The effect of spatial resolution in disease mapping: A simulation study

Garyfallos Konstantinoudis; Christian Kreis; Alexios Karagiannis; Dominic Schuhmacher; Ben Spycher

Institute of Social and Preventive Medicine ISPM

University of Bern

7th September 2017

Background

UNIVERSITÄT BERN

- Mapping of cancers
  - Allocating resources in prevention
  - Elucidating causes
- Childhood leukaemia: 5.4 per 100,000 person years
- Environmental risk factors
- Data availability in Switzerland
- Confidentiality issues
- Methods

Schweizer Kinderkrebsregister
Registre Suisse du Cancer de l'Enfant
Registro Svizzero dei Tumori Pediatrici
Swiss Childhood Cancer Registry

Previous studies

UNIVERSITÄT BERN

- Areal data: Conditional Autoregressive (CAR) models
  - Besag *Ann Inst Statist Math* 1991
  - Faure et al. *European Journal of Cancer Prevention* 2009, Thompson et al. *Cancer Causes & Control* 2007, Manda et al. *Eur J Epidemiol* 2009.
- Precise data: Log Gaussian Cox process (LGCP)
  - Møller et al. *Scand J Stat* 1998
  - Cancer mapping: Lung cancer in Spain (Diggle et al. Stat Sci 2013), Colon and rectum in Minesota (Liang et al. Ann Appl Stat 2008)
  - none for childhood cancers.
- Simulation studies:
  - Lung and stomach cancer (Li et al. *J R Stat Soc C-Appl* 2012)
  - Syphilis (Li et al. *Methods in Medical Research* 2012)
  - LGCP outperforms CAR

Aim

- For rare cancers (for instance childhood leukamia) does LGCP provide additional benefits over CAR models on aggregated data in:
  - Quantifying the risk in space
  - Identifying high risk areas

Methods: Data Availability

UNIVERSITÄT BERN

- Cases
  - Swiss Childhood Cancer Registry (SCCR)
  - >90% coverage since 1985
  - Precise location
- Population
  - Census (1990, 2000, 2010 onwards)
  - Precise location
- Geographical units in Switzerland
  - 26 Cantons
  - 2353 municipalities

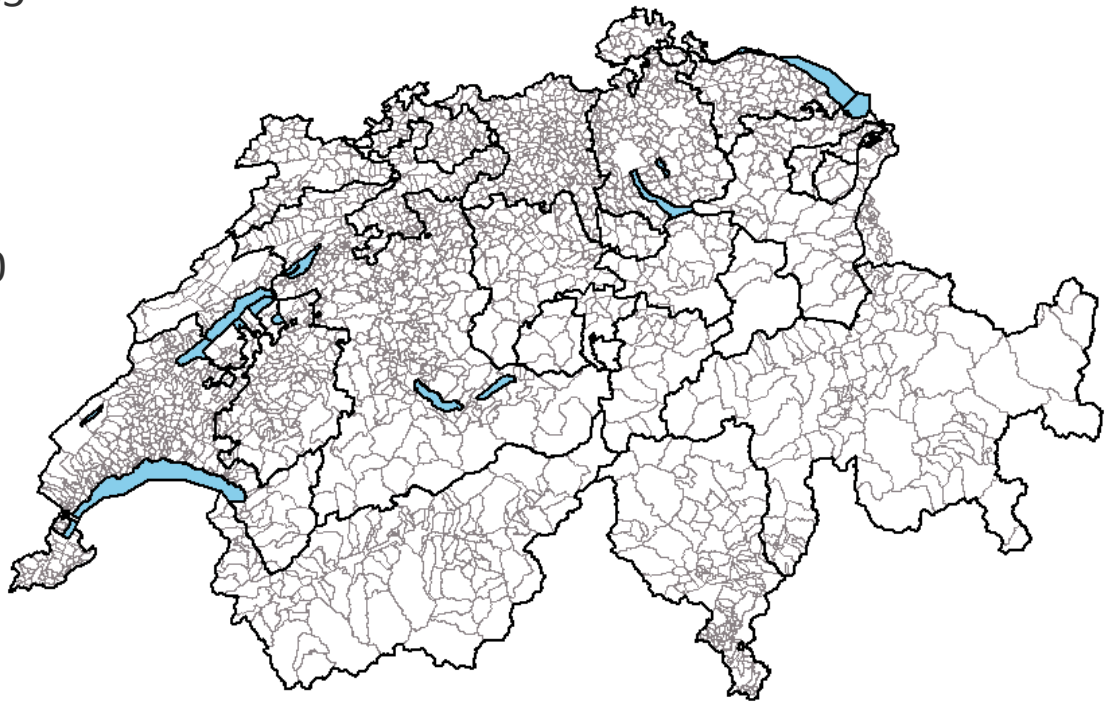


Figure. Geographical units in Switzerland

Methods: Data Simulation

- Canton of ZH (168 municipalities)
- 205,242 (15%) children
- Leukaemia incidence 1985-2015 (n = 334)
- Simulation scenarios
  - 3 high risk areas
  - 2 and 10-fold increase (constant)
  - Circles with 5km radius

