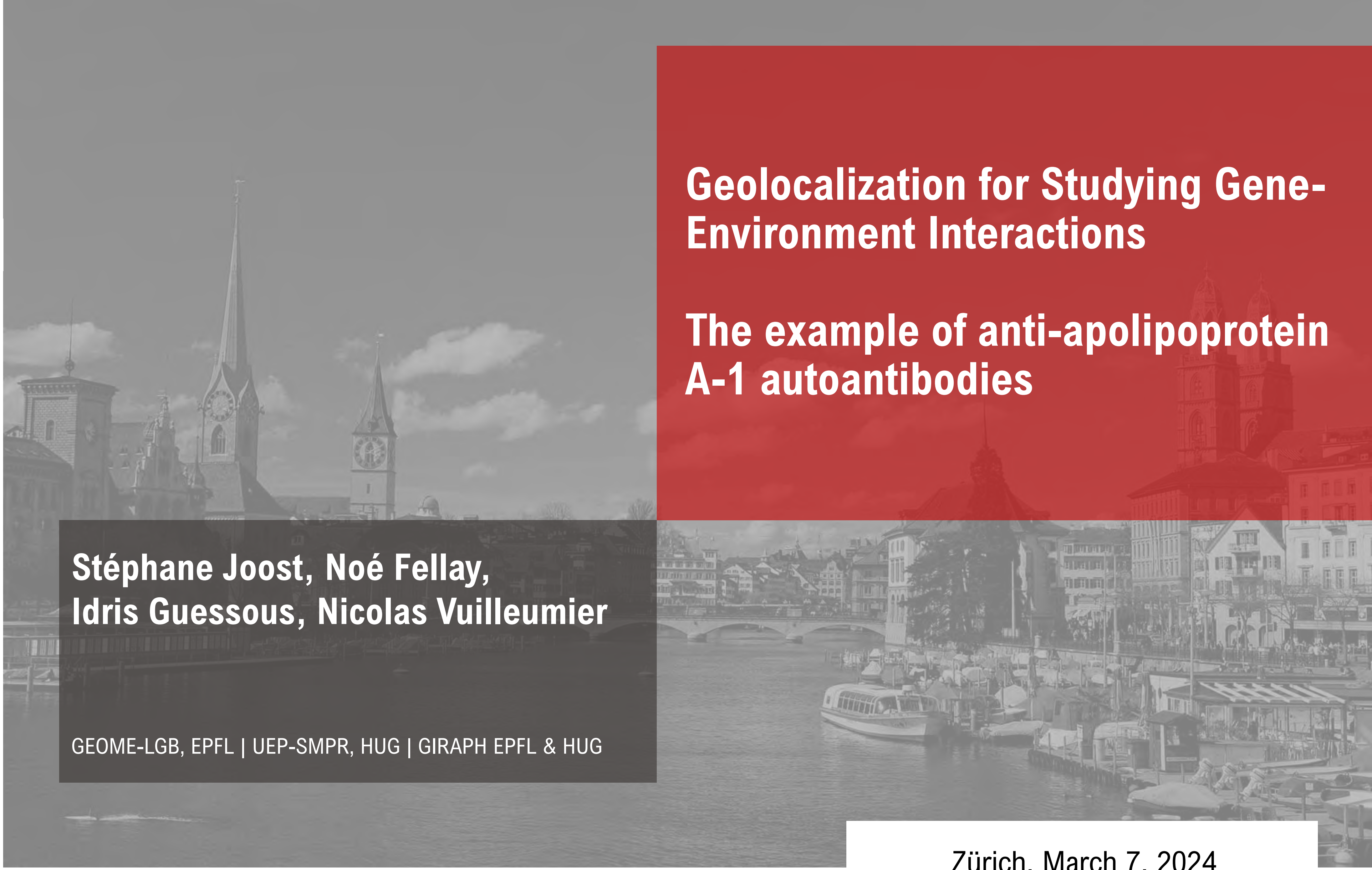
Molecular Diagnostics 2024
Symposium
March 07 + 08, 2024



Geolocalization for Studying Gene-Environment Interactions

The example of anti-apolipoprotein A-1 autoantibodies

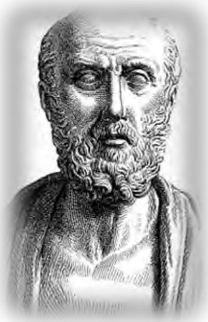
Stéphane Joost, Noé Fellay,
Idris Guessous, Nicolas Vuilleumier

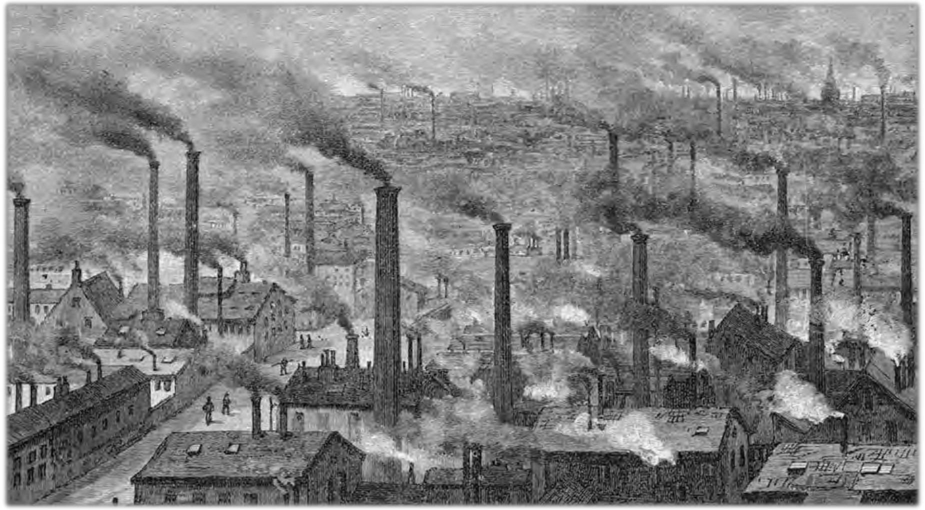
GEOME-LGB, EPFL | UEP-SMPR, HUG | GIRAPH EPFL & HUG

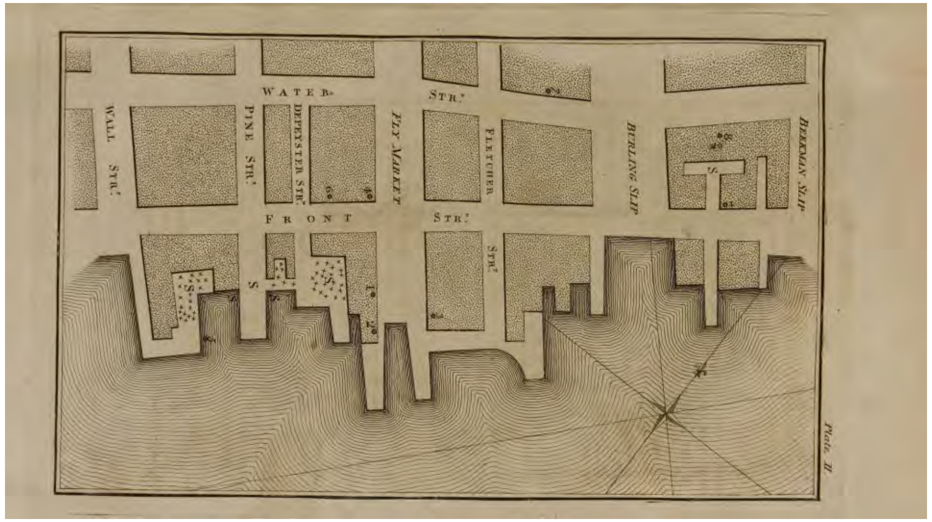
Zürich, March 7, 2024

Health and place

- The study of the relationship between health and place: Hippocrates
- "Airs, waters, and places" influences a whole section of medicine (more than 2,000 years ago)
- Air and water quality must be considered, but also the socio-economic environment and behaviors
- Medical geography: largely based on associations between health and place
- Industrial revolution in the 18th Century: many public health issues
- Emergence of medical cartography
- Need to inform about risk areas where tuberculosis, cholera or yellow fever appear most frequently
- Map by Valentine Seaman, yellow fever cases, 1795



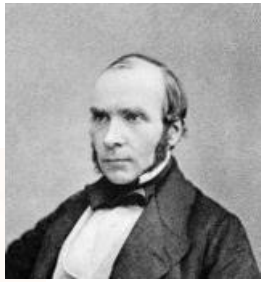


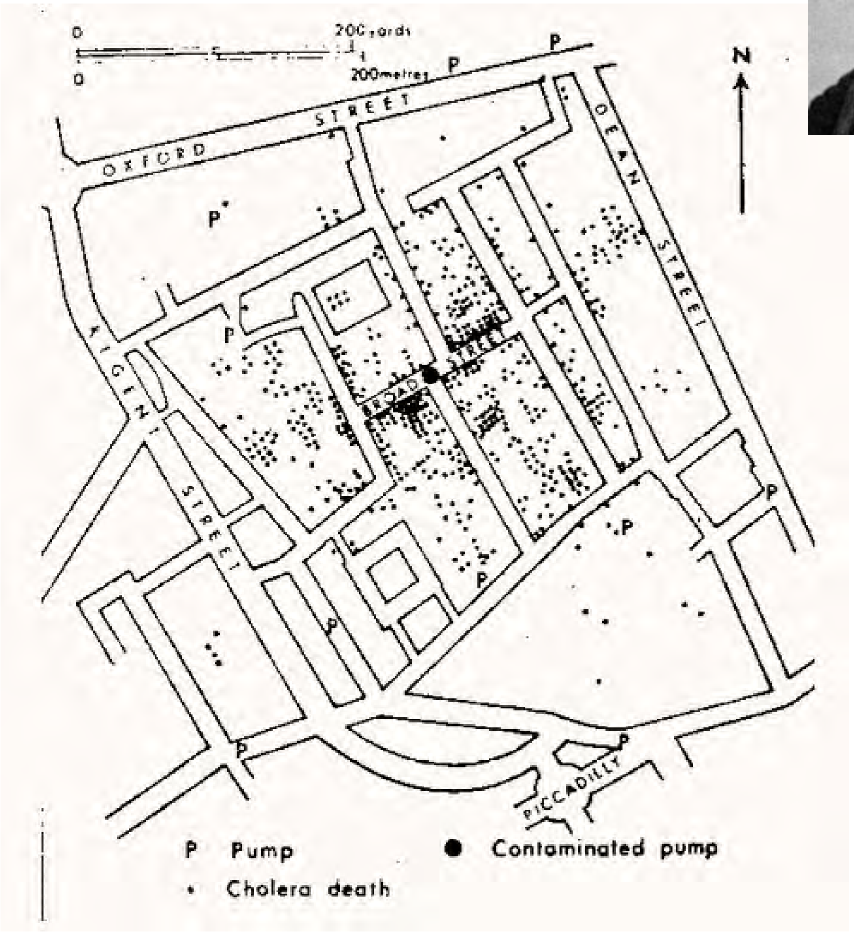


Thomas Shapter

Exeter 1832

John Snow





London, 1854

Jacques May (1896-1975)

- French surgeon
- Notices differences between health status of patients according to the regions from which they came
- Understanding the nature of the relationships between pathogen transmission and geographic factors
- Beginnings of a systematic formulation of **medical geography**
- Cooperation between a medical doctor and a geographer or engineer
- E.g. John Snow (medical doctor) and Edmund Cooper (civil engineer)

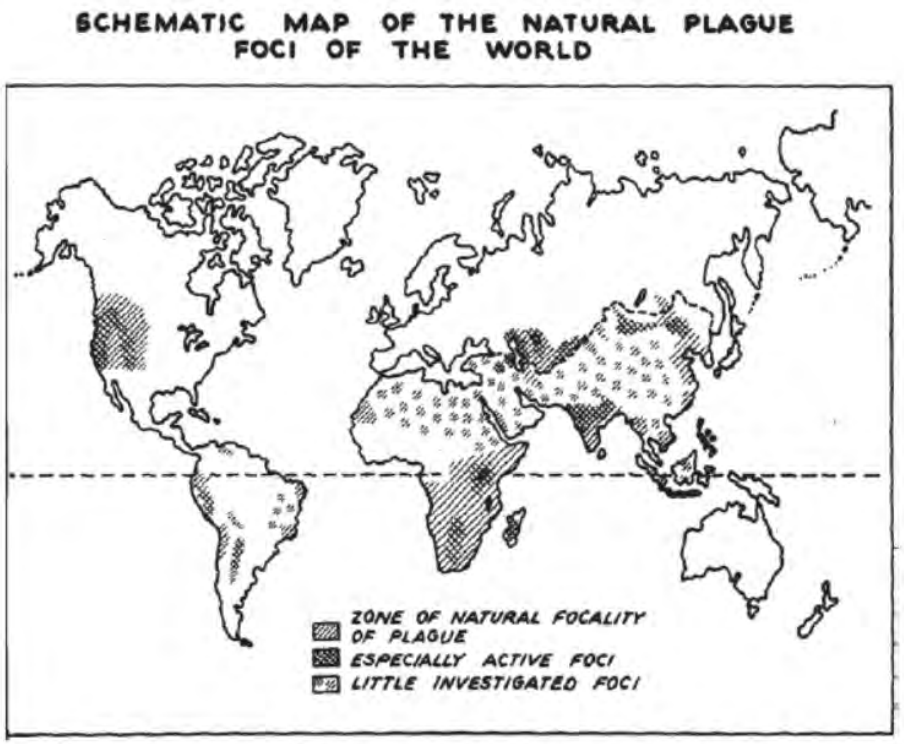


Figure 9.2| May's schematic map of plague foci, distinguishing between environments where plague was known, recurrent, and in some cases, epidemic. Source: May, American Geographical Society (1961).

