A joint modelling approach to relate within-individual variability in a repeatedly measured exposure to a future outcome, allowing for measurement error in the repeated measures

Richard Parker

Harvey Goldstein

Jon Heron

Laura Howe

George Leckie

Graciela Muniz-Terrera

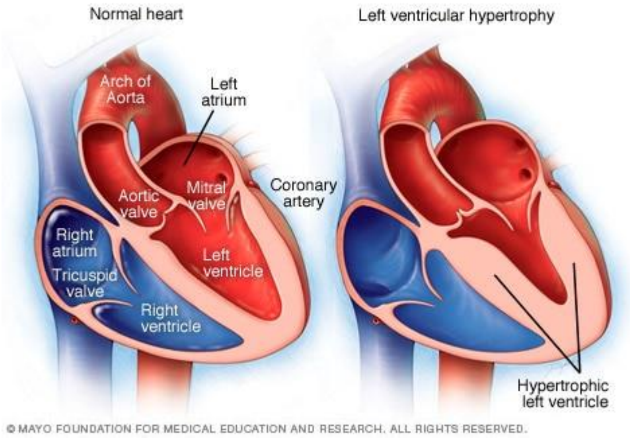
Kate Tilling

Example: does blood pressure (BP) history predict later biomarkers of cardiovascular disease (CVD)?

- Systolic blood pressure (*mmHg*): repeatedly-measured outcome
- Left ventricular mass ($g/m^{2.7}$): later outcome

Pexels

© Mayo Foundation for Medical Education and Research



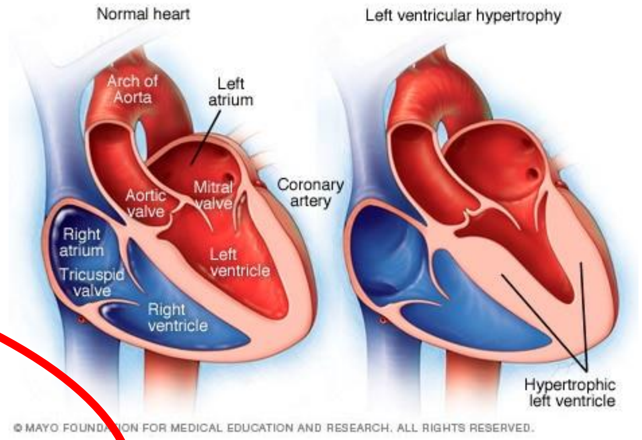
"Left ventricular hypertrophy is both a major maladaptive response to chronic pressure overload and an important risk factor in patients with hypertension."

Katholi & Couri, 2011

Example: does blood pressure (BP) history predict later biomarkers of cardiovascular disease (CVD)?

- Systolic blood pressure (*mmHg*): repeatedly-measured outcome
- Left ventricular mass (*g/m*$^{2.7}$): later outcome

Pexels



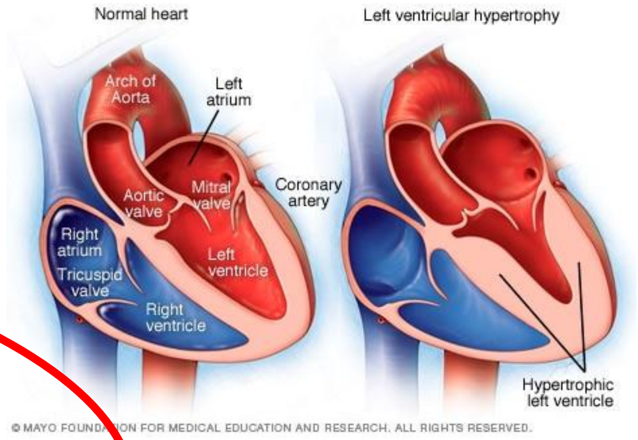
© Mayo Foundation for Medical Education and Research

There's been a long-standing interest in investigating whether **mean, or mean trajectory,** of repeatedly-measured BP predicts later signs of CVD…

Example: does blood pressure (BP) history predict later biomarkers of cardiovascular disease (CVD)?

- Systolic blood pressure (*mmHg*): repeatedly-measured outcome
- Left ventricular mass (*g/m*$^{2.7}$): later outcome

Pexels

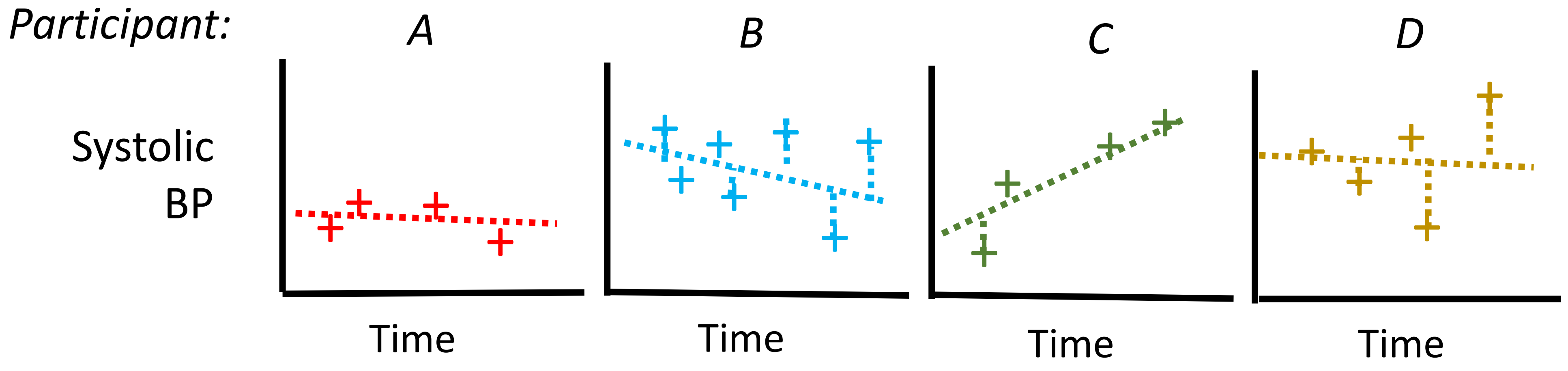


© Mayo Foundation for Medical Education and Research

…but what about **within-individual variability**?

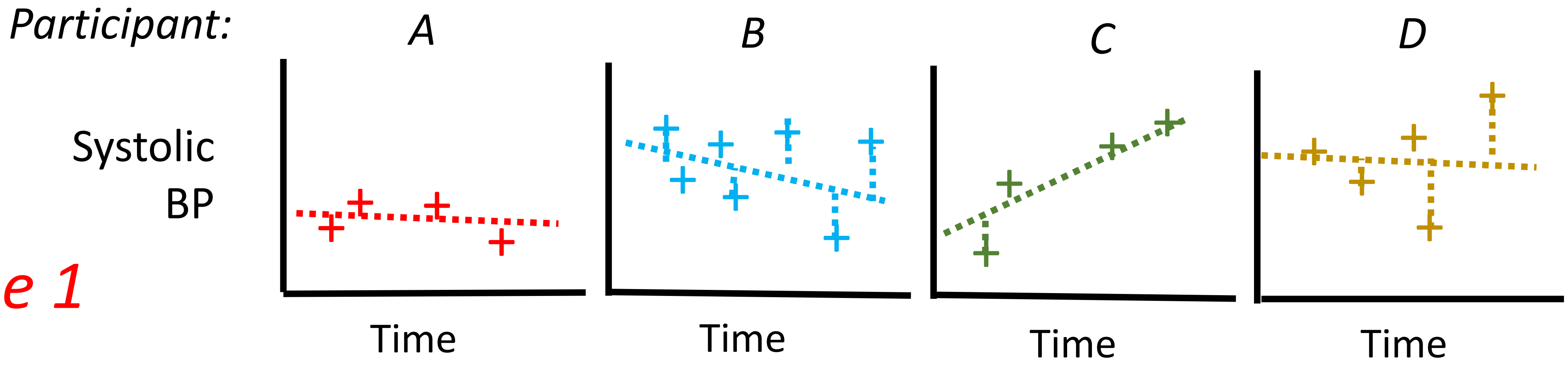
Within-individual variability in BP over the longer-term is an independent CVD risk factor over & above mean blood pressure (e.g. Rothwell, 2010)

