DDZ German Diabetes Center

Leibniz Association

HEINRICH HEINE UNIVERSITÄT DÜSSELDORF

Sequences of prevalence surveys with dependent data – new insights based on spatial statistics

Annika Hoyer, Ralph Brinks

Deutsches Diabetes-Zentrum (DDZ), Leibniz-Zentrum für Diabetes Forschung an der Heinrich-Heine-Universität Düsseldorf,

Institut für Biometrie und Epidemiologie

annika.hoyer@ddz.uni-duesseldorf.de

Overview

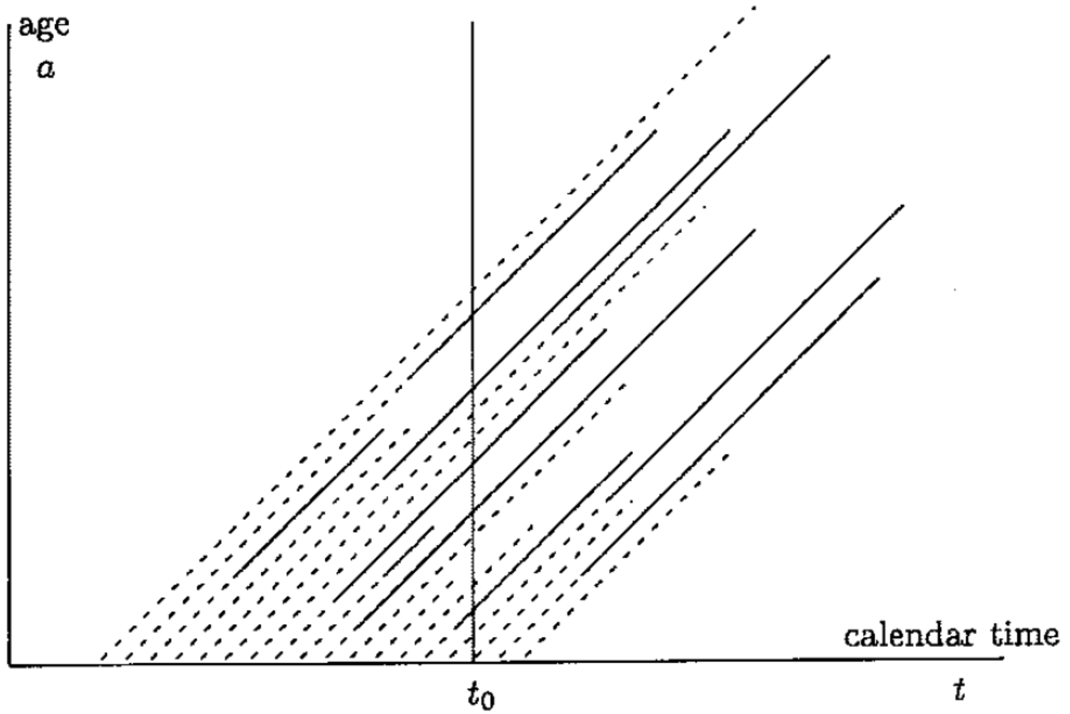
1. Background
2. Methods
3. Example
4. Outlook