A medical doctor and a geographer

Prof Idris Guessous
Head Service of Primary Care Medicine
HUG

Bus santé, Specchio UEP

GIRAPH

Geographic Information Research and Analysis in Population Health

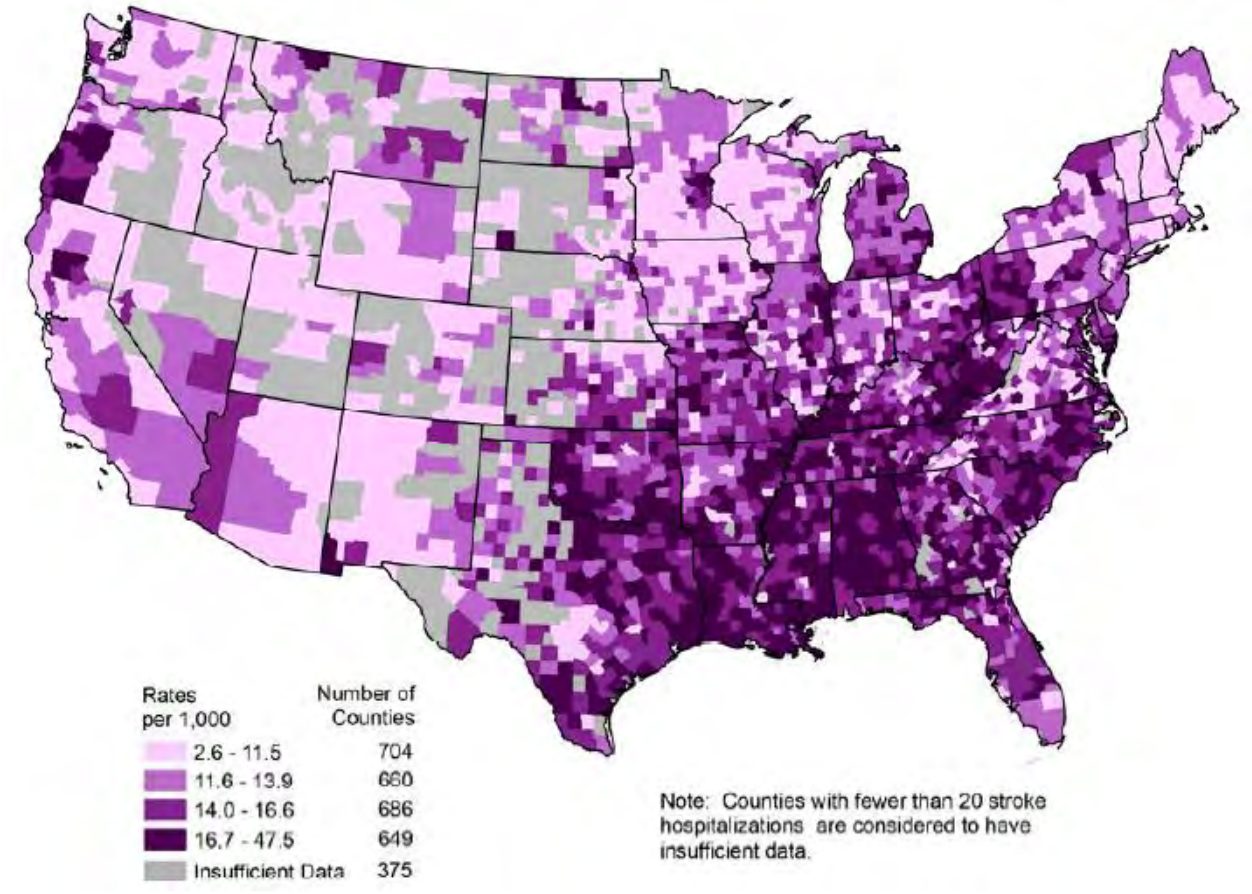
Bus santé, Specchio UEP

Dr Stéphane Joost
Institute of Environmental Engineering, EPFL

Toward more precision in medical geography

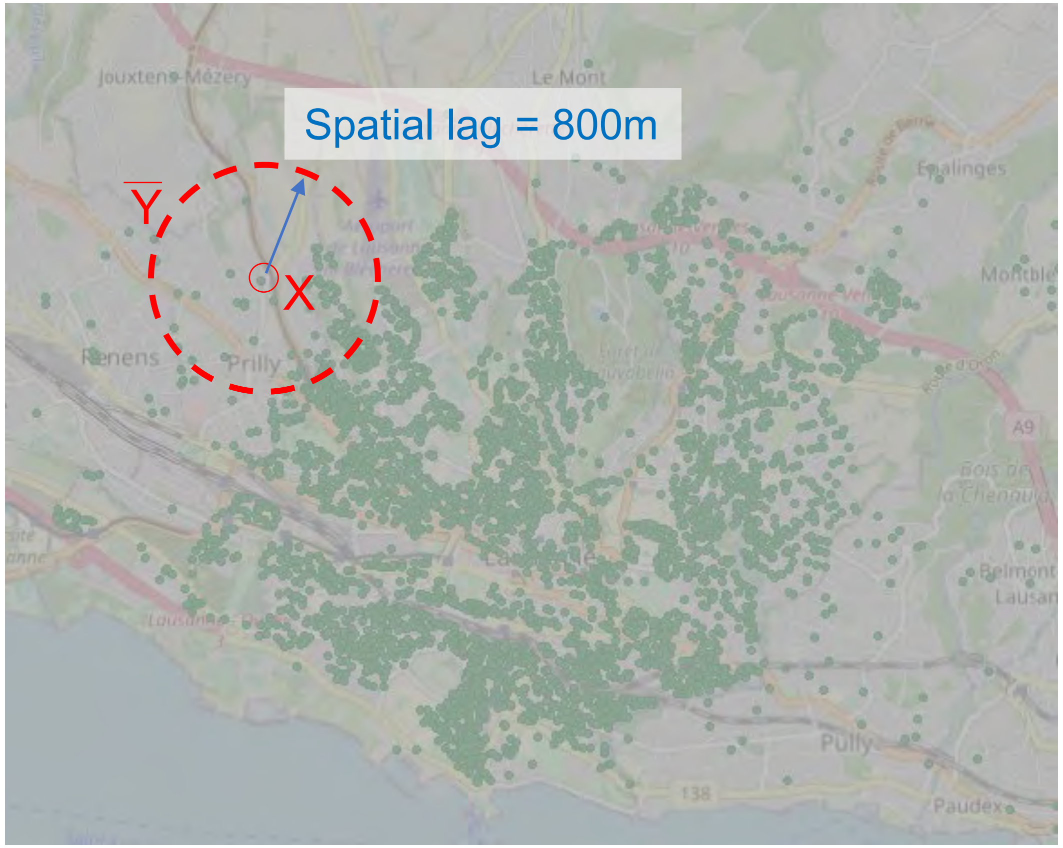
- In spatial epidemiology, geolocated data usually aggregated within administrative units
- Here: stroke hospitalization rates of medicare beneficiaries aged 65 or older in 2005-2006 at the county level
- Poses methodological problems like ecological fallacy or Modifiable Areal Unit Problem (MAUP)
- For what use and effectiveness in public health, in the field of prevention in particular?

- GIRAPH's contribution towards precision public health…



Working with georeferenced individual data

- Participants in medical cohorts precisely located in space
- Method: georeferencing
- X,Y (geographic coordinates) of places of residence (Rue Neuve 14, 1009 Pully)
- First law of geography “Everything is related to everything else, but near things are more related than distant things” (Tobler, 1970)
- Measuring spatial dependence

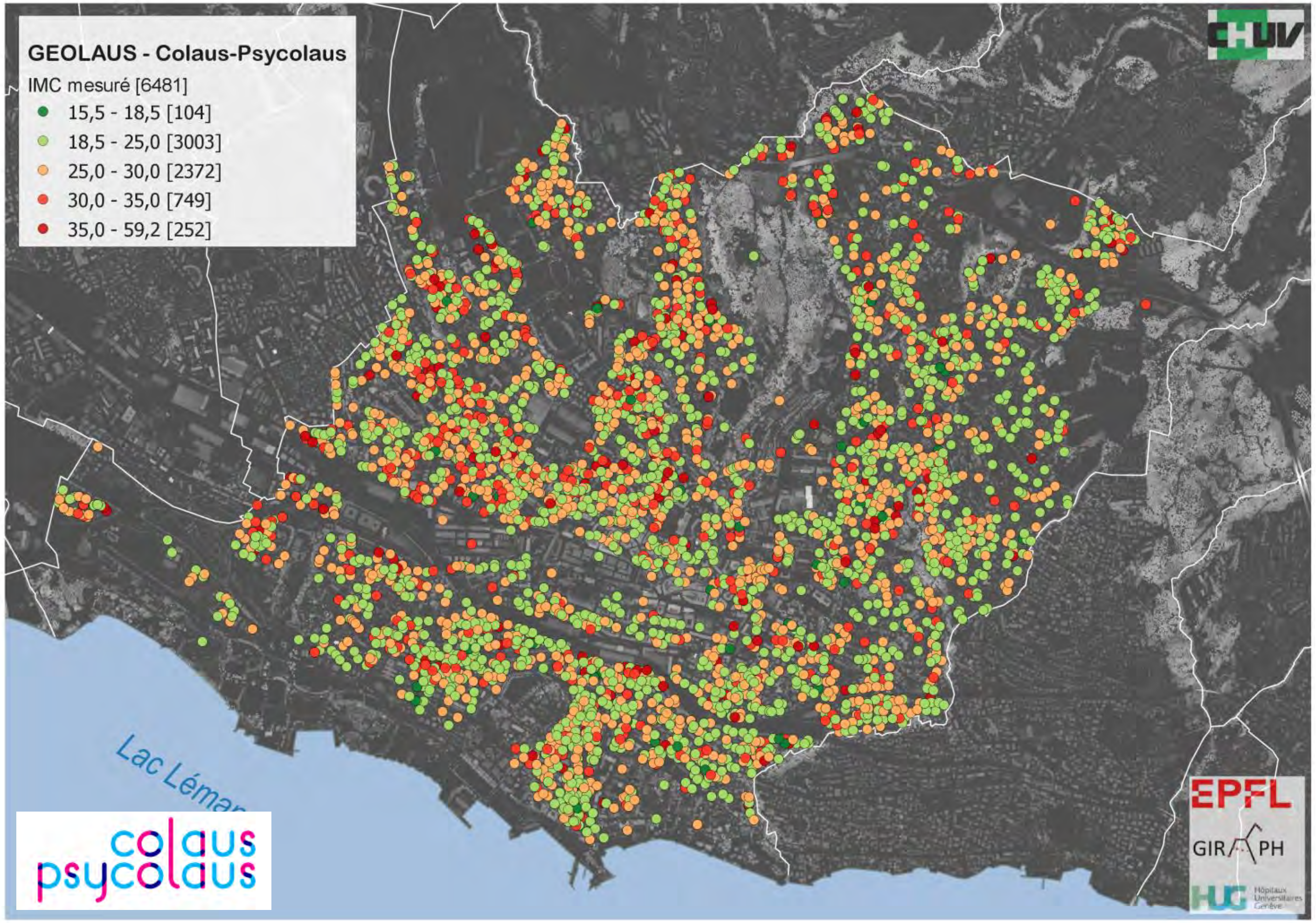


Does X look like $\bar{Y}$?