Typical procedures to investigate association of within-individual variability with later outcome include…



Typical procedures to investigate association of within-individual variability with later outcome include…

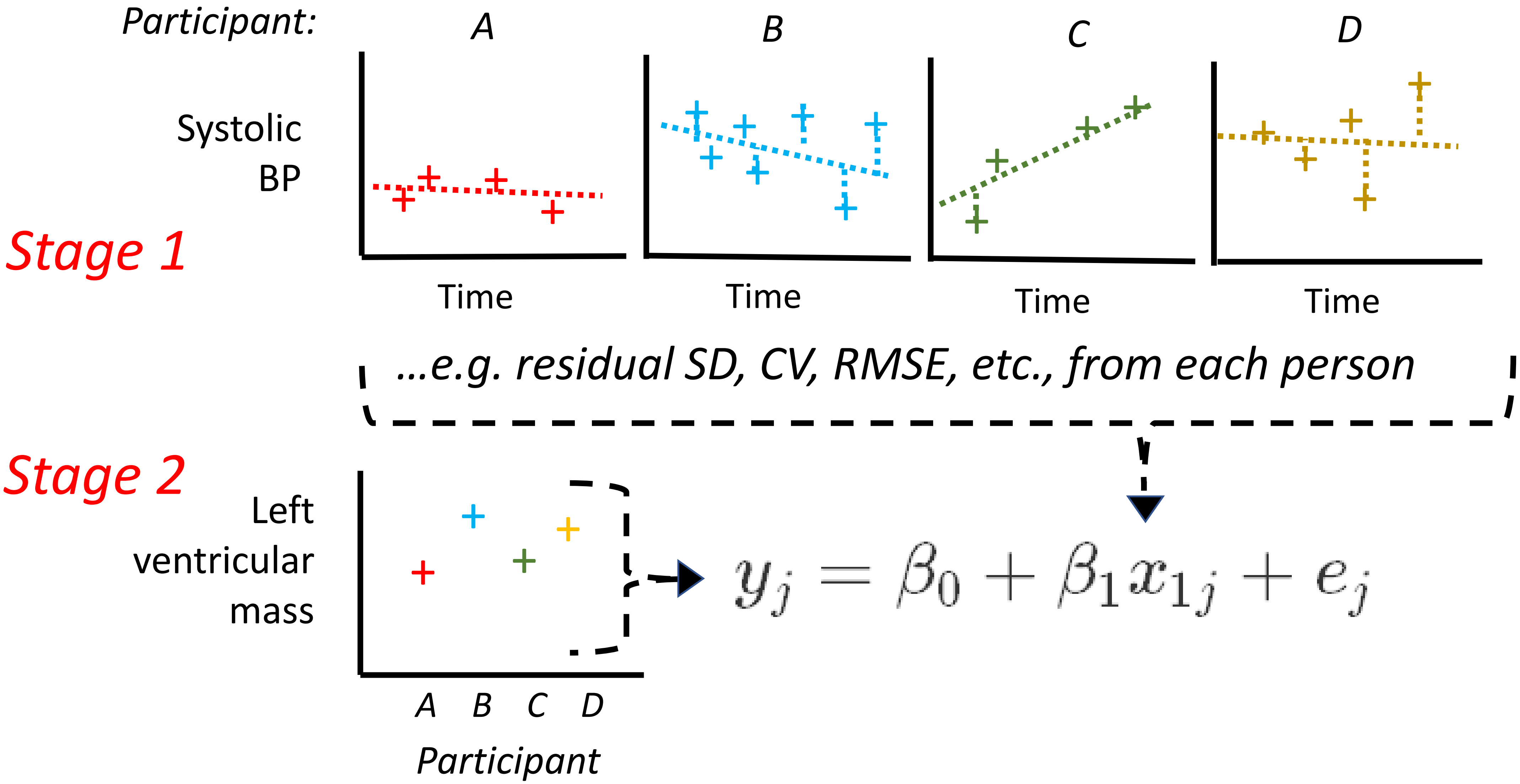
Stage 1



…e.g. residual SD, CV, RMSE, etc., from each person

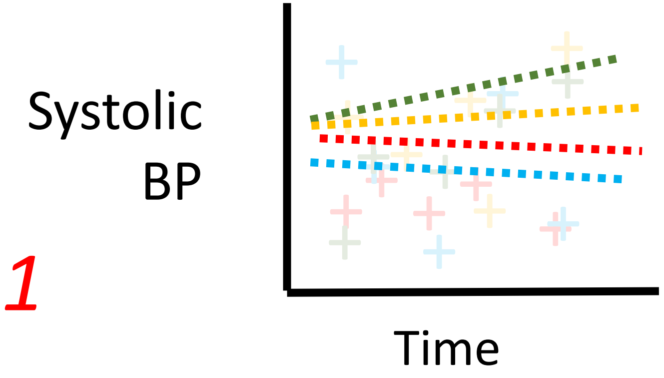
$$x_{1j}$$

Typical procedures to investigate association of within-individual variability with later outcome include…



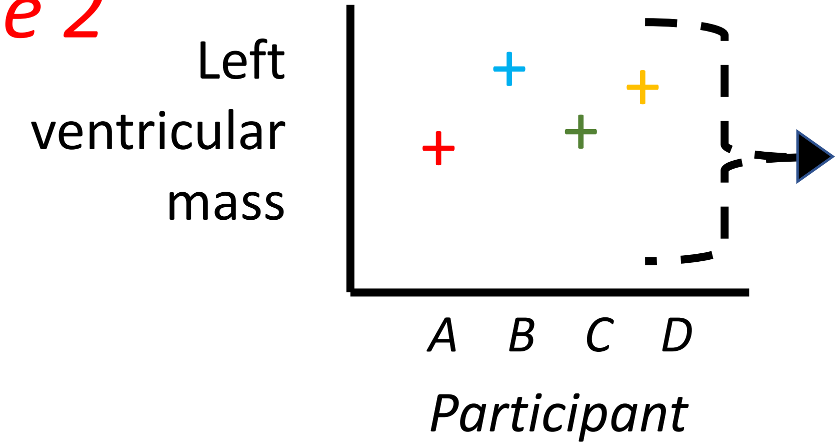
Typical procedures to investigate association of within-individual variability with later outcome include…