BACKGROUND

Background

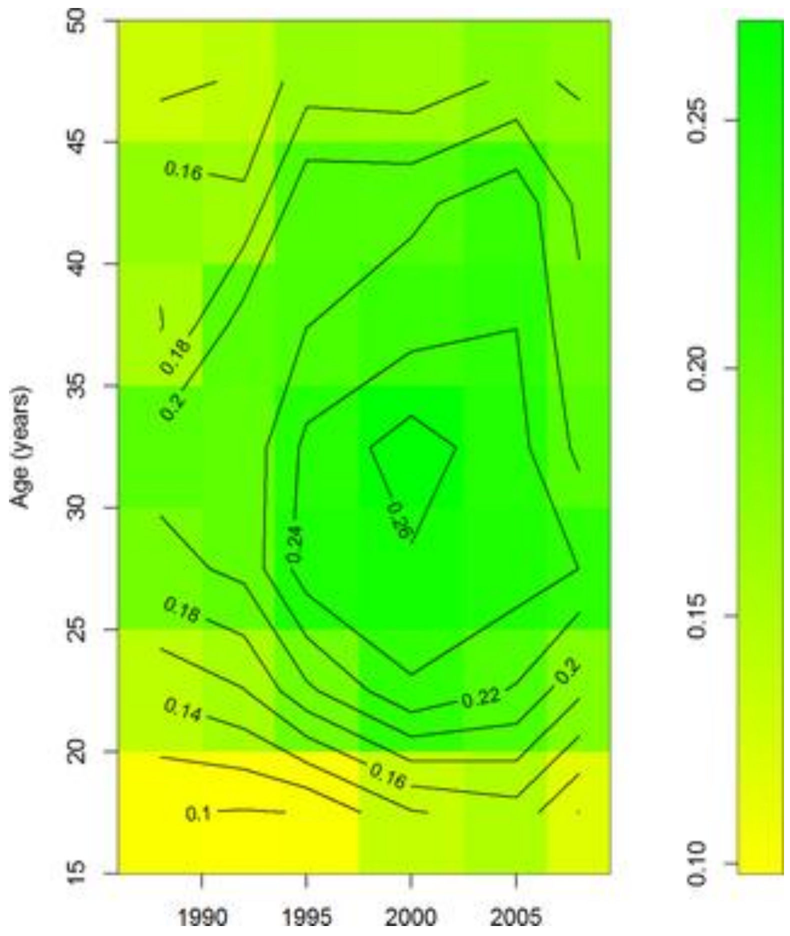
- sequence of surveys about age-specific prevalence of a chronic disease may be represented in a Lexis diagram (Keiding 2006)
- map-like illustration of the relevant variables (calendar time and age)

Lexis diagram



- … : individuals born healthy
- - : individuals becoming (irreversibly) ill

Representation of multiple cross-sections



- Six cross-sections
- Prevalence of each cross-section color-coded

Lexis diagram – even more complex

- sometimes, the data points in the multiple cross-sections are not independent
- contain data from people surveyed repeatedly
- Aim: estimating temporal trends of the age-specific prevalence

→ challenge to deal with that type of data because of existing dependencies

METHODS

Kriging

- Aim: estimation of the (true) surface
- Given data: prevalence($time_i$, age_j)
- Model:

$$Prev(t_i, a_j) = X(t_i, a_j)\beta + Z(t_i, a_j) + \varepsilon(t_i, a_j)$$

with:

- $X(t_i, a_j)$ = matrix of spatially referenced non random variables
- β spatial trend parameter vector
- $Z(t_i, a_j)$ spatial random effect
- $\varepsilon(t_i, a_j)$ measurement error process

In our case

$$Prev = \beta_0 + \beta_1 a^2 + \beta_3 at + \beta_4 t^2 + Z(t_i, a_j) + \varepsilon(t_i, a_j)$$