Without Tobler's law

- No detectable signal through standard thematic mapping

Classification en fonction de l'indice de masse corporelle	
Insuffisance pondérale	< 18.5
Éventail normal	18.5 - 24.9
Surpoids	≥ 25.0
Préobésité	25.0 - 29.9
Obésité	≥ 30.0
Obésité, classe I	30.0 - 34.9
Obésité, classe II	35.0 - 39.9
Obésité, classe III	≥ 40.0



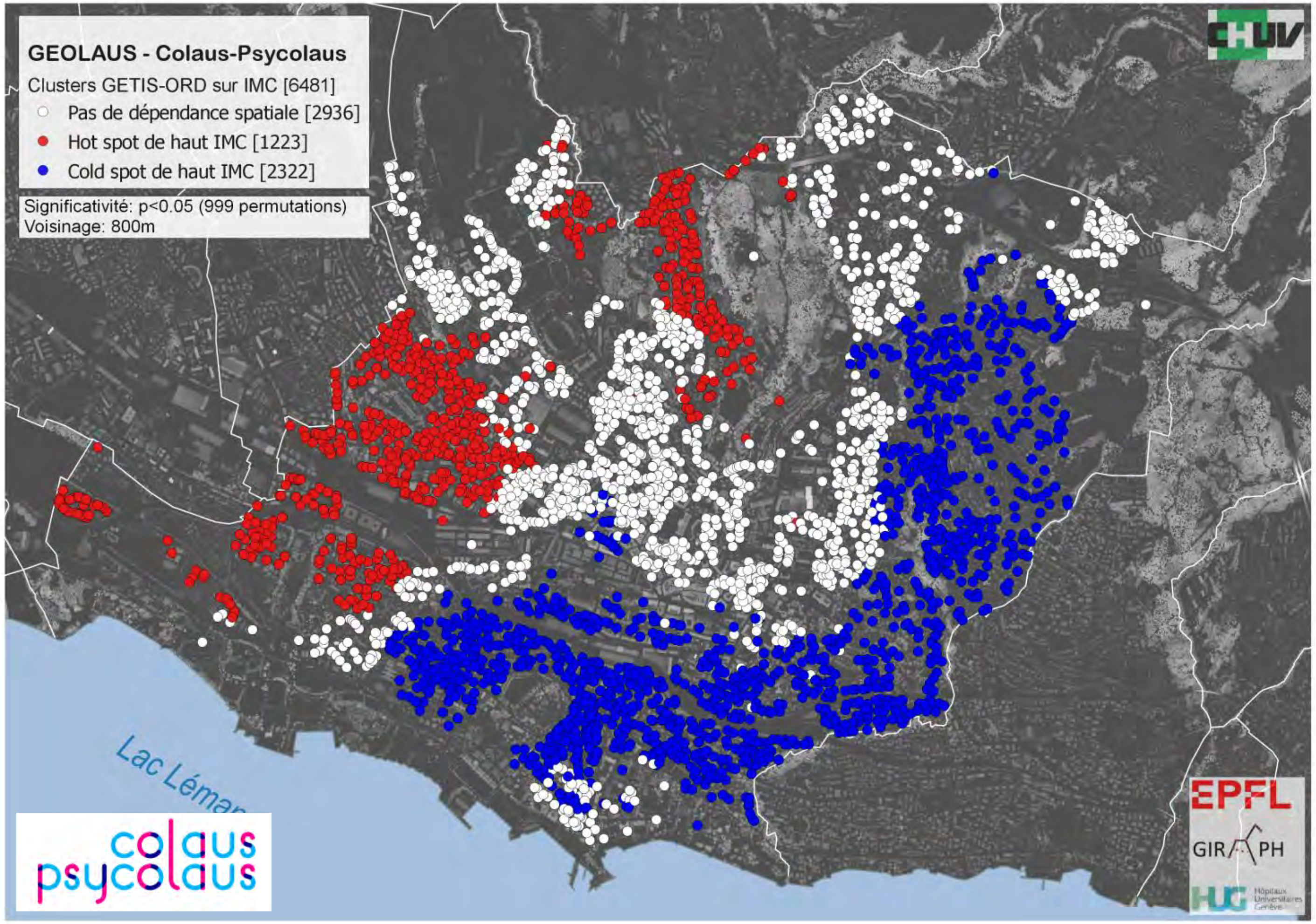
Body Mass Index, Colaus-PsyColaus cohort

Joost et al. 2016

Spatial statistics based on spatial dependence concept

- Specific statistical tools
- Key to make the invisible visible!
- Open doors for spatial data exploration

- Potential to develop targeted prevention actions where needed!

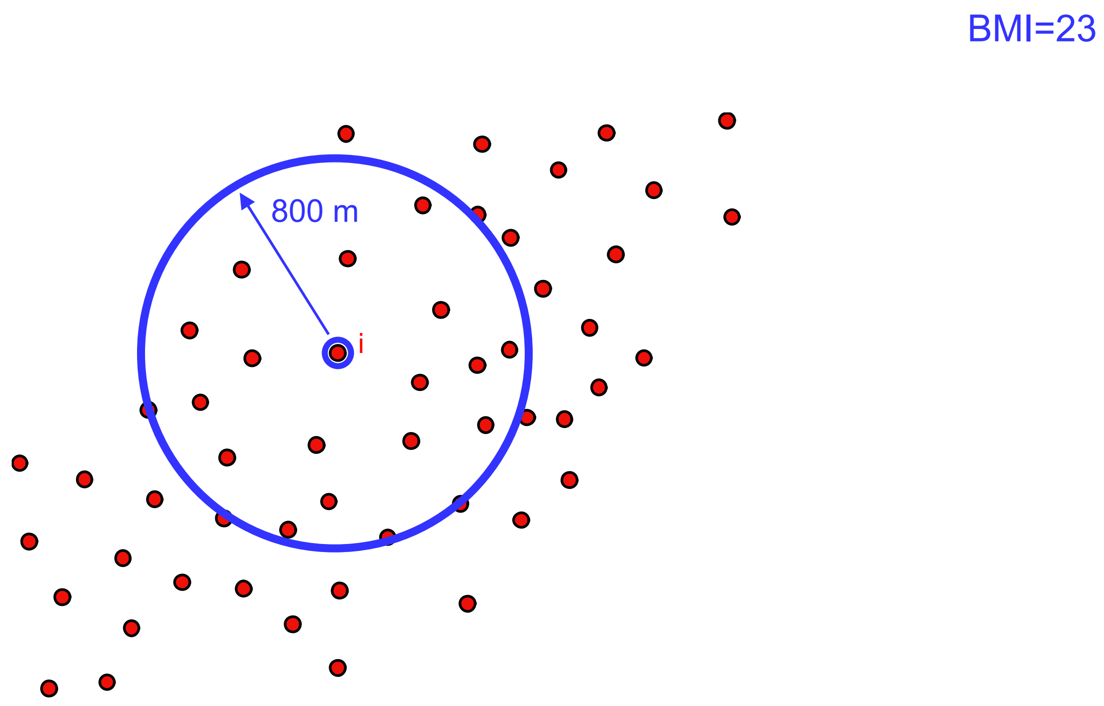


Body Mass Index, Colaus-PsyColaus cohort

Joost et al. 2016

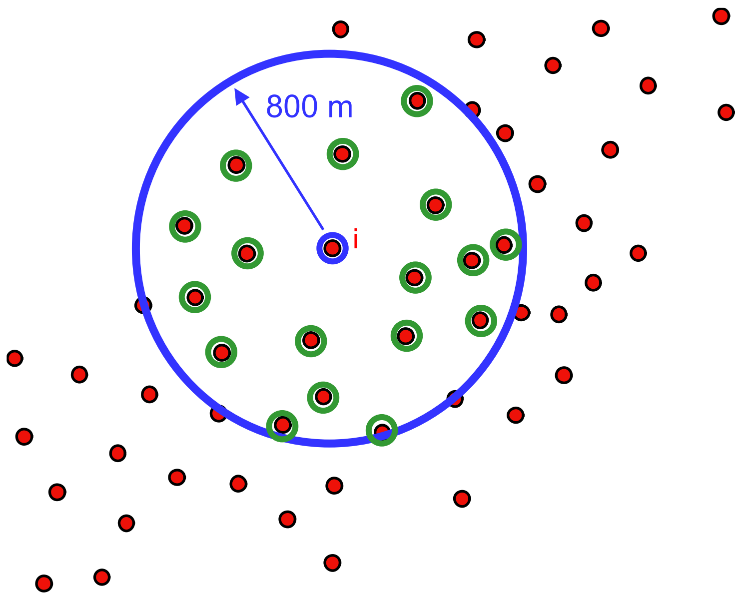
How to measure spatial dependence?

- Use neighborhood relationships
- "How much" an individual resembles his neighbors in a given neighborhood (e.g. 800m)
- Example with body mass index (BMI) and Getis-Ord statistics (Getis & Ord, 1992)
- BMI = $\frac{Weight}{(Height)^2}$



How to measure spatial dependence?

- Use neighborhood relationships
- "How much" an individual resembles his neighbors in a given neighborhood (e.g. 800m)
- Example with body mass index (BMI) and Getis-Ord statistics (Getis & Ord, 1992)
- BMI = $\frac{Weight}{(Height)^2}$

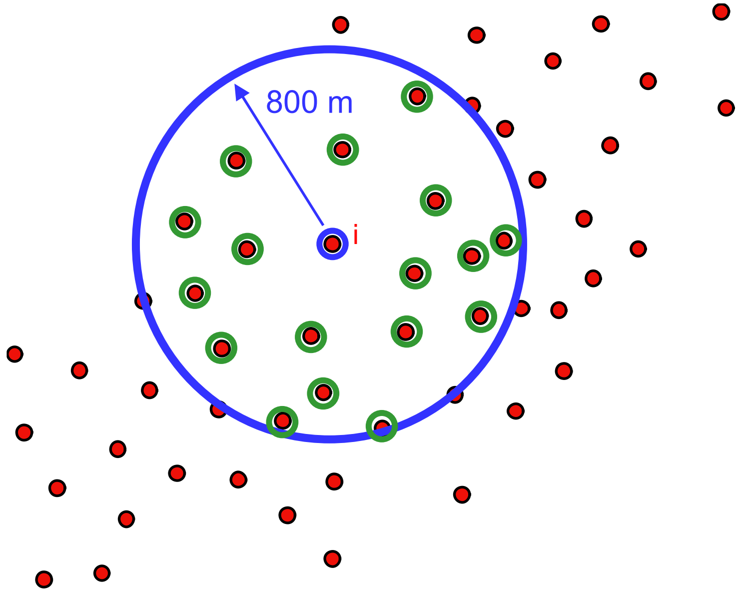


BMI=23
BMI=21
BMI =27
BMI =23
BMI =29
BMI =23
BMI =33
BMI =35
BMI =18
BMI =23
BMI =22
BMI =26
BMI =27
BMI =27
BMI =25
BMI =21

$\overline{BMI}_i = 25.18$

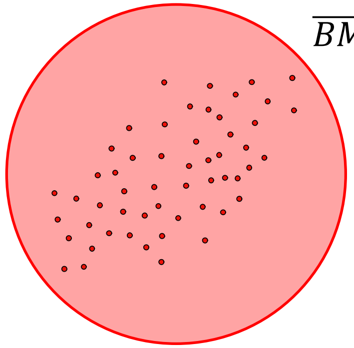
How to measure spatial dependence?

- Use neighborhood relationships
- "How much" an individual resembles his neighbors in a given neighborhood (e.g. 800m)
- Example with body mass index (BMI) and Getis-Ord statistics
- BMI = $\frac{Weight}{(Height)^2}$



BMI=23
BMI=21
BMI =27
BMI =23
BMI =29
BMI =23
BMI =33
BMI =35
BMI =18
BMI =23
BMI =22
BMI =26
BMI =27
BMI =27
BMI =25
BMI =21

$\overline{\text{BMI}}_i$ = 25.18



$\overline{BMI_g} = 22$

As $25.18 > 22$ $\rightarrow$

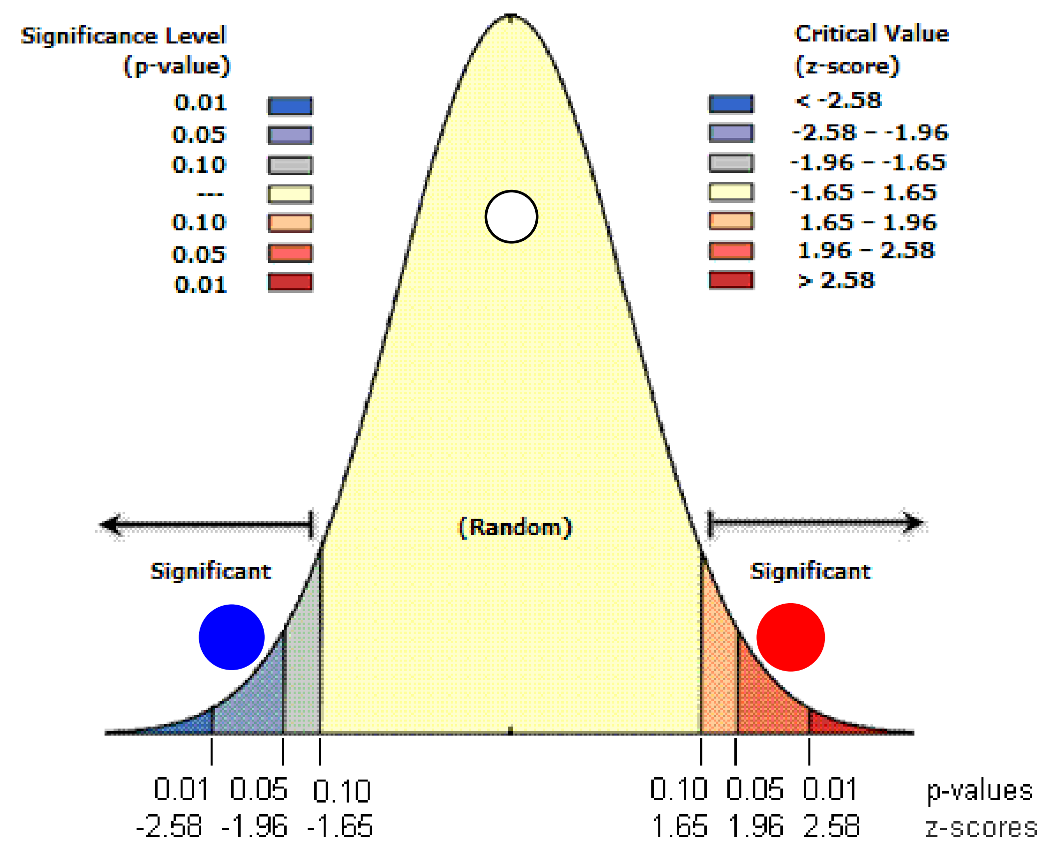
If $IMC_i < 22$ $\rightarrow$

Idem for all points in the dataset

Spatial dependence or random distribution?