Stage 1



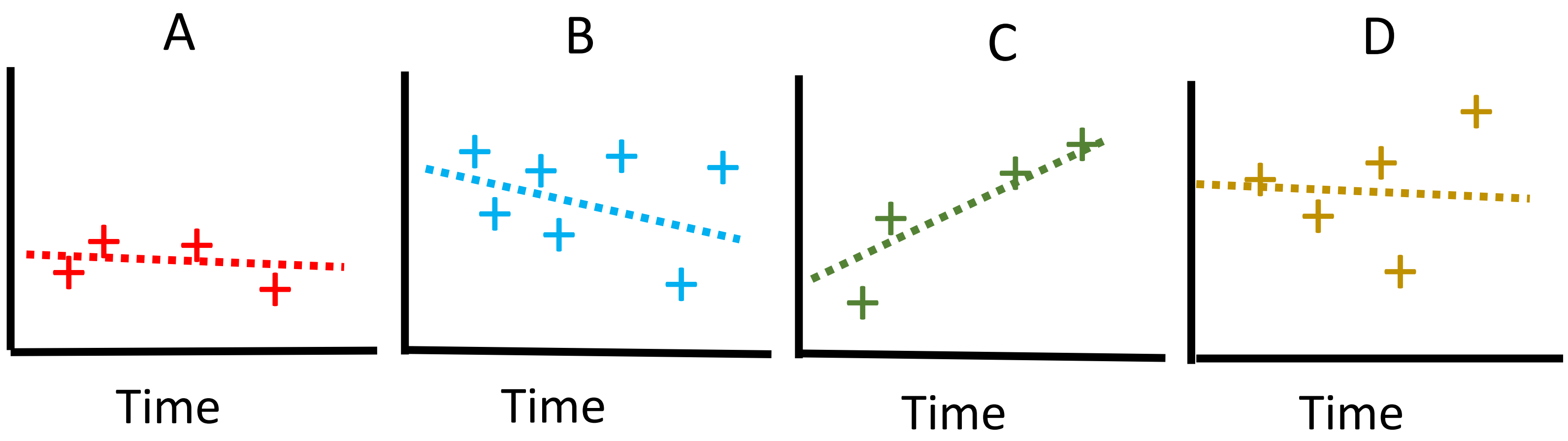
…or e.g. function of residuals from random effects model

Stage 2



$$y_j = \beta_0 + \beta_1 x_{1j} + e_j$$

These 2-stage approaches have important limitations...



Typically a large element of **sampling error** in the estimate of within-individual variability as derived in Stage 1…

…but **information regarding precision of this estimate is lost** between the two stages…

…resulting in regression dilution or attenuation **bias** (towards the null) when fitting model in Stage 2 (akin to measurement error in predictor).

…to address this issue, we use a **joint model, with shared random effects**, to simultaneously estimate within-individual variability in the repeatedly-measured exposure and its association with the later outcome.

Will demonstrate by:

1. Introducing dataset, then stepping through simplified example (with just one covariate)…
2. …concluding by illustrating with results (from more complex models).

ALSPAC Dataset

n = 1,986

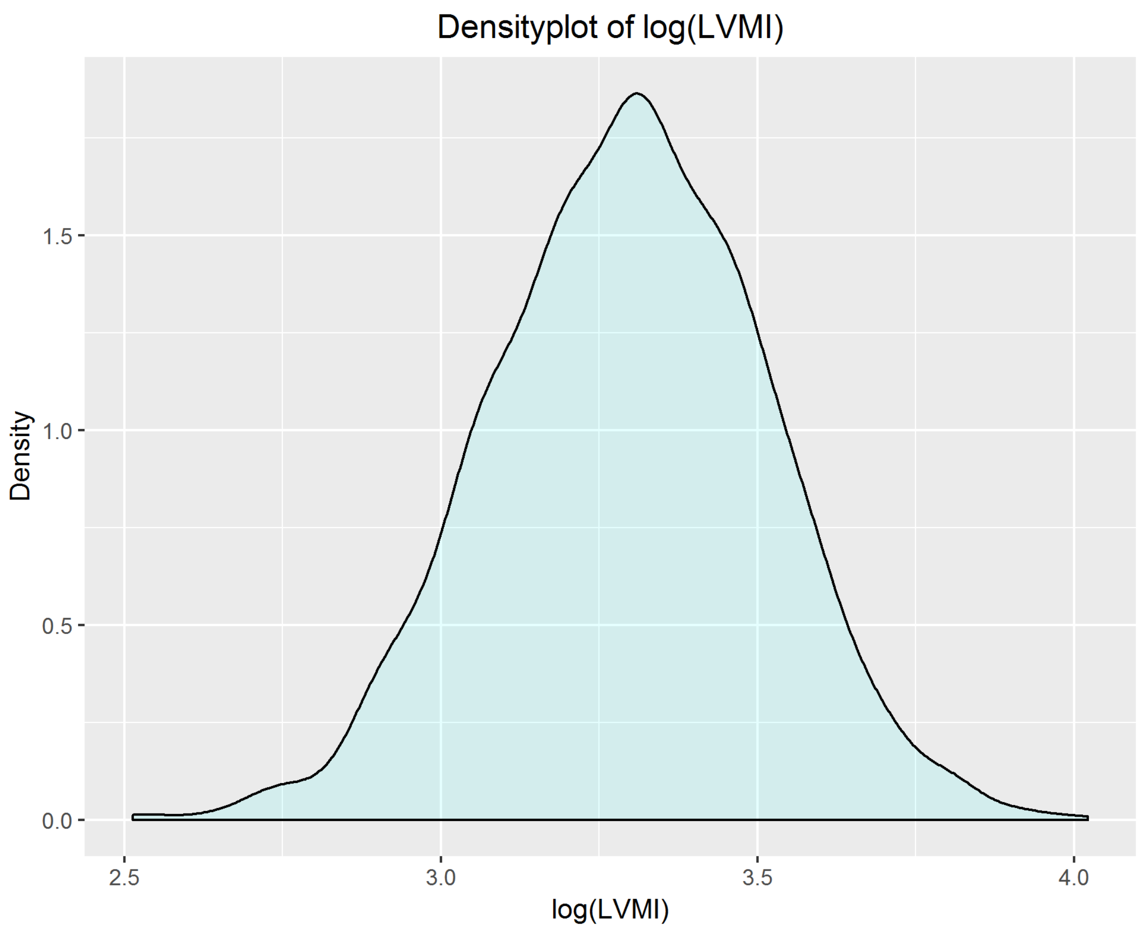
…of the ALSPAC cohort had their systolic blood pressure (SBP) recorded on at least one occasion prior to…

Repeatedly-measured outcome

ALSPAC Dataset

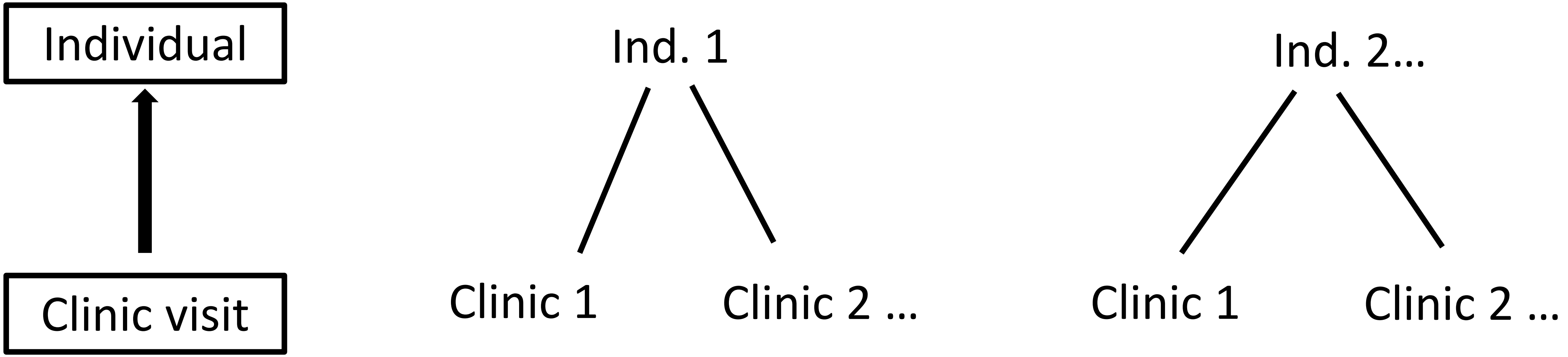
n = 1,986

…having echocardiography at *c.*18 years of age.