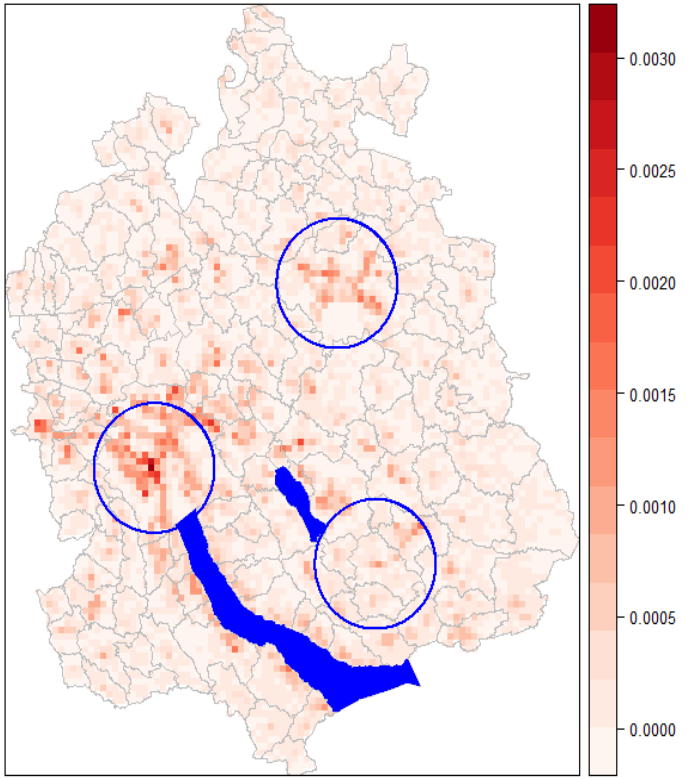
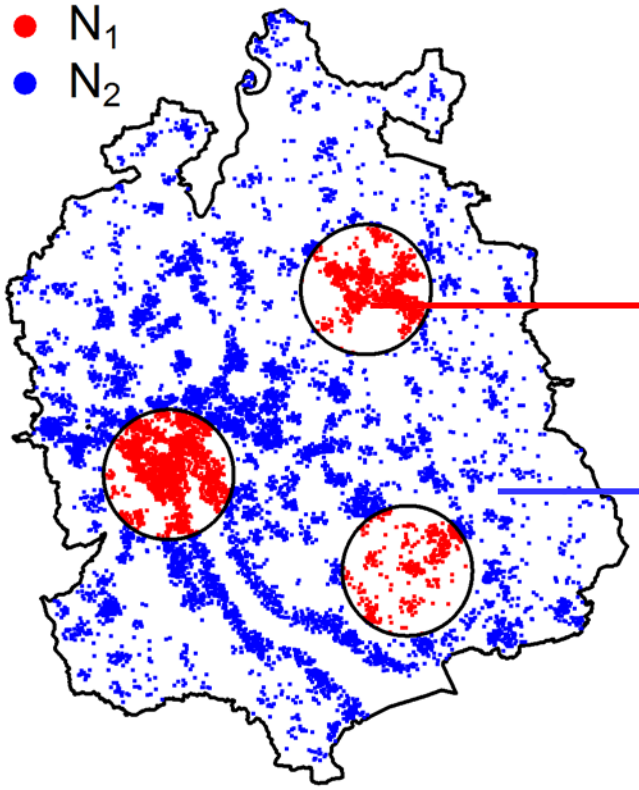


Figure. Population density

Methods: Data simulation



- Let n be the number of leukaemia cases
- N_1 be the number of residents inside the circle
- N_2 be the number of residents outside the circle
- $p = \frac{n}{c \times N_1 + N_2}$

$$X_1 \sim Binomial(c \times p, N_1)$$

$$X_2 \sim Binomial(p, N_2)$$

Methods: Model description

> CAR (BYM) model on municipalities

$$\log(Y_i) = \log(M_i) + \beta_0 + u_i + v_i \,,$$

$$u_{i|-i} \sim N\left(\frac{\sum w_{ij} u_j}{\sum w_{ij}}, \frac{\sigma_1^2}{\sum w_{ij}}\right), v_i \sim N(0, \sigma_2^2), \ \mathrm{i} = 1, \ldots, \mathrm{m}$$

> LGCP model

$$\mathrm{Y(s)} = \lambda(s)\exp\{\beta_0 + u(s)\}$$

$$u(s) \sim GRF(\mu, \kappa)\,, \kappa(|h|) = \sigma^2 \rho_\nu(|h|/\phi),\ \rho_\nu(\cdot) \text{ is Matérn}$$

