0:00 / 10:57

Multilevel models

Professor Andy Field

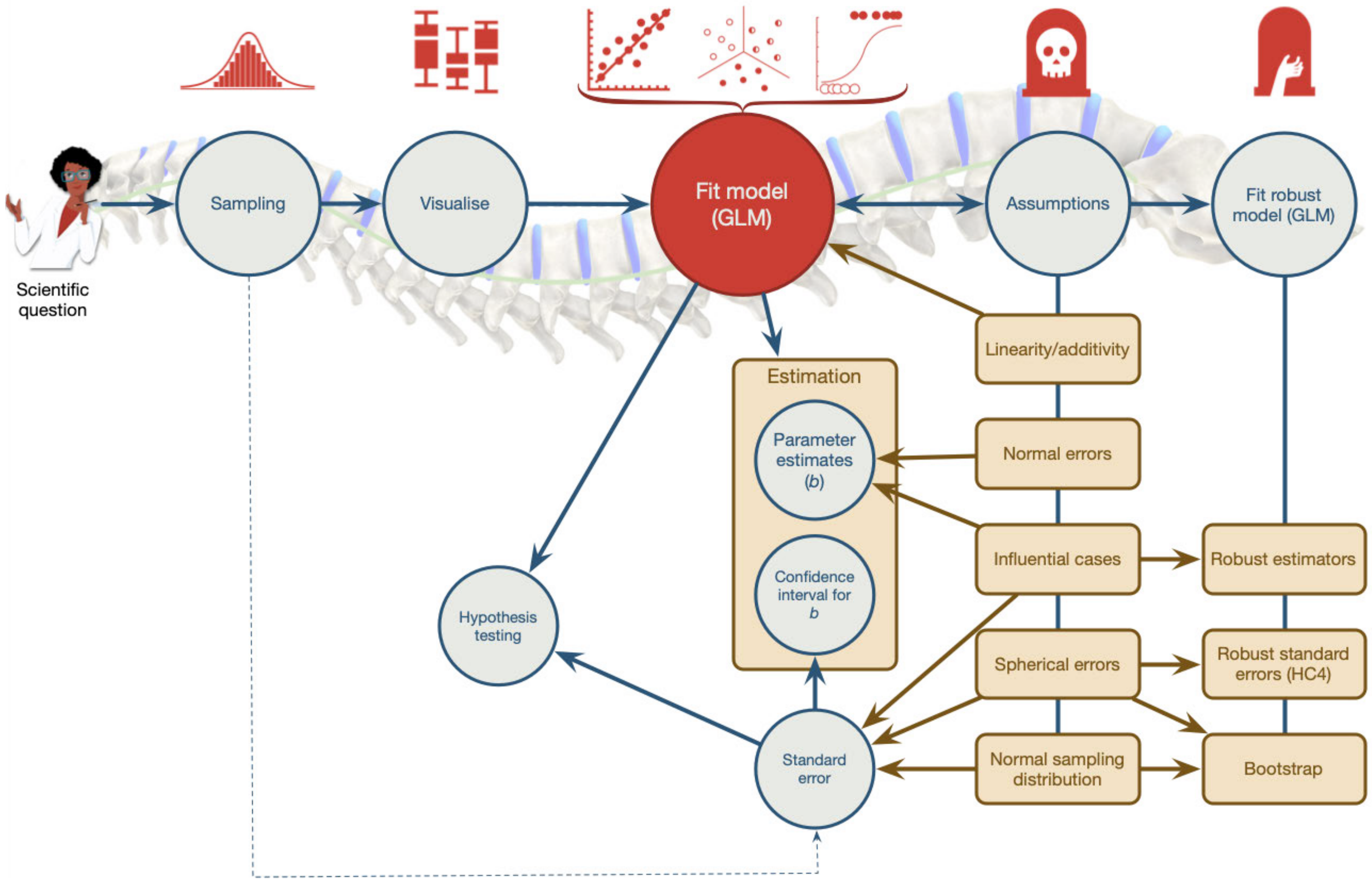
@profandyfield

www.youtube.com/user/ProfAndyField/

www.discoveringstatistics.com

www.milton-the-cat.rocks

www.discovr.rocks



Scientific question
Sampling
Visualise
Fit model (GLM)
Assumptions
Fit robust model (GLM)
Estimation
Parameter estimates (b)
Confidence interval for b
Linearity/additivity
Normal errors
Influential cases
Spherical errors
Normal sampling distribution
Robust estimators
Robust standard errors (HC4)
Bootstrap
Hypothesis testing
Standard error