Later (individual-level) outcome

Example starts with a 2-level model…

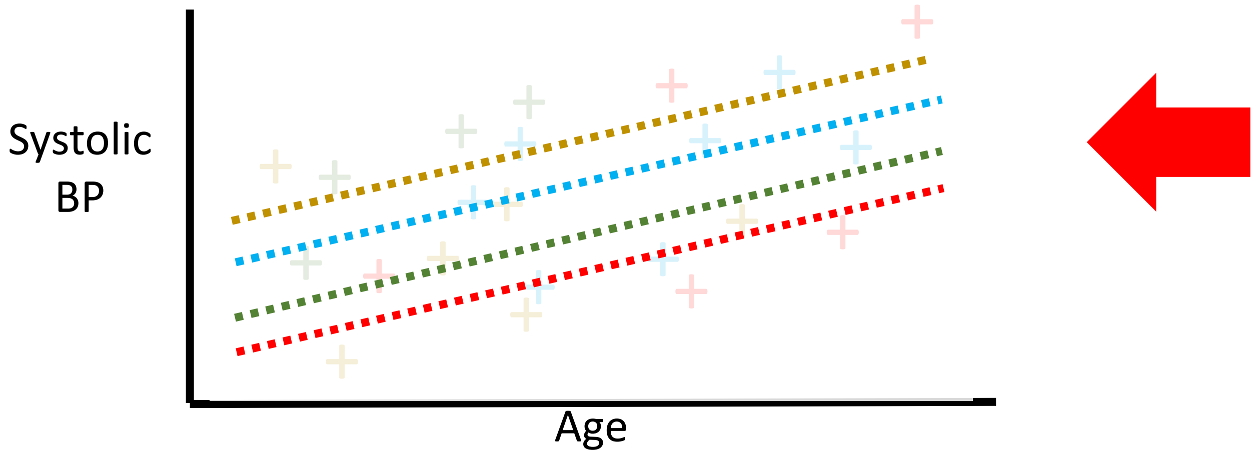


Random (intercept &) slope with complex level 1 variation

$$\text{Clinic BP}_{ij} = \beta_0 + \beta_1 \text{age}_{ij} + u_{0j} + u_{1j}\text{age}_{ij} + e_{ij}$$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim \text{N}\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{u0} & \\ \sigma_{u01} & \sigma^2_{u1} \end{pmatrix}\right]$$

$$e_{ij} \sim \text{N}(0, \sigma^2_{eij}), \quad \ln(\sigma^2_{eij}) = \alpha_0 + \alpha_1 \text{age}_{ij}$$

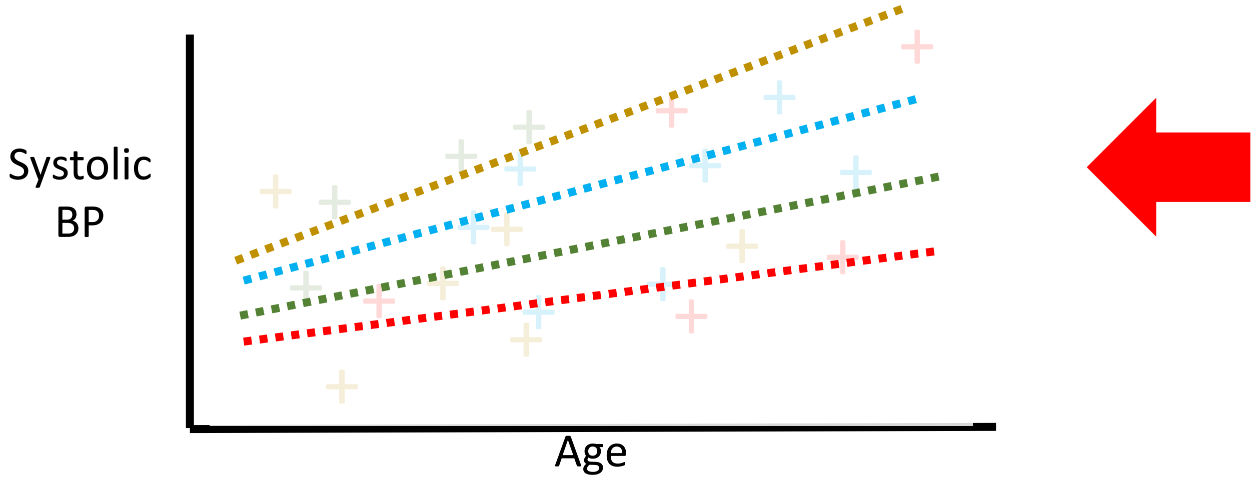


Random (intercept &) slope with complex level 1 variation

$$\text{Clinic BP}_{ij} = \beta_0 + \beta_1 \text{age}_{ij} + u_{0j} + u_{1j}\text{age}_{ij} + e_{ij}$$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim \text{N}\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{u0} & \\ \sigma_{u01} & \sigma^2_{u1} \end{pmatrix}\right]$$

$$e_{ij} \sim \text{N}(0, \sigma^2_{eij}), \quad \ln(\sigma^2_{eij}) = \alpha_0 + \alpha_1 \text{age}_{ij}$$

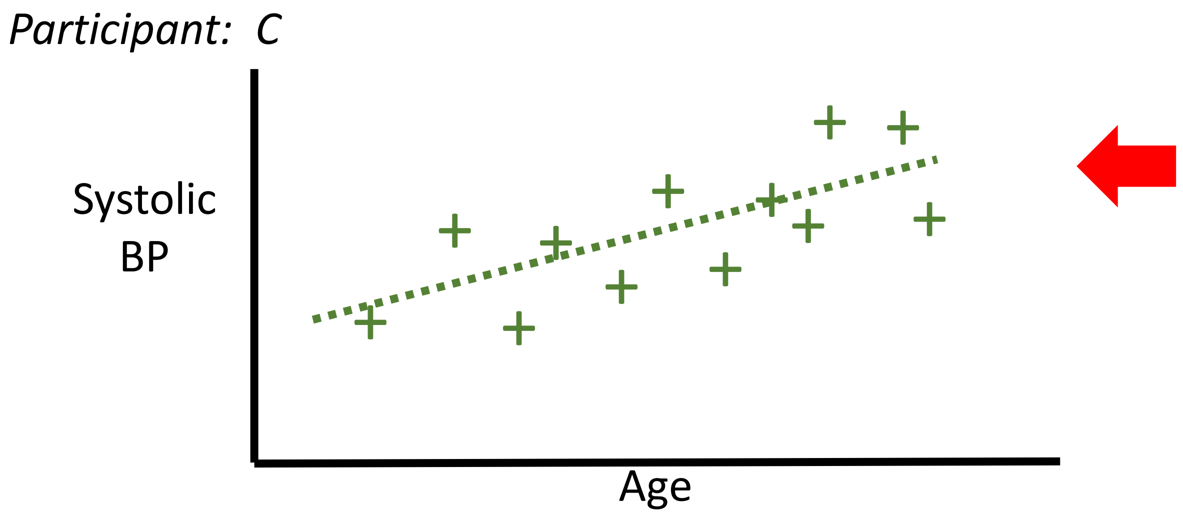


Random (intercept &) slope with complex level 1 variation

$$\text{Clinic BP}_{ij} = \beta_0 + \beta_1 \text{age}_{ij} + u_{0j} + u_{1j}\text{age}_{ij} + e_{ij}$$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim \text{N}\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{u0} & \\ \sigma_{u01} & \sigma^2_{u1} \end{pmatrix}\right]$$

$$e_{ij} \sim \text{N}(0, \sigma^2_{eij}), \quad \ln(\sigma^2_{eij}) = \alpha_0 + \alpha_1 \text{age}_{ij}$$



