thus, we find evidence of overdispersion heterogeneity, and the robust PCA-GMM confidence intervals above should be wider than the non-robust version.

What are the results of the non-robust version of the PCA-GMM method?

```
# non-robust PCA-GMM results using principal components explaining 99.9% of weighted genetic variation
ci <- cbind(pca_res0$liml.nr,pca_res0$liml.nr - qnorm(1 - 0.05/2)*pca_res0$se.liml.nr,pca_res0$liml.nr + qnorm(1
- 0.05/2)*pca_res0$se.liml.nr,2*(1-pnorm(abs(pca_res0$liml.nr/pca_res0$se.liml.nr))),rep(pca.thres[3],2),rep(pca_
res0$factors,2),rep(1-pchisq(pca_res0$Q.nr,pca_res0$factors-(ncol(bx)+1)),2))

colnames(ci) <- c("estimate","lower","upper","p-value","prop. var. expl.","PCs","J test p-value"); rownames(ci) <
- c("bmi","t2dm")
round(ci,3)
```

```
##      estimate  lower upper p-value prop. var. expl. PCs J test p-value
## bmi     1.484  0.828  2.14   0.000             0.95  40          0.003
## t2dm   -0.139 -0.288  0.01   0.068             0.95  40          0.003
```

We find that the non-robust PCA-GMM estimate of the genetically predicted BMI effect is similar but more precise. Importantly, however, the non-robust model (that does not include an overdispersion heterogeneity parameter) does not pass the heterogeneity test (J test p-value = 0.003) which suggests the model may be mis-specified.

Lets now try robust PCA-GMM estimation using only the first 17 principal components, which as we know from above, explain 99% of weighted genetic variation in *GLPR1*.

```
# robust PCA-GMM results using principal components explaining 99% of weighted genetic variation
pca_res0 <- pca_res(2)

ci <- cbind(pca_res0$liml,pca_res0$liml - qnorm(1 - 0.05/2)*pca_res0$se.liml,pca_res0$liml + qnorm(1 - 0.05/2)*pc
a_res0$se.liml,2*(1-pnorm(abs(pca_res0$liml/pca_res0$se.liml))),rep(pca.thres[2],2),rep(pca_res0$factors,2),pca_r
es0$condF,(1-pchisq((pca_res0$condF*(pca_res0$factors-ncol(bx)+1)),pca_res0$factors-ncol(bx)+1)))
colnames(ci) <- c("estimate","lower","upper","p-value","prop. var. expl.","PCs", "cond. F-statistic","cond. F p-v
alue"); rownames(ci) <- c("bmi","t2dm")
round(ci,3)
```

```
##      estimate  lower upper p-value prop. var. expl. PCs cond. F-statistic
## bmi     2.876  0.618 5.135   0.013             0.99  17             3.271
## t2dm   -0.094 -0.417 0.229   0.569             0.99  17             4.476
##      cond. F p-value
## bmi                0
## t2dm               0
```

The estimates are now less precise, but agree with the previous results using a 99.9% threshold. In particular, the 95% confidence interval for the BMI effect using a 99% threshold overlaps the tighter interval under the 99.9% threshold.

Finally, note that the conditional F-statistics are higher under the 99% threshold compared with the 99.9% threshold. This highlights a trade-off: using many principal components allows us to reliably estimate the overdispersion parameter, and provides a more powerful analysis. However, using a fewer number of relevant principal components should boost conditional F-statistics, potentially allowing for estimates that are less biased.