1. Estimation by a linear mixed model (lme)

2. Estimation based on thin plate splines (tps)

EXAMPLE

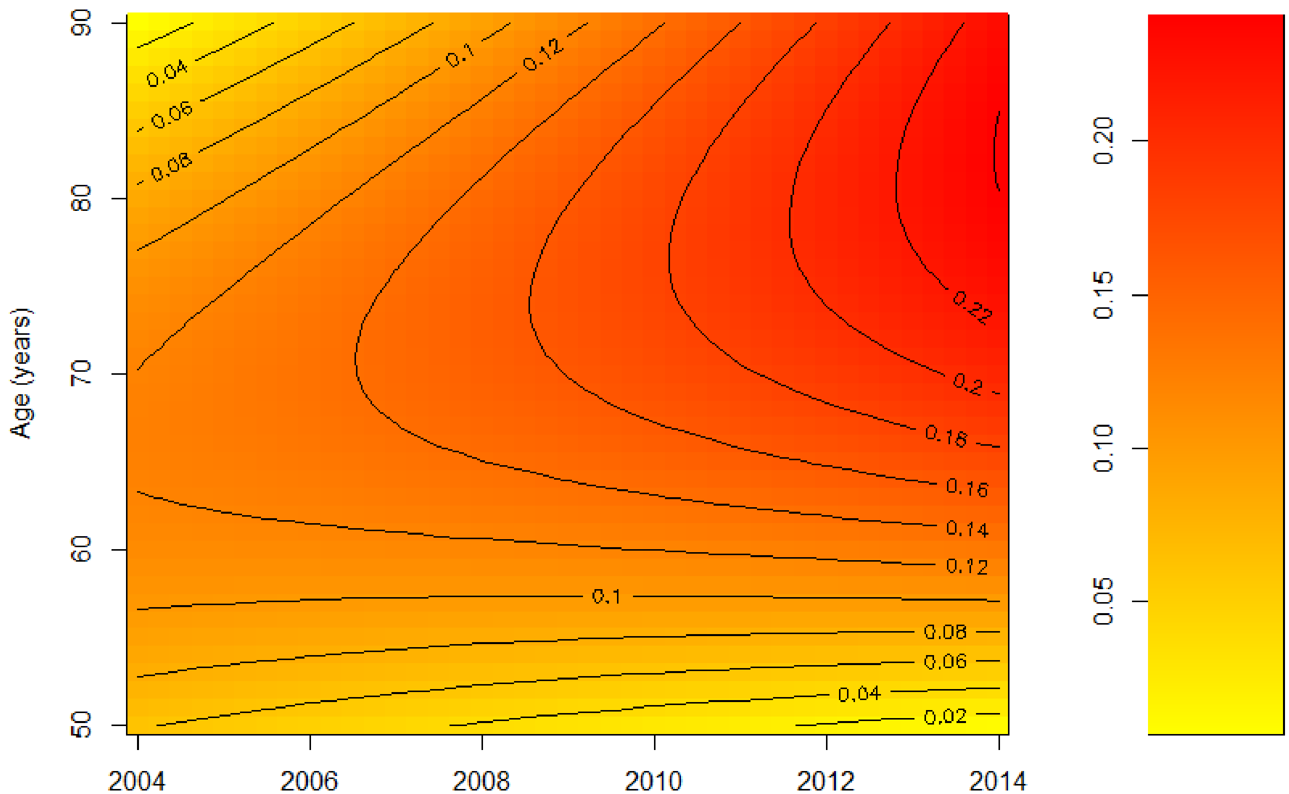
The SHARE study

- SHARE: **S**urvey of **H**ealth, **A**geing and **R**etirement in **E**urope
- Open cohort study conducted in several waves
- Multidisciplinary and cross-national panel database of micro data on health, socio-economic status and social and family networks
- approximately 123,000 individuals aged 50 or older (more than 293,000 interviews)
- Covers 27 European countries and Isreal

