Patients, hands, fingers or joints?

Clustered data

Dieuwke Broekstra, PhD
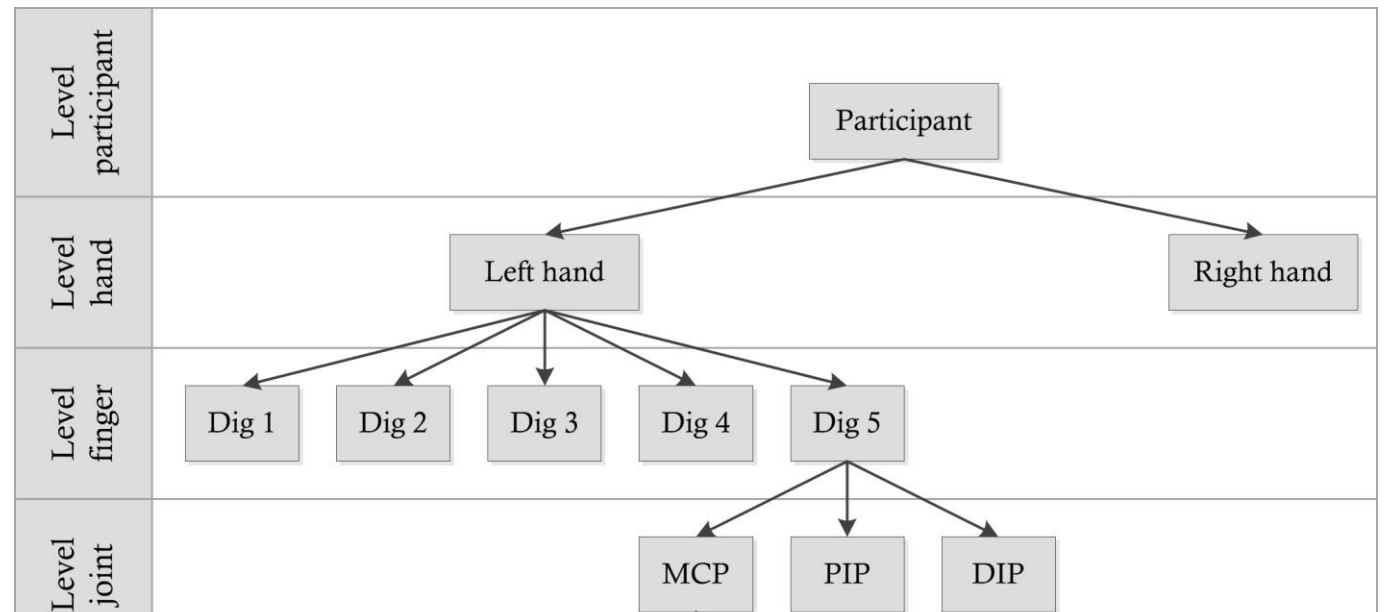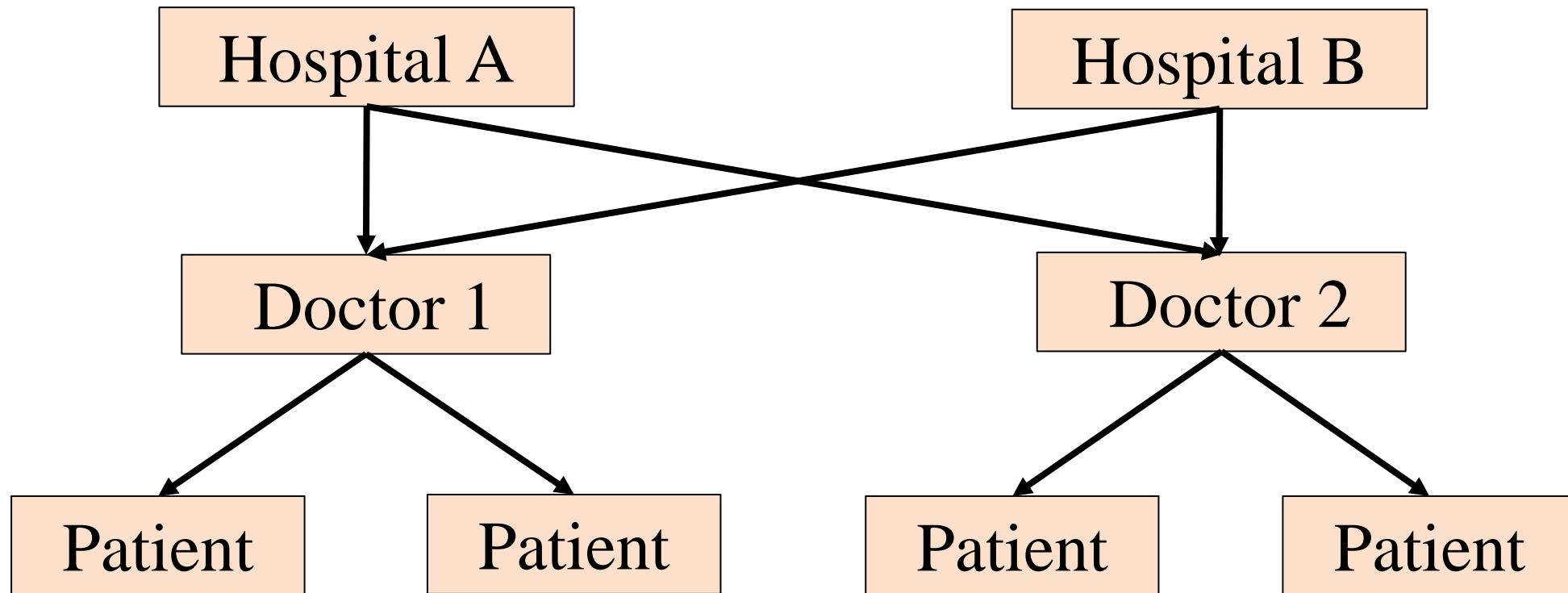Dept. Plastic Surgery