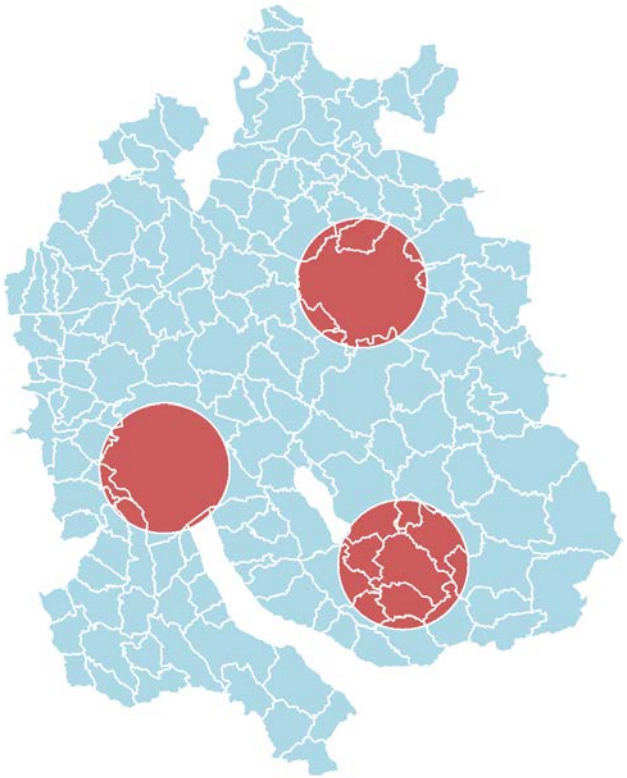
> Inference with Integrated Nested Laplace Approximation (INLA)

Simulation Metrics

- Root mean integrated square error (RMISE):

$$RMISE = \left\{ E\left[\int w(s) \left(\hat{R}(s) - R(s) \right)^2 ds \right] \right\}^{1/2}$$

  - $RMISE_1$ for w = 1
  - $RMISE_2$ for w population intensity weights



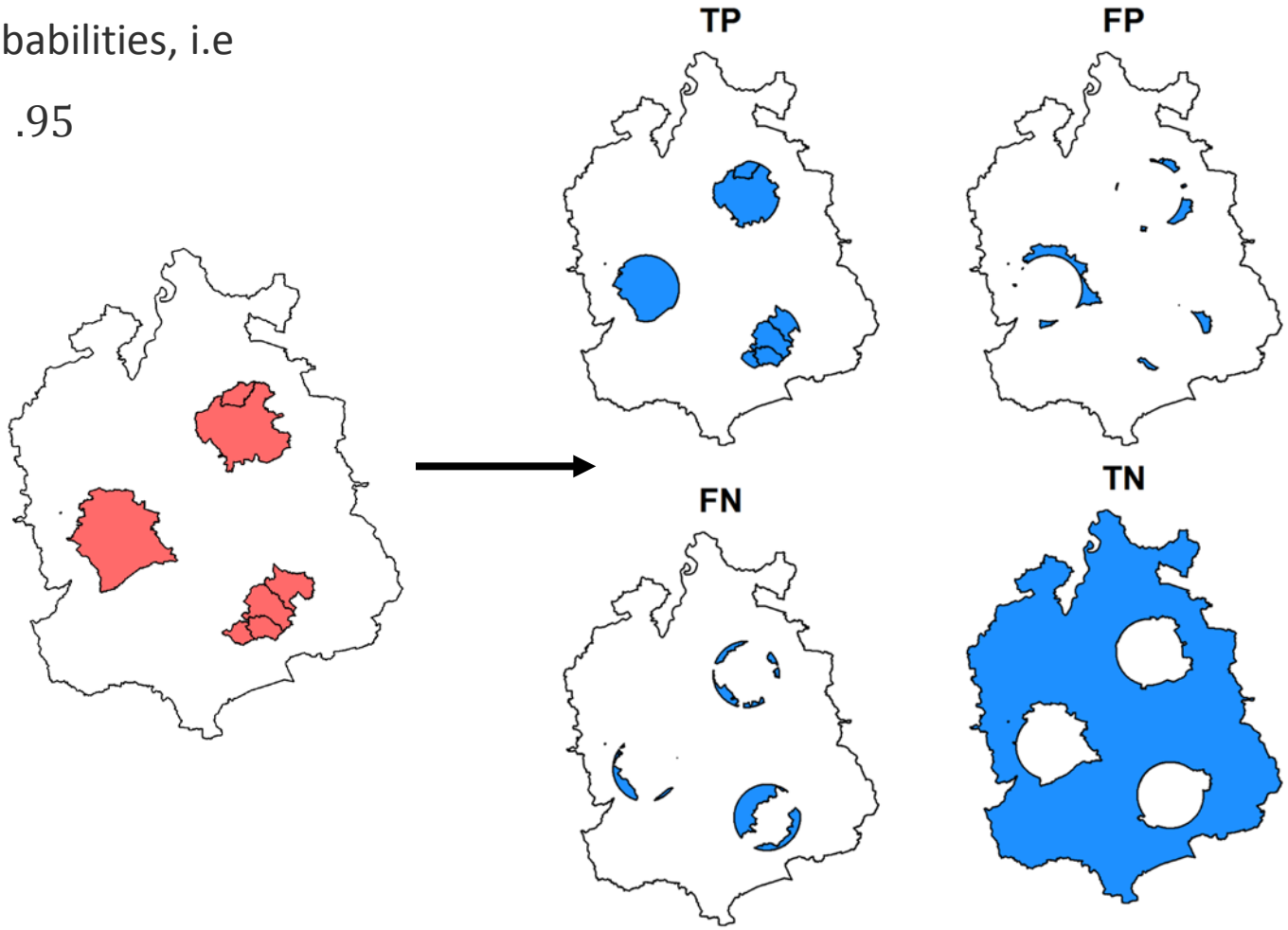
- Sensitivity and Specificity
  - Area based
  - Case based