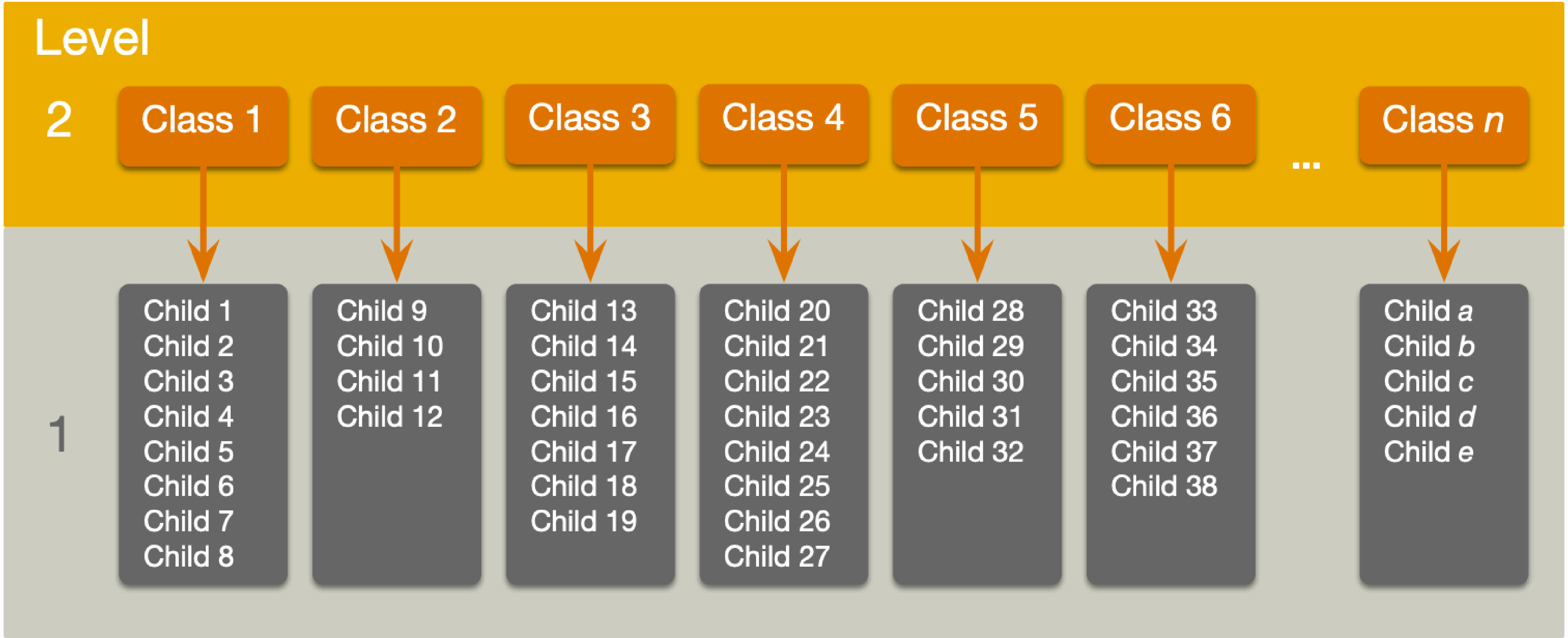
Learning outcomes

- Understand what hierarchical data are
  - Why we can't use the 'normal' GLM
- Understand fixed and random coefficients
- Understand how to build models
- Be able to conduct and interpret models of hierarchical data