Background

Dupuytren's disease

Clustered data

- "Multilevel data"

- "Correlated data"

- "Nested data"
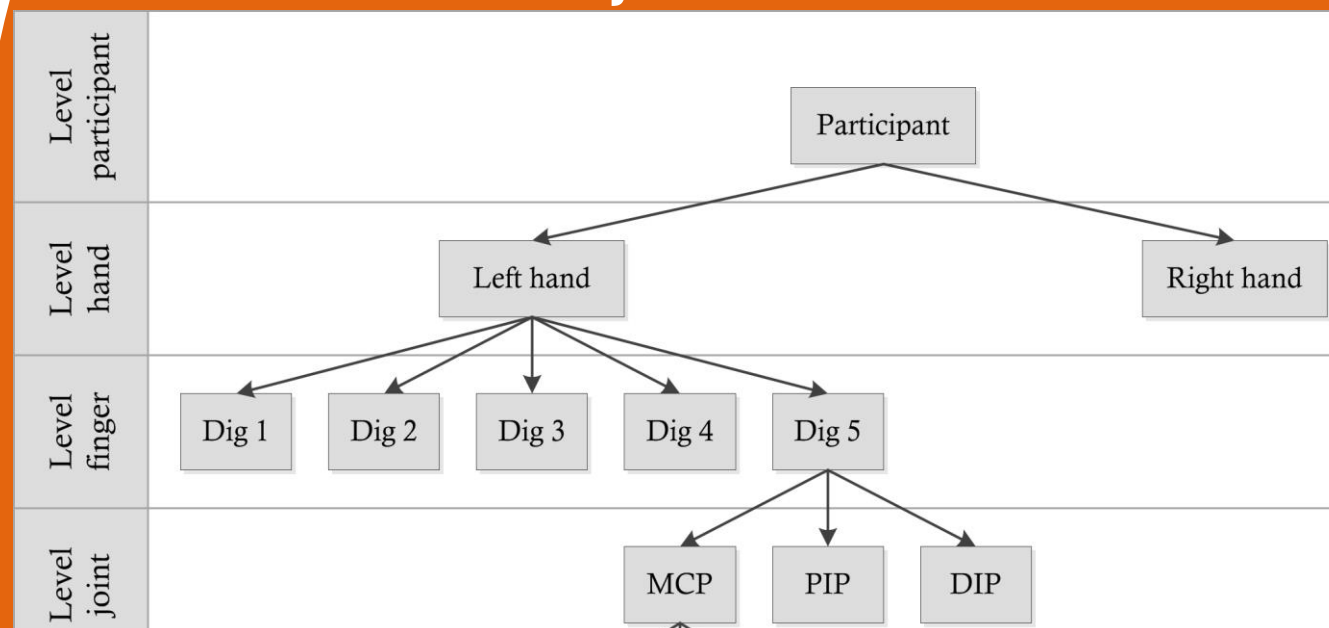
Clustered data

Background

Assumption of independent observations or residuals

Independent t-test

Mann–Whitney U-test

Independent??

