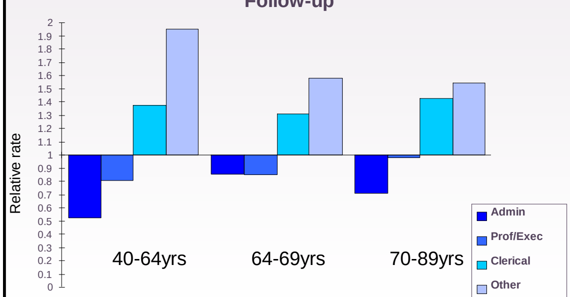
Multilevel Models of Determinants of Health

Hynek Pikhart
University College London

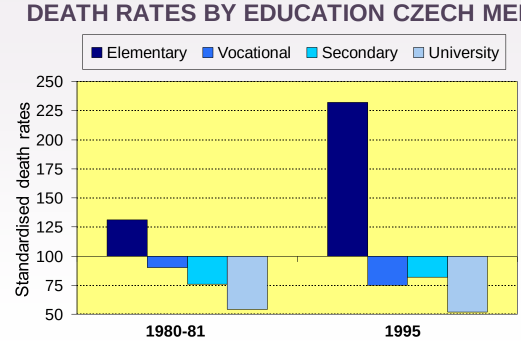
Individual level

Mortality by Grade of Employment Whitehall Men 25 yr Follow-up



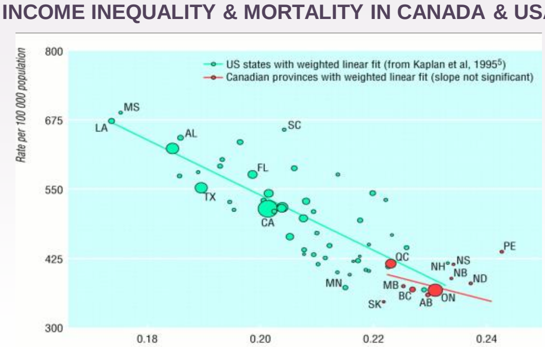
Marmot and Shipley, 1996

DEATH RATES BY EDUCATION CZECH MEN



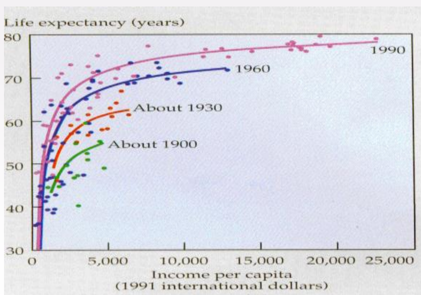
Area level

INCOME INEQUALITY & MORTALITY IN CANADA & USA



Ross et al; BMJ 2000

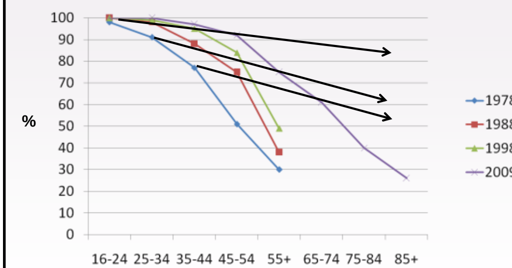
LIFE EXPECTANCY AND INCOME FOR SELECTED COUNTRIES AND TIME PERIODS



World Bank Development Report 1993

7

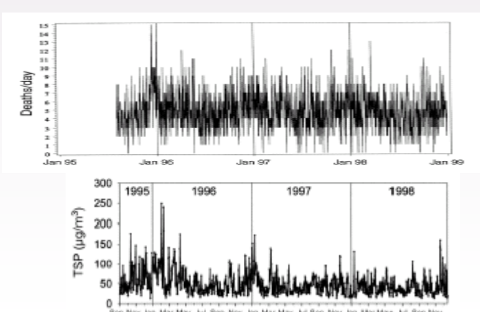
Functional dentition (21+ natural teeth): England



- **Substantial improvement - In two decades half of the oldest population may be able to rely on natural teeth alone**

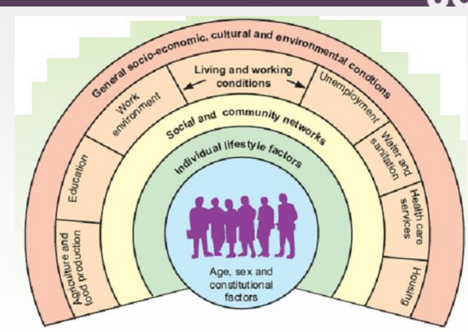
Steele JG, Treasure ET, O'Sullivan I, Morris J, Murray JJ. BDJ 2012; 213: 523-527.