...constant within-individual variation would assume (e.g.) this

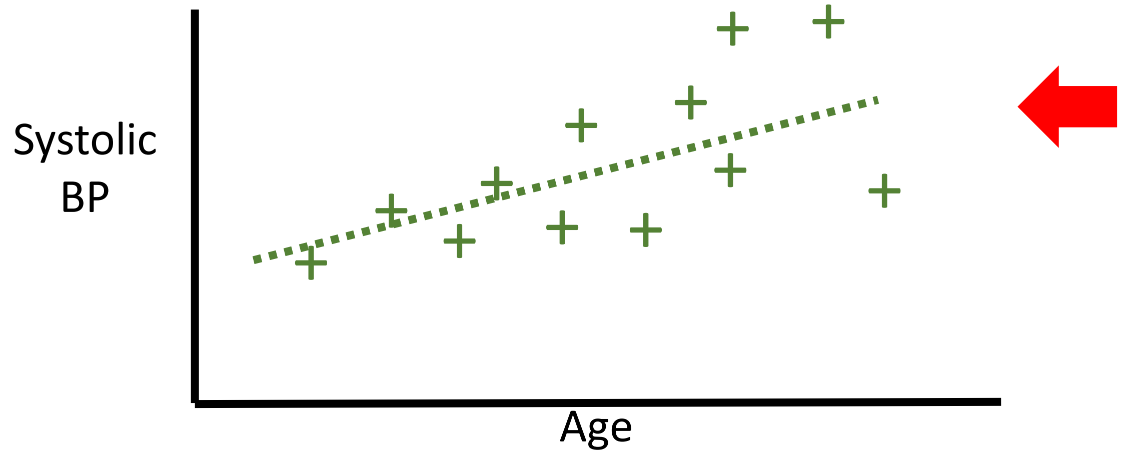
Random (intercept &) slope with complex level 1 variation

$$\text{Clinic BP}_{ij} = \beta_0 + \beta_1 \text{age}_{ij} + u_{0j} + u_{1j}\text{age}_{ij} + e_{ij}$$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim \text{N}\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{u0} & \\ \sigma_{u01} & \sigma^2_{u1} \end{pmatrix}\right]$$

$$e_{ij} \sim \text{N}(0, \sigma^2_{eij}), \quad \ln(\sigma^2_{eij}) = \alpha_0 + \alpha_1 \text{age}_{ij}$$

Participant: C



...but allowing it to be a function of covariates allows for (e.g.) this

Random (intercept &) slope with complex level 1 variation

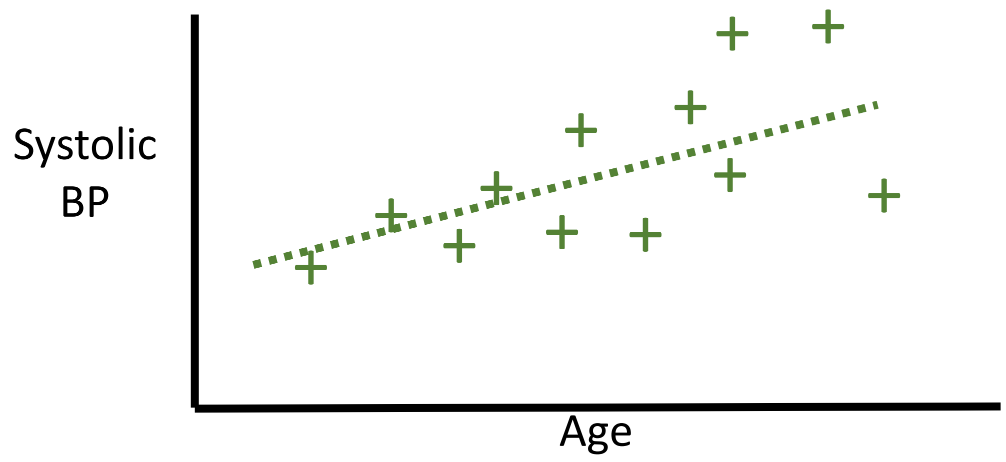
$$\text{Clinic BP}_{ij} = \beta_0 + \beta_1 \text{age}_{ij} + u_{0j} + u_{1j}\text{age}_{ij} + e_{ij}$$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim \text{N}\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{u0} & \\ \sigma_{u01} & \sigma^2_{u1} \end{pmatrix}\right]$$

$$e_{ij} \sim \text{N}(0, \sigma^2_{eij}), \quad \ln(\sigma^2_{eij}) = \alpha_0 + \alpha_1 \text{age}_{ij}$$

(log link ensures within-individual variance remains positive)

Participant: C



...but allowing it to be a function of covariates allows for (e.g.) this

…adding a random effect for within-individual variance

$$\text{Clinic BP}_{ij} = \beta_0 + \beta_1 \text{age}_{ij} + u_{0j} + u_{1j}\text{age}_{ij} + e_{ij}$$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{pmatrix} \sim \text{N}\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{u0} & & \\ \sigma_{u01} & \sigma^2_{u1} & \\ \sigma_{u02} & \sigma_{u12} & \sigma^2_{u2} \end{pmatrix}\right]$$

$$e_{ij} \sim \text{N}(0, \sigma^2_{eij}), \quad \ln(\sigma^2_{eij}) = \alpha_0 + \alpha_1 \text{age}_{ij} + u_{2j}$$

"Are some people more variable than others?"

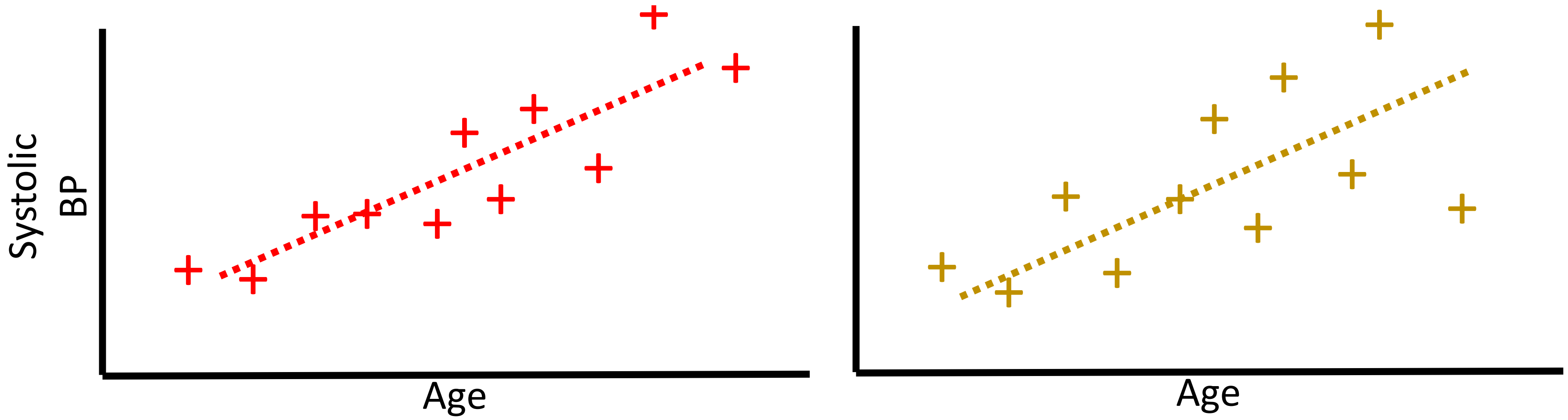
(…having adjusted for other covariates and random effects in the model).