Example

Dupuytren's disease: difference treatments on angular deformities?

Linear regression model:

Variable	Effect estimate	SE	T-value	P
Intercept	42.16	0.77	54.84	<0.001
Age	0.04	0.05	0.75	0.45
Treatment type	4.17	1.12	3.72	<0.001

Example

Dupuytren's disease: difference treatments on angular deformities?

Linear mixed-effect model:

Variable	Effect estimate	SE	T-value	P
Intercept	38.11	3.52	10.82	<0.001
Age	−0.10	0.21	−0.47	0.65
Treatment type	4.97	4.46	1.11	0.28

Overestimation of precision

- Linear regression model:

Variable	Effect estimate	SE	T-value	P
Intercept	42.16	0.77	54.84	<0.001
Age	0.04	0.05	0.75	0.45
Treatment type	4.17	1.12	3.72	<0.001

- Linear mixed-effect model:

Variable	Effect estimate	SE	T-value	P
Intercept	38.11	3.52	10.82	<0.001
Age	−0.10	0.21	−0.47	0.65
Treatment type	4.97	4.46	1.11	0.28

Overestimation of precision

Linear regression model

Linear mixed-effect model

Ntani et al. (2021) BMC Med Res Methodol 21:139.
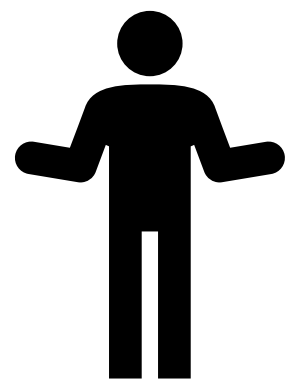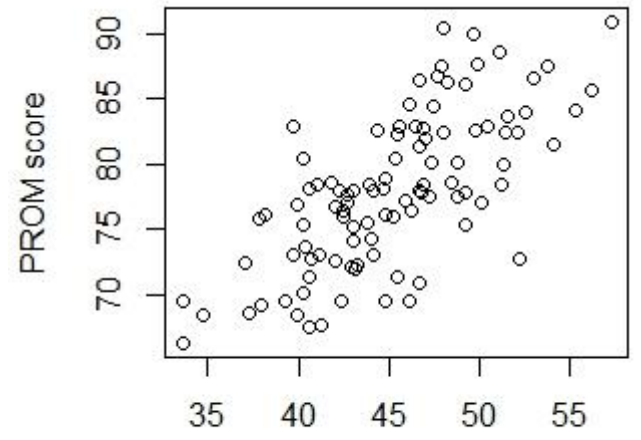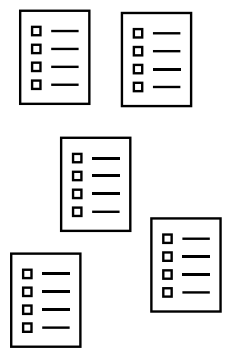
9