Daily deaths and pollution



Wichman et al, HEI research report, 2000

How can society affect our health?



Dahlgren and Whitehead "rainbow"

Source: Dahlgren G and Whitehead M, Health Inequalities, London HMSO 1998

10

CLUSTERING

4.0	4.0	6.0	2.7
3.9	4.3	5.9	2.6
4.8	5.2	6.0	3.0
5.3	5.4	6.4	2.9

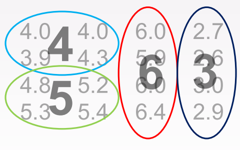
11

CLUSTERING

4.0	4.0	6.0	2.7
3.9	4.3	5.9	2.6
4.8	5.2	6.0	3.0
5.3	5.4	6.4	2.9

12

CLUSTERING



The importance of data structure

"Standard" models not always appropriate

Why do we need additional techniques?

- Often data is **hierarchically structured**
- Real data tend to violate the assumptions of **independence**

Any examples?

What are hierarchical data?

Many kinds of data, including observational data collected in human and biological sciences, have a **hierarchical** or **clustered** structure:

- Children with the same parents tend to be more alike in their physical and mental characteristics than individuals chosen at random from the population at large.
- Individuals may be **nested** within geographical areas or institutions such as schools or employers.
- **Multilevel** data structures also arise in longitudinal studies where an individual's responses over time are correlated with each other.

Examples of natural clustering/grouping

People in households in areas in countries

Pupils within classes within schools

Patients within wards within hospitals

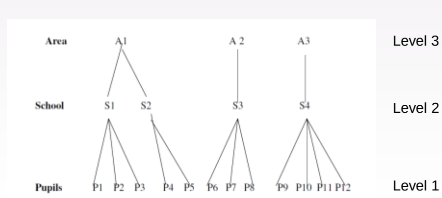
Measurements within people within general practices

Hierarchical / clustered data

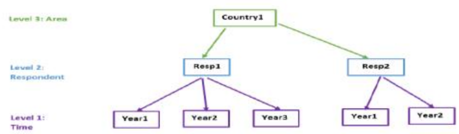
- Hierarchical data:
  - Data clustered / grouped in **space**: different individuals interviewed in the same area (e.g. different pupils within the same school)
  - Data clustered / grouped in **time**: same individuals are measured repeatedly over time (e.g. the same measures of cognitive function gathered at 2-year intervals)
- Observations from hierarchical data structures are correlated as they come from different units that belong to the same group (pupils in classes; persons with repeated measures). These are non-independencies in hierarchical data. This is not taken into account by standard analytical techniques.

Hierarchical structure

Pupils can be classified in schools in areas



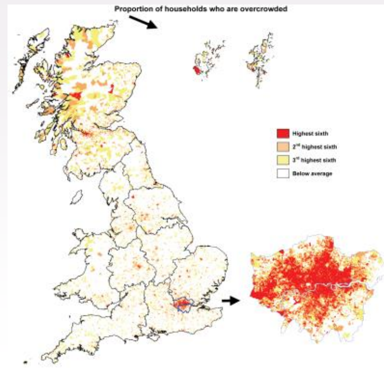
Hierarchical structure



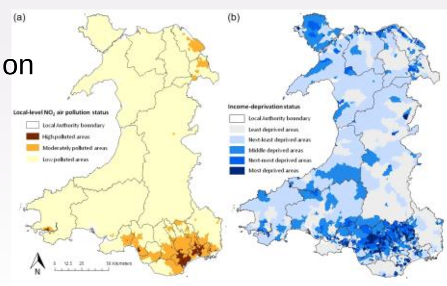
serial measurement **occasions (level 1: *i*)**
clustered within **individuals (level 2: *j*)**
clustered within **area (level 3: *k*)**

19

Housing

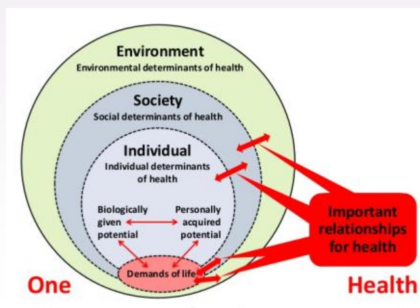


Pollution



From: Air pollution, deprivation and health: understanding relationships to add value to local air quality management policy and practice in Wales, UK
J Public Health (Oxf). 2016;39(3):485-497. doi:10.1093/pubmed/fdw084

Place Matters



The independence assumption
A: Data collection

Survey data rarely come from a Simple Random Sample (SRS)

Surveys often have multi-stage designs

- Cost advantages.
- Often necessary when there is no suitable frame for households (or individuals)

Result: **clustered data**

- i.e. the data collection process generates observations that are **not independent** e.g. clustered by geography / time / household, etc.

