

# Min kamp — med longitudinelle modellar av pasientrapporterte data

**John Roger Andersen**, professor ved Høgskulen på Vestlandet.  
Forskar i Helse Førde. Leiar av forskingsgruppa Folkehelse, livsstil og overvekt.  
Rådgjevar ved fagsenter for pasientrapporterte data.

[johnra@hvl.no](mailto:johnra@hvl.no)

## Utgangspunkt

- Jobbar med mykje med PRO data, særleg før og etter fedmekirurgi
- Preop, 1 år, 2 år, 5 år og 10 år
- Brukar SPSS og Stata
- Her nokre erfaringar og døme som eg vonar kan vere nyttige

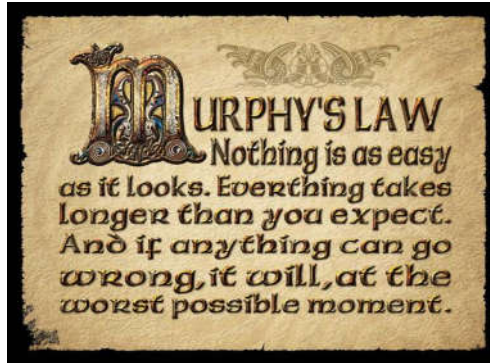
## Grunnleggande prinsipp i analysar av longitudinelle data



**ADVARSEL :** Det er mange tvilsame heimesnikra demoar av korleis du gjer longitudinelle analysar på youtube...

Ta heller eit kurs med praktiske øvingar.

# #1 Det manglar alltid data



I Nurses Health Study (Harvard i USA) feilkopierte dei SF-36 til 120.000 personar utan at det vart oppdaga før det var for seint.

## Manglande data

- Missing completely at random
- Missing at random
- Missing not at random (ikkje lett å vite!)

Gode PRO data krev gode  
rutinar/logistikk og at ein  
er i praksis for å bidra!

Ettersend skjema til pasientar med manglande data.....

Tilsett person som har alt dette som jobb.

#2 Reperert ANOVA og parra t-  
test er ut?

---

*Should we quit using repeated measures analysis of variance?*

---

## Repeated Measures ANOVA, R.I.P.?



Charles E. McCulloch

<https://doi.org/10.1080/09332480.2005.10722732>

## #3 Imputering av missing data –kan det vere løysinga?

(Er det litt skummelt?)



## Ikkje ein dum ide viser det seg!

- Analysar på komplette data = risiko for bias
- Men er frameleis veldig vanleg!
- Multippel imputering (MI) er beste imputeringsmetode..... last value carried forward er bias imputation!
- MI er ikkje “magi” men kan redusere bias og auke power
- **MI ved missing i både x og y variablar, er særleg nyttig**
- Men problem med data som har ulik kovariansstruktur (med meir) står att om ein nyttar repetert ANOVA

## #4 Vi forlet reperert ANOVA for godt og går over til marginale modellar eller mixed effect modellar

**Brukar alle y-data som er tilgjengelege, og modellane er langt meir fleksible og robuste**

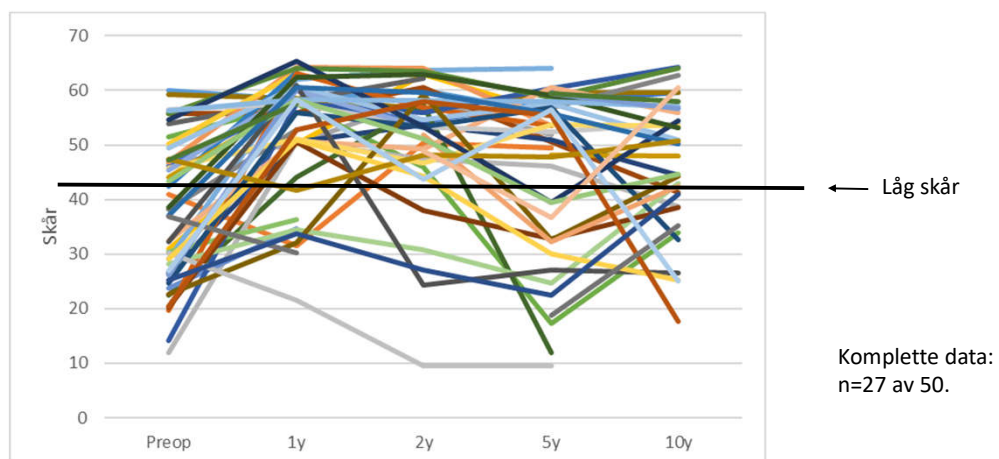
# Men først: bli godt kjent med data – synleggjer dei.....