umcg
Plastic Surgery
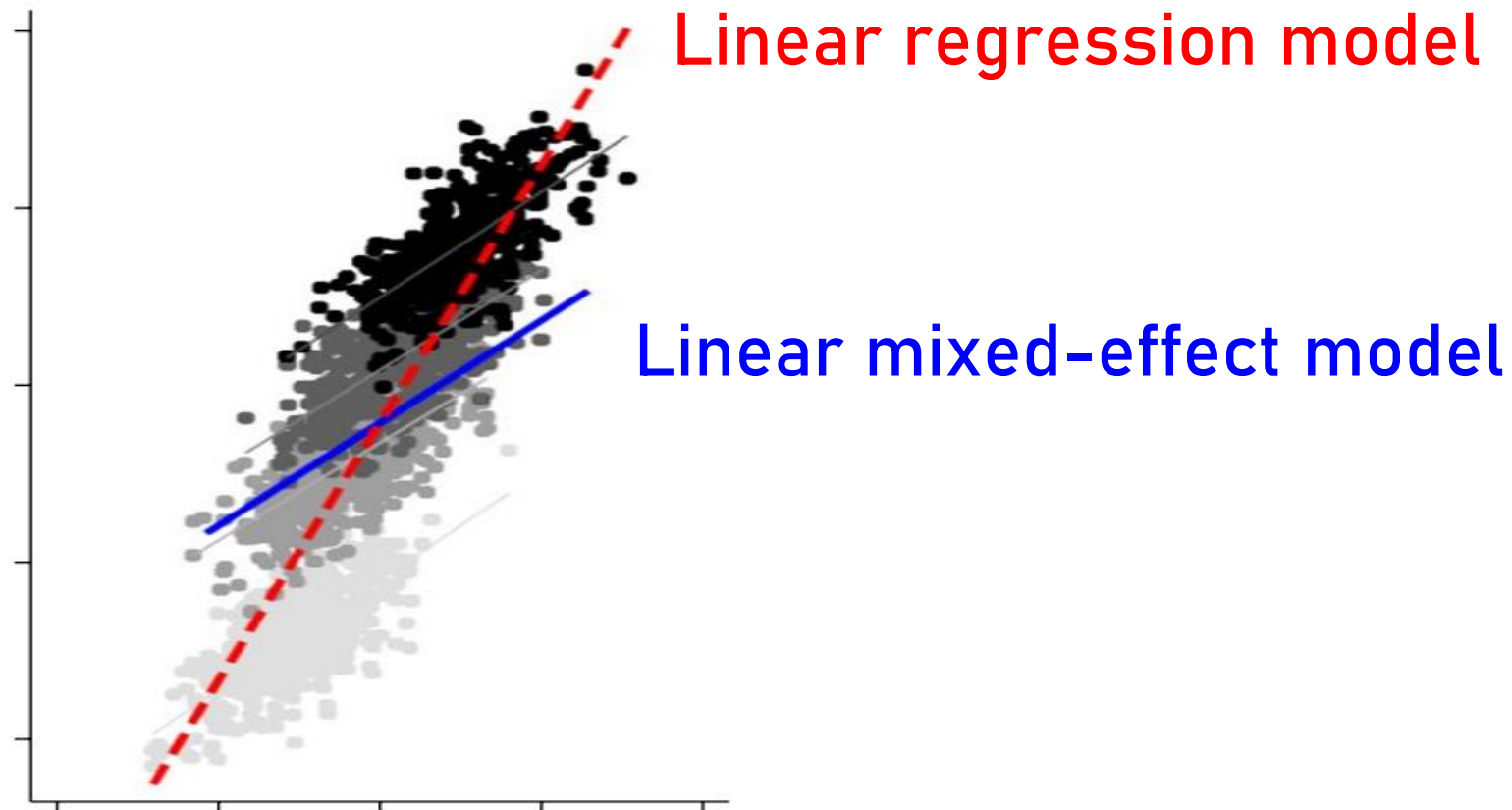
Relevance

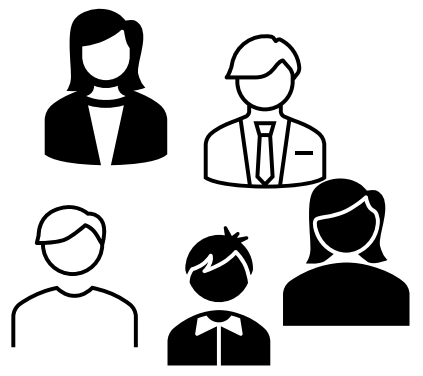Ignoring clustering leads to overestimation of precision
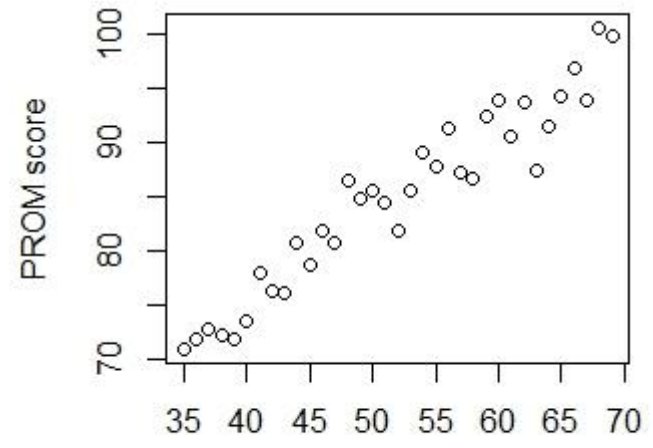
→ Too small standard errors

