Longitudinal relationships between fat mass and physical activity

Kate Tilling, Alex Griffiths, Sam Leary

Outline

- Motivating question and data
- Basic analyses
- Multilevel models

 - Using summary measures as exposures

 - Calculating regression coefficients directly

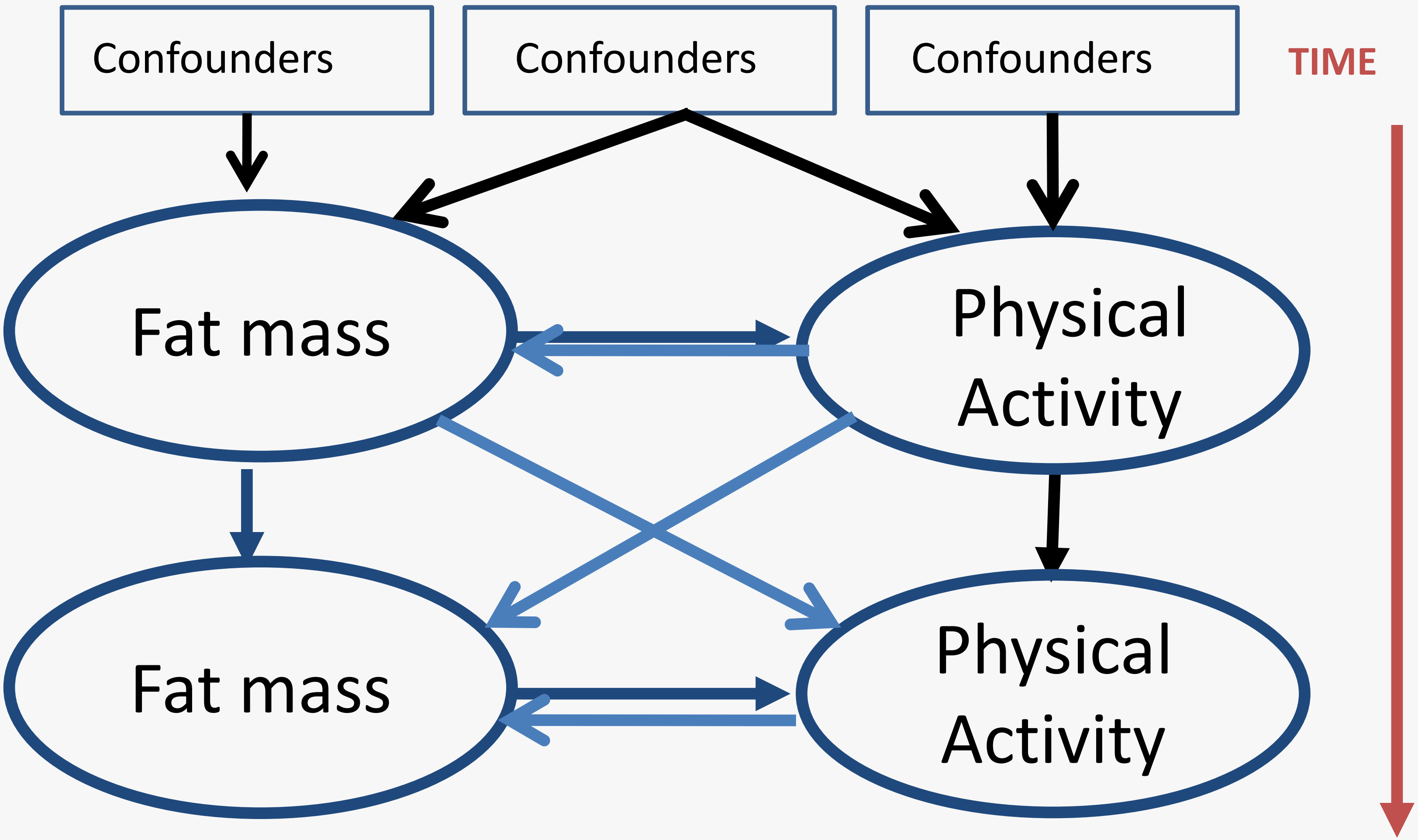
Fat mass and PA

- Understanding long-term impact of physical activity and obesity has public health and clinical importance
- Few studies able to investigate detailed trajectories – need multiple measures of PA and fat mass
- Need correct modelling of complex relationships along the lifecourse

