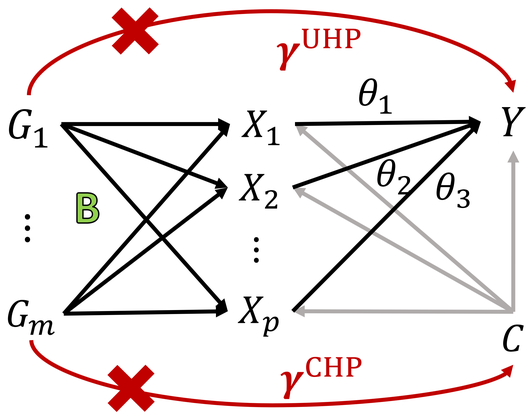
De-biased multivariable Mendelian Randomization and its implementation

Noah Lorincz-Comi (njl96@case.edu), MSc, Yihe Yang, PhD, Gen Li, BS, Xiaofeng Zhu, PhD
Department of Population and Quantitative Health Sciences, School of MedicineCase Western Reserve University

Case Western Reserve University School of Medicine
NIH National Heart, Lung, and Blood Institute
NIH National Human Genome Research Institute

Background



$$(G_1, \ldots, G_m) \xrightarrow{\alpha} Y$$

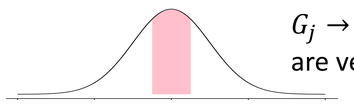
$$\alpha_j = \beta_j^\top \theta + \gamma_j^{UHP} + \gamma_j^{CHP}$$

Estimated from GWAS

Outliers that can bias $\widehat{\theta}$

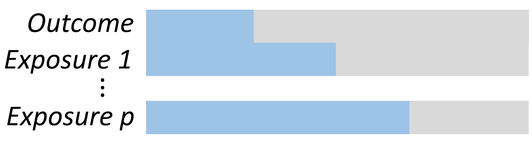
Current methods estimate θ with bias due to

Weak IVs

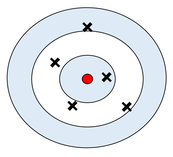


$G_j \to X_k$ effects are very small

GWAS sample overlap



GWAS estimation error



True $G_j \to X_k$ and $G_j \to Y$ effects measured with sampling error

$$(\widehat{\beta}_j, \hat{\alpha}_j) = (\beta_j, \alpha_j) + (\mathbf{u}_j, v_j)$$

MRBEE Software (R & Python*)

Input Data

`data`
- *Merged data set of all standardized exposure, outcome GWAS summary statistics*
  - *effect sizes*
  - *standard errors*
  - *unique SNP identifier (e.g., rsID)*

`IVInds`
- *Row index positions of IVs in* `data`

`leadSNPs`
- *Row index positions of lead outcome SNPs (optional)*

Preparing Data

`biasTerms()`
- *Bias-correction terms*
- *Harmonised alleles*

`prepData()`
- *Simplify data format for all analyses*

Example code

```
bT=biasTerms(data,dNames)
pD=prepData(bT,IVInds,oi)
```

`dNames`: Named list of effect size, SE, and effect allele column name vectors in `data`, e.g.: `list(est=<x>,se=<x>,allele=<x>)`

`oi`: "outcome index"; index position of `dNames` vector elements corresponding to outcome

Performing MR

`MRBEE.IMRP()`
`MRBEE.Median()`
`MRBEE.Mix()`
`MRBEE.IPOD()`
`MRBEE.MLqe()`

(Correct biases in multivariable MR)

Visualizing causality

`causalPlot()`
- *Display causal estimates*

Example code

```
fit<x>=MRBEE.<x>(pD)
causalPlot(fit<x>,eNames)
```

`eNames`: Vector of exposure names, order corresponding to orders in `dNames$<x>` e.g.: `c(<x>, …, <x>)`

Pleiotropy analysis

`Spleio()`
- *Horizontal pleiotropy test*

`genomePleio()`
- *Genome-wide* `Spleio()`

Visualizing pleiotropy

`residualPlot()`
- *Assess pleiotropy in MR*

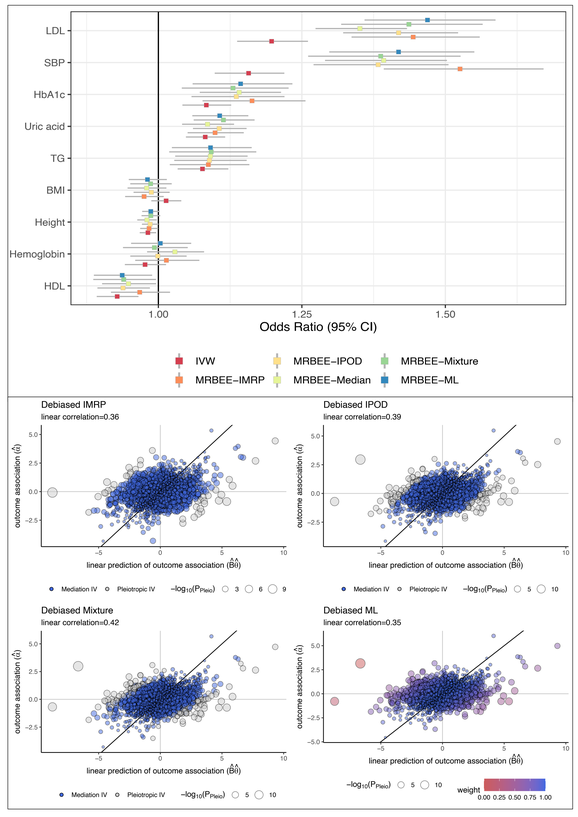
`gwPleioPlot()`
- *Assess pleiotropy in genome*