Part III: Mendelian randomization analyses for tissue-specific *GLP1R* expression on CAD risk

We now investigate the likely tissue at which the effect of GLP1R perturbation on CAD may risk occur. Lets load the harmonized *GLP1R* expression and CAD association data. We set up our inputs for an analysis of 3 tissues relating to brain-caudate, heart-atrial appendage, and pancreas.

```
load("expression.Rdata")
bx.brain <- brain$beta; sx.brain <- brain$se
bx.heart <- heart$beta; sx.heart <- heart$se
bx.pancreas <- pancreas$beta; sx.pancreas <- pancreas$se
bx <- cbind(bx.brain,bx.heart,bx.pancreas); sx <- cbind(sx.brain,sx.heart,sx.pancreas)
by <- cad$Effect; sy <- cad$StdErr
nx <- rep(838,ncol(bx)); ny <- mean(cad$n)
cor.x <- diag(ncol(bx)) # unconditional correlation between exposures set to 0
```

Lets perform robust PCA-GMM estimation using principal components which explain 99.9% of genetic variation in *GLP1R*.

```
# calculate what number of PCAs that explain a certain amount of weighted variation
pca.no <- function(k){
  Phi <- ((rowSums(abs(bx))/sy)%*%t(rowSums(abs(bx))/sy))*ld
  return(which(cumsum(prcomp(Phi,scale=FALSE)$sdev^2/sum((prcomp(Phi,scale=FALSE)$sdev^2)))>k)[1])}

# robust PCA-GMM results using principal components explaining 99.9% of weighted genetic variation
pca_res0 <- PCA_GMM(bx=bx,sx=sx,by=by,sy=sy,ld=ld,cor.x=cor.x,nx=nx,ny=ny,r=pca.no(0.999))

ci <- cbind(pca_res0$liml,pca_res0$liml - qnorm(1 - 0.05/2)*pca_res0$se.liml,pca_res0$liml + qnorm(1 - 0.05/2)*pc
a_res0$se.liml,2*(1-pnorm(abs(pca_res0$liml/pca_res0$se.liml))),rep(0.999,3),rep(pca_res0$factors,3),pca_res0$con
dF,(1-pchisq((pca_res0$condF*(pca_res0$factors-ncol(bx)+1)),pca_res0$factors-ncol(bx)+1)))
colnames(ci) <- c("estimate","lower","upper","p-value","prop. var. expl.","PCs", "cond. F-statistic","cond. F p-v
alue"); rownames(ci) <- c("brain-caudate","heart-atrial appendage","pancreas")
round(ci,3)
```

```
##                       estimate  lower  upper p-value prop. var. expl. PCs
## brain-caudate           -0.067 -0.118 -0.015   0.011            0.999  21
## heart-atrial appendage   0.018 -0.025  0.061   0.418            0.999  21
## pancreas                 0.010 -0.037  0.056   0.678            0.999  21
##                       cond. F-statistic cond. F p-value
## brain-caudate                     1.256           0.201
## heart-atrial appendage            1.387           0.120
## pancreas                          1.441           0.096
```

We find that genetically-predicted gene expression in brain-caudate is associated with CAD risk (estimate -0.065, 95% CI [-0.112, -0.019]), whereas gene expression in the other tissues is not.

What are the results when we use fewer principal components (that explain 99% of genetic variation only)?

```
# robust PCA-GMM results using principal components explaining 99% of weighted genetic variation
pca_res1 <- PCA_GMM(bx=bx,sx=sx,by=by,sy=sy,ld=ld,cor.x=cor.x,nx=nx,ny=ny,r=pca.no(0.99))

ci1 <- cbind(pca_res1$liml,pca_res1$liml - qnorm(1 - 0.05/2)*pca_res1$se.liml,pca_res1$liml + qnorm(1 - 0.05/2)*p
ca_res1$se.liml,2*(1-pnorm(abs(pca_res1$liml/pca_res1$se.liml))),rep(0.99,3),rep(pca_res1$factors,3),pca_res1$con
dF,(1-pchisq((pca_res1$condF*(pca_res1$factors-ncol(bx)+1)),pca_res1$factors-ncol(bx)+1)))
colnames(ci1) <- c("estimate","lower","upper","p-value","prop. var. expl.","PCs", "cond. F-statistic","cond. F p-
value"); rownames(ci1) <- c("brain-caudate","heart-atrial appendage","pancreas")
round(ci1,3)
```

```
##                       estimate  lower  upper p-value prop. var. expl. PCs
## brain-caudate           -0.097 -0.163 -0.031   0.004             0.99  10
## heart-atrial appendage   0.029 -0.014  0.072   0.191             0.99  10
## pancreas                 0.005 -0.029  0.039   0.790             0.99  10
##                       cond. F-statistic cond. F p-value
## brain-caudate                     1.427           0.179
## heart-atrial appendage            2.188           0.025
## pancreas                          2.571           0.008
```

The estimates are therefore very similar using the first 10 principal components.

