Mendelian randomization incorporating uncertainty about pleiotropy

John R Thompson, Cosetta Minelli, Jack Bowden, Fabiola Del Greco M, Dipender Gill, Elinor M Jones, Chin Yang Shapland, Nuala A Sheehan

1 Online Supplement

1.1 OpenBUGS code

The OpenBUGS code given below performs model averaging allowing for uncertainty in the variant-exposure effect sizes as in Tables 3 and 4 of the main paper. The code is written in terms of the precision so that t in the code represents the reciprocal of the variance and itau2=$1/\tau^2$. T is the model indicator that can take values 1,2,3. The data supplied to the program contains N=number of instruments, x[] and y[] , containing the estimates $\hat{\gamma}_j$, $\hat{\Gamma}_j$ respectively, pg[] containing $1/s_j^2$ and vG[] containing S_j^2, the prior model probabilities, p[], and the parameters of the priors.

Under all three models we assume a normal prior for β with mean MB and precision PB. Under M_3 the intercept mu has a normal prior with mean PA and precision PB. Under M_1 and M_2 the intercept is zero so the values of mu are generated from a normal pseudo prior with mean zero and precision 100; this helps mixing between models since these values correspond to the approximate scale on which the simulations were created. The pleiotropy terms under M_2 and M_3 have a precision itau2 that is given a G(GA,GB) prior.

```
model {
   T ~ dcat(p[])
   for( i in 1:N ) {
      z[i] ~ dnorm(m,t)
      x[i] ~ dnorm(z[i],pg[i])
      m[i] <- mu*equals(T,3) + beta*z[i]
      pr[i] <- 1/(vG[i]+step(T-1.5)/itau2)
      y[i] ~ dnorm(m[i],pr[i])
   }
   ta <- 100 - (100-PA)*equals(T,3)
   mu ~ dnorm(MA,ta)
   beta ~ dnorm(MB,PB)
   at <- 1*equals(T,1) + GA*step(T-1.5)
   bt <- 0.001*equals(T,1) + GB*step(T-1.5)
   itau2 ~ dgamma(at,bt)
   m ~ dnorm(0.2,0.1)
   t ~ dgamma(0.1,0.001)
```

```
}
```

The priors on m and t need to be chosen to be appropriate for the particular problem. To remove the adjustment for uncertainty in the measurement of the gene-intermediate estimates the priors on m and t should be dropped and the loop should be replaced by,

```
for( i in 1:N ) {
    m[i] <- mu*equals(T,3) + beta*x[i]
    pr[i] <- 1/(vG[i]+step(T-1.5)/itau2)
    y[i] ~ dnorm(m[i],pr[i])
}
```

1.2 Mixed simulation

Data for the mixed simulation shown in Figure 4 were generated so that there was an equal probability of no pleiotropy (M_1), zero-centred pleiotrop (M_2) or non zero-centred pleiotropy (M_3). The true value of the effect of X on Y, β, was fixed to be 0.5. The effects of the genetic variants on X were generated from a Normal(0.2,0.05) distribution. Under M_2 and M_3 the precision of the pleiotropy was generated from a Gamma(10,0.04) distribution and under M_3 the mean pleiotropy was generated from uniform(-0.2,0.2) distribution. Two-sample data were generated with both sample sizes equal to 50,000.

1.3 Average performance

Tables showing the root mean square errors corresponding to the simulations in Tables 2, 3 and 4.

Table S.1: Root mean square errors for 100 datasets simulated without pleiotropy (under M_1). Corresponding to Table 2

		Analysis Model		
β	Fixed effects	Random effects	MR-Egger	Model Average
Without adjusting for variant-exposure uncertainty				
Small samples (n=5,000)				
0.0	0.029	0.029	0.094	0.029
0.5	0.034	0.034	0.213	0.051
Large samples (n=50,000)				
0.0	0.010	0.010	0.039	0.010
0.5	0.011	0.010	0.051	0.010
Adjusting for variant-exposure uncertainty				
Small samples (n=5,000)				
0.0	0.030	0.030	0.134	0.030
0.5	0.028	0.028	0.134	0.038
Large samples (n=50,000)				
0.0	0.010	0.010	0.041	0.010
0.5	0.011	0.011	0.043	0.011