Simulation Metrics: Sensitivity/Specificity

- Exceedance probabilities, i.e

  $\mathrm{pr}[u(s) > 0] > .95$

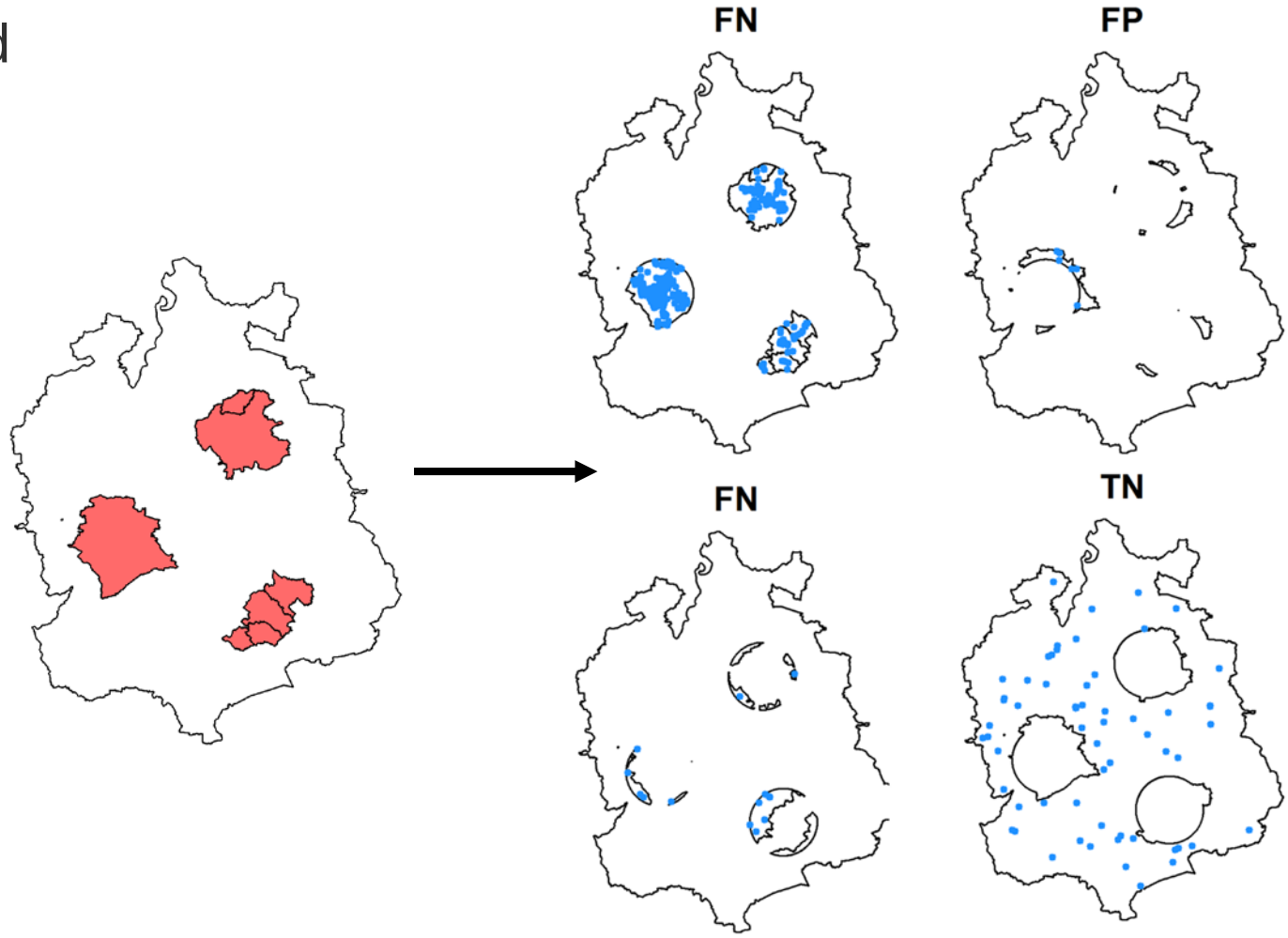
- Area based
- Case based



Simulation Metrics: Sensitivity/Specificity

UNIVERSITÄT BERN