Now lets consider an extended analysis with 10 tissues that were chosen based on *GLP1R* expression levels. In addition to brain-caudate, heart-atrial appendage, and pancreas, we now additionally include thyroid, testis, stomach, nerve, lung, heart-left ventricle, and brain-hypothalamus. Again, lets start by setting up the inputs.

```
bx <- cbind(brain$beta,heart$beta,hypothalamus$beta,left_ventricle$beta,lung$beta,nerve$beta,pancreas$beta,stomac
h$beta,testis$beta,thyroid$beta)
sx <- cbind(brain$se,heart$se,hypothalamus$se,left_ventricle$se,lung$se,nerve$se,pancreas$se,stomach$se,testis$se
,thyroid$se)
by <- cad$Effect; sy <- cad$StdErr
nx <- rep(838,ncol(bx)); ny <- mean(cad$n)
cor.x <- diag(ncol(bx))
cor.x[1,3] <- 0.7; cor.x[3,1] <- 0.7
cor.x[2,4] <- 0.7; cor.x[4,2] <- 0.7
```

We chose 0.7 for the correlation between brain-caudate and brain-hypothalamus, and for the correlation between heart-atrial appendage and heart-left ventricle. The results are not too sensitive to this choice.

Lets compute the results using robust PCA-GMM with principal components that explain 99.9% of genetic variation.

```
# calculate what number of PCAs that explain a certain amount of variation
pca.no <- function(k){
  Phi <- ((rowSums(abs(bx))/sy)%*%t(rowSums(abs(bx))/sy))*ld
  return(which(cumsum(prcomp(Phi,scale=FALSE)$sdev^2/sum((prcomp(Phi,scale=FALSE)$sdev^2)))>k)[1])}

# robust PCA-GMM results using principal components explaining 99.9% of genetic variation
pca.no <- pca.no(0.999)
pca_res <- PCA_GMM(bx=bx,sx=sx,by=by,sy=sy,ld=ld,cor.x=cor.x,nx=nx,ny=ny,r=pca.no)
ci.full <- cbind(pca_res$liml,pca_res$liml-qnorm(1 - 0.05/2)*pca_res$se.liml,pca_res$liml+qnorm(1 - 0.05/2)*pca_r
es$se.liml)
colnames(ci.full) <- c("estimate","lower","upper"); rownames(ci.full) <- c("brain-caudate","heart-atrial appendag
e","brain-hypothalamus","heart-left ventricle","lung","nerve","pancreas","stomach","testis","thyroid")
round(ci.full,3)
```

```
##                       estimate  lower upper
## brain-caudate           -0.108 -0.216 0.000
## heart-atrial appendage  -0.019 -0.185 0.148
## brain-hypothalamus      -0.088 -0.192 0.016
## heart-left ventricle     0.049 -0.072 0.170
## lung                    -0.009 -0.089 0.071
## nerve                   -0.017 -0.191 0.157
## pancreas                 0.043 -0.077 0.164
## stomach                 -0.012 -0.098 0.073
## testis                  -0.036 -0.202 0.129
## thyroid                  0.041 -0.068 0.149
```

We again find evidence that only gene expression in brain-caudate is associated with CAD risk. However, the results of this extended analysis must be interpreted with caution since the relevance of other tissues may be masked by the complexity of the model. Moreover, the conditional F-statistics for this analysis are quite low:

```
# conditional F-statistics for the 10 tissue analysis
condF <- as.matrix(pca_res$condF)
colnames(condF) <- c("cond. F-statistic"); rownames(condF) <- c("brain-caudate","heart-atrial appendage","brain-h
ypothalamus","heart-left ventricle","lung","nerve","pancreas","stomach","testis","thyroid")
round(condF,3)
```

```
##                       cond. F-statistic
## brain-caudate                     0.442
## heart-atrial appendage            0.491
## brain-hypothalamus                0.380
## heart-left ventricle              0.974
## lung                              1.484
## nerve                             0.461
## pancreas                          0.351
## stomach                           0.787
## testis                            0.575
## thyroid                           0.793
```