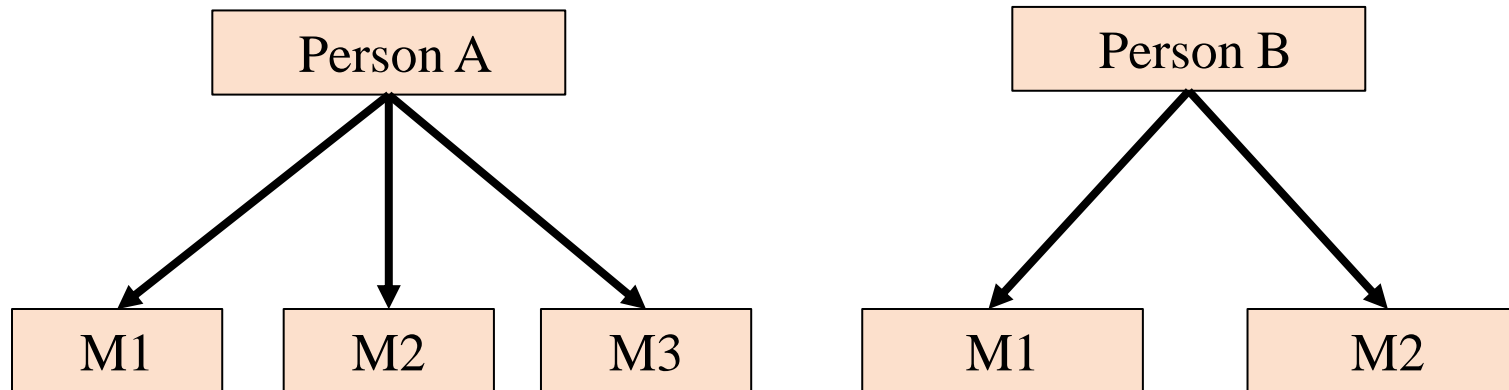→ Reject H0 too easily

Shared factors
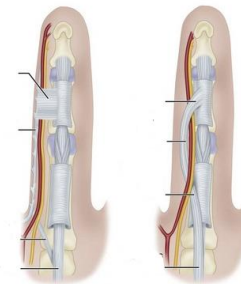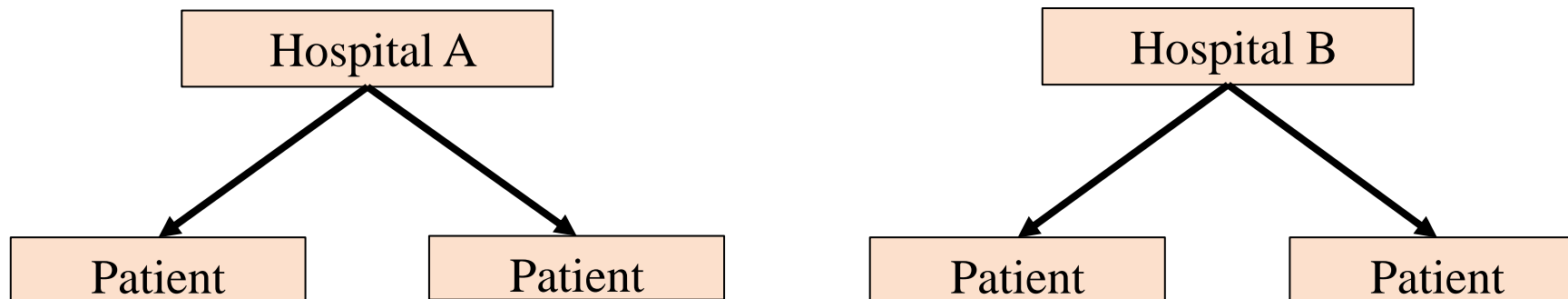
Correlated scores
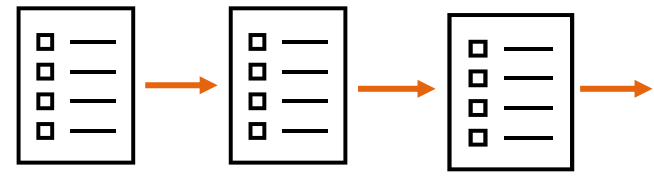
Clustering in clinical studies