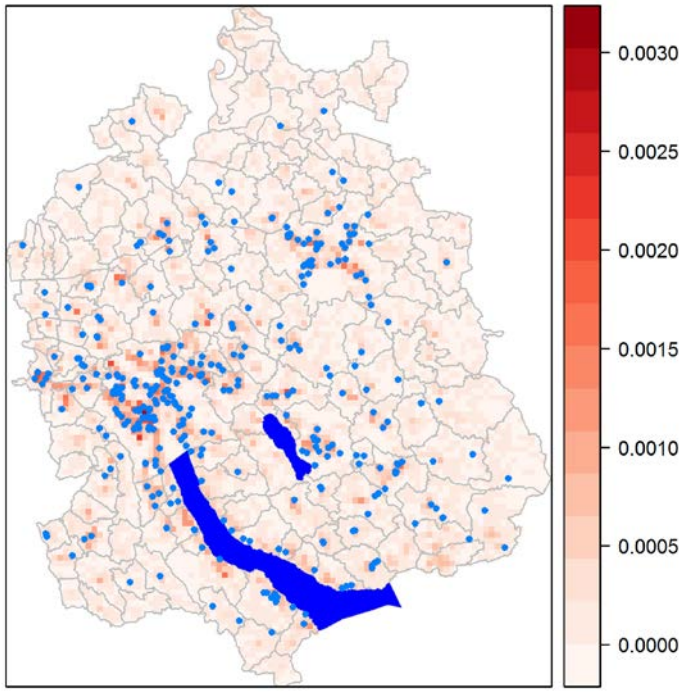
- Case based

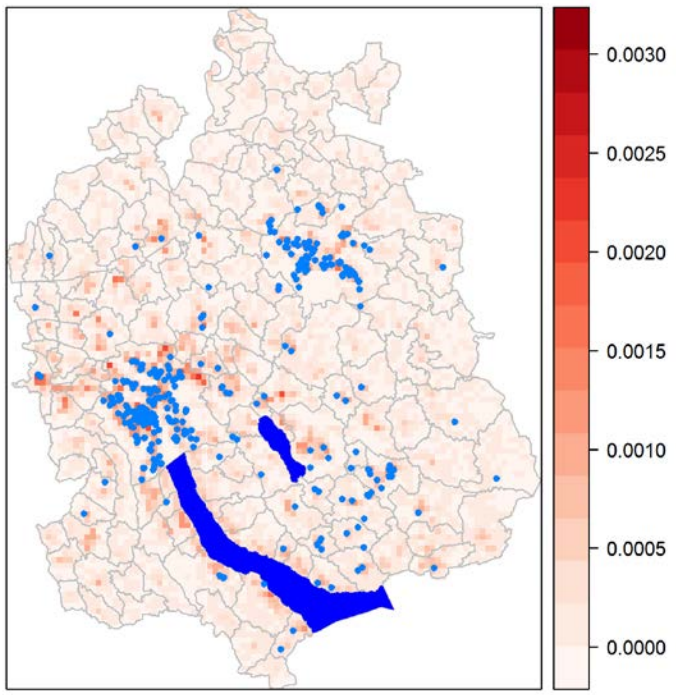


Results

UNIVERSITÄT BERN

> Results will focus of 5km radius and 2-fold (left plot) or 10-fold (right plot) increase within the high risk area