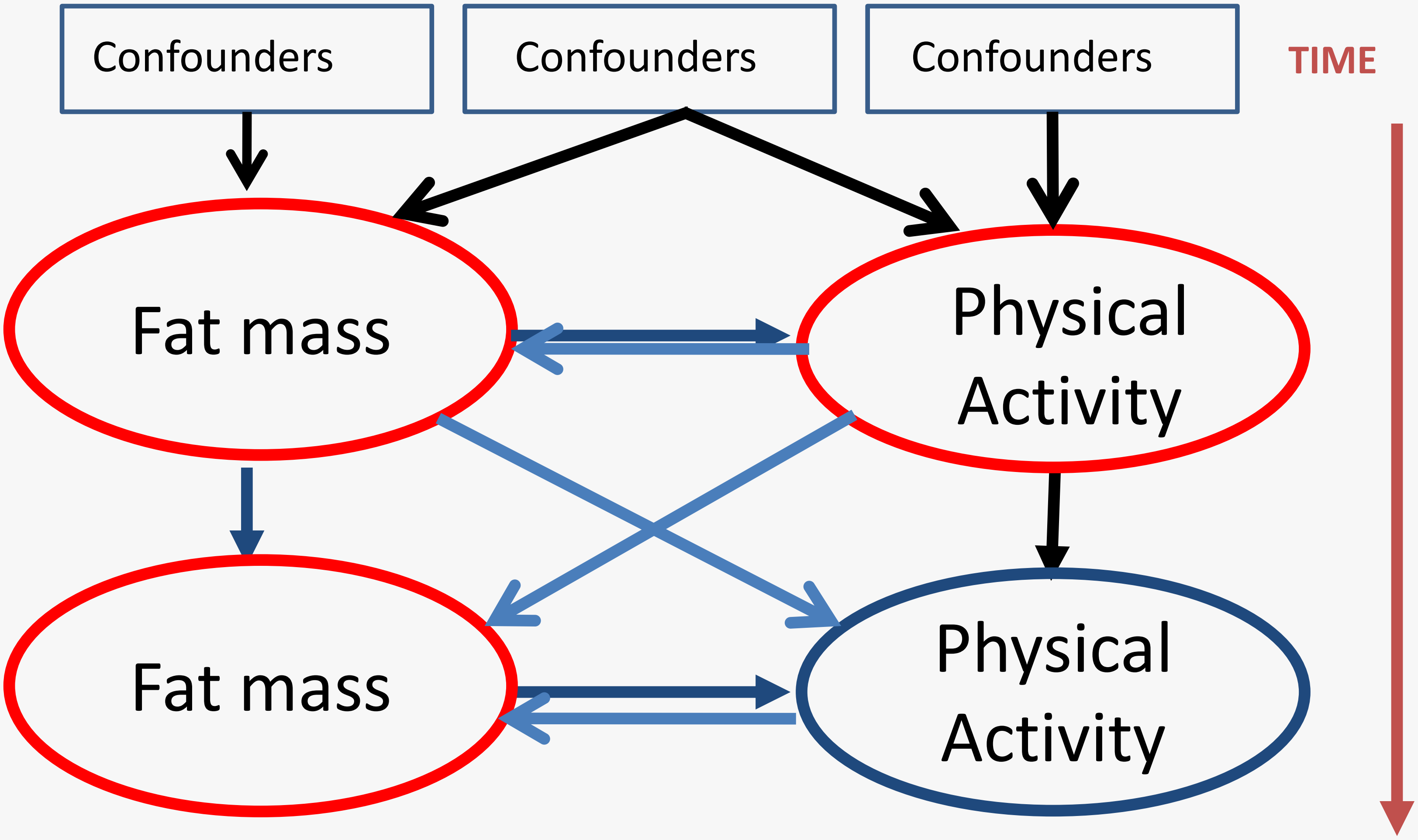
Fat mass and PA in ALSPAC

- PA measured using accelerometer – average CPM (in 100cpm) and minutes MVPA (in 15 min) over period worn
- Fat mass (kg) measured using DXA scan. Log fat mass used.
- Fat mass and PA measured at 11, 13 and 15. Only 11 and 13 used here.