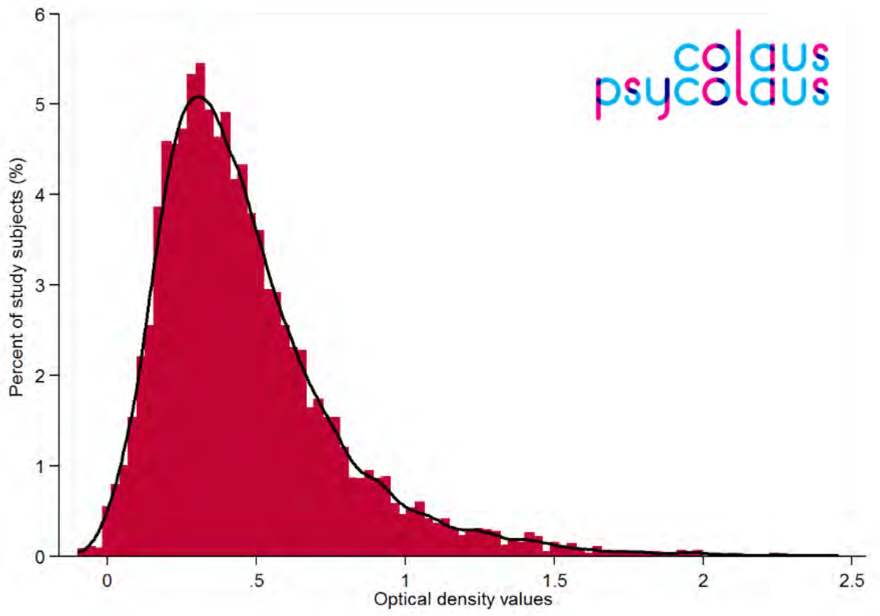
- Null hypothesis: the observed local value is not different from the general mean
- T-test* to compare the means of two groups, random permutations

- **Significantly lower than the general average**, the value depends on its geographical location
- **Random value**, does not depend on its geographical location
- **Significantly higher than the general average**, the value depends on its geographical location



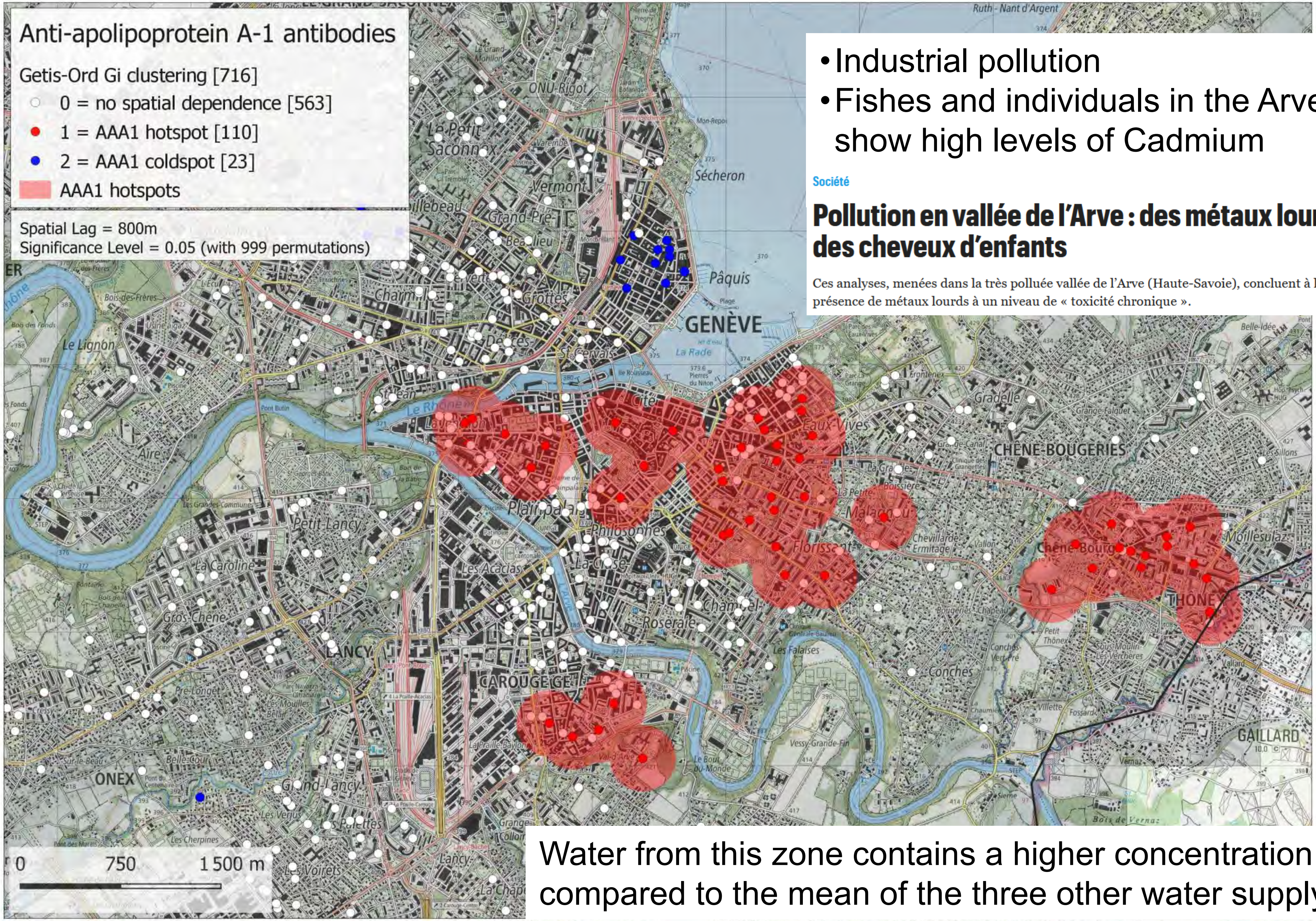
Anti-apolipoprotein A-1 autoantibodies (AAA1)

- Harmful antibodies directed against apolipoprotein A-1 (protein carrying "good cholesterol")
- AAA1 constitute independent cardiovascular risk factors (Antiochos et al. 2016)
- Deleterious biological actions, stimulating inflammation and promoting the formation of atherosclerosis (Vuilleumier et al. 2020)
- AAA1 Leenaards Foundation project (UNIGE, HUG, CHUV and UNIL, 2014): "AAA1 as a prognostic biomarker of cardiovascular risk and new therapeutic target" (Leenaards prize 2014)
- Using data from the **CoLaus-PsyColaus** cohort (N=6700)
- Study the mechanisms through which AAA1 contribute to the pathogenesis of cardiovascular disease
- Investigate a possible genetic predisposition to AAA1
- Determine their frequency in the population
- Better understand how these antibodies act



AAA1 in the geographic space

- Preliminary study in **Geneva** showed an interesting AAA1 spatial pattern (D. de Ridder, N. Vuilleumier, I. Guessous, 2020)
- HUG, UEP Bus Santé data (N=716)
- Geographic clusters along the ARVE river within the same drinking water supply basin



- Industrial pollution
- Fishes and individuals in the Arve valley show high levels of Cadmium

Société

Le Parisien

Pollution en vallée de l'Arve : des métaux lourds détectés sur des cheveux d'enfants

Le 14 janvier 2021

Ces analyses, menées dans la très polluée vallée de l'Arve (Haute-Savoie), concluent à la présence de métaux lourds à un niveau de « toxicité chronique ».

Water from this zone contains a higher concentration of nitrates, compared to the mean of the three other water supply zones

Spatial analysis of AAA1 in Lausanne

- How are AAA1 geographically distributed in the Colaus-PsyColaus population?
- Are they spatially dependent?
- Association with heavy metals?
- Like Cadmium, known as (underestimated) cardiovascular risk factor
- Cadmium promotes inflammation and shown to be bound to Apo-lipoprotein A1 (Li et al. 2019)

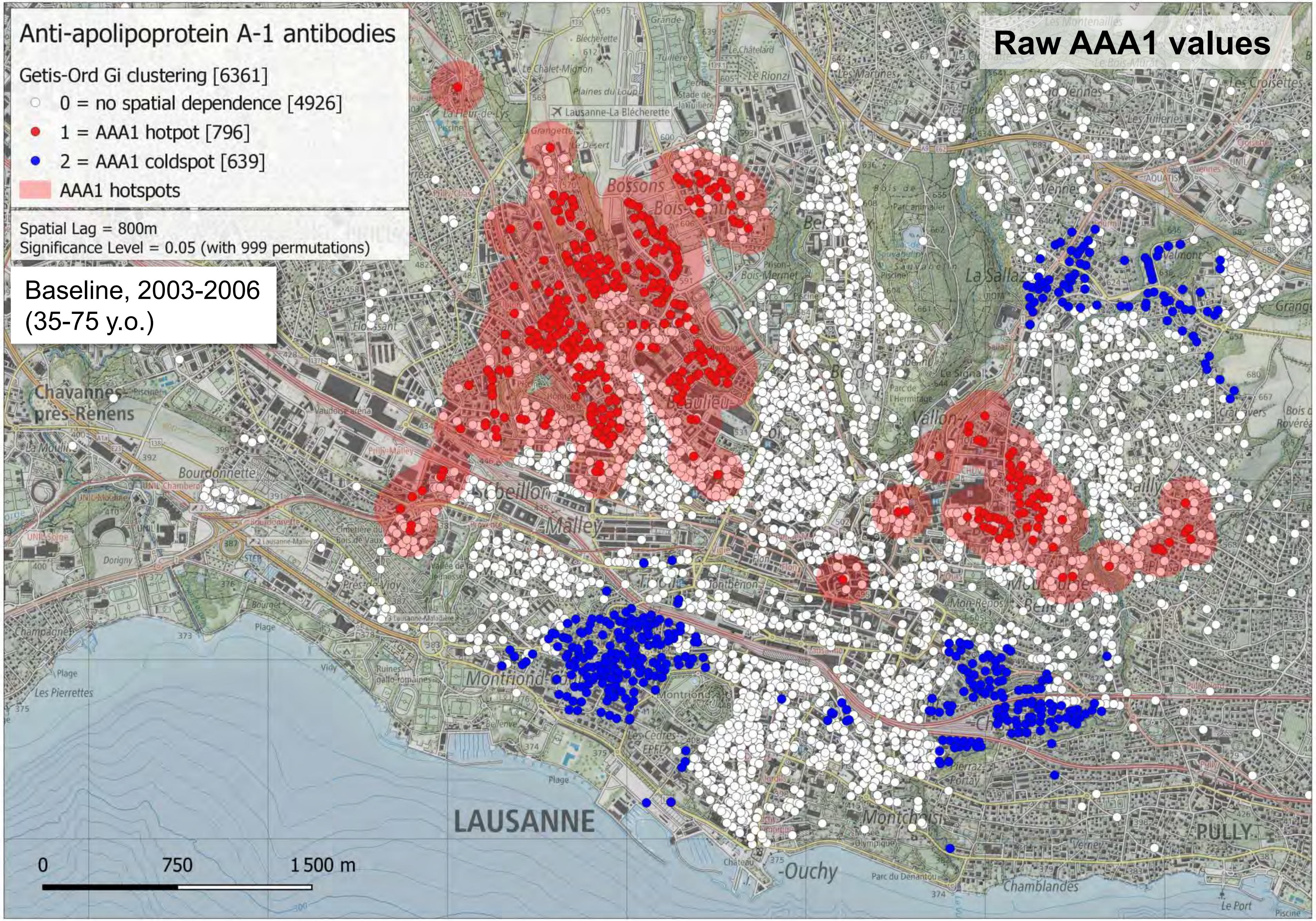
Methodology

- Exploratory Spatial Data Analysis (ESDA)
- Exploratory investigations first and confirmatory analyses afterwards
- Tukey (1980) We Need Both Exploratory and Confirmatory, The American Statistician, Vol. 34, No. 1

- Exploration with spatial statistics described earlier
- Confirmation with Geographically Weighted Regression (GWR) (Brunsdon & Fotheringham 2007)

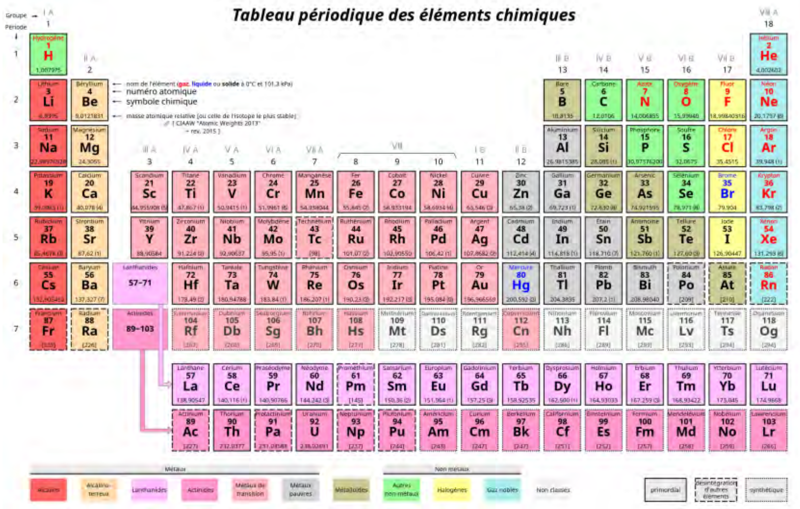
AAA1 spatial footprint

- Are AAA1 spatially distributed at random through Lausanne ?