- Multiple measurements in time of the same persons

Clustering in clinical studies

- Multicenter study

Clustering in clinical studies

- Multiple organs/body parts in the same person
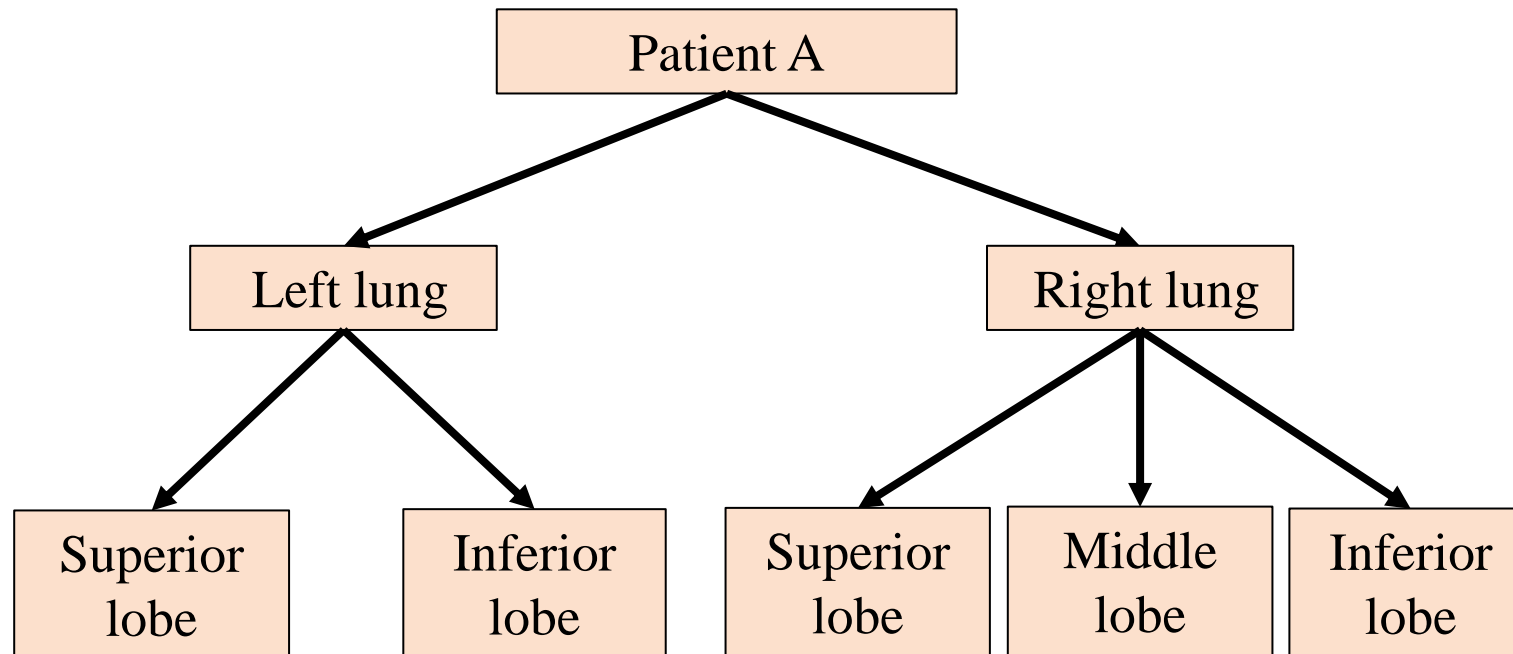
Clustering in clinical studies