Table S.2: Root mean square errors for 100 datasets simulated with zero-centred pleiotropy (under M_2). Corresponding to Table 3

| | Analysis Model | | | |
β	Fixed effects	Random effects	MR-Egger	Model Average
Small pleiotropic variance: MAPR=2%				
Small samples (n=5,000)				
0.0	0.033	0.032	0.168	0.027
0.5	0.036	0.036	0.147	0.032
Large samples (n=50,000)				
0.0	0.012	0.012	0.040	0.011
0.5	0.010	0.010	0.046	0.011
Medium pleiotropic variance: MAPR=10%				
Small samples (n=5,000)				
0.0	0.035	0.035	0.216	0.040
0.5	0.040	0.040	0.194	0.040
Large samples (n=50,000)				
0.0	0.020	0.020	0.103	0.022
0.5	0.025	0.022	0.167	0.023
Large pleiotropic variance: MAPR=19%				
Small samples (n=5,000)				
0.0	0.051	0.049	0.761	0.118
0.5	0.050	0.048	0.736	0.176
Large samples (n=50,000)				
0.0	0.043	0.039	0.160	0.042
0.5	0.046	0.039	0.156	0.042

Table S.3: Root mean square errors for 100 datasets simulated with directional pleiotropy (under M_3). Corresponding to Table 4

|		Analysis Model		
β	Fixed effects	Random effects	MR-Egger	Model Average
Small pleiotropic mean, small variance: MAPR=7%				
Small samples (n=5,000)				
0.0	0.072	0.071	0.150	0.064
0.5	0.073	0.071	0.137	0.061
Large samples (n=50,000)				
0.0	0.063	0.063	0.041	0.061
0.5	0.066	0.065	0.044	0.062
Small pleiotropic mean, medium variance: MAPR=11%				
Small samples (n=5,000)				
0.0	0.076	0.076	0.213	0.067
0.5	0.082	0.080	0.181	0.069
Large samples (n=50,000)				
0.0	0.066	0.065	0.120	0.062
0.5	0.073	0.066	0.156	0.062
Large pleiotropic mean, small variance: MAPR=27%				
Small samples (n=5,000)				
0.0	0.257	0.257	0.152	0.240
0.5	0.249	0.248	0.137	0.232
Large samples (n=50,000)				
0.0	0.253	0.252	0.043	0.064
0.5	0.254	0.252	0.045	0.097
Large pleiotropic mean, medium variance: MAPR=27%				
Small samples (n=5,000)				
0.0	0.262	0.261	0.215	0.235
0.5	0.268	0.266	0.166	0.242
Large samples (n=50,000)				
0.0	0.255	0.251	0.111	0.200
0.5	0.265	0.255	0.157	0.207