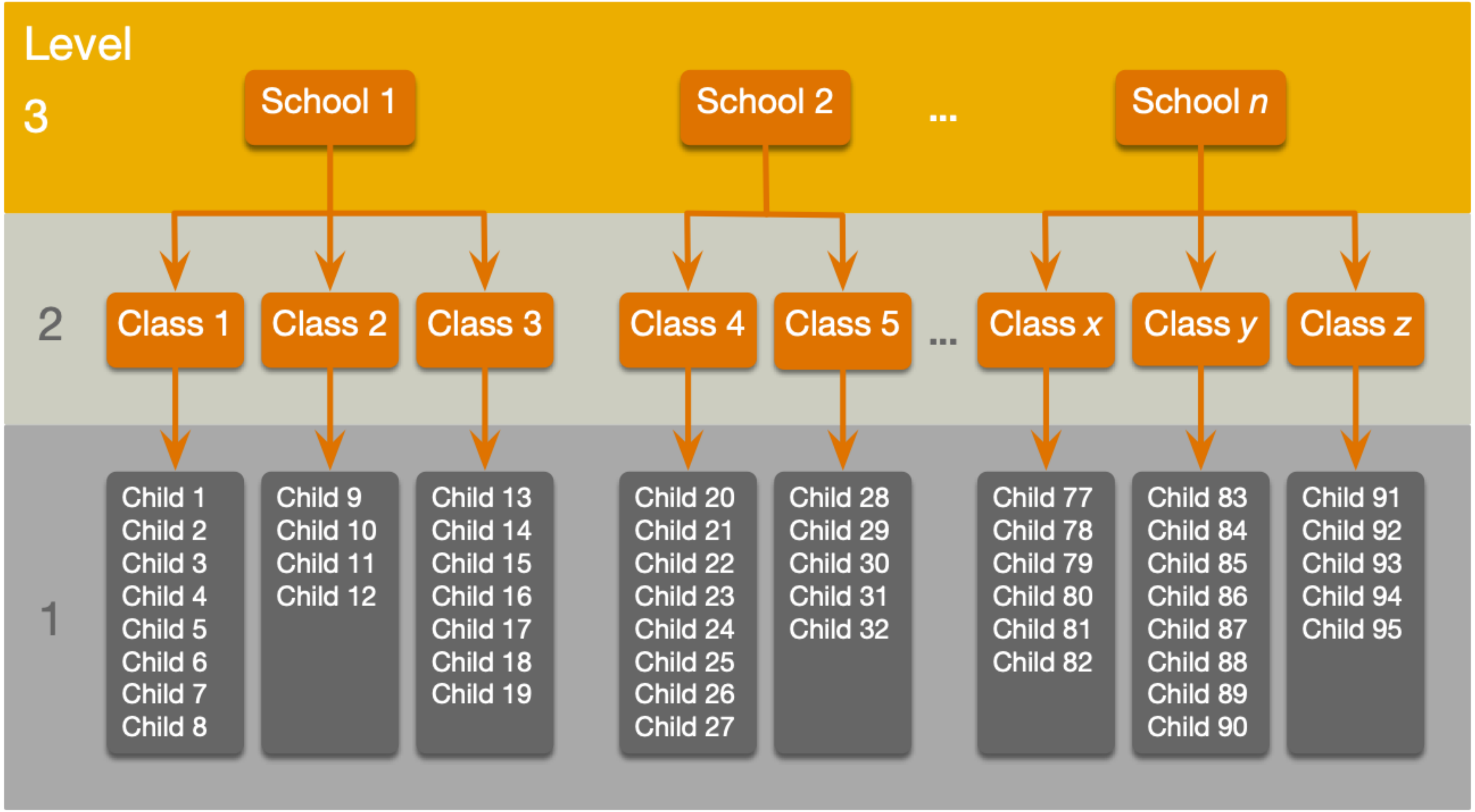
Hierarchical data

- Data structures are often hierarchical
    - Children nested within classrooms
    - Observations nested within people
    - Employees nested within organisations
    - Patients nested within hospitals
    - Patients nested within teams nested within hospitals
    - Service users nested within clinicians nested within hospitals nested within NHS trusts!
    - Zombies nested within rehabilitation clinics 😉

A two-level hierarchy



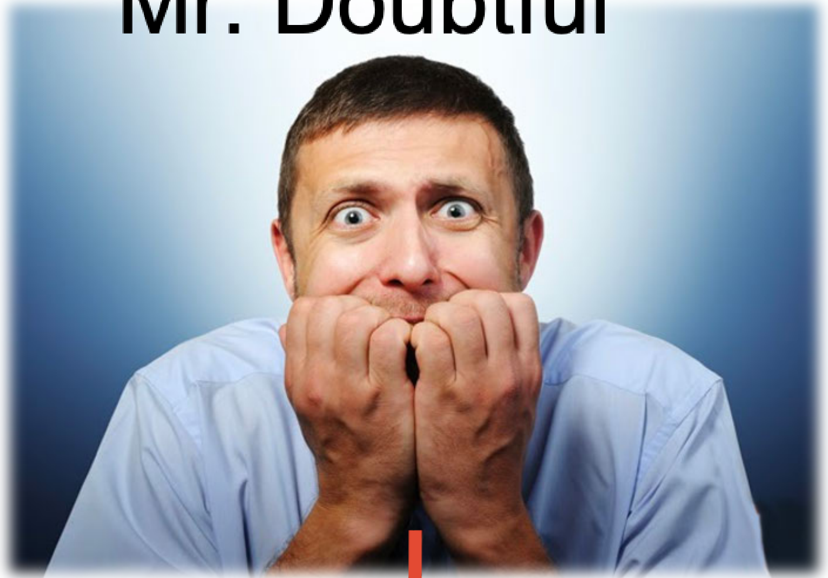
A three-level hierarchy



Why hierarchies matter

- Data from the same context will be more similar than data from different contexts
  - Children in the same class will perform more similarly than children from different classes
  - People treated in the same clinics should be more similar in response than those treated at different clinics
- Lack of independence
  - Violates the assumption of spherical errors (specifically, independence)
  - Biases SEs, CIs and *p* -values

Mr. Doubtful



Ms. Confident

Benefits of multilevel models

- Modelling variability in effects across contexts
    - Model the variability in intercepts
    - Model the variability in slopes
- Model violations of the assumption of spherical errors
    - Model differences in the variability of errors
    - Model relationships between errors
        - (Linear model for repeated observations – two weeks time!)
- Missing data
    - MLMs cope well with missing data (in general)