Example code

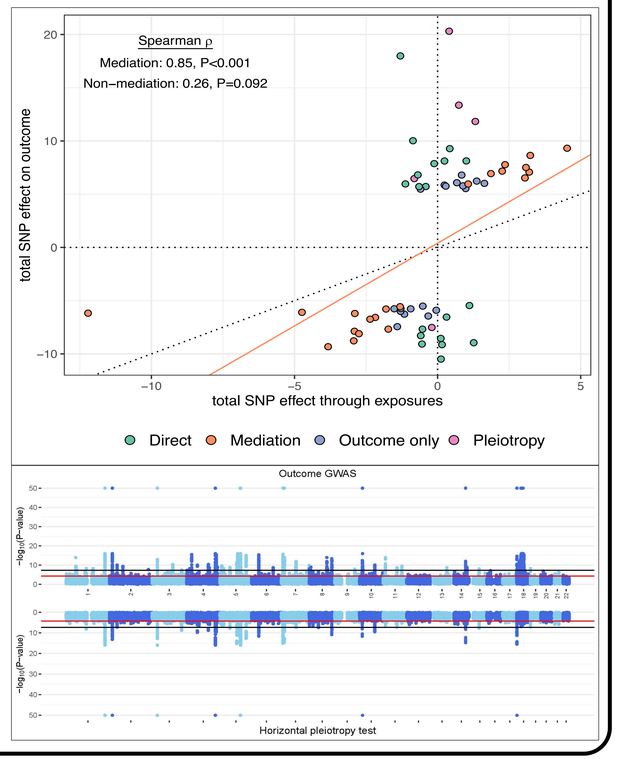
```
sP<x>=Spleio(pD,fit<x>)
gP<x>=genomePleio(pD0,fit<x>)
residualPlot(fit<x>)
gwPleioPlot(gP<x>,leadSNPs,CHR_BP)
```

```
pD0=prepData(bT,oi=oi)
leadSNPs: Vector of row indices in data of lead SNPs in outcome GWAS tagging independent loci
CHR_BP: Vector of positions as <CHR>_<BP>
```

Causal effect estimates and model fit[1]



Classifying outcome loci using horizontal pleiotropy [1]



Methods

MR using **B**ias-corrected **E**stimating **E**quations

$$S_{\text{MRBEE}}(\boldsymbol{\theta}) = \mathbf{0} = \sum_{j=1}^{m} \rho\left(\hat{\alpha}_j - \widehat{\boldsymbol{\beta}}^\top \boldsymbol{\theta}\right) - \Gamma(\boldsymbol{\theta}, \boldsymbol{\Sigma})$$

$$\overset{[4]}{=} \sum_{j=1}^{m} w_j \widehat{\boldsymbol{\beta}}(\hat{\alpha}_j - \widehat{\boldsymbol{\beta}}^\top \boldsymbol{\theta}) - \Gamma(\boldsymbol{\theta}, \boldsymbol{\Sigma})$$

ρ: Loss function robust to (γ^{CHP}, γ^{UHP})

- `MRBEE.IMRP()`: Removes UHP/CHP IVs
- `MRBEE.Median()`: assumes <50% UHP/CHP IVs
- `MRBEE.Mix()`: Mixture of valid & UHP/CHP IVs
- `MRBEE.IPOD()`: Penalized outliers
- `MRBEE.MLqe()`: Likelihood-weighted IVs

Compare methods using residual plots

$\Gamma(\boldsymbol{\theta}, \boldsymbol{\Sigma})$: Bias-correction terms

- $\mathrm{m}^{-1}\Gamma(\boldsymbol{\theta}, \boldsymbol{\Sigma}) \overset{[2]}{=} \mathrm{Cov}(\mathbf{u}_j)\boldsymbol{\theta} - \mathrm{Cov}(v_j, \mathbf{u}_j)$
- Estimated using full GWAS summary statistics[3]
- Necessary for valid statistical inference[2]

Conclusion

MRBEE software performs

a) Bias-correction for $\Gamma(\boldsymbol{\theta}, \boldsymbol{\Sigma})$
   - Commonly nonzero in practice

b) Bias-protection against UHP/CHP
   - Many methods user can choose
   - Chosen method is setting-specific

Download software

R: https://github.com/noahlorinczcomi/MRBEE
Python*: In progress

References

[1]: Lorincz-Comi et al (2023) **PMID: 36778480**
[2]: Yang et al (2023) ***arXiv:2301.05130***
[3]: Zhu et al (2015) **PMID: 25500260**
[4]: She & Owen (2011) ***JASA 106*(494), 626-639**