Fitting a model with all ten tissues as risk factors can result in imprecise estimates that are difficult to interpret due to multicollinearity. Instead, the MR-BMA method (Zuber et al. 2020, Nat. Commun.) fits models with each risk factor in turn, all pairs of risk factors, all triples of risk factors, and so on. Each model, representing a particular combination of risk factors, receives a posterior model probability based on its goodness-of-fit; the model that best explains the genetic associations with the outcome will receive the greatest posterior probability.

Additionally, each risk factor is assigned a marginal inclusion probability, calculated as the sum of the posterior model probabilities for all models including that risk factor.

The MR-BMA method can be applied to standardized PCA-transformed multivariable associations relating to each of the 10 tissues. Lets calculate these inputs.

```
# calculate principal components
bx <- cbind(brain$beta,heart$beta,hypothalamus$beta,left_ventricle$beta,lung$beta,nerve$beta,pancreas$beta,stomac
h$beta,testis$beta,thyroid$beta)
sy <- cad$StdErr; p <- nrow(bx)
Phi <- ((rowSums(abs(bx))/sy)%*%t(rowSums(abs(bx))/sy))*ld
r <- which(cumsum(prcomp(Phi,scale=FALSE)$sdev^2/sum((prcomp(Phi,scale=FALSE)$sdev^2)))>0.999)[1]
lambda <- sqrt(p)*prcomp(Phi,scale=FALSE)$rotation[,1:r]
evec <- eigen((t(lambda)%*%lambda))$vectors
eval <- eigen((t(lambda)%*%lambda))$values
lambda <- lambda%*%(solve(evec%*%diag(sqrt(eval))%*%t(evec)))
dim(lambda) <- c(p,r)

# transform to PCA-standardised multivariable betas weighted
pca.transform <- function(beta,se,n,ld){
a <- 1/((n*se^2)+beta^2); ay <- 1/((1131832*cad$StdErr^2)+cad$Effect^2)
A <- (sqrt(a)%*%t(sqrt(a)))*ld; Ay <- (sqrt(ay)%*%t(sqrt(ay)))*ld
B <- a*beta; By <- ay*cad$Effect
A.f <- t(lambda)%*%A%*%lambda; B.f <- as.vector(t(lambda)%*%B)
Ay.f <- t(lambda)%*%Ay%*%lambda; By.f <- as.vector(t(lambda)%*%By)
pca.var.y <- solve(Ay.f)*(1-as.numeric(t(By.f)%*%solve(Ay.f)%*%By.f)); pca.var.y <- pca.var.y*(1/(1131832-r+1))
inv.pca.var.y <- solve(pca.var.y)
evec <- eigen(inv.pca.var.y)$vectors; eval <- eigen(inv.pca.var.y)$values; inv.pca.var.sq.y <- evec%*%diag(sqrt(e
val))%*%t(evec)
pca.beta <- inv.pca.var.sq.y%*%as.vector(solve(A.f)%*%B.f)
return(pca.beta)
}


brain <- pca.transform(brain$beta,brain$se,838,ld)
heart <- pca.transform(heart$beta,heart$se,838,ld)
hypothalamus <- pca.transform(hypothalamus$beta,hypothalamus$se,838,ld)
left_ventricle <- pca.transform(left_ventricle$beta,left_ventricle$se,838,ld)
lung <- pca.transform(lung$beta,lung$se,838,ld)
nerve <- pca.transform(nerve$beta,nerve$se,838,ld)
pancreas <- pca.transform(pancreas$beta,pancreas$se,838,ld)
stomach <- pca.transform(stomach$beta,stomach$se,838,ld)
testis <- pca.transform(testis$beta,testis$se,838,ld)
thyroid <- pca.transform(thyroid$beta,thyroid$se,838,ld)
cad <- pca.transform(cad$Effect,cad$StdErr,1131832,ld)
```