Deretter kan du velje den beste modellen og evt søke råd.

## Døme

- 50 pasientar før og etter fedmekirurgi (preop, 1, år, 2 år, 5 år og 10 år)
- Mental helse (MCS) frå SF-36
- Høgare skår = betre skår

## Spagettiplot for MCS over tid



## Snitt: listwise versus alle verdier

### Summary of Estimated Means

	OBLMCS1	OBLMCS2	OBLMCS3	OBLMCS4	OBLMCS5
Listwise	37,466	53,660	50,142	46,799	45,366
All Values	37,789	52,500	50,848	45,889	46,229



## Standardavvik: listwise versus alle verdier

**Summary of Estimated Standard Deviations**

	OBLMCS1	OBLMCS2	OBLMCS3	OBLMCS4	OBLMCS5
Listwise	14,5853	8,9035	10,2749	13,5658	12,5389
All Values	12,8100	10,9019	11,8908	15,4864	12,4216

## Kva er kovariasjonsstrukturen?

	OBLMCS1	OBLMCS2	OBLMCS3	OBLMCS4	OBLMCS5
OBLMCS1	212,7309				
OBLMCS2	58,4257	79,2728			
OBLMCS3	52,0059	39,2069	105,5741		
OBLMCS4	38,6880	58,4807	97,8357	184,0312	
OBLMCS5	87,5553	22,1975	57,3762	80,2794	157,2237

	OBLMCS1	OBLMCS2	OBLMCS3	OBLMCS4	OBLMCS5
OBLMCS1	1				
OBLMCS2	,450	1			
OBLMCS3	,347	,429	1		
OBLMCS4	,196	,484	,702	1	
OBLMCS5	,479	,199	,445	,472	1

Tilbake til  
utgangspunktet?

## Marginale modellar

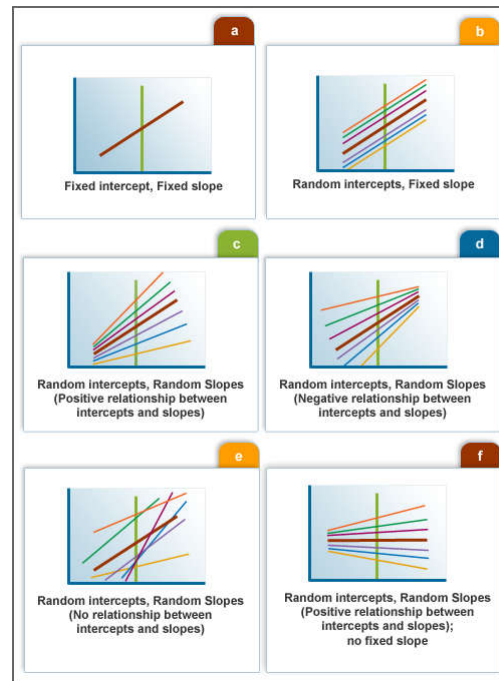
- Modell med generelle estimeringslikninger (GEE)
- **Modell med maximum likelihood estimering (MLE) (min favoritt enkelt....)**
- Når endringar i gjennomsnitt med 95% CI er det du vil ha.
- Funkar også fint til regresjonsanalysar for prediksjon m.m.
- Viktig å velje riktig kovariasjonsstruktur (ustrukturert vil ofte fungere fint om modellen ikkje er for kompleks), men bruker mange fridomsgrader.