…adding a random effect for within-individual variance

$$\text{Clinic BP}_{ij} = \beta_0 + \beta_1 \text{age}_{ij} + u_{0j} + u_{1j}\text{age}_{ij} + e_{ij}$$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{pmatrix} \sim \text{N}\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{u0} & & \\ \sigma_{u01} & \sigma^2_{u1} & \\ \sigma_{u02} & \sigma_{u12} & \sigma^2_{u2} \end{pmatrix}\right]$$

$$e_{ij} \sim \text{N}(0, \sigma^2_{eij}), \quad \ln(\sigma^2_{eij}) = \alpha_0 + \alpha_1 \text{age}_{ij} + u_{2j}$$



…adding a random effect for within-individual variance

$$\text{Clinic BP}_{ij} = \beta_0 + \beta_1 \text{age}_{ij} + u_{0j} + u_{1j}\text{age}_{ij} + e_{ij}$$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{pmatrix} \sim \text{N}\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{u0} & & \\ \sigma_{u01} & \sigma^2_{u1} & \\ \sigma_{u02} & \sigma_{u12} & \sigma^2_{u2} \end{pmatrix}\right]$$

$$e_{ij} \sim \text{N}(0, \sigma^2_{eij}), \quad \ln(\sigma^2_{eij}) = \alpha_0 + \alpha_1 \text{age}_{ij} + u_{2j}$$

Hedeker et al. (2008): mixed-effects location scale model

- *random scale effects*
- *random location effects*

…adding a random effect for within-individual variance

$$\text{Clinic BP}_{ij} = \beta_0 + \beta_1 \text{age}_{ij} + u_{0j} + u_{1j}\text{age}_{ij} + e_{ij}$$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{pmatrix} \sim \text{N}\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{u0} & & \\ \sigma_{u01} & \sigma^2_{u1} & \\ \sigma_{u02} & \sigma_{u12} & \sigma^2_{u2} \end{pmatrix}\right]$$

$$e_{ij} \sim \text{N}(0, \sigma^2_{eij}), \quad \ln(\sigma^2_{eij}) = \alpha_0 + \alpha_1 \text{age}_{ij} + u_{2j}$$

Hedeker et al. (2008): mixed-effects location scale model

- *random scale effects*
- *random location effects*

…adding a random effect for within-individual variance

$$\text{Clinic BP}_{ij} = \beta_0 + \beta_1 \text{age}_{ij} + u_{0j} + u_{1j}\text{age}_{ij} + e_{ij}$$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{pmatrix} \sim \text{N}\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{u0} & & \\ \sigma_{u01} & \sigma^2_{u1} & \\ \sigma_{u02} & \sigma_{u12} & \sigma^2_{u2} \end{pmatrix}\right]$$

$$e_{ij} \sim \text{N}(0, \sigma^2_{eij}), \quad \ln(\sigma^2_{eij}) = \alpha_0 + \alpha_1 \text{age}_{ij} + u_{2j}$$

Hedeker et al. (2008): mixed-effects location scale model

- *random scale effects*
- *random location effects*

…adding a random effect for within-individual variance

$$\text{Clinic BP}_{ij} = \beta_0 + \beta_1 \text{age}_{ij} + u_{0j} + u_{1j}\text{age}_{ij} + e_{ij}$$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{pmatrix} \sim \text{N}\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{u0} & & \\ \sigma_{u01} & \sigma^2_{u1} & \\ \sigma_{u02} & \sigma_{u12} & \sigma^2_{u2} \end{pmatrix}\right]$$

$$e_{ij} \sim \text{N}(0, \sigma^2_{eij}), \quad \ln(\sigma^2_{eij}) = \alpha_0 + \alpha_1 \text{age}_{ij} + u_{2j}$$

Hedeker et al. (2008): mixed-effects location scale model

- *random scale effects*
- *random location effects*

Don Hedeker presented this work at the MLM conference in 2009!

…adding the later (individual-level) outcome: joint model

$$\text{Clinic BP}_{ij} = \beta_0 + \beta_1 \text{age}_{ij} + u_{0j} + u_{1j}\text{age}_{ij} + e_{ij}$$

$$\log(\text{LVMI})_j = \gamma_0 + \gamma_1 u_{0j} + \gamma_2 u_{1j} + \gamma_3 u_{2j} + u_{3j}$$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \\ u_{3j} \end{pmatrix} \sim \text{N}\left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{u0} & & & \\ \sigma_{u01} & \sigma^2_{u1} & & \\ \sigma_{u02} & \sigma_{u12} & \sigma^2_{u2} & \\ 0 & 0 & 0 & \sigma^2_{u3} \end{pmatrix}\right]$$

$$e_{ij} \sim \text{N}(0, \sigma^2_{eij}), \quad \ln(\sigma^2_{eij}) = \alpha_0 + \alpha_1 \text{age}_{ij} + u_{2j}$$

…adding the later (individual-level) outcome: joint model

$$\text{Clinic BP}_{ij} = \beta_0 + \beta_1 \text{age}_{ij} + u_{0j} + u_{1j}\text{age}_{ij} + e_{ij}$$

$$\log(\text{LVMI})_j = \gamma_0 + \gamma_1 u_{0j} + \gamma_2 u_{1j} + \gamma_3 u_{2j} + u_{3j}$$

Later (individual-level) outcome

$$\begin{pmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \\ u_{3j} \end{pmatrix} \sim \text{N}\left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{u0} & & & \\ \sigma_{u01} & \sigma^2_{u1} & & \\ \sigma_{u02} & \sigma_{u12} & \sigma^2_{u2} & \\ 0 & 0 & 0 & \sigma^2_{u3} \end{pmatrix}\right]$$

$$e_{ij} \sim \text{N}(0, \sigma^2_{eij}), \quad \ln(\sigma^2_{eij}) = \alpha_0 + \alpha_1 \text{age}_{ij} + u_{2j}$$

…adding the later (individual-level) outcome: joint model

$$\text{Clinic BP}_{ij} = \beta_0 + \beta_1 \text{age}_{ij} + u_{0j} + u_{1j}\text{age}_{ij} + e_{ij}$$

$$\log(\text{LVMI})_j = \gamma_0 + \gamma_1 u_{0j} + \gamma_2 u_{1j} + \gamma_3 u_{2j} + u_{3j}$$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \\ u_{3j} \end{pmatrix} \sim \text{N}\left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{u0} & & & \\ \sigma_{u01} & \sigma^2_{u1} & & \\ \sigma_{u02} & \sigma_{u12} & \sigma^2_{u2} & \\ 0 & 0 & 0 & \sigma^2_{u3} \end{pmatrix}\right]$$

