# Introduction to Multivariate Genetic Analysis

#### Kate Morley and Frühling Rijsdijk

21st Twin and Family Methodology Workshop, March 2008

# **Aim and Rationale**

- Aim: to examine the source of factors that make traits correlate or co-vary
- Rationale:
  - Traits may be correlated due to shared genetic factors (A) or shared environmental factors (C or E)
  - Can use information on multiple traits from twin pairs to partition covariation into genetic and environmental components

# **Example 1**

- Why do traits correlate/covary?
- How can we explain the association?
  - Additive genetic factors (r<sub>G</sub>)
  - Shared environment (r<sub>c</sub>)
  - Non-shared environment (r<sub>E</sub>)
- Kuntsi et al. (2004) Am J Med Genet B, 124:41



# Example 2

- Associations between phenotypes over time
  - Does anxiety in childhood lead to depression in adolescence?
- How can we explain the association?
  - Additive genetic factors (a<sub>21</sub>)
  - Shared environment (c<sub>21</sub>)
  - Non-shared environment (e<sub>21</sub>)
  - How much is not explained by prior anxiety?
- Rice et al. (2004) BMC Psychiatry 4:43



# **Sources of Information**

- As an example: two traits measured in twin pairs
- Interested in:
  - Cross-trait covariance within individuals
  - Cross-trait covariance between twins
  - MZ:DZ ratio of cross-trait covariance between twins

# **Observed Covariance Matrix**

		Twin 1		Twin 2	
		Phenotype 1	Phenotype 2	Phenotype 1	Phenotype 2
Twin 1	Phenotype 1	Variance P1			
	Phenotype 2	Covariance P1-P2	Variance P2		
-win 2	Phenotype 1	Within-trait P1	Cross-trait	Variance P1	
F	Phenotype 2	Cross-trait	Within-trait P2	Covariance P1-P2	Variance P2

# **Observed Covariance Matrix**

		Twin 1		Twin 2		
		Phenotype 1	Phenotype 2	Phenotype 1	Phenotype 2	
		Within-twin	covariance			
Twin 1	Phenotype 1	Variance P1				
	Phenotype 2	Covariance P1-P2	Variance P2			
				Within-twin	covariance	
Twin 2	Phenotype 1	Within-trait P1	Cross-trait	Variance P1		
	Phenotype 2	Cross-trait	Within-trait P2	Covariance P1-P2	Variance P2	

# **Observed Covariance Matrix**

		Twin 1		Twin 2		
		Phenotype 1	Phenotype 2	Phenotype 1	Phenotype 2	
		Within-twin	covariance			
/in 1	Phenotype 1	Variance P1				
F	Phenotype 2	Covariance P1-P2	Variance P2			
		Cross-twin	covariance	Within-twin	covariance	
Twin 2	Phenotype 1	Within-trait P1	Cross-trait	Variance P1		
	Phenotype 2	Cross-trait	Within-trait P2	Covariance P1-P2	Variance P2	

## **SEM: Cholesky Decomposition**



#### **SEM: Cholesky Decomposition**



# **SEM: Cholesky Decomposition**



# Why Fit This Model?

- Covariance matrices must be positive definite
- If a matrix is positive definite, it can be decomposed into the product of a triangular matrix and its transpose:

▲ A = X\*X'

- Many other multivariate models possible
  - Depends on data and hypotheses of interest

#### **Cholesky Decomposition**

# **Path Tracing**



![](_page_14_Figure_1.jpeg)

![](_page_15_Figure_1.jpeg)

![](_page_16_Figure_1.jpeg)

![](_page_17_Figure_1.jpeg)

![](_page_18_Figure_1.jpeg)

![](_page_19_Figure_1.jpeg)

![](_page_20_Figure_1.jpeg)

![](_page_21_Figure_1.jpeg)

![](_page_22_Figure_1.jpeg)

![](_page_23_Figure_0.jpeg)