Now we can run the MR-BMA method. We choose a prior probability of p = 0.1 for each tissue, representing a prior expectation that one tissue is the true causal tissue.

```
exposure.id <- c("caudate","atrial appendage","hypothalamus","left ventricle","lung","nerve","pancreas","stomach"
,"testis","thyroid")
pca.bx <- cbind(brain,heart,hypothalamus,left_ventricle,lung,nerve,pancreas,stomach,testis,thyroid)
colnames(pca.bx) <- exposure.id; pca.by <- cad; snps <- ld_dat$rsid

source("summary_mvMR_BF.R"); source("summary_mvMR_SSS.R")
```

```
##
## Attaching package: 'combinat'
```

```
## The following object is masked from 'package:utils':
##
##     combn
```

```
## hash-2.2.6.2 provided by Decision Patterns
```

```
amd_nmr_input=new("mvMRInput", betaX = pca.bx, betaY = pca.by, snps=snps, exposure=exposure.id, outcome = "cad")
BMA_output=summarymvMR_SSS(amd_nmr_input,kmin=1,kmax=10, prior_prob=0.1, max_iter=100)
```

The top 5 models are:

```
BMA_pp <- function(k){sort(BMA_output@pp,decreasing=TRUE)[1:5][[k]]}; BMA_pp <- sapply(1:5,BMA_pp)
BMA_models <- function(k){exposure.id[as.numeric(strsplit(names(sort(BMA_output@pp,decreasing=TRUE)),split=",")[[
k]])]}
sapply(1:5,BMA_models)
```

```
## [1] "caudate" "testis"  "lung"    "nerve"   "stomach"
```

…and their associated posterior probabilities are:

```
round(BMA_pp,3)
```

```
## [1] 0.451 0.110 0.107 0.067 0.064
```

Therefore, the MR-BMA method which compares evidence for models containing different combinations of risk factors, indicates a clear preference for the model containing brain-caudate and no other tissue. The posterior probability for this model was 45.1%. The next highest ranking models contained testis only (posterior probability 11.0%), lung only (10.7%), nerve only (6.7%), and then stomach only (6.4%).

Finally, lets find the marginal inclusion probabilities.

```
# marginal inclusion probabilities
marg.inc <- cbind(exposure.id[order(BMA_output@pp_marginal, decreasing=TRUE)],round(BMA_output@pp_marginal[order(
BMA_output@pp_marginal, decreasing=TRUE)],3))
colnames(marg.inc) <- c("tissue","marg. inc. prob.")
marg.inc
```

```
##       tissue             marg. inc. prob.
##  [1,] "caudate"          "0.471"
##  [2,] "testis"           "0.118"
##  [3,] "lung"             "0.112"
##  [4,] "nerve"            "0.072"
##  [5,] "stomach"          "0.067"
##  [6,] "hypothalamus"     "0.06"
##  [7,] "thyroid"          "0.045"
##  [8,] "atrial appendage" "0.03"
##  [9,] "pancreas"         "0.029"
## [10,] "left ventricle"   "0.025"
```

Overall, the marginal inclusion probability for brain-caudate was 47.1%, indicating that brain-caudate was selected as a causal risk factor in models with a total posterior probability of 47.1%. The next ranking tissues by marginal inclusion probability were testis (11.8%) and lung (11.2%).

Summary

In conclusion, our results in parts I, II, and III suggest that there is an association between genetically-predicted BMI and CAD risk in a multivariable model including both BMI and T2D, suggesting that the mechanistic pathway from *GLP1R* to CAD risk passes predominantly via BMI rather than T2D. We also demonstrated an association between genetically-predicted *GLP1R* gene expression in the brain and CAD risk in a multivariable model including *GLP1R* gene expression in the brain, heart, and pancreas. This was further validated in a multivariable model including gene expression in a wide range of tissues.