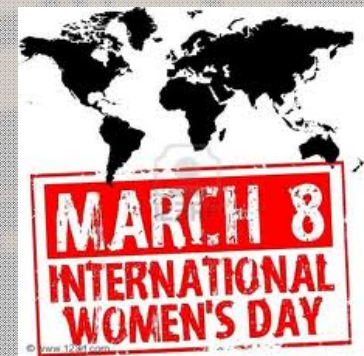


Including covariates in your model

Sarah Medland



Before I forget...

- <http://www.statmethods.net/>



R is an elegant and comprehensive statistical and graphical programming language. Unfortunately, it can also have a [steep learning curve](#). I created this website for both current R users, and experienced users of other statistical packages (e.g., SAS, SPSS, Stata) who would like to transition to R. My goal is to help you quickly access this language in your work.

I assume that you are already familiar with the [statistical methods](#) covered and instead provide you with a roadmap and the code necessary to get started quickly, and orient yourself for future learning. I designed this web site to be an easily accessible reference. Look at the [sitemap](#) to get an overview.



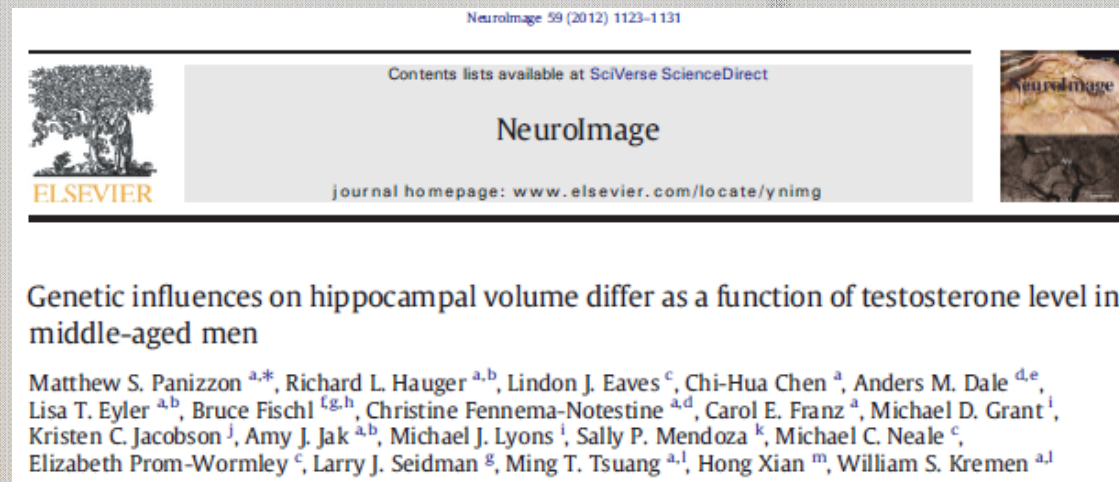
The approach up till now...

- Account for sex by using different means matrices for males and females
- Ignore other covariates
- This is a very bad idea...
 - Be guided by the literature Age, Sex, Age², SES, birthweight...



Can you include a covariate that is in itself influenced by genes?

- Eg correcting for total brain size when estimating the heritability of specific brain regions



- Depends on your research question
 - Is it a moderator or a confounder

Accounting or correcting for a covariate

- Most common method
 - Add a correction in the form of a linear regression to the mean
 - If the covariate is binary code it as 0 vs 1

- $\text{expMean} = \text{intercept} + \beta * \text{covariate}$

Mean for the group coded 0

Coded 0/1

Unstandardised regression B/
the deviation of group 1 from
group 0

Coded 0/1

$$\text{expMean} = \text{intercept} + \beta * \text{covariate}$$

Mean for the group coded 0

Unstandardised regression B/
the deviation of group 1 from
group 0

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	20.598	.127		162.104	.000
	sex1	.318	.045	.162	7.084	.000
	age	.033	.005	.143	6.244	.000

a. Dependent Variable: bmi1

Setting this up in openMx...

```
intercept  <- mxMatrix( type="Full", nrow=1, ncol=ntv, free=TRUE,
                        values= 20, label="mean", name="Mean" )

# Matrix for moderating/interacting variable
defSex    <- mxMatrix( type="Full", nrow=1, ncol=2, free=FALSE,
                        labels=c("data.sex1","data.sex2"), name="Sex" )

# Matrices declared to store linear Coefficients for covariate
B_Sex    <- mxMatrix( type="Full", nrow=1, ncol=1, free=TRUE,
                        values= .01, label="betaSex", name="bSex" )

meanSex   <- mxAlgebra( bSex%*%Sex, name="SexR")

expMean   <- mxAlgebra( Mean + SexR + AgeR, name="expMean")

defs     <- list( intercept, defSex, B_Sex, meanSex)
```

Setting this up in openMx...

```
intercept <- mxMatrix( type="Full", nrow=1,  
ncol=ntv, free=TRUE, values= 20, label="mean",  
name="Mean" )
```

1*2 matrix

Containing 2 elements

[mean mean]

Start value =20

[20 20]



Setting this up in openMx...

```
# Matrix for moderating/interacting variable  
defSex <- mxMatrix( type="Full", nrow=1, ncol=2,  
free=FALSE, labels=c("data.sex1","data.sex2"),  
name="Sex")
```

1*2 matrix

Containing 2 elements – the values of sex1 and sex2

$$[\textit{sex1} \quad \textit{sex2}]$$

This matrix is repopulated for each family with the actual values of sex1 and sex2

$$[\textit{sex1} \quad \textit{sex2}]$$


Setting this up in openMx...

```
B_Sex <- mxMatrix( type="Full", nrow=1, ncol=1,  
free=TRUE, values= .01, label="betaSex",  
name="bSex" )
```

1*1 matrix

Containing 1 element – the unstandardise regression beta for sex on bmi

$[B_{sex}]$

This element will be estimated and has a start value of .01

$[.01]$

Setting this up in openMx...

```
meanSex <- mxAlgebra( bSex%*%Sex,  
name="SexR")
```

$$[B_{sex}] * [sex1 \quad sex2] = [B_{sex \cdot sex1} \quad B_{sex \cdot sex2}]$$

```
expMean <- mxAlgebra( Mean + SexR + AgeR,  
name="expMean")
```

$$[mean \quad mean] + [B_{sex \cdot sex1} \quad B_{sex \cdot sex2}]$$
$$= [mean + B_{sex \cdot sex1} \quad mean + B_{sex \cdot sex2}]$$

Eg:

$$= [20 + (.5 * 0) \quad 20 + (.5 * 1)]$$

$$= [20 \quad 20.5]$$

Lets give it a go...

1. twinACE.R
2. twinACECovSex.R

Does it make a difference to the fit?

3. twinACECovSexAge.R

Lets build a script together?