Anti-apolipoprotein A-1 antibodies
Getis-Ord Gi clustering [6361]
0 = no spatial dependence [4926]
1 = AAA1 hotpot [796]
2 = AAA1 coldspot [639]
AAA1 hotspots
Spatial Lag = 800m
Significance Level = 0.05 (with 999 permutations)
Baseline, 2003-2006
(35-75 y.o.)
Raw AAA1 values
LAUSANNE
0
750
1 500 m

Spatial coïncidence with heavy metals

- A total of 24 heavy metals were investigated in urine's participants
- Interesting patterns only in cadmium, lead and arsenic

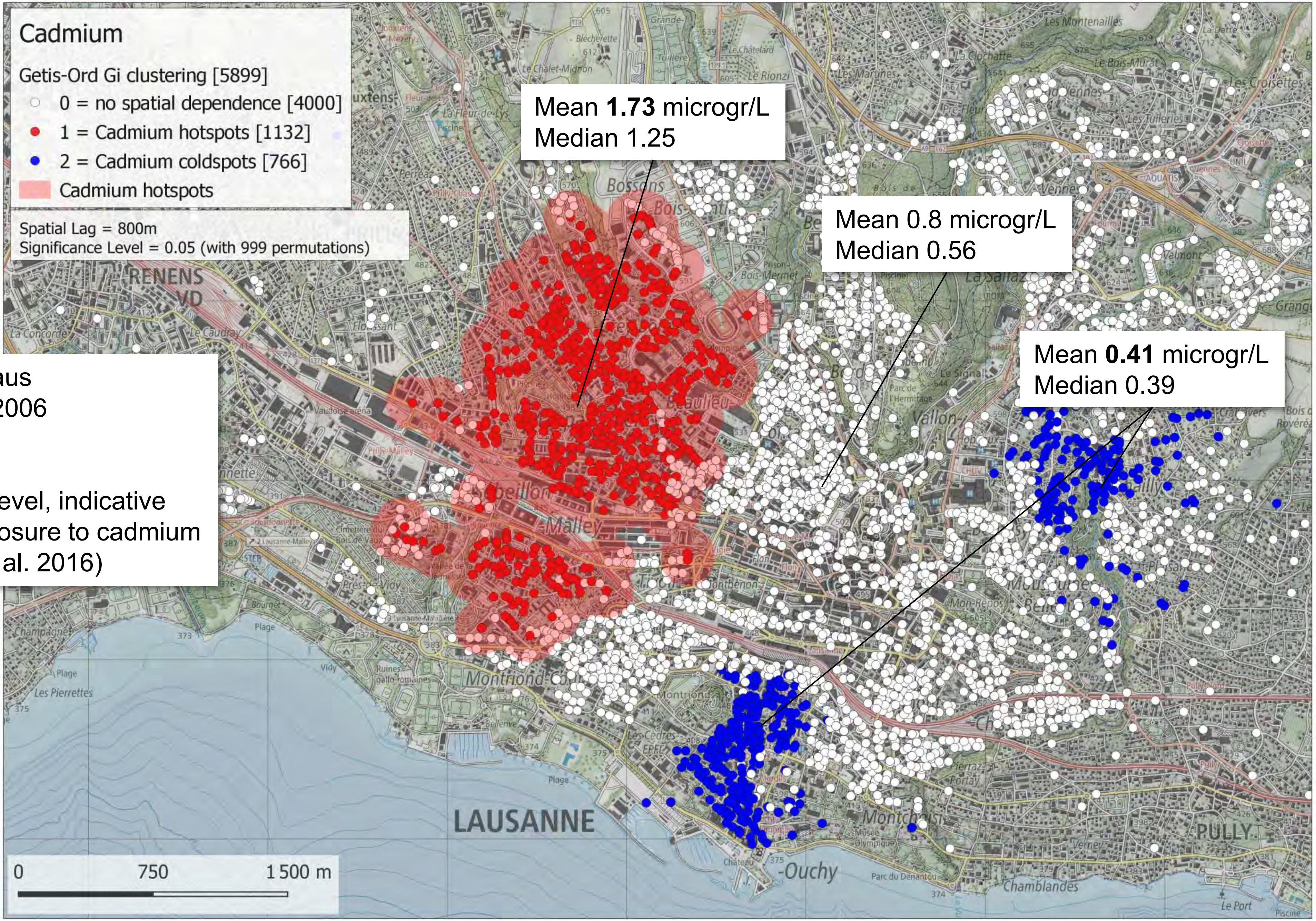


Variable Properties - metauxlourds

variable name	type	parent group
Li_std	real	
Be_std	real	
Al_std	real	
V_std	real	
Cr_std	real	
Mn_std	real	
Co_std	real	
Ni_std	real	
Cu_std	real	
Zn_std	real	
As_std	real	
Se78_std	real	
Se82_std	real	
Mo_std	real	
Ag_std	real	
Cd_std	real	
Sn_std	real	
Sb_std	real	
Hg201_std	real	
Tl_std	real	
Pb208_std	real	
Bi_std	real	
Fe_std	real	
Iode_std	real	

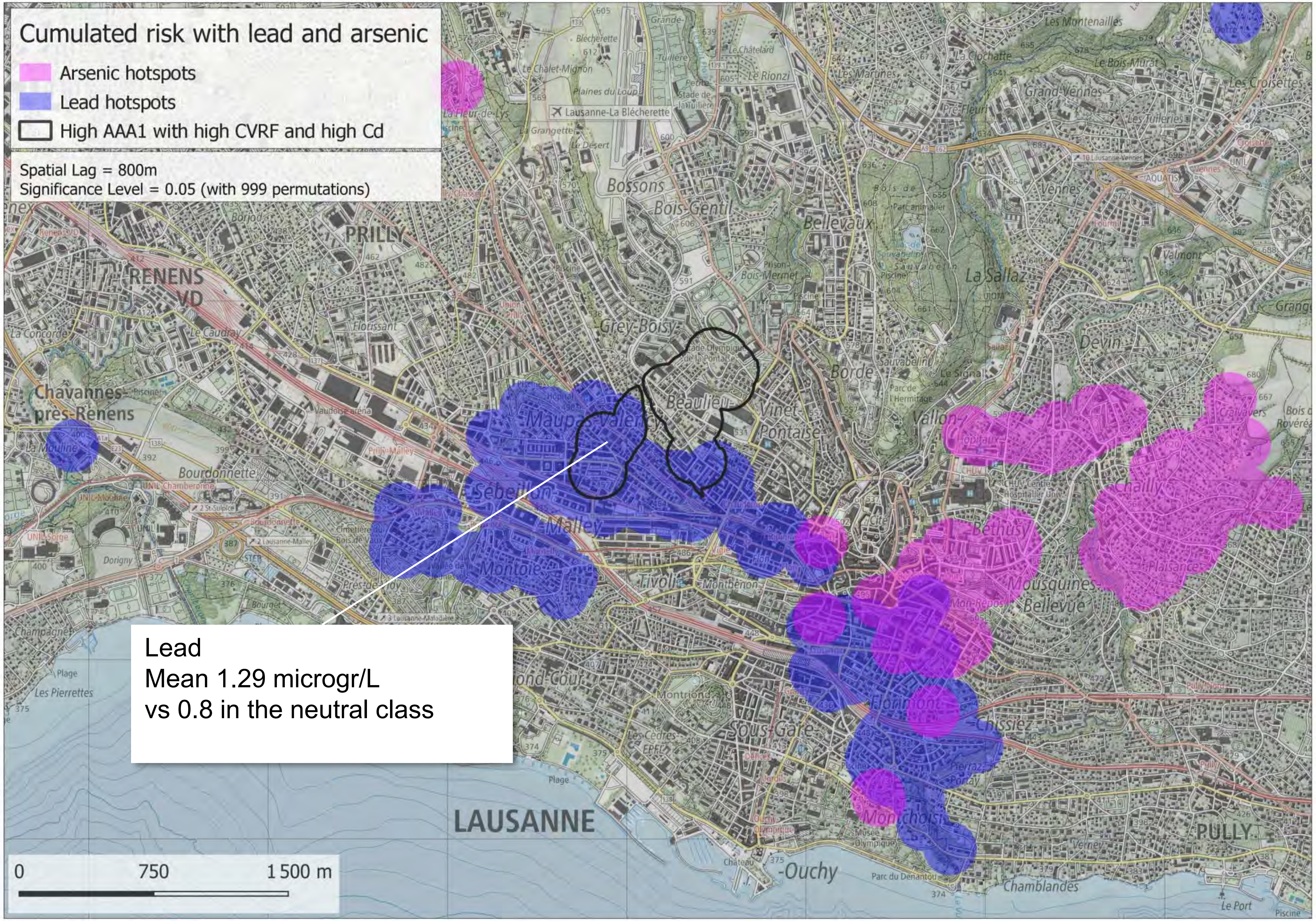
Cadmium spatial footprint

- Is cadmium spatially distributed at random through Lausanne ?



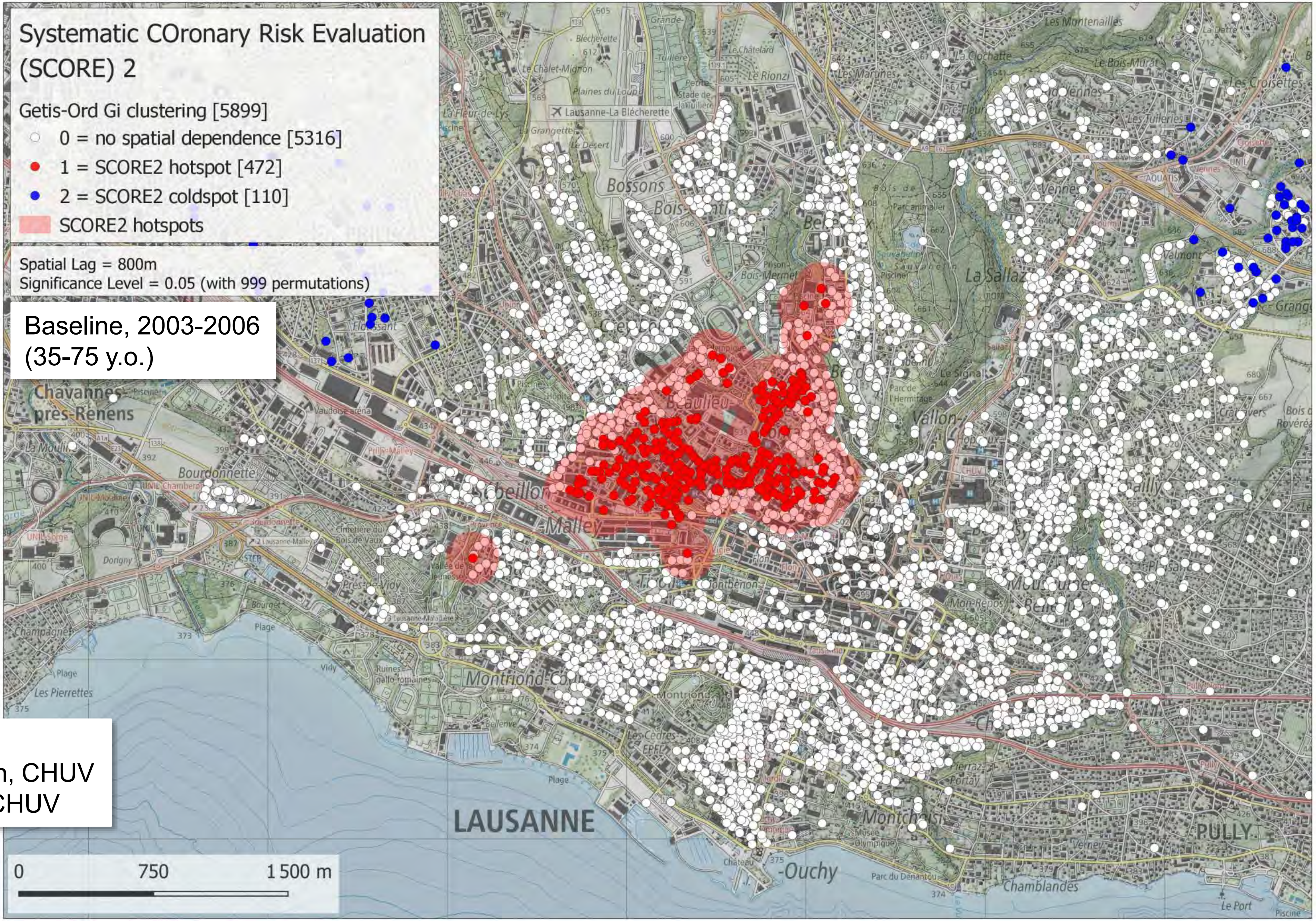
CoLaus-PsyColaus
Baseline, 2003-2006
(35-75 y.o.)