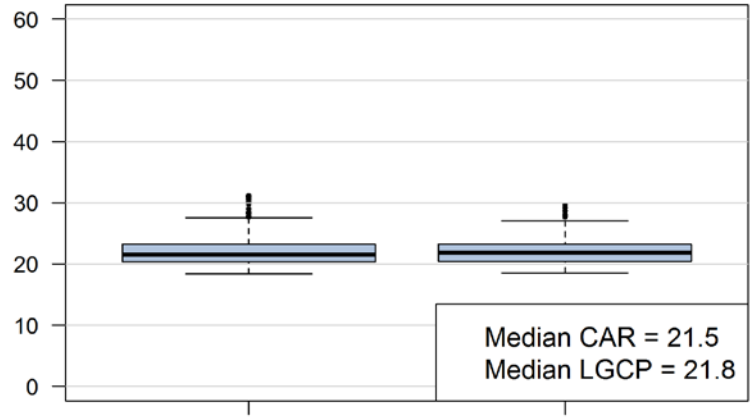


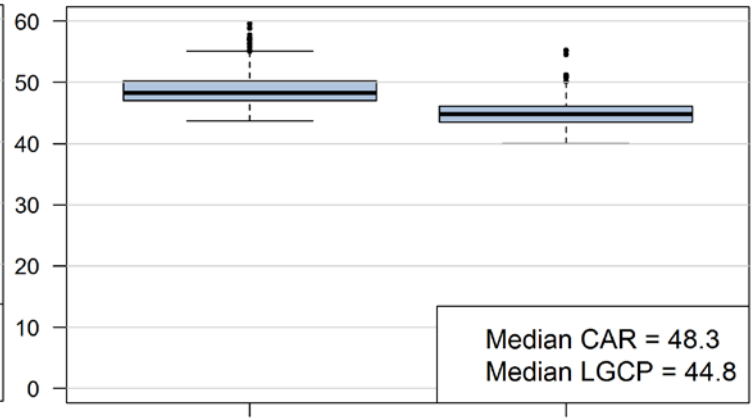
Results: RMISE

2-fold increase 5km radius

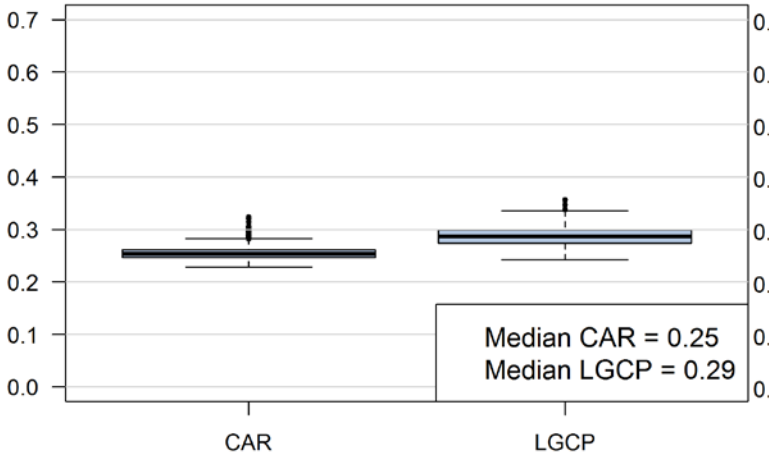
10-fold increase 5km radius

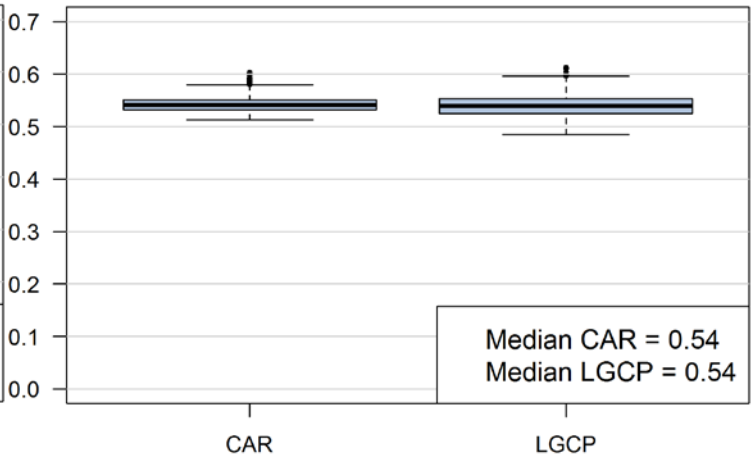
$RMISE_1$





$RMISE_2$





Results: Sensitivity/Specificity (Areas)