- Multiple studies in a meta-analysis
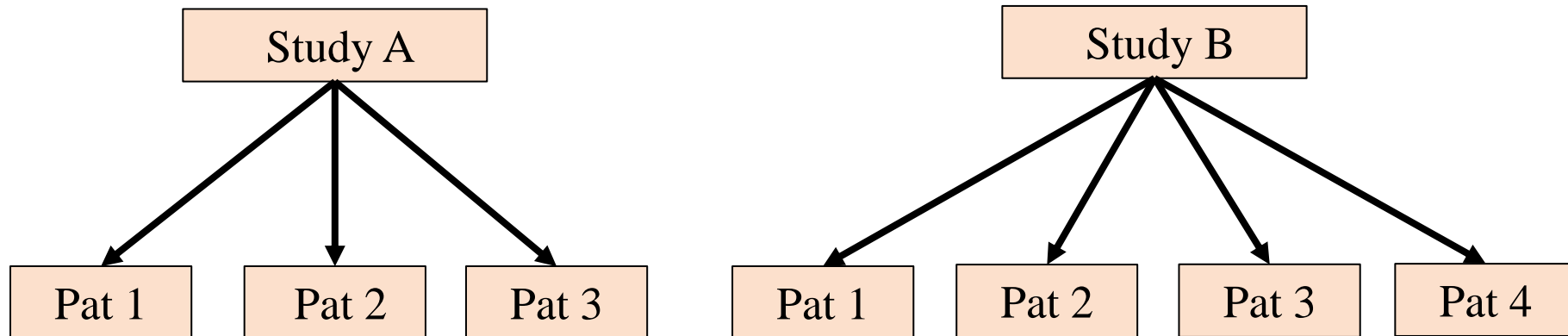
Clustering in clinical studies

- Multiple observers in an agreement study
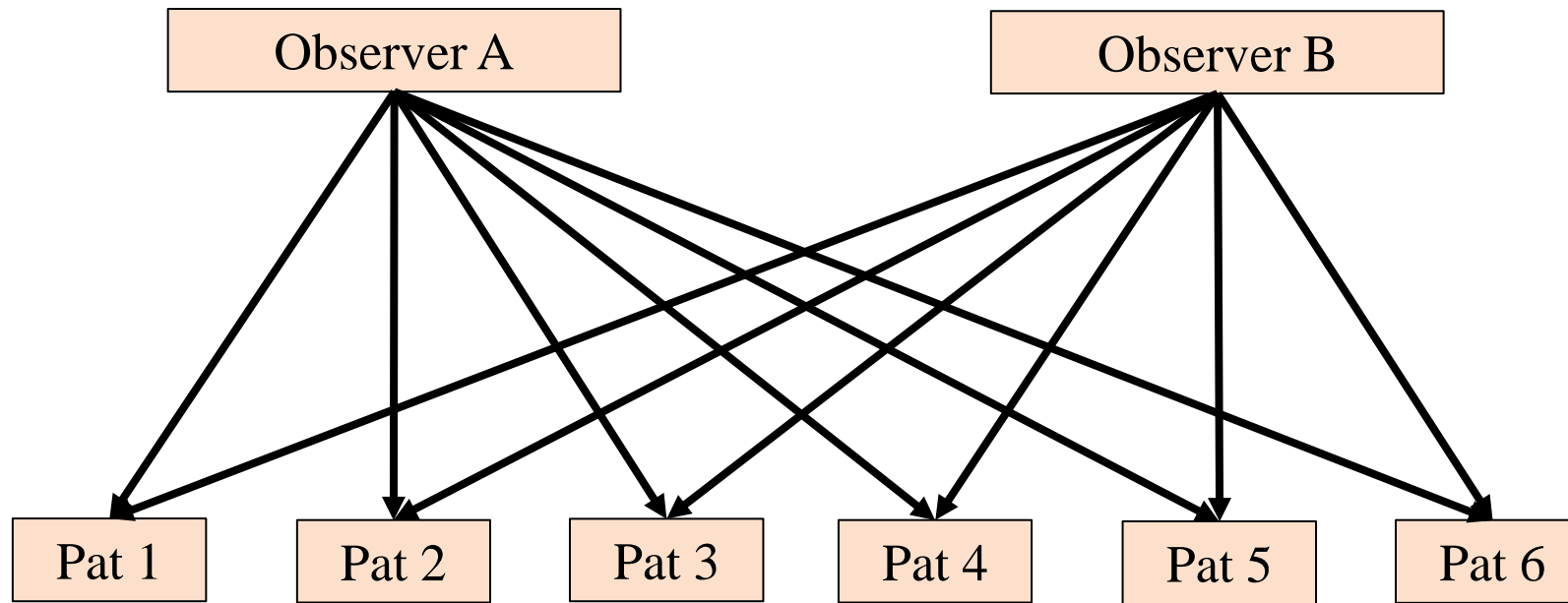
How to handle this

1. Minimize levels in your dataset
2. Use analysis techniques that can handle clustering
3. Define the amount of clustering and decide based on that what to do