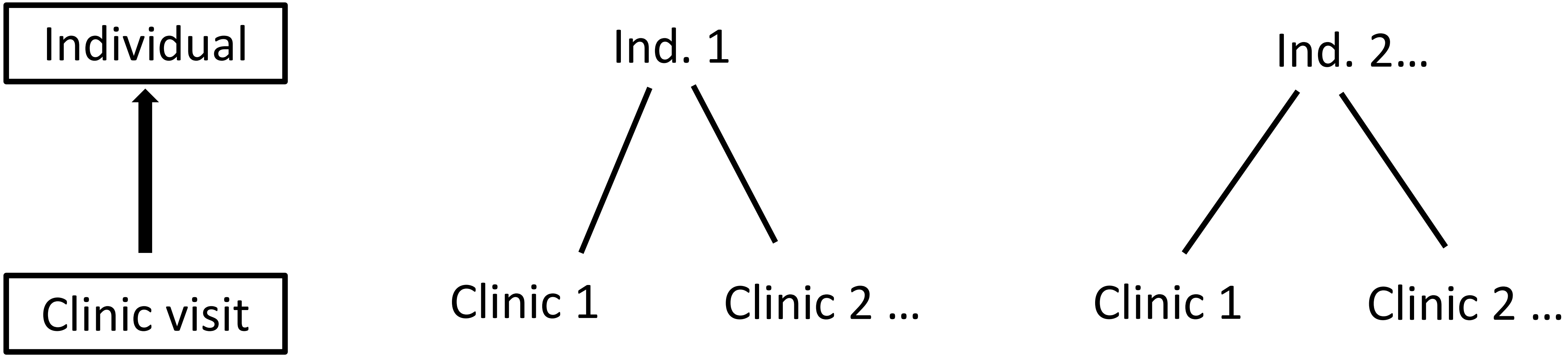
$$e_{ij} \sim \text{N}(0, \sigma^2_{eij}), \quad \ln(\sigma^2_{eij}) = \alpha_0 + \alpha_1 \text{age}_{ij} + u_{2j}$$

Later (individual-level) outcome

Do estimates of the mean, slope and/or within-individual variability predict later outcome?

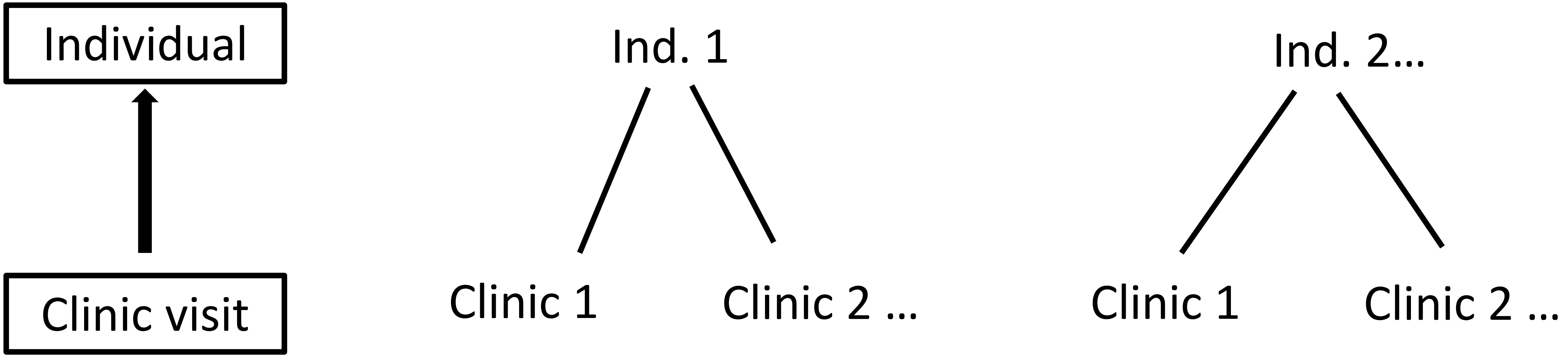
…uncertainty in estimates of random effects incorporated into estimates of their effect on later outcome.

Example starts with a 2-level model…



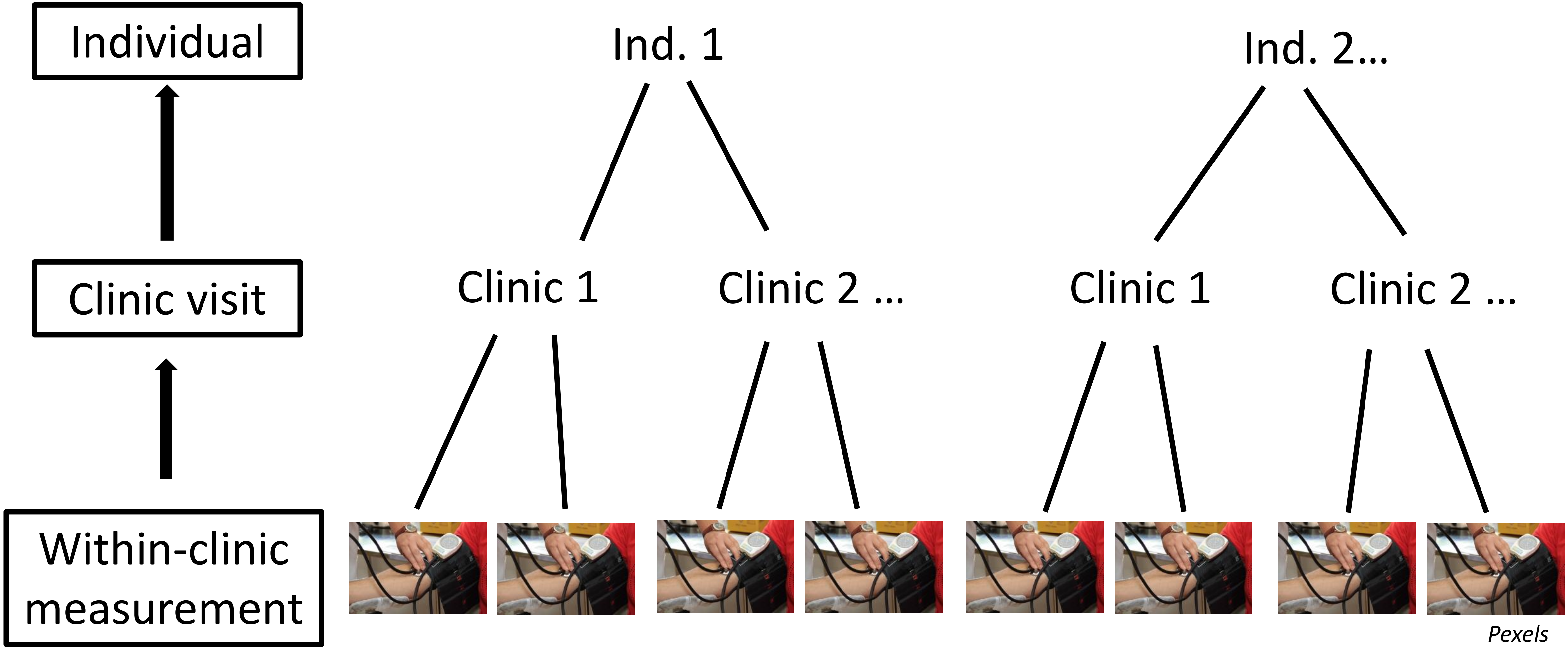
NB: Clinic BP outcome in these initial models is the mean of two measurements taken in that clinic

Example starts with a 2-level model…



…but we have access to those two individual measurements, so…

…can expand to 3 levels, distinguishing within-clinic (short term) from between-clinic (longer term) within-individual variability