		Twin 1		Twin 2		
		Phenotype 1	Phenotype 2	Phenotype 1	Phenotype 2	
		Within-twin	covariance			
vin 1	Phenotype 1	a <sub>11</sub> <sup>2</sup> +c <sub>11</sub> <sup>2</sup> +e <sub>11</sub> <sup>2</sup>				
F	Phenotype 2	a <sub>11</sub> a <sub>21</sub> +c <sub>11</sub> c <sub>21</sub> + e <sub>11</sub> e <sub>21</sub>	$a_{22}^{2}+a_{21}^{2}+c_{22}^{2}+c_{21}^{2}+c_{21}^{2}+c_{21}^{2}+c_{21}^{2}$			
		Cross-twin	covariance	Within-twin	covariance	
win 2	Phenotype 1	1/.5a <sub>11</sub> <sup>2</sup> +c <sub>11</sub> <sup>2</sup>		a <sub>11</sub> <sup>2</sup> +c <sub>11</sub> <sup>2</sup> +e <sub>11</sub> <sup>2</sup>		
	Phenotype 2	1/.5a <sub>11</sub> a <sub>21</sub> + C <sub>11</sub> C <sub>21</sub>	$\frac{1}{5a_{22}^{2}+1}$	a <sub>11</sub> a <sub>21</sub> +c <sub>11</sub> c <sub>21</sub> + e <sub>11</sub> e <sub>21</sub>	$a_{22}^{2}+a_{21}^{2}+c_{22}^{2}+c_{21}^{2$	

		Twin 1		Twin 2		
		Phenotype 1	Phenotype 2	Phenotype 1	Phenotype 2	
		Within-twin	covariance			
Twin 1	Phenotype 1	Variance P1				
	Phenotype 2	Covariance P1-P2	Variance P2			
		Cross-twin	covariance	Within-twin	covariance	
Twin 2	Phenotype 1	Within-trait P1	Cross-trait	Variance P1		
	Phenotype 2	Cross-trait	Within-trait P2	Covariance P1-P2	Variance P2	

![](_page_26_Figure_1.jpeg)

		Twin 1		Twin 2		
		Phenotype 1	Phenotype 2	Phenotype 1	Phenotype 2	
		Within-twin	covariance			
/in 1	Phenotype 1	Variance P1		Covariance P2 the sam	e of P1 and ne across	
Ļ	Phenotype 2	Covariance P1-P2	Variance P2	twins and z	zygosity groups	
		Cross-twin covariance		Within-twir	n covariance	
Twin 2	Phenotype 1	Within-trait P1	Cross-trait	Variance F1		
	Phenotype 2	Cross-trait	Within-trait P2	Covariance P1-P2	Variance P2	

		Twin 1		Twin 2		
		Phenotype 1	Phenotype 2	Phenotype 1	Phenotype 2	
		Within-twin	covariance			
Twin 1	Phenotype 1	Variance P1		Cross-tw within ea	vin covariance ach trait differs	
	Phenotype 2	Covariance P1-P2	Variance P2	by zygos	sity	
		Cross-twin covariance		Within-twi	n covariance	
Twin 2	Phenotype 1	Within-trait P1	Cross-trait	Variance P1		
	Phenotype 2	Cross-trait	Within-trait P2	Covariance P1-P2	Variance P2	

		Twin 1		Twin 2		
		Phenotype 1	Phenotype 2	Pł	nenotype 1	Phenotype 2
		Within-twin	covariance			
Twin 1	Phenotype 1	Variance P1		Cross-twin cross-tra covariance differs b zygosity		n cross-trait e differs by
	Phenotype 2	Covariance P1-P2	Variance P2			
		Cross-twin covariance		V	Vithin-twin	covariance
Twin 2	Phenotype 1	Within-trait P1	Cross-trait		Variance P1	
	Phenotype 2	Cross-trait	Within-trait P2	C	Covariance P1-P2	Variance P2

![](_page_30_Figure_1.jpeg)

![](_page_31_Figure_1.jpeg)

![](_page_32_Figure_1.jpeg)

![](_page_33_Figure_1.jpeg)

![](_page_34_Figure_1.jpeg)

## Summary

- Within-individual cross-trait covariance implies common aetiological influences
- Cross-twin cross-trait covariance implies common aetiological influences are familial
- Whether familial influences genetic or environmental shown by MZ:DZ ratio of cross-twin cross-trait covariances

#### **Cholesky Decomposition**

## **Specification in Mx**

#### **Mx Parameter Matrices**

#define nvar 2 Begin Matrices; X lower nvar nvar free Y lower nvar nvar free Z lower nvar nvar free G Full 1 nvar free H Full 1 1 fix End Matrices; Begin Algebra; A=X\*X'; C=Y\*Y'; E=Z\*Z'; P=A+C+EEnd Algebra;

- ! Genetic coefficients
  - ! C coefficients
  - ! E coefficients
  - ! Means
  - ! 0.5 for DZ A covar
  - ! A var/cov
  - ! C var/cov
  - ! E var/cov

# **Within-Twin Covariance**

![](_page_38_Figure_1.jpeg)