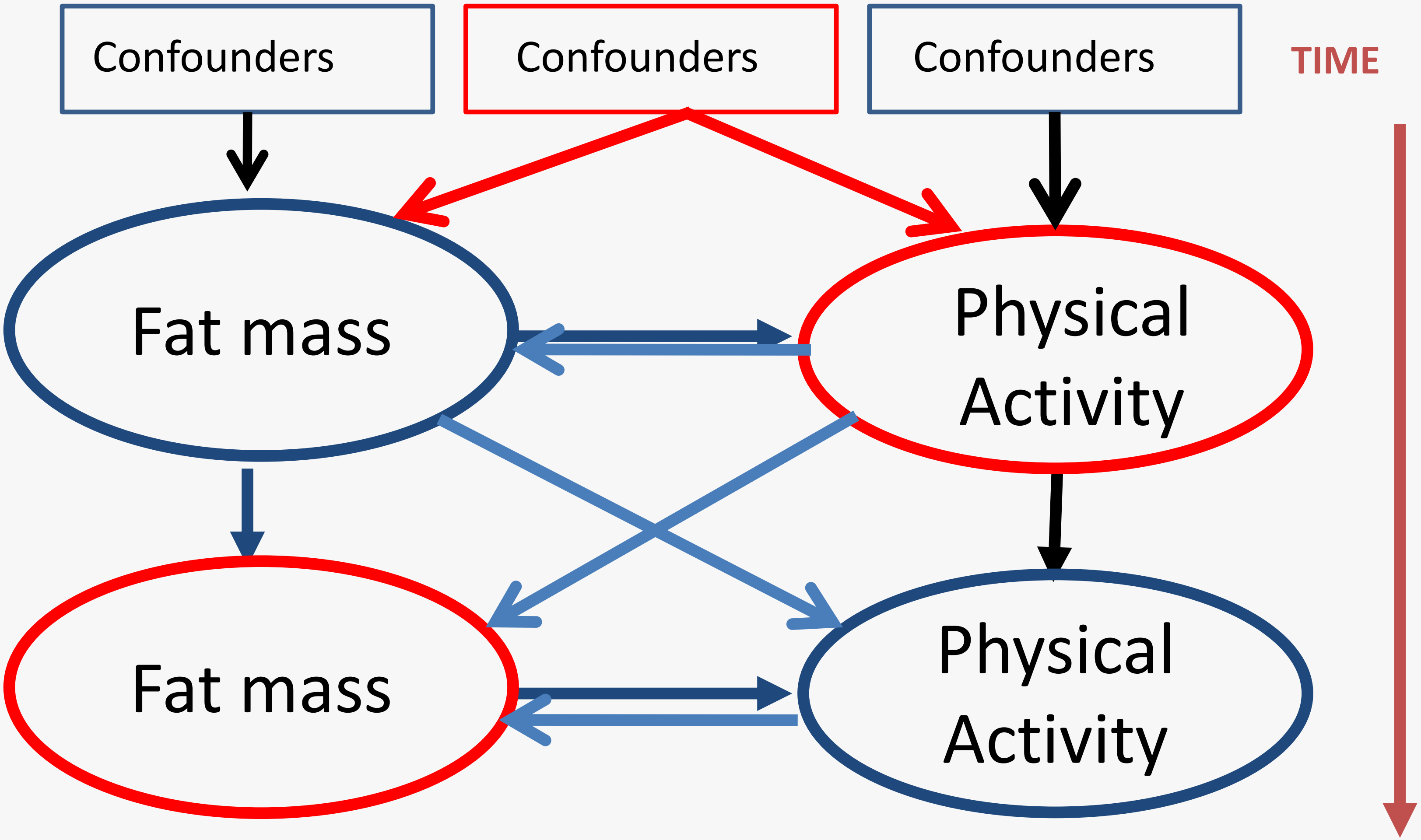
Fat mass and Physical Activity



Fat mass and Physical Activity



Fat mass and Physical Activity



Time-varying confounding

Outcome = fat mass at final age (e.g.16y)