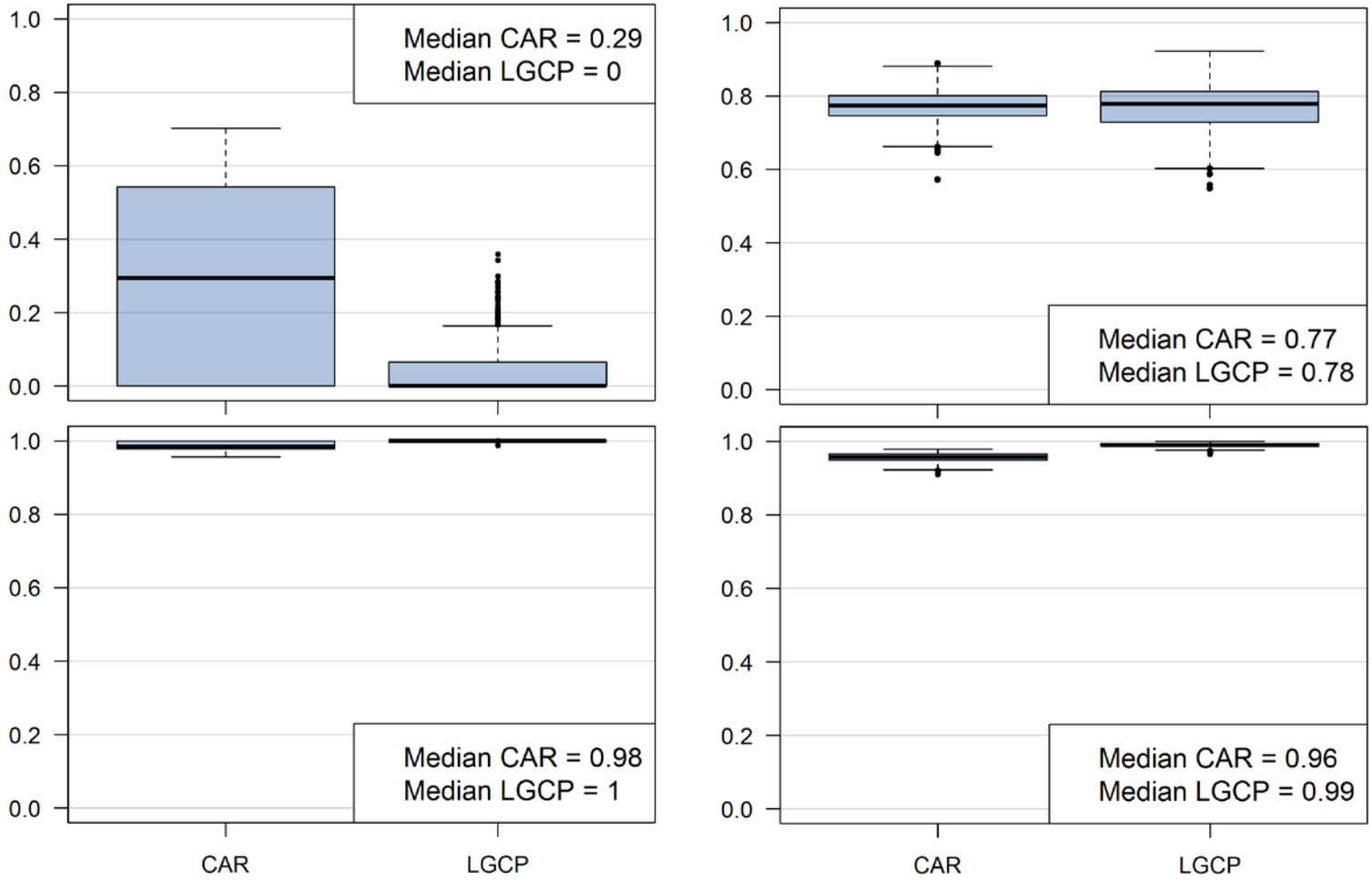
UNIVERSITÄT BERN

2-fold increase 5km radius

10-fold increase 5km radius

Sensitivity

Specificity



Results: Specificity/Specificity (individuals)