Urine cadmium level, indicative of long-term exposure to cadmium (Vacchi-Suzzi et al. 2016)


Cumulated risk with lead and arsenic
Arsenic hotspots
Lead hotspots
High AAA1 with high CVRF and high Cd
Spatial Lag = 800m
Significance Level = 0.05 (with 999 permutations)
Lead
Mean 1.29 microgr/L
vs 0.8 in the neutral class
0
750
1 500 m
PRILLY
RENENS VD
Chavannes-près-Renens
Bossons
Bois-Gentil
Bellevaux
La Sallaz
Grey-Bois
Beaulieu
Vinet-Pontaise
Borde
Devin
Vennes
Grand-Vennes
Sébeillon
Malley
Montoie
Bourdonnette
Tivoli
Montbenon
Bethusy
Mousquines-Bellevue
Florimont
Chissiez
Sous-Gare
Montchoisi
LAUSANNE
Ouchy
PULLY
Chamblandes
Les Pierrettes

Cardiovascular risk

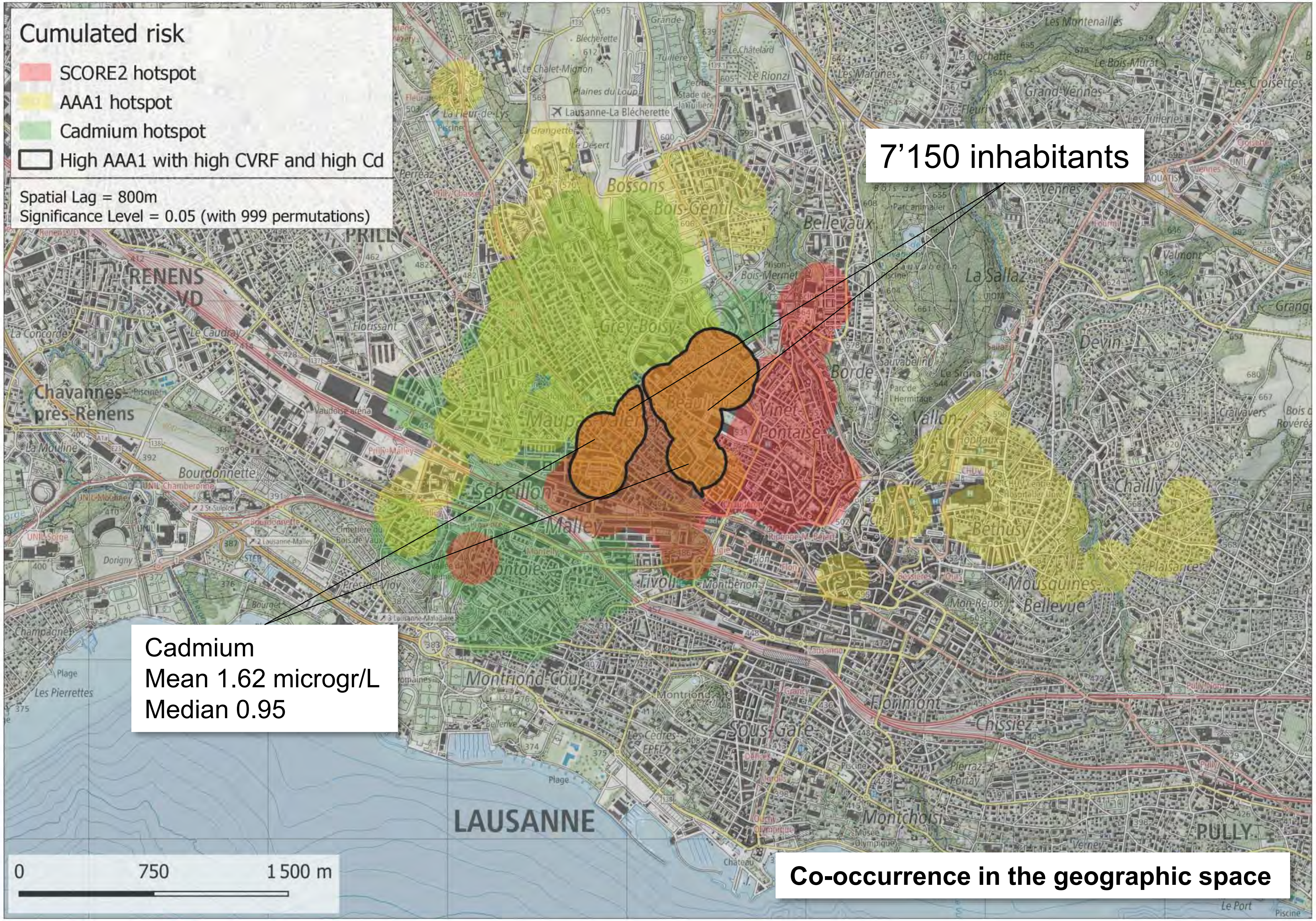
- Is the Systematic Coronary Risk Evaluation (SCORE 2) index spatially dependent?



SCORE2 data
Guillaume Jordan, CHUV
Julien Vaucher, CHUV

Cumulated risk (based on spatial coincidence only)

- Are there high-risk areas cumulating membership to AAA1,Cadmium and CVRF(SCORE2) hotspots?


Cumulated risk
SCORE2 hotspot
AAA1 hotspot
Cadmium hotspot
High AAA1 with high CVRF and high Cd
Spatial Lag = 800m
Significance Level = 0.05 (with 999 permutations)
7’150 inhabitants
Cadmium
Mean 1.62 microgr/L
Median 0.95
0 750 1 500 m
Co-occurrence in the geographic space

Beyond spatial exploration

- Geospatial analyses suggest that cadmium could elicit an AAA1 response
- Pilot study involving two groups of male C57BL6J mice subjected to high-fat diet
- Exposed to either normal drinking water or water with 10 ppm Cd for 10 weeks
- Cadmium induced a significant AAA1 response, while no effect on lipid profile, liver test, nor myocardial necrosis markers

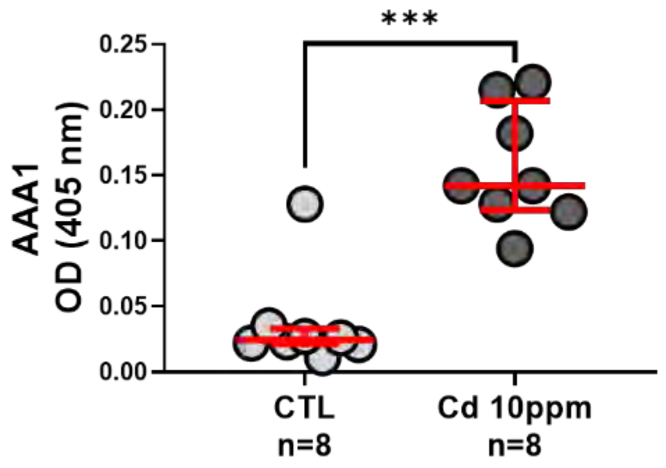
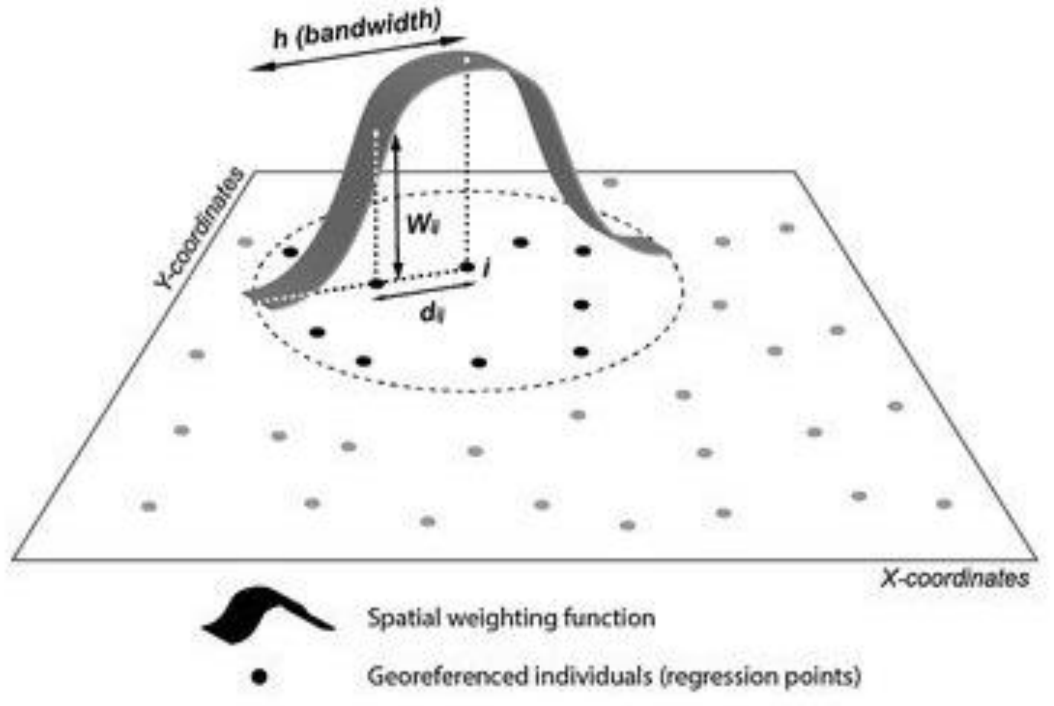


Figure 5. Cadmium induces the production of AAA1 in mice. Results are expressed as individual values and median with interquartile range. *P*-value calculated with Mann-Whitney test.****P*=0.0008.

Spatial relationship between cadmium and AAA1

- Confirmatory analysis
- Geographically Weighted Regression (GWR)
- Spatially explicit
- Independent variables are given more weight locally than further apart
- 1 equation per regression point

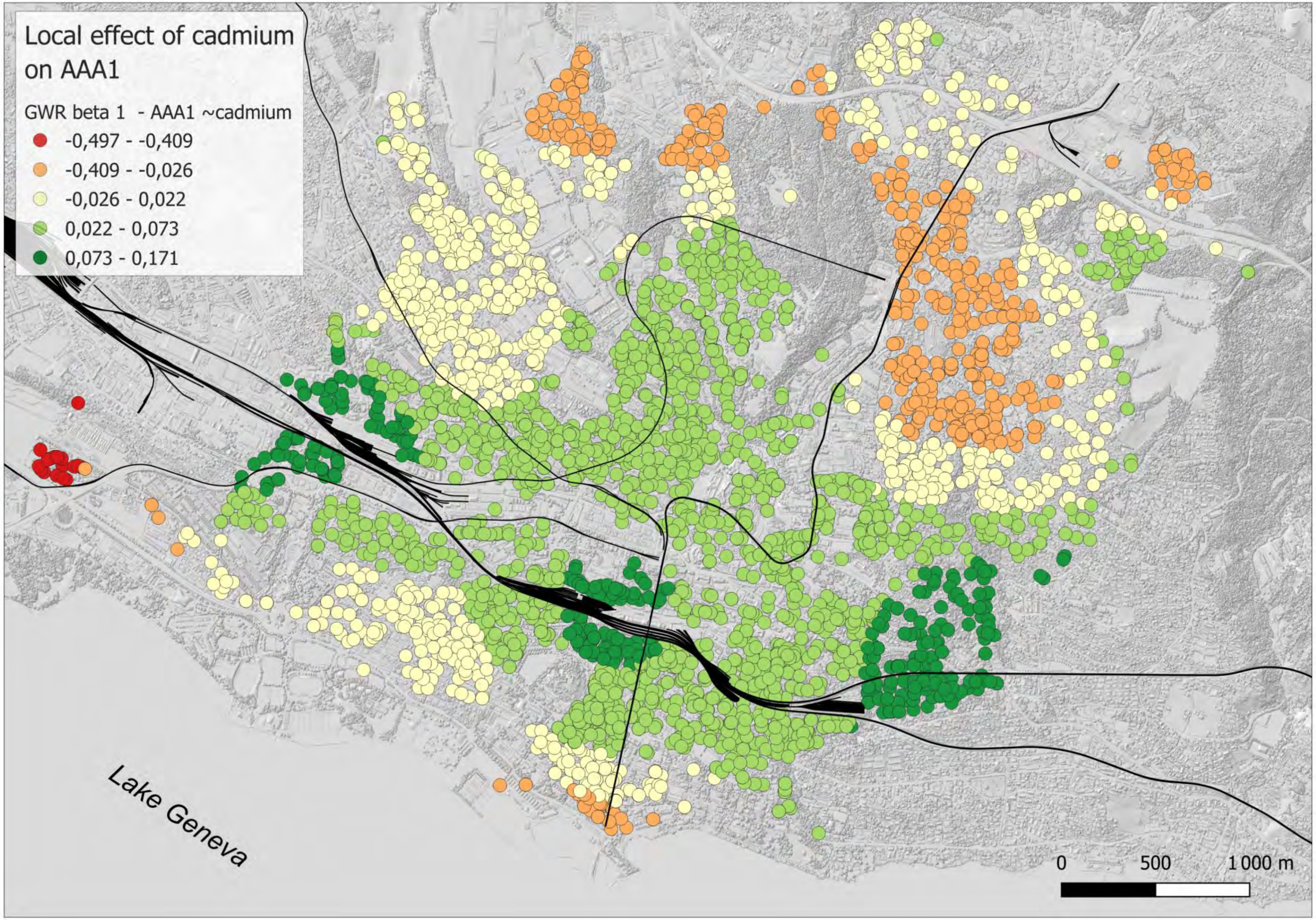


Feuillet et al. 2015

GWR results

Local beta1 coefficients (local effect)