Exposure = physical activity

Fat mass is a time-varying confounder if:

- Fat mass is associated with later PA
- PA is associated with later fat mass
- Fat mass is associated with outcome

Time-varying confounding

Usual methods will be biased

If have >2 measures:

- G-estimation (only for binary outcomes and exposures)
- Marginal Structural Model (weight by probability of having the observed PA pattern – binary exposures)
- Structural Equation Model (initialisation bias?)

Models for change

$FM_t = \alpha_{FM} + \beta_{FM} PA_t + \gamma_{FM}(X) + \varepsilon_{FMt}$ (1)

$FM_{t-1} = \alpha_{FM} + \beta_{FM} PA_{t-1} + \gamma_{FM}(X) + \varepsilon_{FMt-1}$ (2)

Take differences:

$FM_t - FM_{t-1} = \alpha_{FM} + \beta_{FM} PA_t + \gamma_{FM}(X) + \varepsilon_{FMt} - (\alpha_{FM} + \beta_{FM} PA_{t-1} + \gamma_{FM}(X) + \varepsilon_{FMt-1})$

$= \beta_{FM}(PA_t - PA_{t-1}) + \varepsilon$

Fat mass, cpm and MVPA

	N	Mean	SD/IQR
Total PA 11 (counts/min) [1]	4143	576	472, 707
Total PA 13 (counts/min) [1]	2874	512	403, 645
MVPA (min/day) 11yr[1]	4143	20	12, 31
MVPA (min/day) 13yr[1]	2874	21	12, 33
Age at PA measurement 11 (yr)	4143	11.8	0.2
Age at PA measurement 13 (yr)	2874	13.9	0.2
Fat mass 11 (kg) [1]	4367	9.8	6.7, 15.0
Fat mass 13 (kg) [1]	3889	12.0	7.6, 17.7

Log fat mass, cpm and MVPA

For boys

Regress fat at 11 on MVPA at 11 =-0.13 (95% CI -0.16, -0.11)
N=1964

Regress fat at 13 on MVPA at 11 = -0.11 (95% CI -0.14, -0.09)
N=1638

Adjusting also for fat at 11: = 0.01 (95% CI -0.003, 0.02)
N=1619

But this last analysis is now examining change in FM related to MVPA at 11 – makes less sense?

Changes in log fat mass and MVPA

For boys

Regress Δfat on ΔMVPA =-0.03 (95% CI -0.04, -0.02)

N=1130

Adjusting also for fat at 11 = -0.02 (95% CI -0.04, -0.01)

N=1130

But this last analysis is really also examining the association between fat at 13 and change in MVPA – makes less sense?

Problems with simple analyses

1. Can only use people with no missing data
2. Measurements are not taken at exactly the same ages for all individuals
3. Measurement error may vary over time
4. These analyses tell you nothing about how fat mass and physical activity change over time
5. Fat mass at 11 and 13 highly correlated (0.86): PA at 11 and 13 less so (0.37).

Need an analysis which takes into account nested structure of observations.

Extension to simple analyses

Use average and difference of Fat Mass or PA as exposures

Low correlations between average and change in FM (0.11) and PA (0.05)

Regress change in FM on MVPA average (0.006, $p=0.6$), MVPA change (-0.03, $p<0.01$) and FM average (0.04, $p=0.02$)