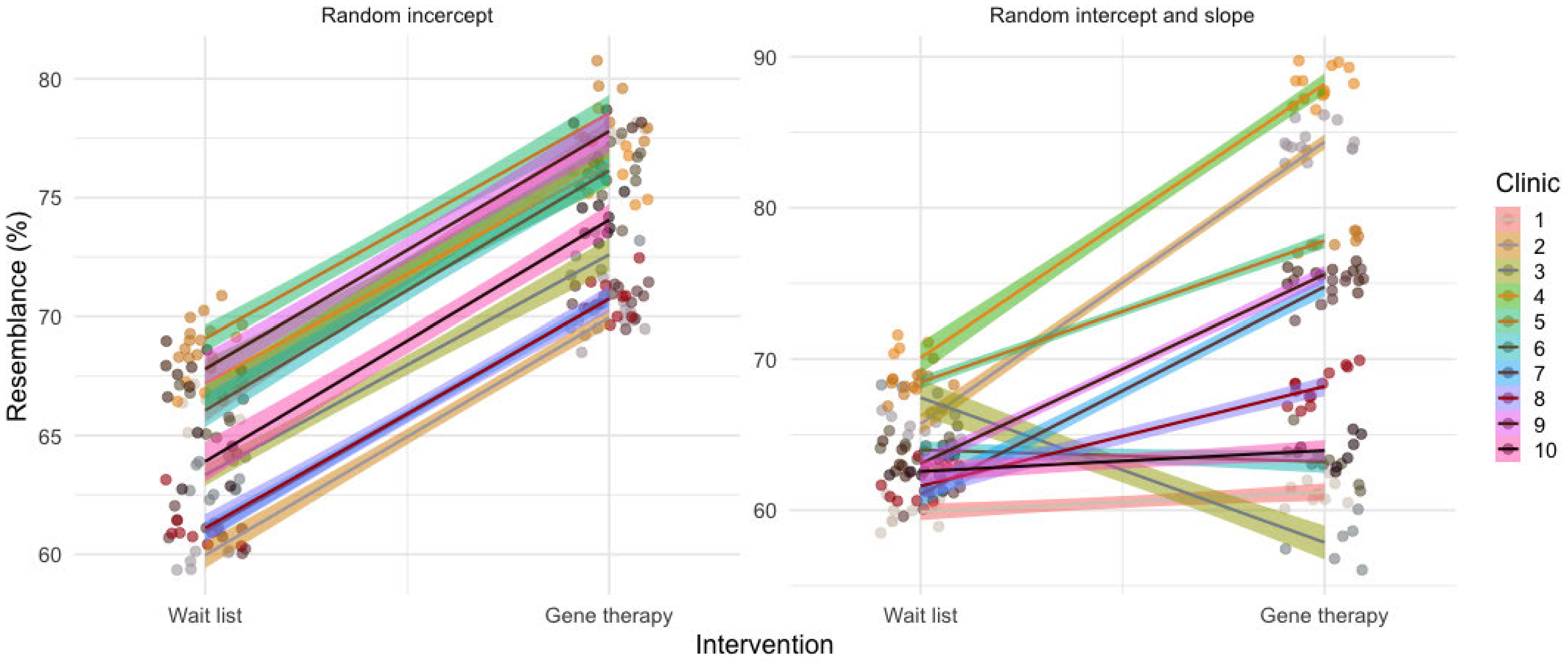
A rehabilitative example

- id_clin: Which of 10 clinics the zombie attended to have the intervention (1 to 10)
- id: Zombie id
- intervention: Was the participant assigned to waiting list (0) or gene therapy (1)?
- gep_level: Genetic Enhancement Program level.
  - Were they originally assigned to the low (0) or high (1) level genetic enhancement program at JIG:SAW?
- tse_months: Time Since Enhancement (months).
  - How long ago did they undergo the original genetic enhancement program (which zombified them)
- resemblance: Resemblance of their face to their pre-zombie state (0% to 100%)