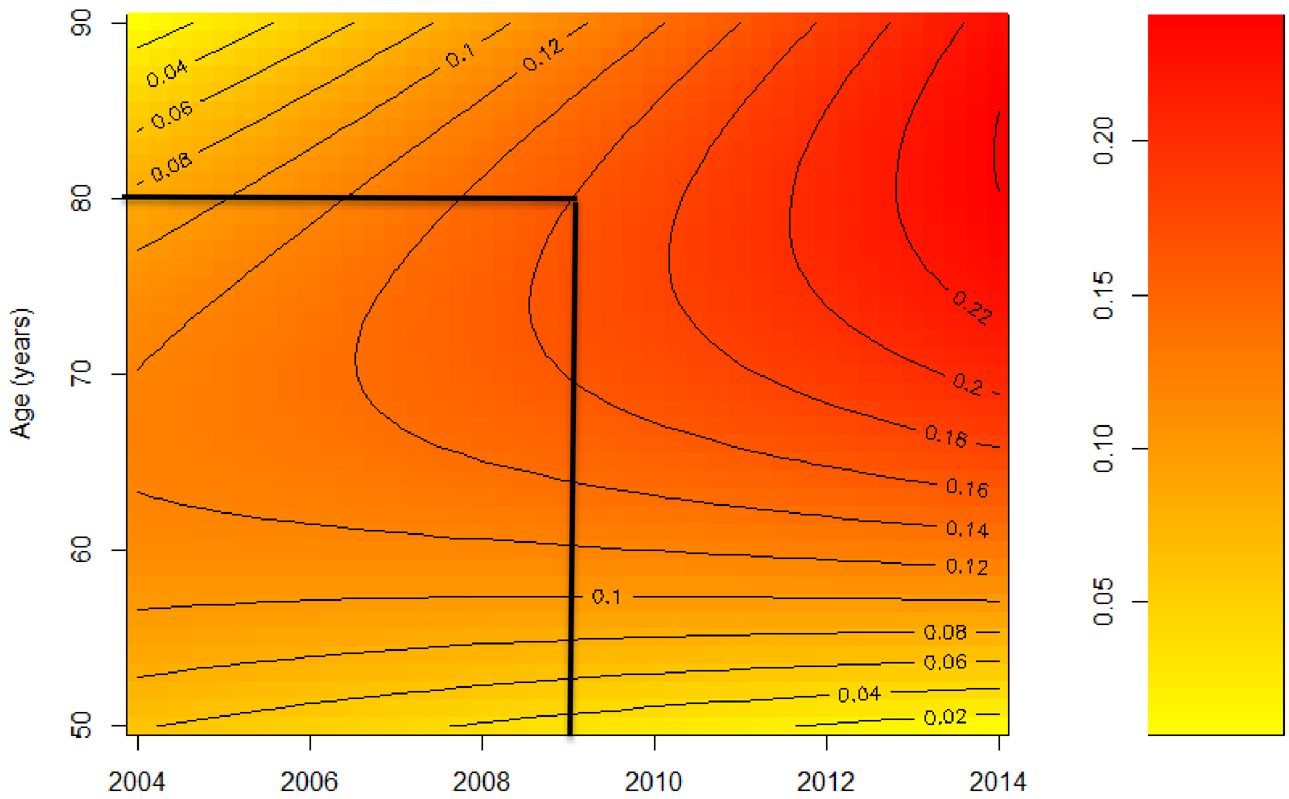
The SHARE study

- Open cohort study → persons are surveyed repeatedly
- Valid model should account for potential dependencies
- Our aim: estimate trends in the prevalence of type 2 diabetes mellitus in Germany

Results – based on the lme



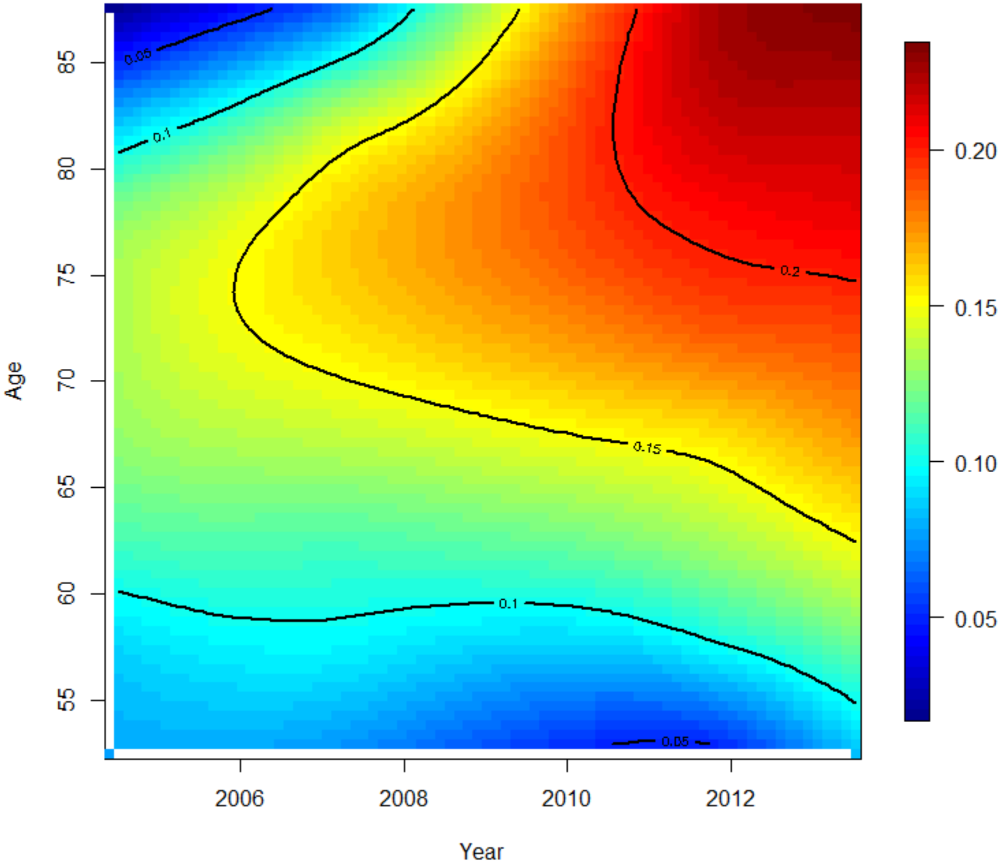
Results – based on the lme



An 80 years old man has a prevalence 16% in 2009.

DDZ German Diabetes Center

Results – based on tps



OUTLOOK

Outlook

- Model prevalences directly → using a generalized linear mixed model
- Investigate different types of „spatial“ correlations
- Performing a simulation study to compare models