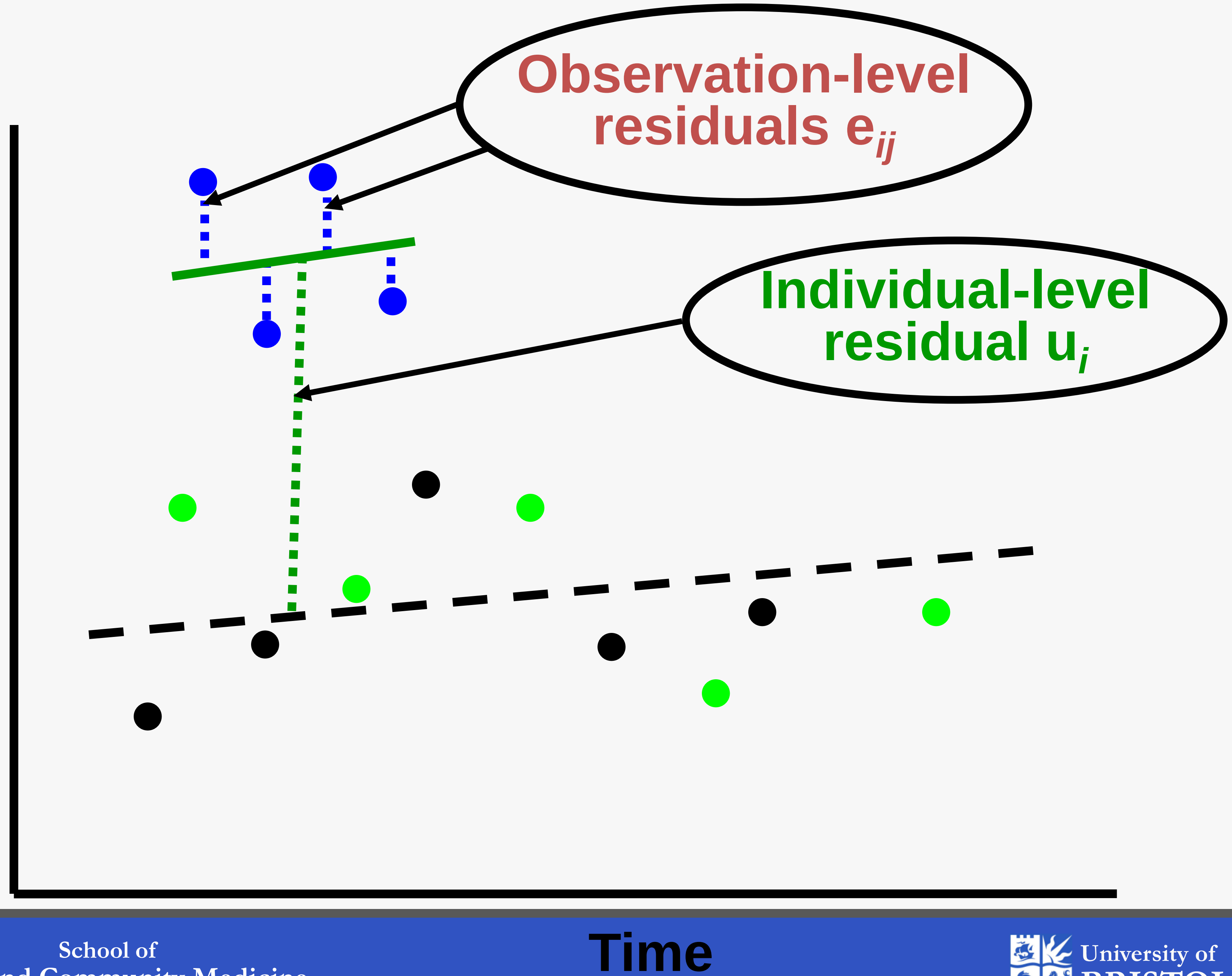
Random-effects (multilevel) models

- Level 1 (measurement on individual) indexed by *j*
- Level 2 (individual) indexed by *i*

$y_{ij} = a + bx_{ij} + \mathbf{u}_i + \mathbf{e}_{ij}$

- **Level 2** (individual) residual $\mathbf{u}_i$ represents the difference between the average regression line and the mean for that individual

- After accounting for the individual residuals, observations within individuals are assumed to be statistically independent