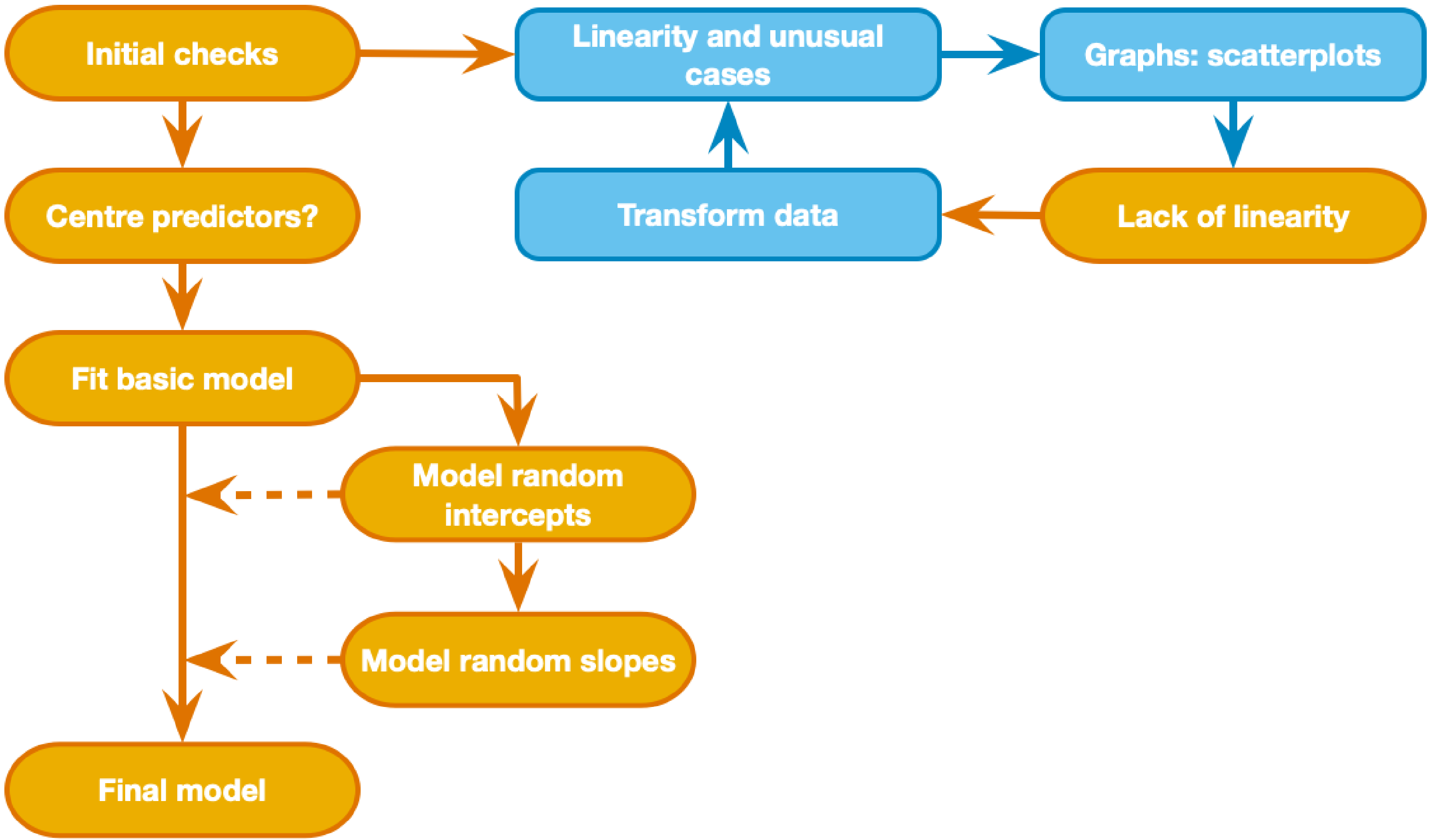
Fixed and random coefficients

- Intercepts and slopes can be fixed or random
  - In OLS regression they are fixed
  - Fixed coefficients
  - Intercepts/slopes are assumed to be the same across different contexts (in this case clinics)
- Random coefficients
  - Intercepts/slopes are allowed to vary across different contexts (in this case clinics)

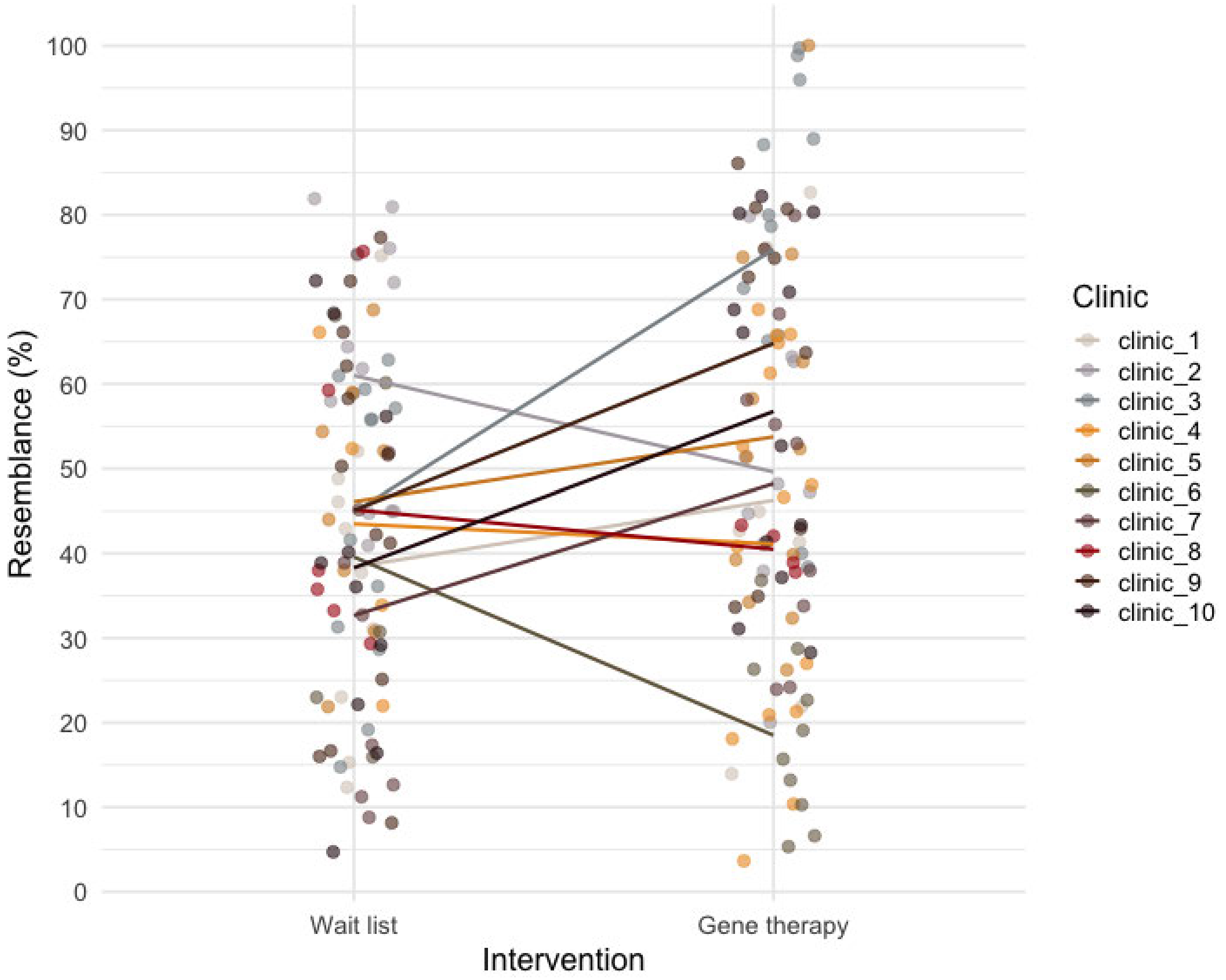
Intercepts and slopes



The process


Initial checks
Linearity and unusual cases
Graphs: scatterplots
Centre predictors?
Transform data
Lack of linearity
Fit basic model
Model random intercepts
Model random slopes
Final model