Minimize levels in dataset

- Aggregate data

- Randomly select one organ / body in analysis

→ Drawback: not using all information available

Analysis techniques for clustered data

- Mixed–effects models

- All information can be used

- Clustering at different levels can be analysed

- → Drawback: complex

Define amount of clustering

- Intraclass correlation coefficient (ICC)

- Variance components

Variable	Effect estimate	SE	T-value	P
Intercept	42.16	0.77	54.84	<0.001
Age	0.04	0.05	0.75	0.45
Treatment type	4.17	1.12	3.72	<0.001

Variable	Effect estimate	SE	T-value	P
Intercept	38.11	3.52	10.82	<0.001
Age	−0.10	0.21	−0.47	0.65
Treatment type	4.97	4.46	1.11	0.28

ICC = 0.34

Other things to keep in mind

- Take clustering into account in sample size calculation

- Crossed and nested effects are handled in a different manner
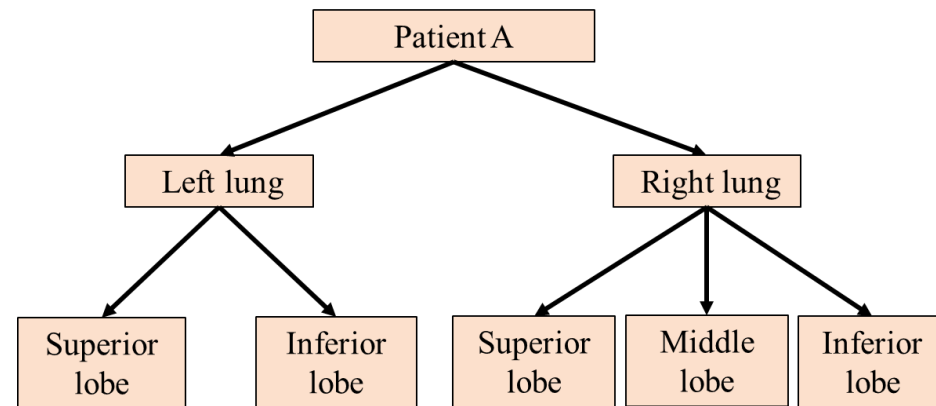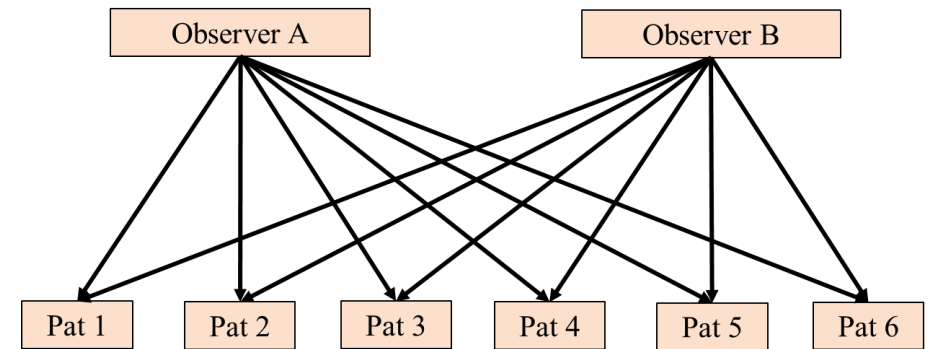
Next "Help! Statistics" lectures

- January – No sessions

- February 7th – Sacha la Bastide "Intro to longitudinal data analysis"

Thank you

Photo by Robin Mathlener on Unsplash

d.c.broekstra@umcg.nl

@DCBroekstra

WeCollaborate!

An initiative to facilitate statistics collaboration

Did you ever get stuck with your analysis because you feel like you are lacking knowledge? Do you feel confident in a specific modelling technique and would be willing to help others in learning more?

We created a Discord server for you to ask questions, provide help, communicate and facilitate collaborations.

The server: Text & Voice channels

\# information-and-rules → a must read

\# choose-your-role → Seeking help / Providing help

\#ᵗ needing-help → describe your project & the benefits for the helper

\#ᵗ providing-help → describe with which statistics/type of analysis you can help with

For whom is the server?
BCN and GSMS
- Research Master Students
- PhD Students
- Post-Docs
- Faculty members

SIGN UP! ☺

NOW IT'S UP TO YOU!
→ Sign up via the QR code (your first + last name + your department)
→ choose your role → ask/provide help!

TIP!
To receive notifications, **bookmark** the server in your browser or **download** the desktop/mobile app!