23

The independence assumption
B: Structures in the population

Even if we have collected data in an unclustered way there is still 'natural' clustering in the population, as we have already remarked.

We want to take a principled approach - build a model that represents the population from which the data was taken.

Therefore the **impact of clustering should be taken into account** and may itself be of substantive interest.

24

The independence assumption C: Dangers

The independence assumption is unrealistic

- for example we expect positive correlation between exam results of pupils from the same school

Ignoring correlation wrongly estimates standard errors

- because we assume an overly simplistic model structure leading to an **overstatement (sometimes understatement) of statistical significance**.

Consequently we might believe our conclusions to be statistically significant when in fact they are not, or vice versa.

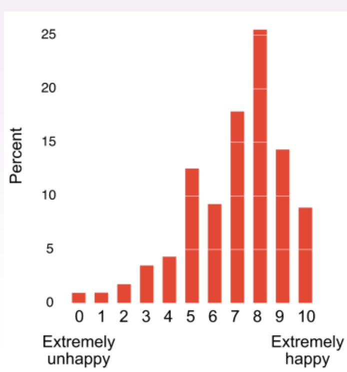
25

Example: The European Social Survey

- Cross-sectional, biennial study of European countries (2002-2012).
- Freely available online www.europeansocialsurvey.org
- Wide range of topics (core and rotating modules)
- Example here comes from the 2010 wave (27 countries)

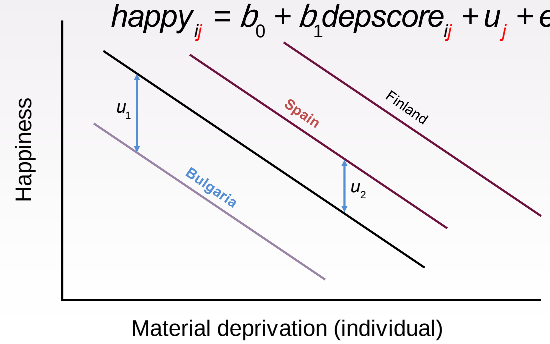
Happiness and material deprivation

"Taking all things together, how happy would you say you are?"

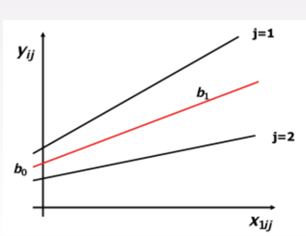


Random intercept: parallel slopes

$$happy_{ij} = b_0 + b_1 depscore_{ij} + u_j + e_{ij}$$



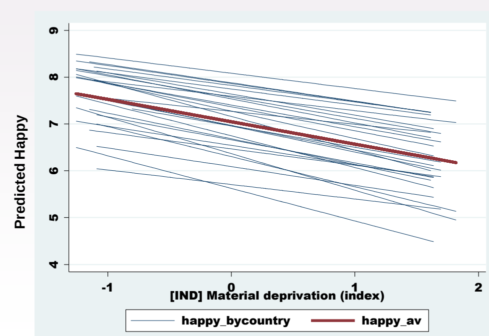
We can extend the model by allowing the gradient / slope as well as the intercept to vary with cluster



Random Intercept: heterogeneity = trajectories with intercept above or below average (b_0)
Random Slope: heterogeneity = trajectories with slope above or below average (b_1)

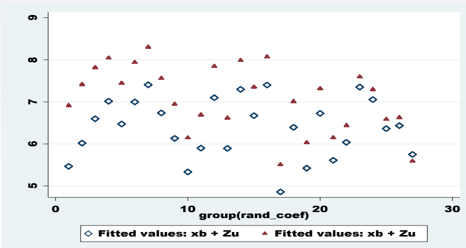
29

Predicted happiness by deprivation score



Predicted Happy by country

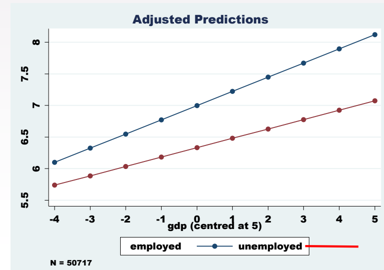
Unemployed: Diamonds, Employed: Triangles



Countries ranked by unemployed-employed difference:
Larger differences on left.