Level
2
Clinic 1
Clinic 2
...
Clinic 9
Clinic 10
1
Zombie 1
Zombie 2
Zombie 3
Zombie 4
Zombie 5
Zombie 6
Zombie 7
Zombie 8
Zombie 9
Zombie 10
Zombie 11
Zombie 12
Zombie 13
Zombie 14
Zombie 15
Zombie 16
Zombie 17
Zombie 18
Zombie 19
Zombie 20
Zombie 21
Zombie 22
Zombie 23
Zombie 24
Zombie 25
Zombie 26
Zombie 27
Zombie 28
Zombie 29
Zombie 30
Zombie 31
Zombie 32
Zombie 33
Zombie 34
Zombie 35
Zombie 36
Zombie 37
Zombie 38
...
Zombie 146
Zombie 147
Zombie 148
Zombie 149
Zombie 150
Zombie 151
Zombie 152
Zombie 153
Zombie 154
Zombie 155
Zombie 156
Zombie 157
Zombie 158
Zombie 159
Zombie 160
Zombie 161
Zombie 162
Zombie 163
Zombie 164
Zombie 165
Zombie 166
Zombie 167
Zombie 168
Zombie 169
Zombie 170
Zombie 171
Zombie 172
Zombie 173
Zombie 174
Zombie 175
Zombie 176
Zombie 177
Zombie 178
Zombie 179
Zombie 180
Zombie 181
Zombie 182
Zombie 183
Zombie 184
Zombie 185
Zombie 186
Zombie 187
Zombie 188
Zombie 189
Zombie 190


Resemblance (%)
100
90
80
70
60
50
40
30
20
10
0
Wait list
Gene therapy
Intervention
Clinic
clinic_1
clinic_2
clinic_3
clinic_4
clinic_5
clinic_6
clinic_7
clinic_8
clinic_9
clinic_10

The models (using familiar symbols)

- Fixed intercepts and slopes

$$\text{resem}\hat{\text{b}}\text{lance}_i = \hat{b}_0 + \hat{b}_1\text{Intervention}_i + e_i$$

- Random intercepts but fixed slopes

$$\text{resem}\hat{\text{b}}\text{lance}_{ij} = \hat{b}_{0j} + \hat{b}_1\text{Intervention}_{ij} + e_{ij}$$
$$\hat{b}_{0j} = \hat{b}_0 + \hat{u}_{0j}$$

- Random intercepts and slopes

$$\text{resem}\hat{\text{b}}\text{lance}_{ij} = \hat{b}_{0j} + \hat{b}_{1j}\text{Intervention}_{ij} + e_{ij}$$
$$\hat{b}_{0j} = \hat{b}_0 + \hat{u}_{0j}$$
$$\hat{b}_{1j} = \hat{b}_1 + \hat{u}_{1j}$$

The models (using common notation)

- Fixed intercepts and slopes

$$\widehat{\text{resemblance}}_i = \hat{\gamma}_0 + \hat{\gamma}_1\text{Intervention}_i + e_i$$

- Random intercepts but fixed slopes

$$\widehat{\text{resemblance}}_{ij} = \hat{\gamma}_{0j} + \hat{\gamma}_1\text{Intervention}_{ij} + e_{ij}$$
$$\hat{\gamma}_{0j} = \hat{\gamma}_0 + \hat{u}_{0j}$$

- Random intercepts and slopes

$$\widehat{\text{resemblance}}_{ij} = \hat{\gamma}_{0j} + \hat{\gamma}_{1j}\text{Intervention}_{ij} + e_{ij}$$
$$\hat{\gamma}_{0j} = \hat{\gamma}_0 + \hat{u}_{0j}$$
$$\hat{\gamma}_{1j} = \hat{\gamma}_1 + \hat{u}_{1j}$$

Adding predictors

- Add whether original enhancement was low or high intensity (gep_level)

$$
\begin{aligned}
\widehat{\text{resemblance}}_{ij} &= \hat{\gamma}_{0j} + \hat{\gamma}_{1j}\text{Intervention}_{ij} + \hat{\gamma}_{2}\text{gep\_level}_{ij} + e_{ij} \\
\hat{\gamma}_{0j} &= \hat{\gamma}_{0} + \hat{u}_{0j} \\
\hat{\gamma}_{1j} &= \hat{\gamma}_{1} + \hat{u}_{1j}
\end{aligned}
$$

- Add time since genetic enhancement (tse_months)

$$
\begin{aligned}
\widehat{\text{resemblance}}_{ij} &= \hat{\gamma}_{0j} + \hat{\gamma}_{1j}\text{Intervention}_{ij} + \hat{\gamma}_{2}\text{gep\_level}_{ij} + \hat{\gamma}_{3}\text{tse\_months}_{ij} + e_{ij} \\
\hat{\gamma}_{0j} &= \hat{\gamma}_{0} + \hat{u}_{0j} \\
\hat{\gamma}_{1j} &= \hat{\gamma}_{1} + \hat{u}_{1j}
\end{aligned}
$$

Assessing the fit and comparing models

- Assessing fit
    - AIC
    - BIC
- Models should be built up gradually
    - Start with fixed coefficients
    - Change 1 aspect of the model and compare to the previous with the change in the -2 log likelihood

Model 1:

$$\hat{\text{resemblance}}_i = \hat{\gamma}_0 + e_i$$

```
rehab_base <- nlme::gls(resemblance ~ 1,
                        data = rehab_tib,
                        method = "ML"
                        )
```

Model 2

$$\hat{\text{resemblance}}_{ij} = \hat{\gamma}_0 + \hat{u}_{0j} + e_{ij}$$

```
rehab_ri <- nlme::lme(resemblance ~ 1,
                      random = ~1|id_clin,
                      data = rehab_tib,
                      method = "ML"
                      )
```