Pexels

…can expand to 3 levels, distinguishing within-clinic (short term) from between-clinic (longer term) within-individual variability

$$\text{BP}_{ijk} = \beta_0 + \beta_1 \text{age}_{jk} + v_{0k} + v_{1k}\text{age}_{jk} + u_{jk} + e_{ijk}$$

$$\log(\text{LVMI})_k = \gamma_0 + \gamma_1 v_{0k} + \gamma_2 v_{1k} + \gamma_3 v_{2k} + v_{3k}$$

$$\begin{pmatrix} v_{0k} \\ v_{1k} \\ v_{2k} \\ v_{3k} \end{pmatrix} \sim \text{N}\left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{v0} & & & \\ \sigma_{v01} & \sigma^2_{v1} & & \\ \sigma_{v02} & \sigma_{v12} & \sigma^2_{v2} & \\ 0 & 0 & 0 & \sigma^2_{v3} \end{pmatrix}\right]$$

$$u_{jk} \sim \text{N}(0, \sigma^2_{ujk}), \quad \ln(\sigma^2_{ujk}) = \alpha_0 + \alpha_1 \text{age}_{jk} + v_{2k}$$

$$e_{ijk} \sim \text{N}(0, \sigma^2_e)$$

…can expand to 3 levels, distinguishing within-clinic (short term) from between-clinic (longer term) within-individual variability

$$\mathrm{BP}_{ijk} = \beta_0 + \beta_1 \mathrm{age}_{jk} + v_{0k} + v_{1k}\mathrm{age}_{jk} + u_{jk} + e_{ijk}$$

$$\log(\mathrm{LVMI})_k = \gamma_0 + \gamma_1 v_{0k} + \gamma_2 v_{1k} + \gamma_3 v_{2k} + v_{3k}$$

$$\begin{pmatrix} v_{0k} \\ v_{1k} \\ v_{2k} \\ v_{3k} \end{pmatrix} \sim \mathrm{N}\left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{v0} & & & \\ \sigma_{v01} & \sigma^2_{v1} & & \\ \sigma_{v02} & \sigma_{v12} & \sigma^2_{v2} & \\ 0 & 0 & 0 & \sigma^2_{v3} \end{pmatrix}\right]$$

$$u_{jk} \sim \mathrm{N}(0, \sigma^2_{ujk}), \quad \ln(\sigma^2_{ujk}) = \alpha_0 + \alpha_1 \mathrm{age}_{jk} + v_{2k}$$

$$e_{ijk} \sim \mathrm{N}(0, \sigma^2_e)$$

…can expand to 3 levels, distinguishing within-clinic (short term) from between-clinic (longer term) within-individual variability

$$\mathrm{BP}_{ijk} = \beta_0 + \beta_1 \mathrm{age}_{jk} + v_{0k} + v_{1k}\mathrm{age}_{jk} + u_{jk} + e_{ijk}$$

$$\log(\mathrm{LVMI})_k = \gamma_0 + \gamma_1 v_{0k} + \gamma_2 v_{1k} + \gamma_3 v_{2k} + v_{3k}$$

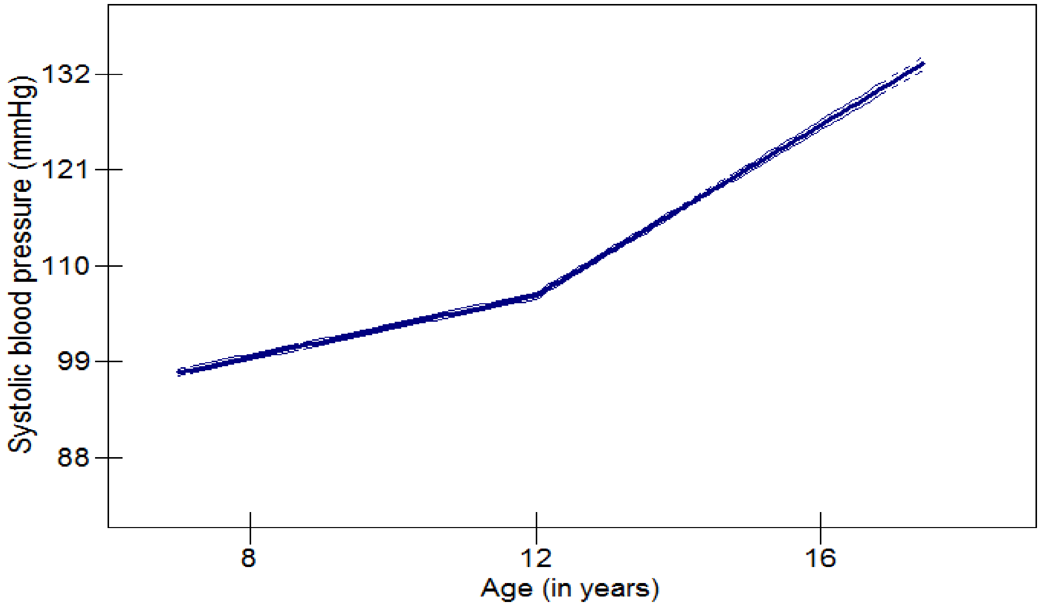
$$\begin{pmatrix} v_{0k} \\ v_{1k} \\ v_{2k} \\ v_{3k} \end{pmatrix} \sim \mathrm{N}\left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{v0} & & & \\ \sigma_{v01} & \sigma^2_{v1} & & \\ \sigma_{v02} & \sigma_{v12} & \sigma^2_{v2} & \\ 0 & 0 & 0 & \sigma^2_{v3} \end{pmatrix}\right]$$

$$u_{jk} \sim \mathrm{N}(0, \sigma^2_{ujk}), \quad \ln(\sigma^2_{ujk}) = \alpha_0 + \alpha_1 \mathrm{age}_{jk} + v_{2k}$$

$$e_{ijk} \sim \mathrm{N}(0, \sigma^2_e)$$

Fitting the models…

- Bayesian estimation in Stan (via rstan)
- Model age via linear spline with a knot point at 12 years of age in fixed part of model *(after: Staley et al, 2015; O'Keeffe et al., 2018)*
- Covariates:
    - Age
    - Sex
    - Weight
    - Height



Fitting the models…

- **Evidence of differences between individuals in their extent of within-individual variability**; *SD on the log scale = 0.40 (0.27, 0.50).*
- Positive correlation between random intercept and random within-individual variability term: i.e. **people with higher BP tend to have more fluctuation in their BP**; *Cor(v_0, v_2) = 0.48 (0.31, 0.69).*
- On average, **greater within-individual variability in BP** (ln(σ^2_{ujk})):
  - **at older ages**; *0.12 (0.06, 0.17)*
  - **in females**; *0.17 (0.05, 0.29)*
  - **for heavier log(bodyweights)**; *0.56 (0.19, 0.93)*

NB: estimates given as: mean (95% credible interval)

Fitting the models…

What about the later outcome, log(LVMI)?

*NB these estimates are * 10^{-2}*

- **Higher within-individual variability predicted greater log(LVMI)**; β = 0.47 (-0.03, 1.07)
- …**but not** when the random intercept and random slope terms were also included as exposures in the linear model for log(LVMI):
  - Random intercept: β *= 0.07 (0.02, 0.14)*
  - Random slope: β *= 0.40 (0.08, 0.78)*
  - Random within-individual variability: β *= -0.85 (-2.77, 0.22)*

NB: estimates given as: mean (95% credible interval)

Further work

Applying this joint modelling approach to other topics: e.g. within-individual variability in cognitive functioning at older ages, and relation to dementia…

…with such psychometric measurements, relationship between random intercept and random scale effects may be non-linear, due to bounded scale: need to model this appropriately.

Any questions?

Want to run a sub study?
Questions for 2019 questionnaire?
Get in touch: alspac-exec@bristol.ac.uk

Find out about future plans for
ALSPAC 2019 – 2024
www.bristol.ac.uk/alspac/renewal