Predicted Happy score

Predicted happy score by country GDP for employed and unemployed individuals



Longitudinal data - growth curve models

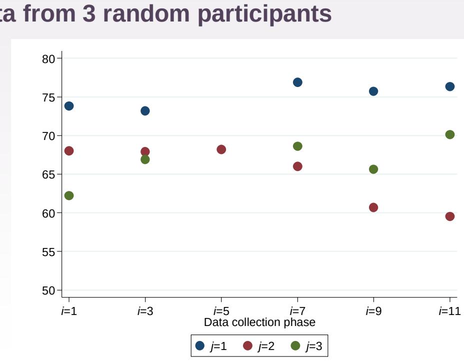
- We may be interested in overall trajectories of change in observations at time points labelled *i* for individuals labelled *j*
- We can fit random effects models to longitudinal data –**growth curve models**
- Growth curve models fit different intercepts and/or slopes for each individual allowing for variation between and within individuals
- Advantages:
  - Account for dependent observations
  - Possible to estimate individual trajectories

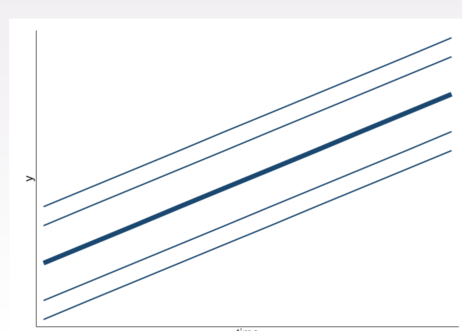
Example

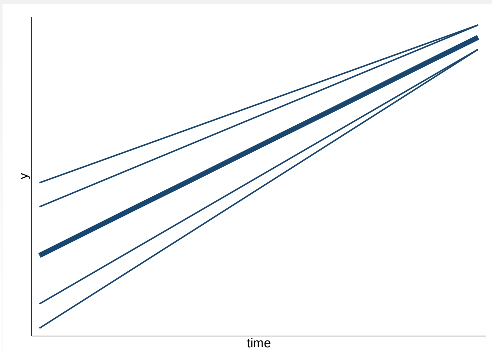
Data from Whitehall II study

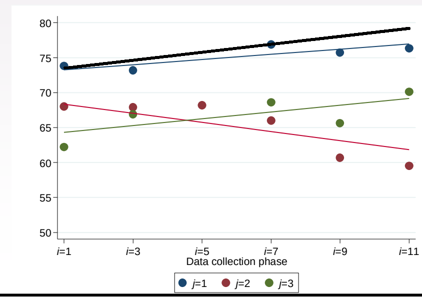
- Over 10,000 participants at start of study
- First wave of data collection in 1985
- Participants aged 35-55 when recruited
- Repeated data collection
- Outcome of interest: weight
- Data collected in waves 1,3,5,7,9,11

Data from 3 random participants

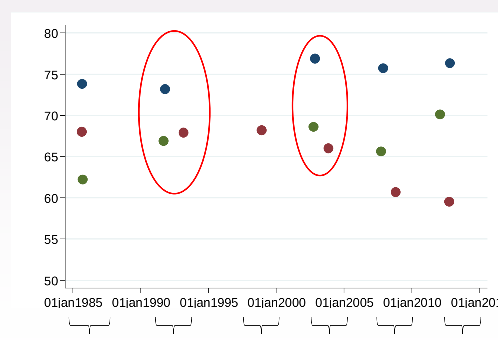




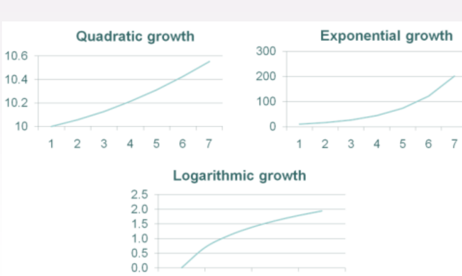




Different time specification



Non-linear trajectories



Summary

- Multilevel models essential in situation when
  - Non-randomly selected data
  - Hierarchical structure of data
  - Contextual variables
  - Repeated measures over the time
- Methods to deal with such data different from "standard" regression methods
- Random data – very rare in real situations – be careful when evaluating existing evidence