Observation-level residuals e_{ij}
Individual-level residual u_i
Time

Random-slopes model

$$y_{ij} = (a + u_{0i}) + (b+u_{1i})t_{ij} + e_{ij}$$

- Effect of *t* (time) now varies between individuals
- The model now estimates:
  - The regression coefficients a and b
  - The between-individual variance in the intercept
  - The between-individual variance in the slope
  - The *covariance* between the intercept and slope
- Obvious extensions to multiple covariates, including non-linear growth curves
- May also include level 1 random time effects

Multivariate multilevel model

- Can model >1 outcome sumultaneously
- Outcome measures (*k*) are nested within occasions (*j*), nested within individuals (*i*)

$$y_{ijk} = (a_k + u_{0ik}) + (b_k + u_{1ik})t_{ijk} + e_{ijk}$$

- This model estimates:
  - The coefficients a_k and b_k for each outcome
  - The between-individual variances in the intercepts
  - The between-individual variances in the slopes
  - The *covariance* between the intercepts and slopes, both within and between outcomes

Fat mass, cpm and MVPA

$$\text{resp}_{1jk} = \beta_{0jk}\text{cons.fat}_{ijk} + \beta_{5k}\text{dxaage.fat}_{ijk} + 2.460(0.148)\text{height.fat}_{ijk} + -6.029(0.390)\text{heightsq.fat}_{ijk}$$

$$\beta_{0jk} = 2.241(0.015) + v_{0k} + u_{0jk}$$

$$\beta_{5k} = -0.086(0.010) + v_{5k}$$

$$\text{resp}_{2jk} = \beta_{1jk}\text{cons.cpm}_{ijk} + \beta_{3k}\text{csaage.cpm}_{ijk}$$