AAA1 adjusted for age, sex, education, smoking



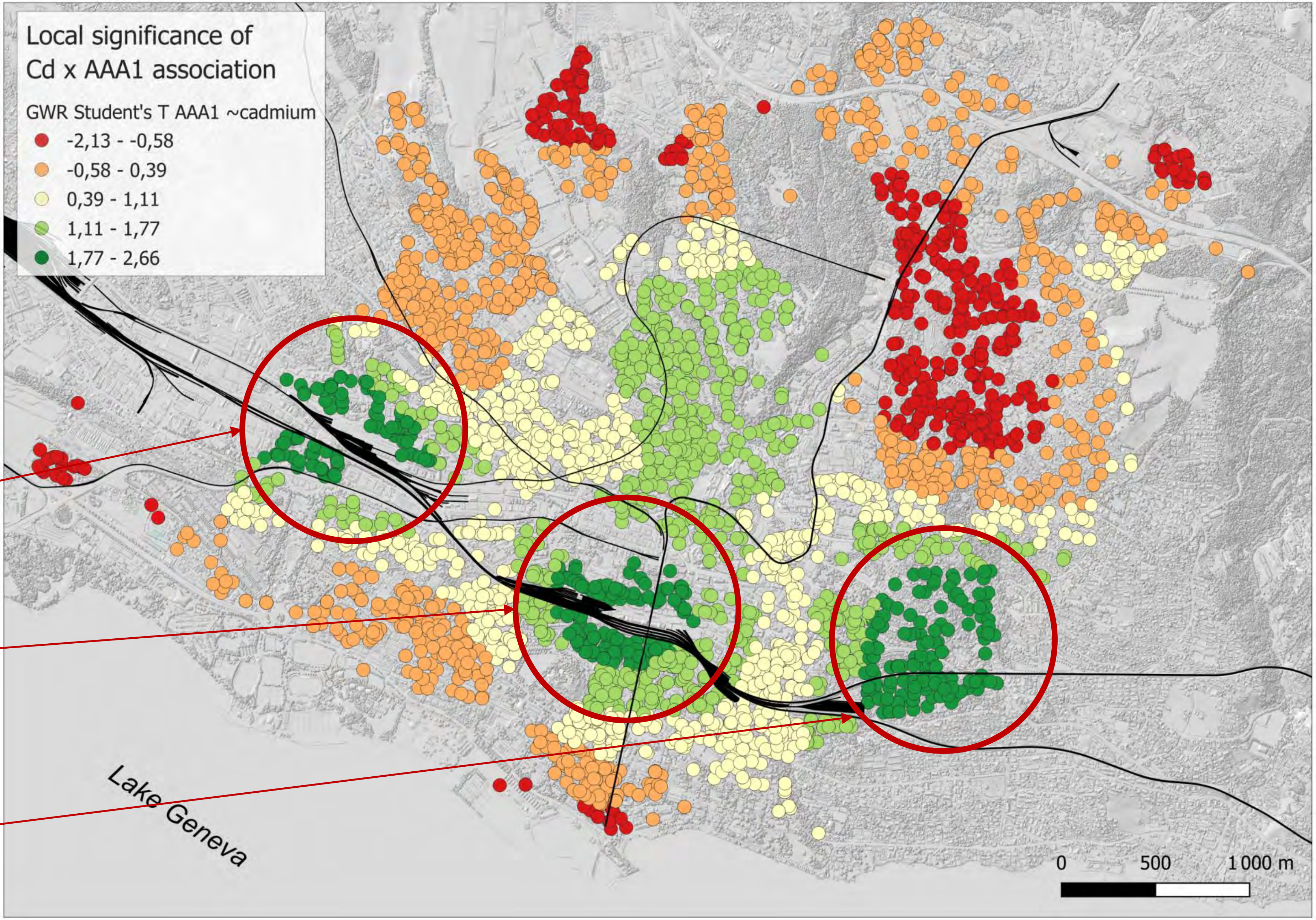
GWR results

Local Student's T (local significance)

Sébeillon marshalling yard

Lausanne main SBB station

Chandieu marshalling yard and train repair warehouses



Railway activities

- Potential sources of cadmium
- Friction with catenaries
- Brake systems
- Local industrial activities
- Altogether

- Rail transport increases cadmium levels in soil along railroad lines (Wilkomirski 2001; Ma et al. 2009)

The genetics of AAA1

Two genomic regions of interest

- 9 SNPs belonging to the Fc receptor-like 3 (FCRL3) gene associated with anti-apoA-1 IgG levels, with the lead SNP **RS64**27397 (Antiochos et al. 2017a)
- Significant interaction between anti-apoA-1 IgG and **RS25**69190 allele status with regards to Coronary Artery Disease (CAD) risk (Antiochos et al. 2017b)