1.4 Impact of the priors
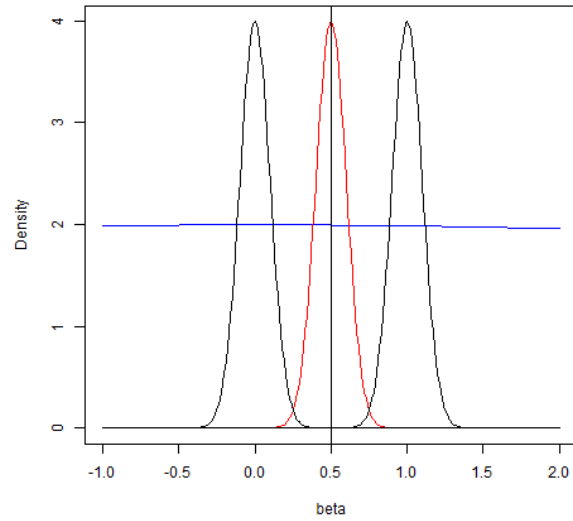
Figure S.1: Realistically vague prior for β times 50 (blue), correctly centred informative prior (red), over and under centred priors (black). True parameter value shown as a vertical line.
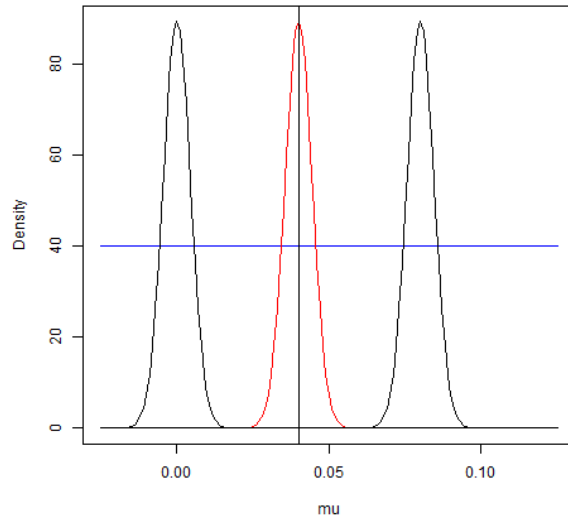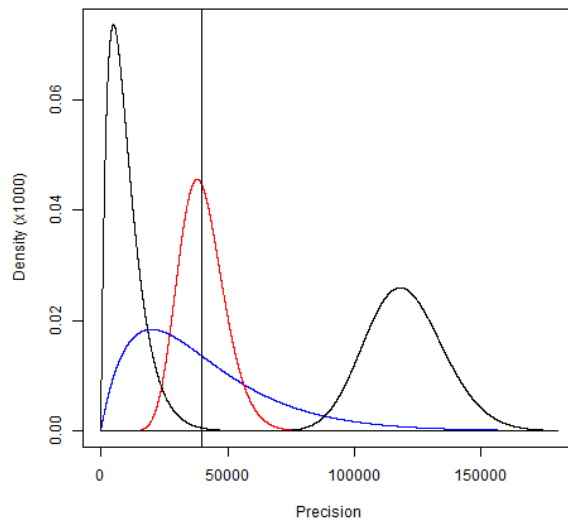
Figure S.2: Realistically vague prior for μ times 1000 (blue), correctly centred informative prior (red), over and under centred priors (black). True parameter value shown as a vertical line.

Figure S.3: Realistically vague prior for the precision (blue), correctly centred informative prior (red), over and under centred priors (black). True parameter value shown as a vertical line.

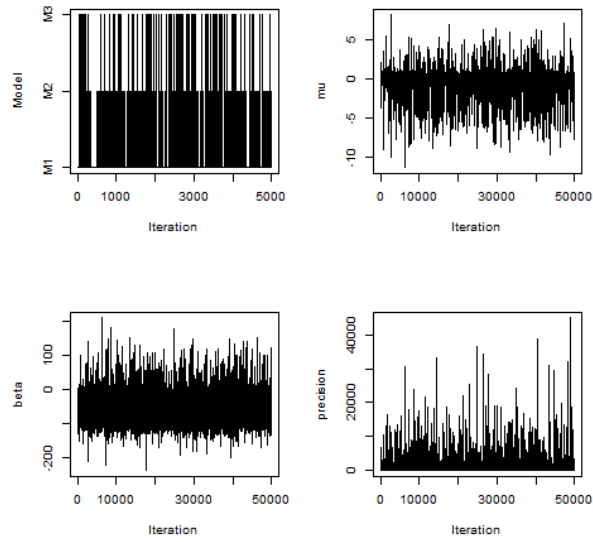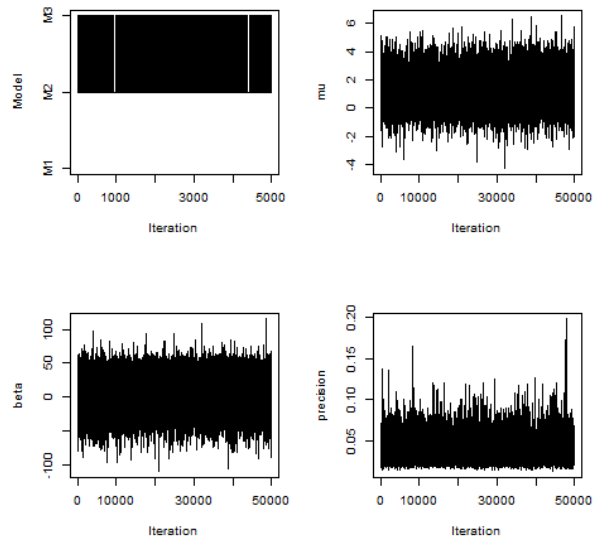## 1.5 Trace plots for age at menarche and lung function

Figure S.4: Trace plots for adolescent women.

Figure S.5: Trace plots for adult women.