$$\beta_{1jk} = 6.819(0.054) + v_{1k} + u_{1jk}$$

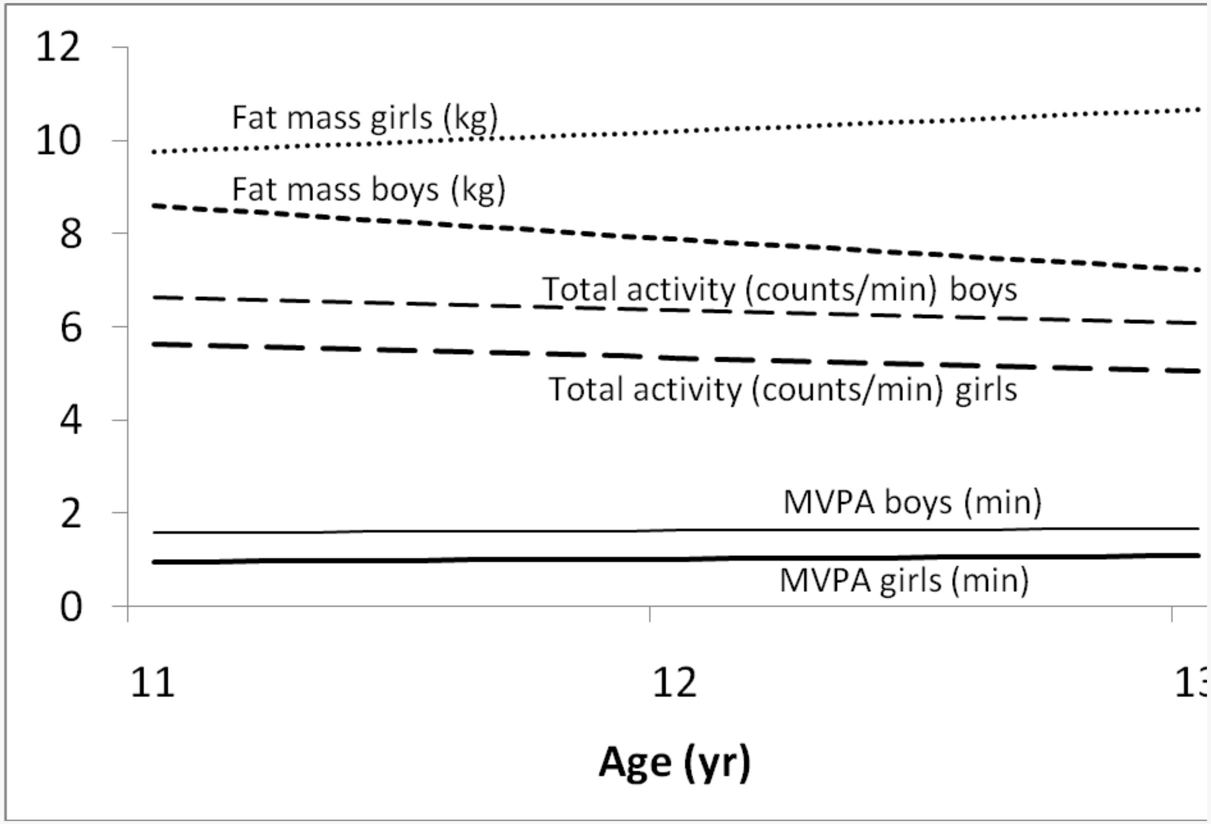
$$\beta_{3k} = -0.275(0.026) + v_{3k}$$

$$\text{resp}_{3jk} = \beta_{2jk}\text{cons.mvpa}_{ijk} + \beta_{4k}\text{csaage.mvpa}_{ijk} + 0.044(0.005)\text{mins.mvpa}_{ijk}$$

$$\beta_{2jk} = 1.549(0.055) + v_{2k} + u_{2jk}$$

$$\beta_{4k} = 0.035(0.017) + v_{4k}$$

Average trajectories



Log fat mass, cpm and MVPA

Variance of fat at age 11

Variance of MVPA at age 11

$$\Omega_v = \begin{bmatrix} 0.249(0.019) & & & & & \\ -0.299(0.031) & 4.519(0.260) & & & & \\ -0.217(0.020) & 2.926(0.106) & 1.936(0.105) & & & \\ 0.057(0.015) & -1.250(0.107) & -0.892(0.046) & 0.731(0.056) & & \\ 0.045(0.010) & -0.891(0.047) & -0.599(0.044) & 0.522(0.023) & 0.362(0.024) & \\ 0.003(0.007) & 0.034(0.010) & 0.022(0.007) & -0.016(0.004) & -0.011(0.003) & 0.008(0.004) \end{bmatrix}$$

Covariance of fat and MVPA at age 11

$$\text{Regress fat at 11 on MVPA at 11} = \frac{\text{(covariance fat and MVPA at 11)}}{\text{(variance MVPA at 11)}}$$

$$= -0.217/1.936$$

$$= -0.112$$

So fat is lower by 11% in boys with 15 mins more MVPA at 11

Uses data from all 2187 boys

Log fat mass, cpm and MVPA

Need confidence intervals around these regression coefficients

1) WINBUGS
   Fit same model in WINBUGS, estimate distribution of the regression coefficients directly

2) Prediction of individual estimates
   Use individual random-effects to predict log fat mass, MVPA and cpm at exactly 11 and 13 years.
   Use these predicted values as exposures/outcomes in usual regression models

Log fat mass, cpm and MVPA

Coefficients

		WinBUGS		Regressing random effects	
Outcome	**Exposure**	**Beta**	**95% CI**	**Beta**	**95%CI**
Fatness at 13	MVPA at 11	-0.1275	-0.1544, -0.1008	-0.1254	-0.1530, -0.0978
Fatness at 13	MVPA at 13	-0.0954	-0.1152, -0.0763	-0.0985	-0.1248, -0.0723
Δ Fatness 11-13	Δ MVPA 12-14	-0.0246	-0.0376, -0.0113	-0.0254	-0.0395, -0.0113

Log fat mass, cpm and MVPA

Problem of specifying the model remains.

Could use the multilevel model to predict PA and FM at specific times, then use these in a structural equation model.

Latent class analysis to identify possible subgroups?

Summary - advantages of ML models

- Summarise patterns of fat gain and changes in PA
- Not restricted to measures at arbitrary times
- Allow for measurement error, collinearity between repeated measures, and missing data
- Exposures can be used in other models