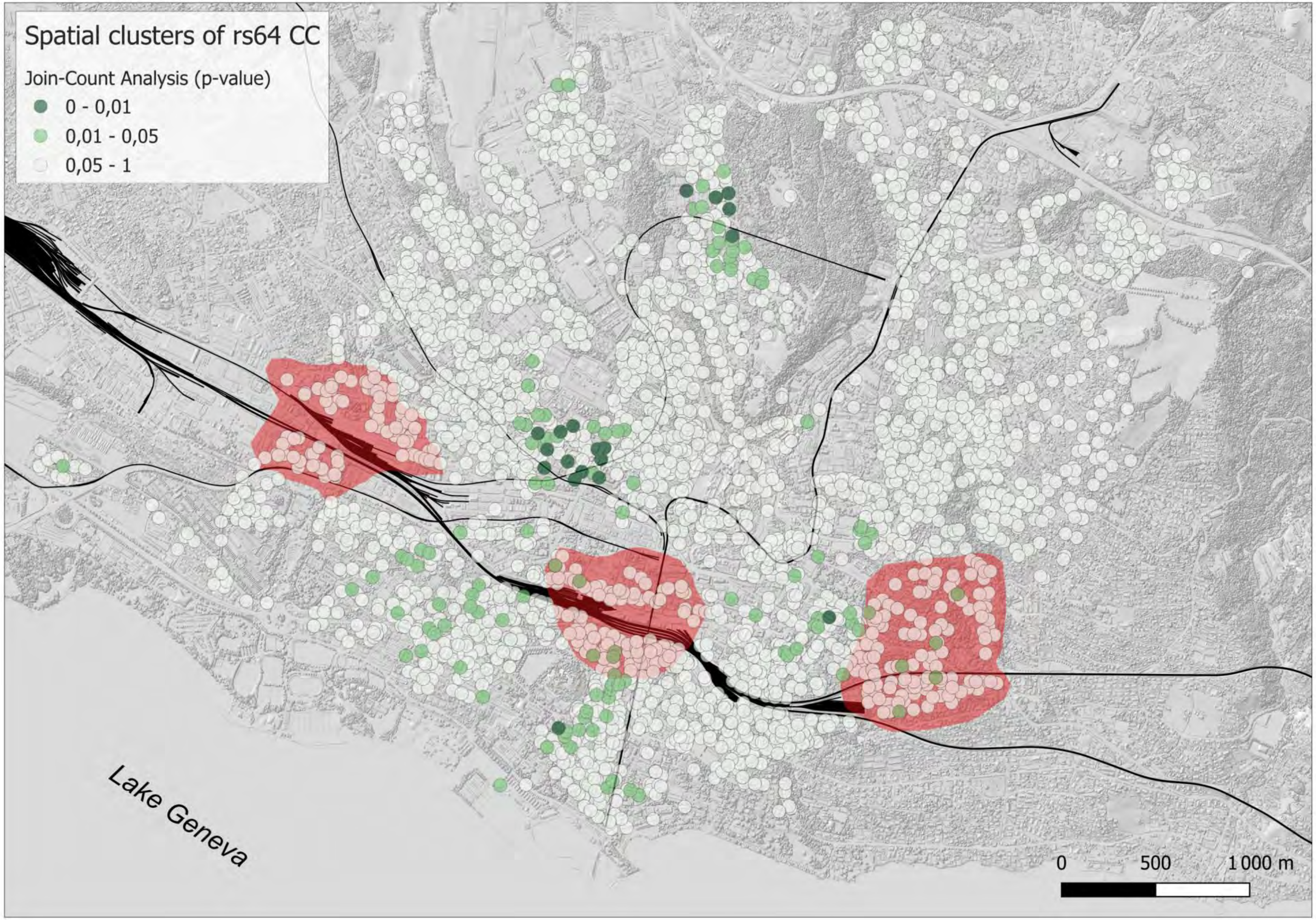
- Exploratory only…

AAA1 Genetics

Spatial clusters

Join-Count Analysis
Spatial lag = 800m

rs64 CC

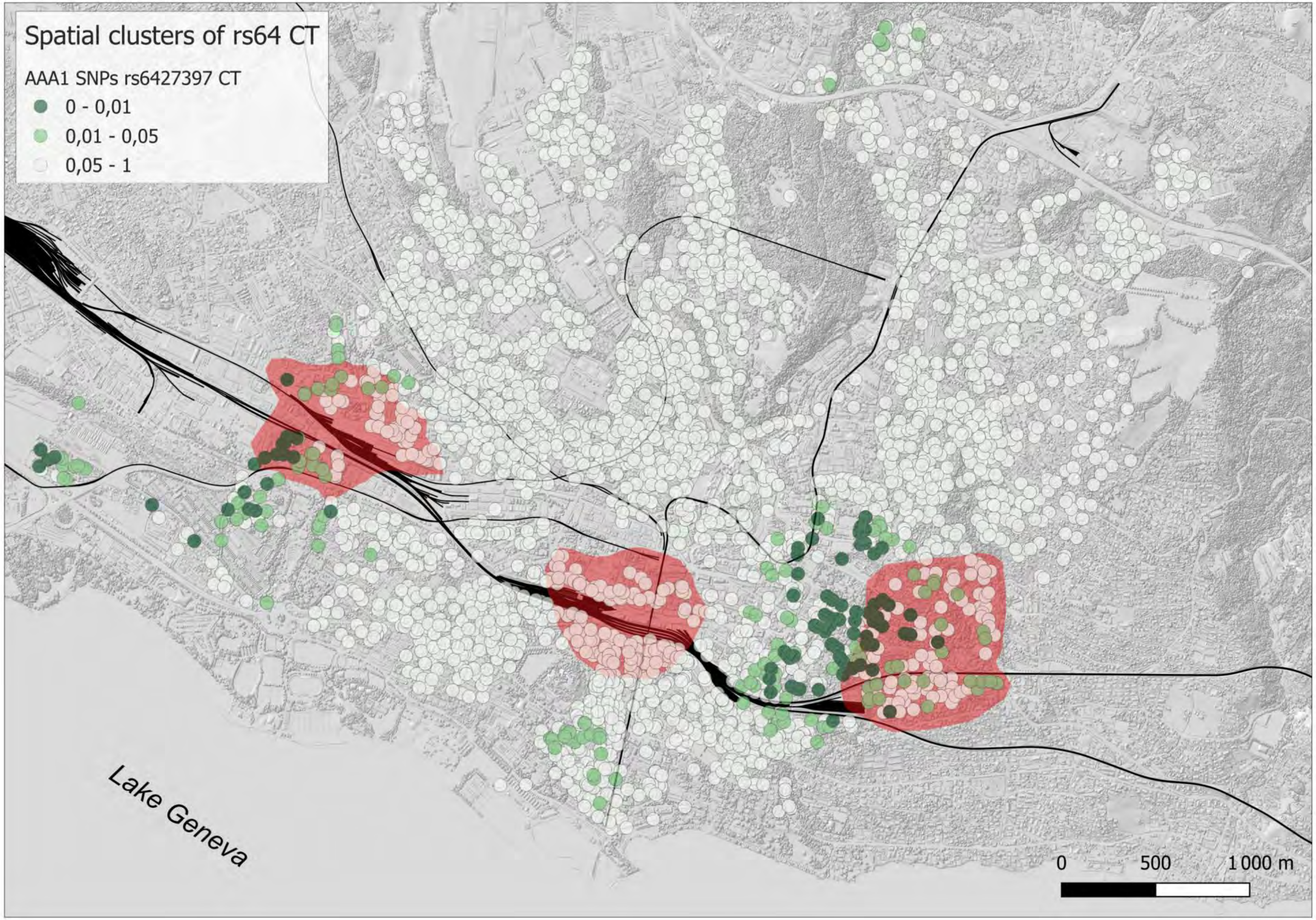


AAA1 Genetics

Spatial clusters

Join-Count Analysis
Spatial lag = 800m

rs64 CT

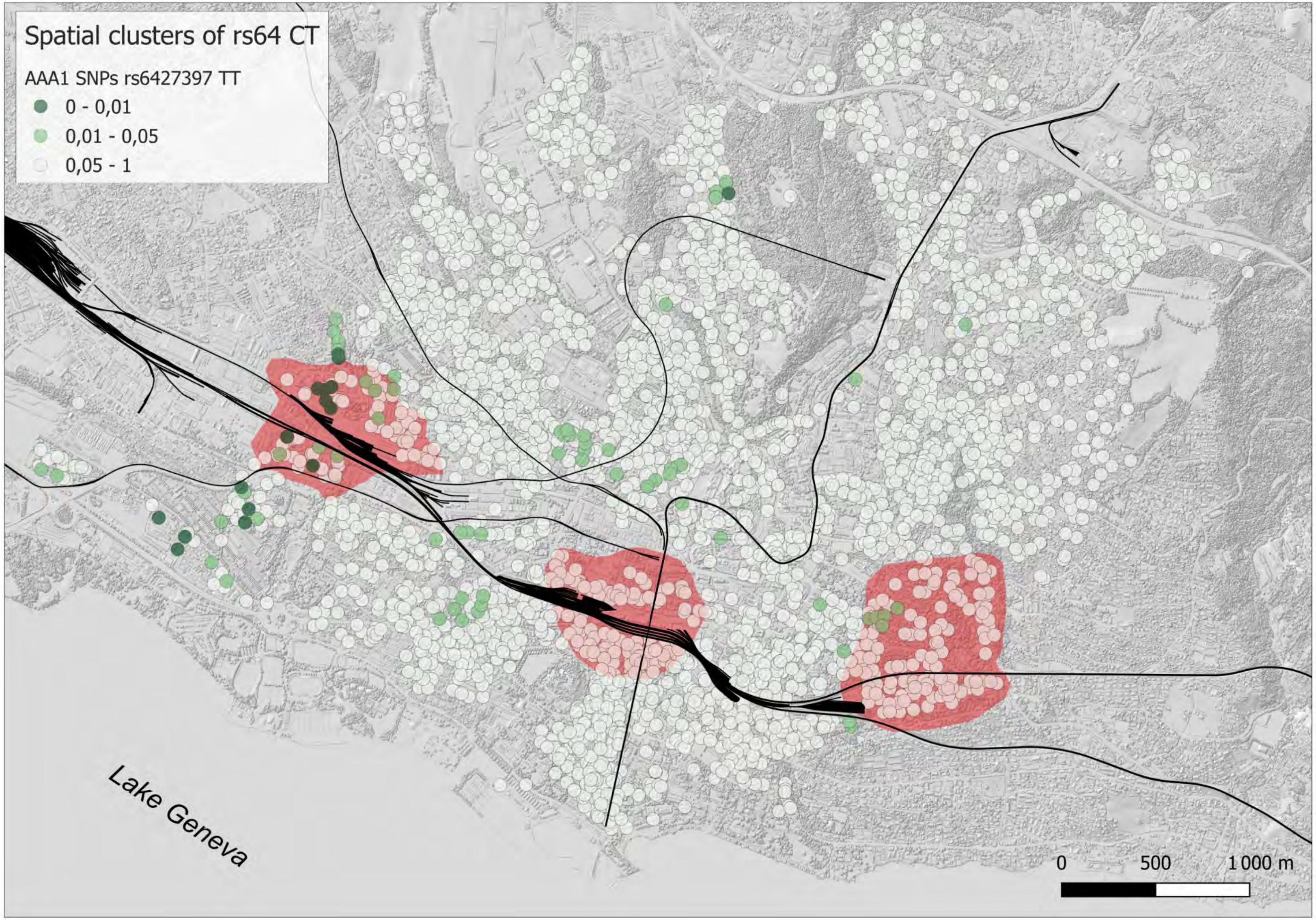


AAA1 Genetics

Spatial clusters

Join-Count Analysis
Spatial lag = 800m

rs64 TT



AAA1 Genetics

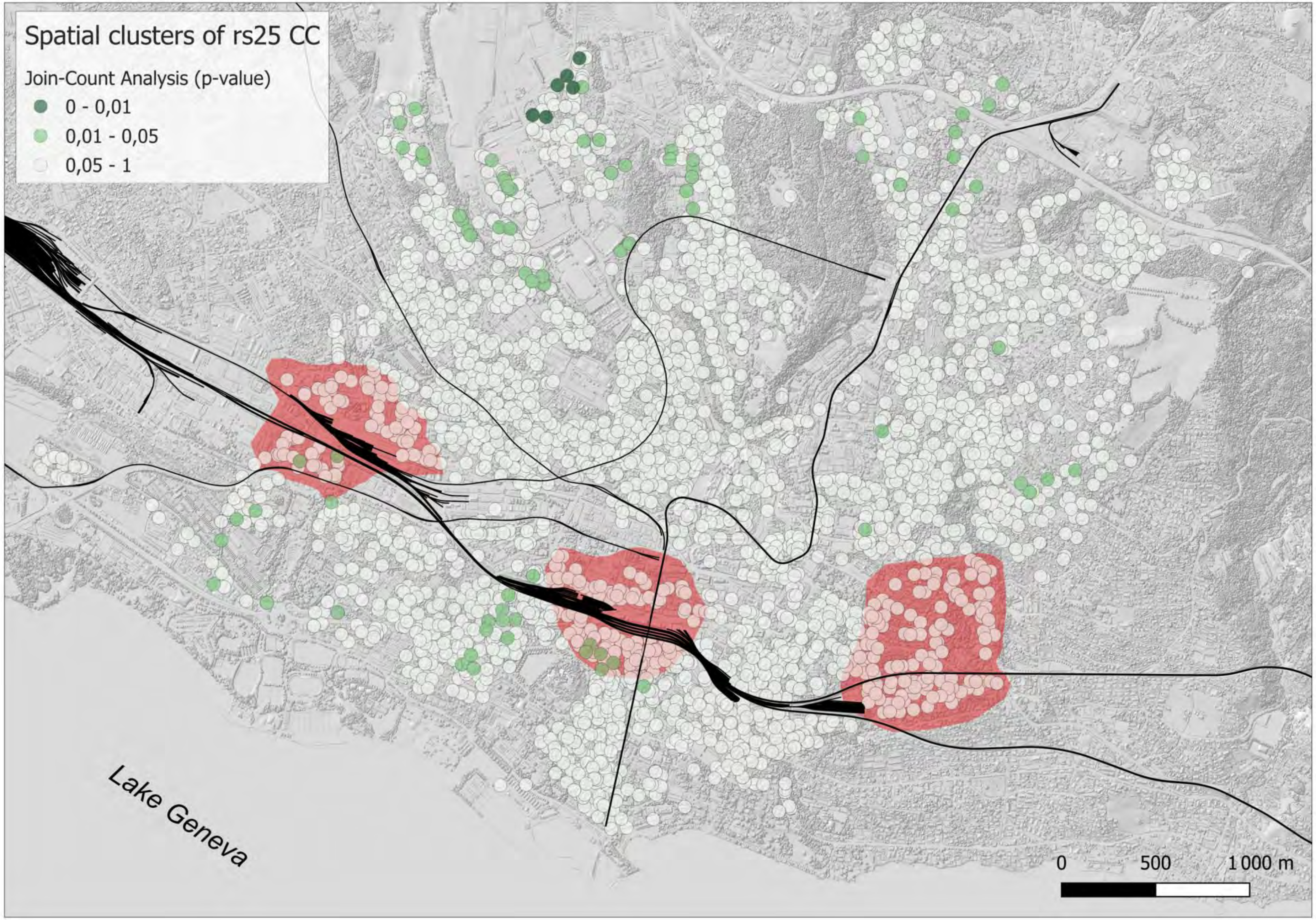
Spatial clusters

Join-Count Analysis
Spatial lag = 800m

rs25 CC

Individuals with the CC genotype had a 2.27 times higher risk of total Coronary Artery Disease (CAD) compared to the reference group

(Antiochos et al. 2017)



AAA1 Genetics

Spatial clusters

Join-Count Analysis
Spatial lag = 800m

rs25 CT

Although not statistically significant, individuals with the CT genotype showed a trend towards an increased risk of total CAD compared to the reference group

(Antiochos et al. 2017)

AAA1 Genetics

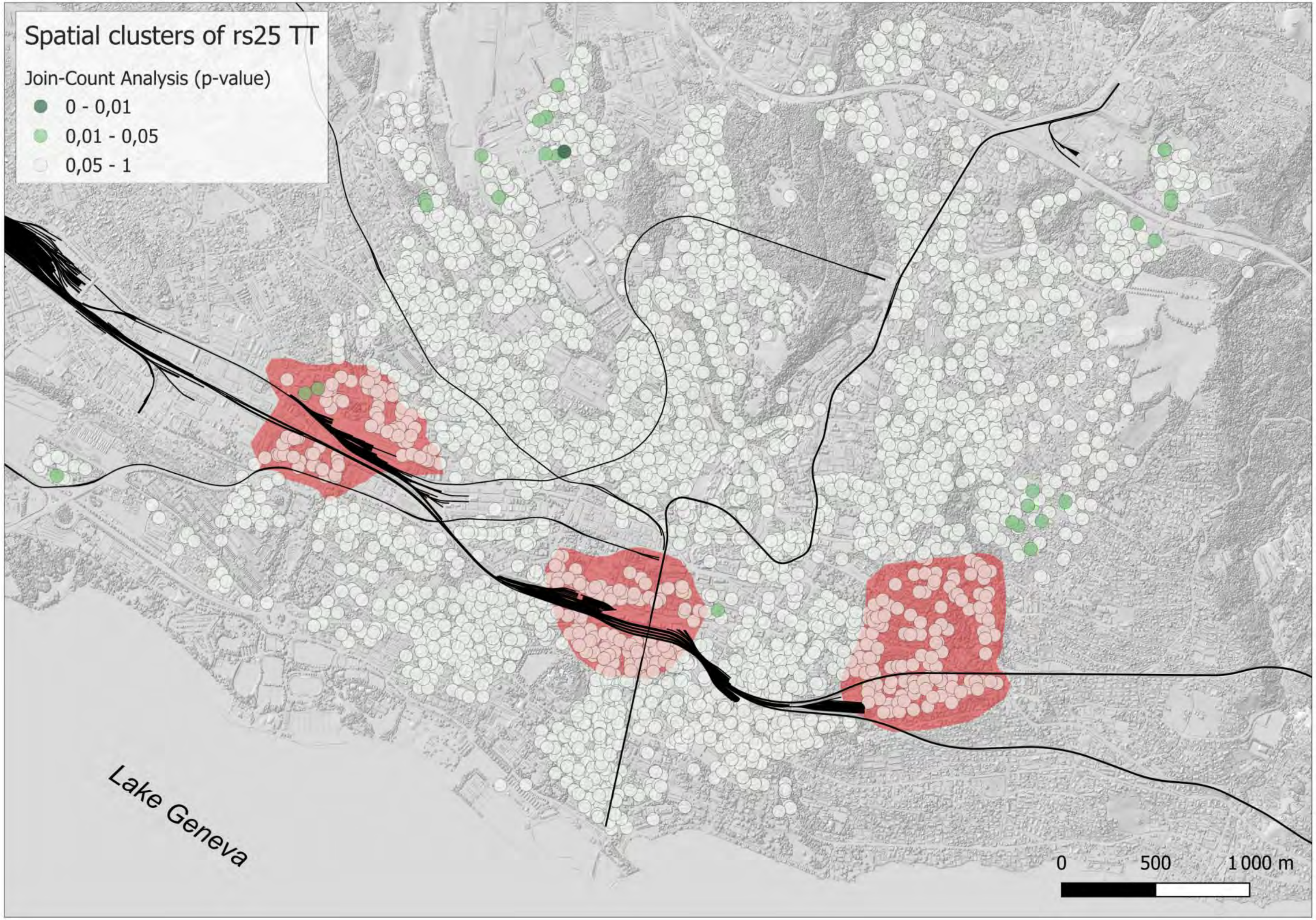
Spatial clusters

Join-Count Analysis
Spatial lag = 800m

rs25 TT

Individuals with the TT genotype had a lower adjusted risk for total CAD compared to the reference group. However, this result was not statistically significant.

(Antiochios et al. 2017)



Conclusion

- Report on the public health policy of the canton of Vaud 2018-2022, Measure No. 4
- → the use of **geographic information** in the development of public health and prevention policy programs must be intensified
- Geospatial approaches can be used systematically to identify neighborhoods exposed to environmental risks and to corresponding diseases (Ladoy 2024)
- Many studies in VD and GE: on BMI (Guessous et al. 2014; Joost et al. 2016), nutrition (Joost et al. 2019; de Ridder et al. 2021), sleep disorders (Joost et al. 2018), physical activity (Vallarta-Robledo et al. 2022), infectious diseases (Ladoy et al. 2021), etc.
- But not on clinical chemistry yet (e.g. heavy metals and associated diseases)

Conclusion

- How to use the results of such studies?
- AAA1 → in case high-risk areas identified here are confirmed, health authorities should transform the identification of clusters into **targeted prevention actions**

→ Environmental remediation if cause identified in the general external environment (Vrijheid, 2014)

→ Development of CVRF prevention program for the target population focused on modification of behavior (tobacco, diet, physical activity, etc.) and psychosocial factors (WHO, 2007)

Thank you for your attention !