Model 3

$$\widehat{\text{resemblance}}_{ij} = \hat{\gamma}_0 + \hat{\gamma}_1 \text{Intervention}_{ij} + \hat{u}_{0j} + e_{ij}$$

```r
rehab_int <- nlme::lme(resemblance ~ intervention,
                       random = ~1|id_clin,
                       data = rehab_tib,
                       method = "ML"
                       )
```

Model 4

$$\widehat{\text{resemblance}}_{ij} = \hat{\gamma}_0 + \hat{\gamma}_1 \text{Intervention}_{ij} + \hat{u}_{0j} + \hat{u}_{1j} + e_{ij}$$

```r
rehab_rs <- nlme::lme(resemblance ~ intervention,
                      random = ~intervention|id_clin,
                      data = rehab_tib,
                      method = "ML"
                      )
```

Model 5

$$\widehat{\text{resemblance}}_{ij} = \hat{\gamma}_0 + \hat{\gamma}_1\text{Intervention}_{ij} + \hat{\gamma}_2\text{gep\_level}_{ij} + e_{ij} + \hat{u}_{0j} + \hat{u}_{1j}$$

```
rehab_gep <- nlme::lme(resemblance ~ intervention + gep_level,
                       random = ~intervention|id_clin,
                       data = rehab_tib,
                       method = "ML"
                       )
```

Model 6

$$\widehat{\text{resemblance}}_{ij} = \hat{\gamma}_0 + \hat{\gamma}_1\text{Intervention}_{ij} + \hat{\gamma}_2\text{gep\_level}_{ij} + \hat{\gamma}_3\text{tse\_months}_{ij} + e_{ij} + \hat{u}_{0j} + \hat{u}_{1j}$$

```
rehab_tse <- nlme::lme(resemblance ~ intervention + gep_level + tse_months,
                       random = ~intervention|id_clin,
                       data = rehab_tib,
                       method = "ML"
                       )
```

Comparing models using R

```
anova(rehab_base, rehab_ri, rehab_int, rehab_rs, rehab_gep, rehab_tse)
```

Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
1	2	1724.75	1731.25	-860.38		NA	NA
2	3	1715.50	1725.24	-854.75	1 vs 2	11.25	0.00
3	4	1710.38	1723.36	-851.19	2 vs 3	7.12	0.01
4	6	1702.11	1721.59	-845.05	3 vs 4	12.27	0.00
5	7	1703.08	1725.81	-844.54	4 vs 5	1.03	0.31
6	8	1103.52	1129.49	-543.76	5 vs 6	601.56	0.00

Model parameters (Fixed effects)

```
broom.mixed::tidy(rehab_tse, effects = "fixed")
```

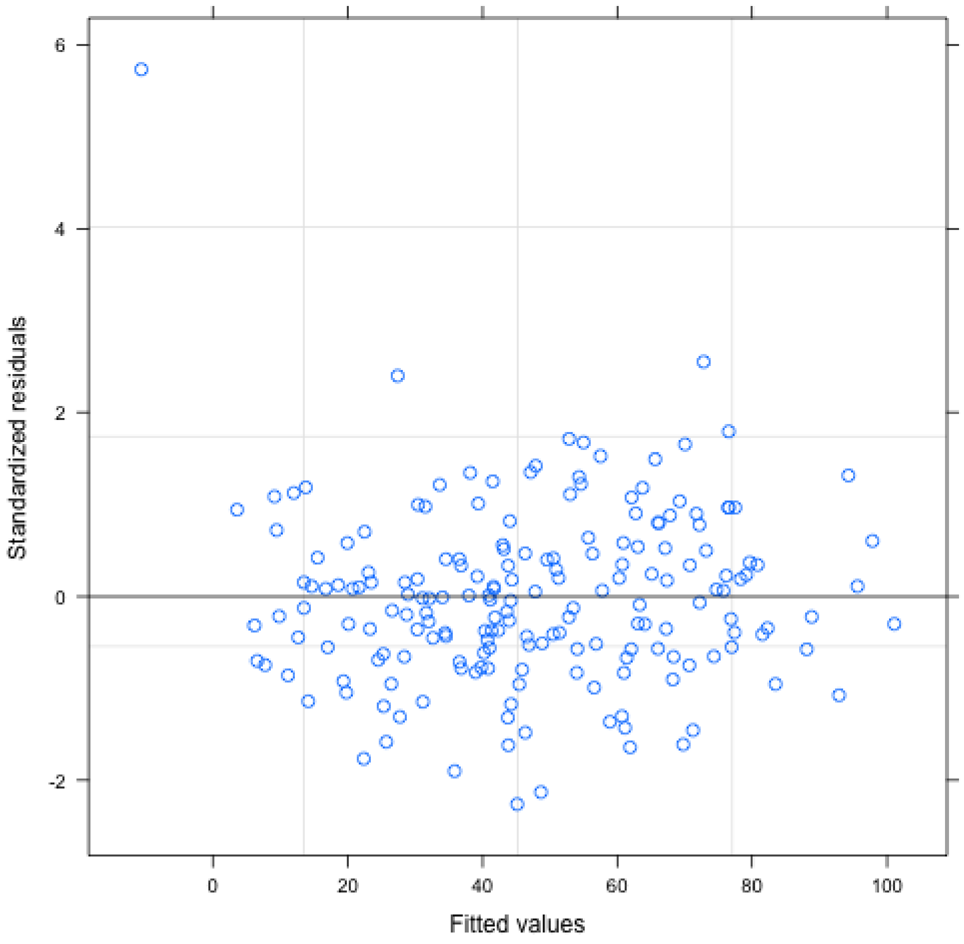
term	estimate	std.error	df	statistic	p.value
(Intercept)	80.933	1.599	177	50.611	0.000
interventionGene therapy	6.794	2.706	177	2.511	0.013
gep_levelHigh intensity	-4.297	0.548	177	-7.844	0.000
tse_months	-1.977	0.028	177	-69.634	0.000