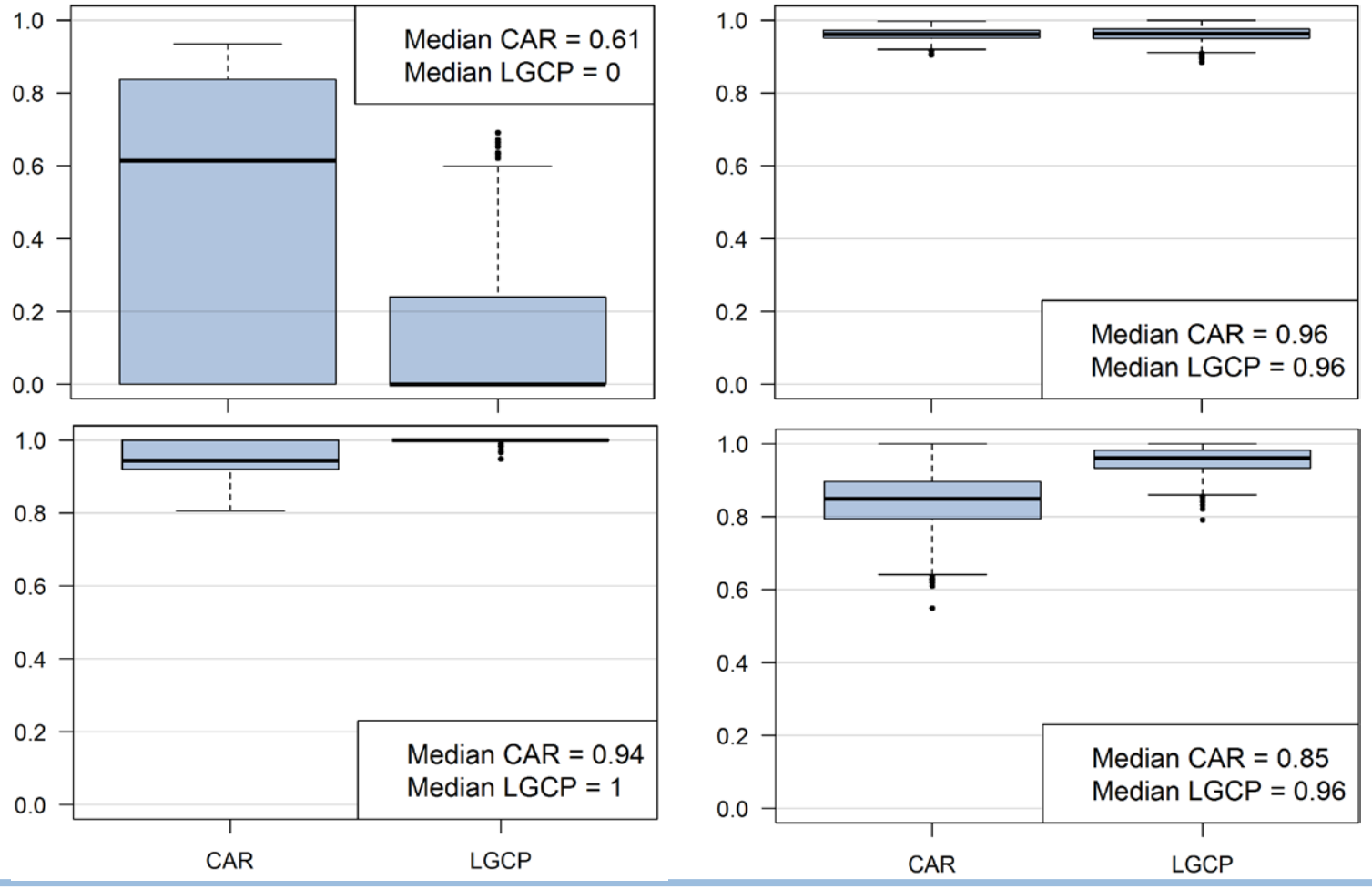
UNIVERSITÄT BERN

2-fold increase 5km radius

10-fold increase 5km radius

Sensitivity

Specificity



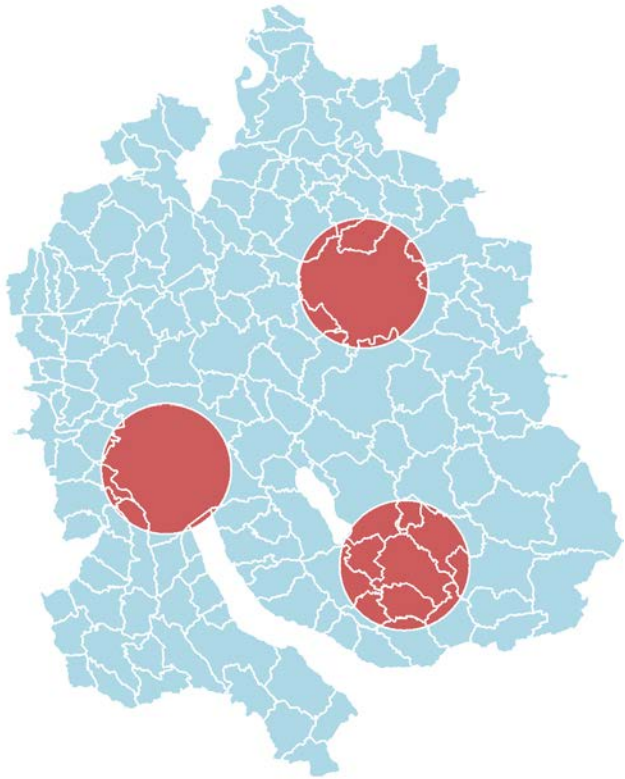
Summary of the results

UNIVERSITÄT BERN

- RMISE:
  - Similar performance on low risk increase
  - LGCP better on higher risk increase
- Sensitivity/Specificity
  - CAR higher sensitivity in lower risk increase
  - LGCP very high specificity in both scenarios

Discussion

- Stable risk might
- ZH and WH had similar size as the municipalities of ZH and WH
- Threshold on the exceedance questionable
- On going project, more scenarios



Take home message

- The evidence so far suggests that for exploratory purposes, if the aetiology of the disease in unknown, results from both methods will be beneficial.

Thank you for your attention!

Funded by:

- Swiss cancer research (4012–08-2016)
- Swiss national science foundation (PZ00P3_147987)

KREBSFORSCHUNG SCHWEIZ
RECHERCHE SUISSE CONTRE LE CANCER
RICERCA SVIZZERA CONTRO IL CANCRO

FNSNF
FONDS NATIONAL SUISSE
SCHWEIZERISCHER NATIONALFONDS
FONDO NAZIONALE SVIZZERO
SWISS NATIONAL SCIENCE FOUNDATION