## Mixed effect modellar

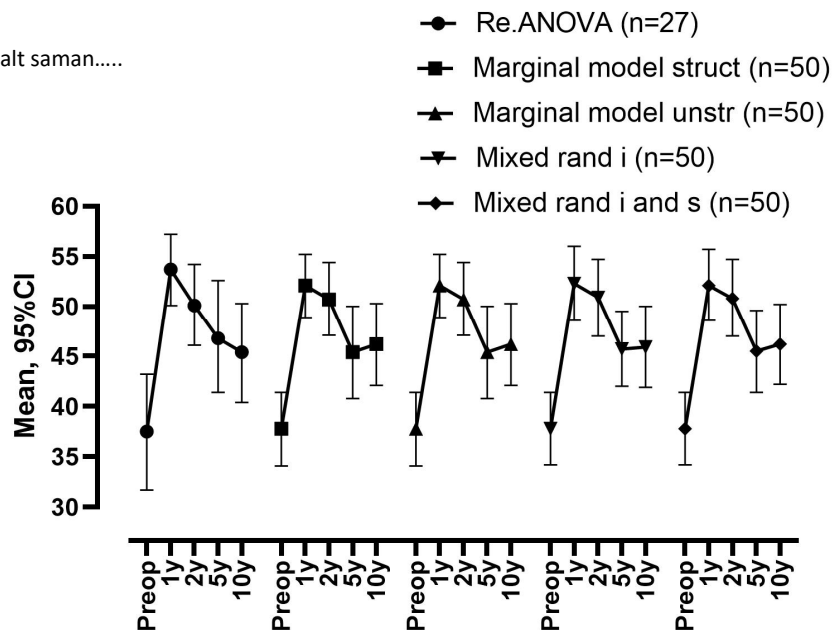
- Supernyttig når kluster må takast omsyn til (sjukehus etc)
- Gir ICC (kor mange prosent er forklart med sjukehus? etc)
- Gir høve til rike analysar
- **Random intercept og slope – kva er det?**
- Random intercect (individuell utgangspunkt)
- Radom slope (individuell utvikling). Unstructured funkar bra som kovarinsstruktur. Sjå på model fit.
- Det hender modellane ikkje konvergerar... prøv noko mindre komplekst...

a. Passar til marginale modellar.  
Mixed effect modellar passer til alt.

<http://www.esourceresearch.org/tabid/334/Default.aspx>

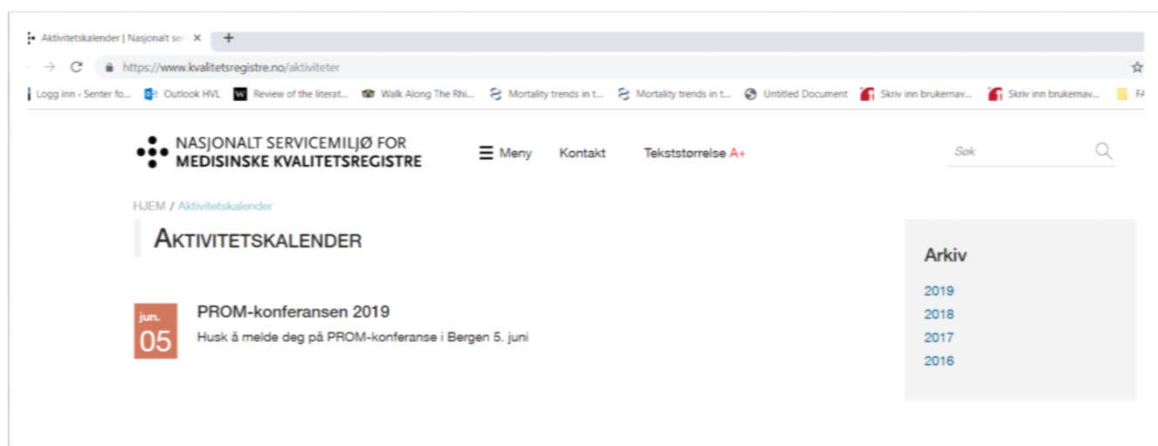


Vi prøvar alt saman.....



## Oppsummering

- Gode rutinar er råd nr 1 for å unngå missing data
- Multippel imputering er valet om ein vil imputere data
- Om du tenkjer på å bruke noko annan imputeringsmetode enn multippel imputering kan du vere på veg inn i trøbbel.
- Marginale modellar og mixed-effect modellar er handterar manglande y-data. Når kun y data (pluss tid) er tema treng du ikkje imputering.
- Mixed effect modellar er særleg fleksible og kan brukast for å justere for cluster-effektar og for å vise ulike responsar over tid m.m.
- Spør for all del erfarne fagfolk. Då kan du i alle fall bli forvirra på eit høgare nivå og det er også framgang!



Følg med på tilboda til Fagsenteret.  
 Kom med innspel til kurs!  
 Ta kontakt!