Path Tracing:  

$$\Sigma_A = \begin{bmatrix} a_{11}^2 & a_{11}a_{21} \\ a_{21}a_{11} & a_{21}^2 + a_{22}^2 \end{bmatrix}$$

### Within-Twin Covariance

![](_page_39_Figure_1.jpeg)

![](_page_39_Figure_2.jpeg)

#### Within-Twin Covariance

![](_page_40_Figure_1.jpeg)

![](_page_41_Figure_0.jpeg)

$$\Sigma_P = \begin{bmatrix} +c_{11}^2 + e_{11}^2 \\ +c_{21}c_{11} + e_{11}e_{21} \end{bmatrix} + c_{11}c_{21} + e_{11}e_{21} \\ +c_{21}^2 + c_{22}^2 + e_{21}^2 + e_{22}^2 \end{bmatrix}$$

![](_page_42_Figure_1.jpeg)

![](_page_43_Figure_1.jpeg)

Within-traits  $P11-P12 = 0.5a_{11}^2$  $P21-P22 = 0.5a_{22}^2+0.5a_{21}^2$ 

![](_page_44_Figure_1.jpeg)

Within-traits P11-P12 =  $0.5a_{11}^2$ P21-P22 =  $0.5a_{22}^2$ + $0.5a_{21}^2$ 

Cross-traits P11-P22 =  $0.5a_{11}a_{21}$ P21-P12 =  $0.5a_{21}a_{11}$ 

# Additive Genetic Cross-Twin Covariance (DZ)

![](_page_45_Figure_1.jpeg)

# Additive Genetic Cross-Twin Covariance (MZ)

![](_page_46_Figure_1.jpeg)

$$1\otimes \Sigma_A = 1\otimes \mathbf{X}*\mathbf{X}' = \left[egin{array}{cc} a_{11}^2 & a_{11}a_{21}\ a_{21}a_{11} & (a_{21}^2+a_{22}^2) \end{array}
ight]$$

# Common Environment Cross-Twin Covariance

![](_page_47_Figure_1.jpeg)

$$1 \otimes \Sigma_C = 1 \otimes \mathbf{Y} * \mathbf{Y}' = \begin{bmatrix} c_{11}^2 & c_{11}c_{21} \\ c_{21}c_{11} & (c_{21}^2 + c_{22}^2) \end{bmatrix}$$

## **Covariance Model for Twin Pairs**

A+C | A+C+E /

- MZ:
  - Covariance
     A+C+E
     A+C\_
- DZ:
  - Covariance A+C+E | H@A+C\_
     H@A+C | A+C+E /

#### N.B. H Full 1 1 Fixed = 0.5

# Obtaining Standardised Estimates

# **Correlated Factors Solution**

![](_page_50_Figure_1.jpeg)

- Each variable decomposed into genetic/environmental components
- Correlations across variables estimated
- Results from Cholesky can be converted to this model

#### **Covariance to Correlation**

$$r_{12} = \frac{\sigma_{12}^2}{\sqrt{\sigma_{11}^2 \times \sigma_{22}^2}} = \frac{1}{\sqrt{\sigma_{11}^2}} * \sigma_{12}^2 * \frac{1}{\sqrt{\sigma_{22}^2}}$$

Using matrix algebra notation:

$$\begin{bmatrix} 1 & r_{12} \\ r_{21} & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{\sigma_{11}^2}} & 0 \\ 0 & \frac{1}{\sqrt{\sigma_{22}^2}} \end{bmatrix} * \begin{bmatrix} \sigma_{11}^2 & \sigma_{12}^2 \\ \sigma_{21}^2 & \sigma_{22}^2 \end{bmatrix} * \begin{bmatrix} \frac{1}{\sqrt{\sigma_{11}^2}} & 0 \\ 0 & \frac{1}{\sqrt{\sigma_{22}^2}} \end{bmatrix}$$

# **Genetic Correlations**

$$\begin{split} \Sigma_A &= \begin{bmatrix} a_{11}^2 & a_{11}a_{21} \\ a_{21}a_{11} & a_{21}^2 + a_{22}^2 \end{bmatrix} \\ &= \begin{bmatrix} \sigma_{A_{11}}^2 & \sigma_{A_{12}}^2 \\ \sigma_{A_{21}}^2 & \sigma_{A_{22}}^2 \end{bmatrix} \\ ^1_{r_G \ 1} \ \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{\sigma_{A_{11}}^2}} & 0 \\ 0 & \frac{1}{\sqrt{\sigma_{A_{22}}^2}} \end{bmatrix} * \begin{bmatrix} \sigma_{A_{11}}^2 & \sigma_{A_{22}}^2 \\ \sigma_{A_{21}}^2 & \sigma_{A_{22}}^2 \end{bmatrix} * \begin{bmatrix} \frac{1}{\sqrt{\sigma_{A_{11}}^2}} & 0 \\ 0 & \frac{1}{\sqrt{\sigma_{A_{22}}^2}} \end{bmatrix} \end{split}$$

