

Neighbourhood sampling and neighbourhood characteristics

Odile Sauzet

Bielefeld School of Public Health
Bielefeld University

March 1, 2024



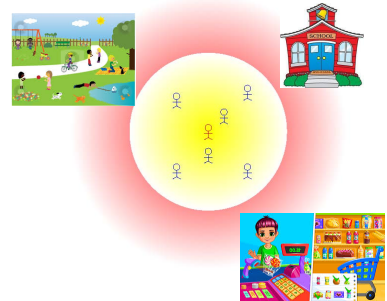
Outline

- 1 Background: neighbourhood effects on health
- 2 Three possible use for neighbourhood sampled data
- 3 Limitations and discussion

- The study of neighbourhood effects on health has shown that ecological "neighbourhood" factors have an effect on health.
- Large interest in neighbourhood level social determinant of health
→ Social capital
- Evidence of a mediating role of perceived neighbourhood social cohesion of the effect of environmental stressors (noise, air pollution) and built environment on health
- Presentation Maria Schäffer

- Typical methodology: neighbourhood level characteristics are measured ecologically at some administrative unit level (in the US: electoral wards) and use multilevel modelling to adjust for unmeasured neighbourhood level characteristics
- This has been largely criticised as not necessarily the relevant exposition scale for the outcome of interest
- Basile Chaix has proposed to consider ego-centred neighbourhood which has been operationalised as "correlation neighbourhood" by OS

- What is a neighbourhood?
- This question is of very little concern in the literature
- Bernard et al have proposed a theoretical framework which includes several dimensions in particular "local sociability"
- Cattell reported qualitative evidence of a two-scale neighbourhood with relevance for health: the immediate surrounding where social contact occurs and the neighbourhood of infrastructure (schools, hospitals ...)

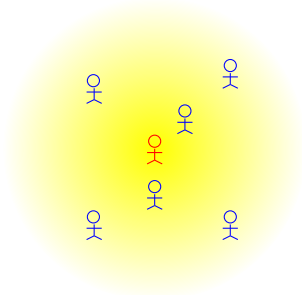


Who are the neighbours?

- Those with whom one has social contact
- Those with whom one has weak ties
- Those who give the neighbourhood its characteristics: define the norms, the feeling of security, the trust etc.

- In order to empirically study the effect of neighbourhood social mechanisms we need data on neighbours
- Neighbours are defined as those who live within a correlation neighbourhood
- Here we argue for more neighbourhood sampling
- Three examples of where neighbourhood sampling is required

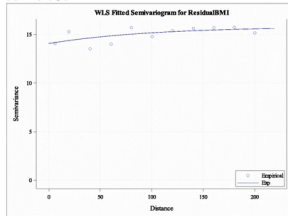
Application I: correlation neighbourhood



- The contextual phenomenon expresses itself as the **spatial correlation** of individual health status.
- Not dependent on administrative boundaries
- Ego-centred approach
- Complement existing approaches including multilevel modelling → modelling of residuals

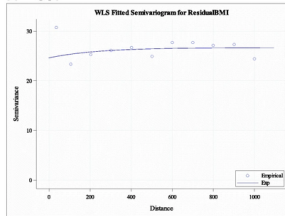
Breckenkamp et al 2021: application to cardiovascular risk factors using three different studies

RECORD Study
N = 7137



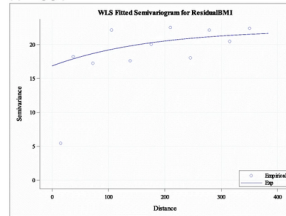
Parameter Estimates
Nugget: 14.07
Partial sill: 1.74
Range: 100.07

BaBi Study
N = 807



Parameter Estimates
Nugget: 24.62
Partial sill: 2.08
Range: 237.11

DHS
N = 964



Parameter Estimates
Nugget: 17.04
Partial sill: 4.98
Range: 175.10

Weighted least squares fitted semi-variograms for residual BMI (adjusted for gender, age, income and educational achievement) of the RECORD, BaBi and DHS studies

Quality parameters for the estimation of the semi-variogram

	Record Study	Babi Study	DHS
Population density	3763	1290	2091
inhabitants per qm ²	(Unité urbaine de Paris)		
Number of pairs (BMI data)			
0 – 20 m	362	8	7
0 – 50 m	1210	78	56
0 – 100 m	3788	234	195
Variance			
BMI	17.51	29.56	21.11
Residual BMI	16.52	27.82	18.89
Sill (estimated)	15.81	26.70	22.02
Variance			
SBP	304.57	154.96	452.72
Residual SBP	255.69	151.39	343.87
Sill (estimated)	243.21	148.44	353.70

Application II: Interpolation

- Data on neighbourhood characteristics are not typically collected in epidemiological studies.
- They are however useful in the study of small-area health inequalities and may be available in social surveys.
- We proposed to use Kriging based on semi-variogram models to predict values at non-observed locations with the aim of obtaining indicators of neighbourhood characteristics of epidemiological study participants. (See Poster Emily Finne)

Interpolation

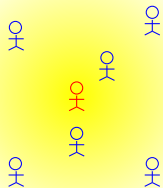
- The spacial data available for Kriging is usually sparse at small distance
 - even if we performed a neighbourhood sampling the sparsity would remain compared to the typical experimental situation for which these methods have been developed
- We perform a simulation study to assess the feasibility and usability of the method
- Case study using data from the RECORD study
- For lack of adequate data we have not applied the method for a real research question

no. sampled points	distance (radius)	Points within radius: M (SD)
500	250	0.3 (0.6)
	500	1.5 (1.5)
	756	3.1 (2.6)
	1000	5.1 (3.8)
	1250	7.4 (5.4)
1000	250	0.8 (1.0)
	500	2.9 (2.4)
	756	6.1 (4.6)
	1000	10.4 (7.4)
	1250	15.4 (10.4)
2000	250	1.8 (1.9)
	500	6.4 (5.0)
	756	12.8 (9.7)
	1000	20.7 (15.1)
	1250	30.7 (21.3)
5000	250	4.3 (3.7)
	500	15.6 (11.2)
	756	31.5 (22.1)
	1000	51.7 (35.6)
	1250	76.2 (51.2)
7090	250	6.1 (4.9)
	500	22.2 (15.8)
	756	45.1 (31.8)
	1000	73.9 (51.2)
	1250	108.8 (73.7)

RECORD Data: Sample size in the correlation neighbourhood for different random sampling

A proportion of predicted values in the correct quintile or its neighbours of 90% was obtained for a sample of 2000 or more.

Application III: neighbourhood social cohesion



- We can estimate a "perceived neighbourhood social cohesion" from survey data
- Ego-centred and estimated using the Kriging method of Application II
- Defining a measure of neighbourhood social cohesion which comes as close as an objective measure as it get
- Could be used as alternative to the individual perception of social cohesion

Discussion points

- What does this mean for survey data?
- How to keep requirement of representativeness
- More neighbourhood sampling would mean less geographical diversity given same financial mean
- How to perform ego-centred sampling?