Model parameters (random effects)

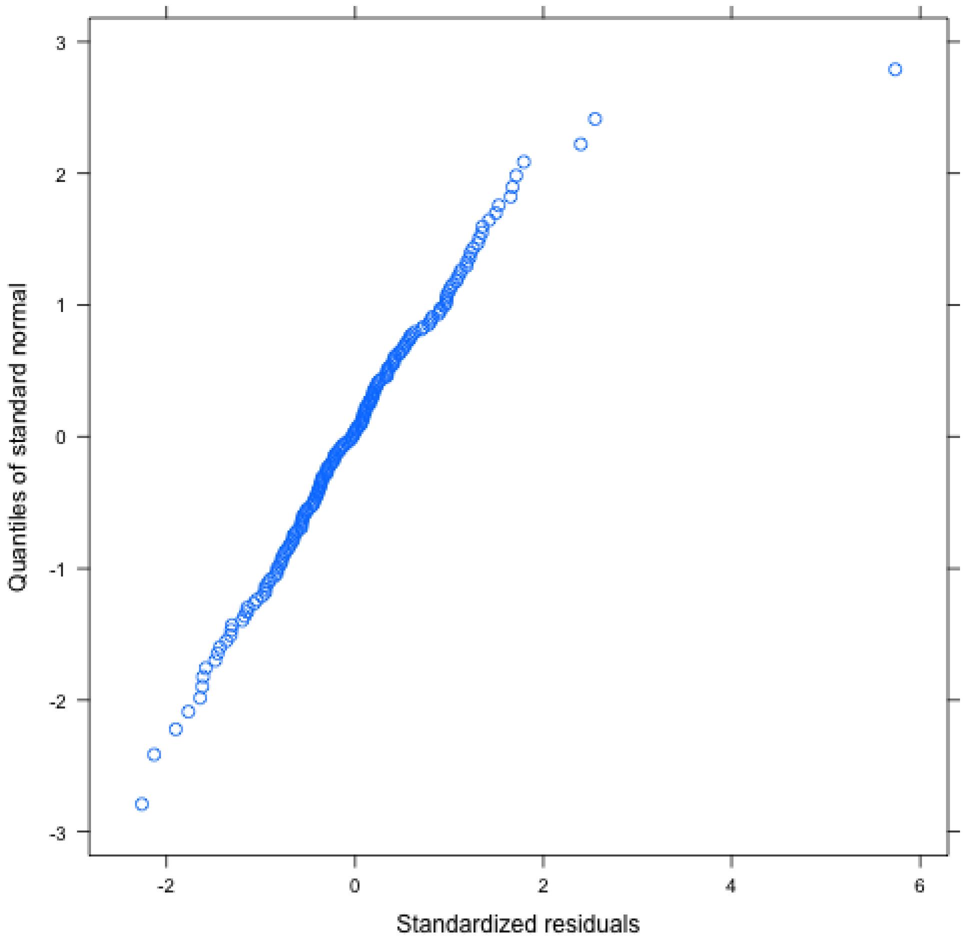
```
broom.mixed::tidy(rehab_tse, effects = "ran_pars")
```

effect	group	term	estimate
ran_pars	id_clin	sd_(Intercept)	4.47
ran_pars	id_clin	cor_interventionGene therapy.(Intercept)	0.66
ran_pars	id_clin	sd_interventionGene therapy	8.28
ran_pars	Residual	sd_Observation	3.60

```
plot(rehab_tse)
```




```
qqnorm(rehab_tse)
```



Robust model parameters

```
rehab_tse_rob <- robustlmm::rlmer(
  resemblance ~ intervention + gep_level + tse_months + (intervention|id_clin),
  data = rehab_tib)
broom.mixed::tidy(rehab_tse_rob, effects = "fixed")
```

effect	term	estimate	std.error	statistic
fixed	(Intercept)	81.70	1.60	51.11
fixed	interventionGene therapy	7.01	3.25	2.16
fixed	gep_levelHigh intensity	-4.58	0.49	-9.29
fixed	tse_months	-1.99	0.03	-77.87

Robust random effects

```
broom.mixed::tidy(rehab_tse_rob, effects = "ran_pars")
```

effect	group	term	estimate
ran_pars	id_clin	sd__(Intercept)	4.47
ran_pars	id_clin	cor__(Intercept).interventionGene therapy	0.55
ran_pars	id_clin	sd__interventionGene therapy	9.83
ran_pars	Residual	sd__Observation	3.18

To sum up …

- Data can be hierarchical and this hierarchical structure can be important.
    - The OLS linear model simply ignores the hierarchy.
- Hierarchical models are just a fancy linear model in which you estimate the variability in the slopes and intercepts within contexts
    - i.e. slopes and intercepts can be random variables (allowed to vary) rather than fixed (assumed to be equal in different situations).
- Start with a model that ignores the hierarchy and then add in random intercepts and slopes to see if they improve the fit of the model.