# **Specification in Mx**

$$\begin{bmatrix} 1 & r_G \\ r_G & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{\sigma_{A_{11}}^2}} & 0 \\ 0 & \frac{1}{\sqrt{\sigma_{A_{22}}^2}} \end{bmatrix} * \begin{bmatrix} \sigma_{A_{11}}^2 & \sigma_{A_{12}}^2 \\ \sigma_{A_{21}}^2 & \sigma_{A_{22}}^2 \end{bmatrix} * \begin{bmatrix} \frac{1}{\sqrt{\sigma_{A_{11}}^2}} & 0 \\ 0 & \frac{1}{\sqrt{\sigma_{A_{22}}^2}} \end{bmatrix}$$

and 
$$\mathbf{I}.\mathbf{A} = \begin{bmatrix} \sigma_{A_{11}}^2 & \mathbf{0} \\ \mathbf{0} & \sigma_{A_{22}}^2 \end{bmatrix}$$

# **Interpreting Results**

- High genetic correlation = large overlap in genetic effects on the two phenotypes
- Does it mean that the phenotypic correlation between the traits is largely due to genetic effects?
  - No: the substantive importance of a particular r<sub>G</sub> depends the value of the correlation and the value of the √σ<sub>A</sub><sup>2</sup> paths i.e. importance is also determined by the heritability of each phenotype

![](_page_55_Figure_0.jpeg)

![](_page_55_Figure_1.jpeg)

Proportion of  $r_P$  due to additive genetic factors:

$$(\sqrt{h_{P1}^2} * r_G * \sqrt{h_{P2}^2})/r_P$$

 $(\sqrt{0.63} * -0.525 * \sqrt{0.33}) / -0.29 = 0.8357$ 

#### **Standardised Results**

- Begin algebra;
  - K = A & P | C & P | E & P;

End algebra;

% is the Mx operator for element division

![](_page_56_Figure_5.jpeg)

## **Example Mx Output**

- Matrix K:
  - Additive Genetic Component of ADHD = 63%, for IQ = 33%
  - % of covariance between ADHD and IQ due to A = 84%

![](_page_57_Figure_4.jpeg)

## Interpretation of Correlations

Consider two traits with a phenotypic correlation of 0.40 :

#### $h_{P1}^2 = 0.7$ and $h_{P2}^2 = 0.6$ with $r_G = .3$

- Correlation due to additive genetic effects = ?
- Proportion of phenotypic correlation attributable to additive genetic effects = ?

 $h_{P1}^2 = 0.2$  and  $h_{P2}^2 = 0.3$  with  $r_G = 0.8$ 

- Correlation due to additive genetic effects = ?
- Proportion of phenotypic correlation attributable to additive genetic effects = ?

# Interpretation of Correlations

Consider two traits with a phenotypic correlation of 0.40 :

#### $h_{P1}^2 = 0.7$ and $h_{P2}^2 = 0.6$ with $r_G = .3$

- Correlation due to additive genetic effects = 0.19
- Proportion of phenotypic correlation attributable to additive genetic effects = 0.49

 $h_{P1}^2 = 0.2$  and  $h_{P2}^2 = 0.3$  with  $r_G = 0.8$ 

- Correlation due to additive genetic effects = 0.20
- Proportion of phenotypic correlation attributable to additive genetic effects = 0.49

Weakly heritable traits can still have a large portion of their correlation attributable to genetic effects.

#### More Variables...

![](_page_60_Figure_1.jpeg)

#### More Variables...

![](_page_61_Figure_1.jpeg)

#### **Mx Parameter Matrices**

- #define nvar 3
- Begin Matrices;
- X lower nvar nvar free
- Y lower nvar nvar free
- Z lower nvar nvar free
- G Full 1 nvar free
- H Full 1 1 fix
- End Matrices;
- Begin Algebra;
- A=X\*X';
- C=Y\*Y';
- E=Z\*Z';
- P=A+C+E
- End Algebra;

- ! Genetic coefficients
  - ! C coefficients
  - ! E coefficients
  - ! Means
  - ! 0.5 for DZ A covar

- ! Gen var/cov
- ! C var/cov
- ! E var/cov

#### **Expanded Matrices**

![](_page_63_Figure_1.jpeg)