

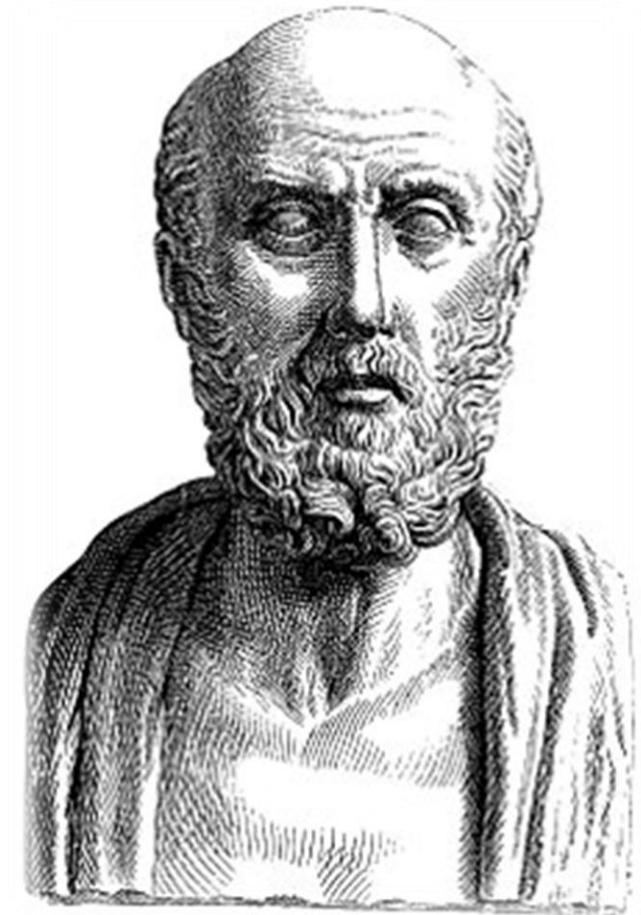
Exploratory data analysis in environmental health

Stéphane Joost & Mayssam Nehme

**Study of the relationship between health and place
& the concept of exposome**

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



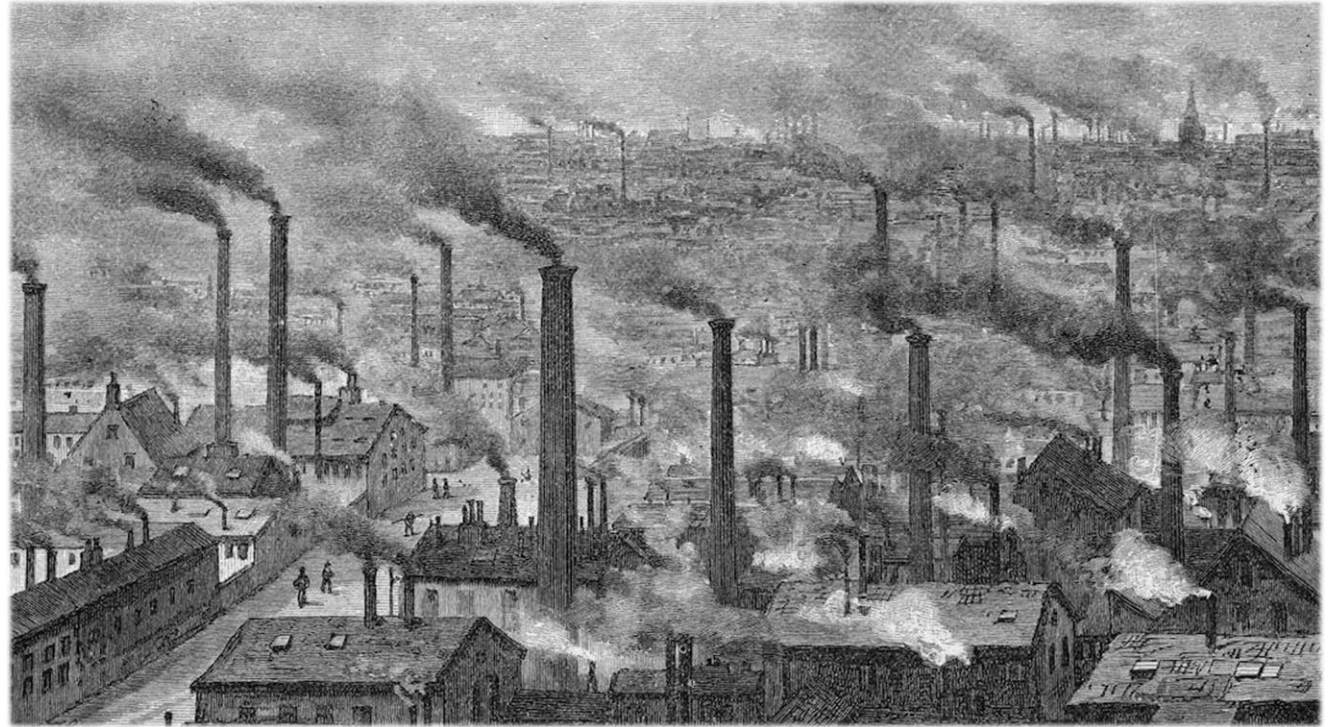
Disease and lockdown areas

- Plague
- 1690 – 1692 in the province of Bari, (Filippo Arrieta)
- The map shows areas most affected by plague
- It highlights boundaries of a quarantine zone imposed to prevent its spread to neighboring towns and other provinces



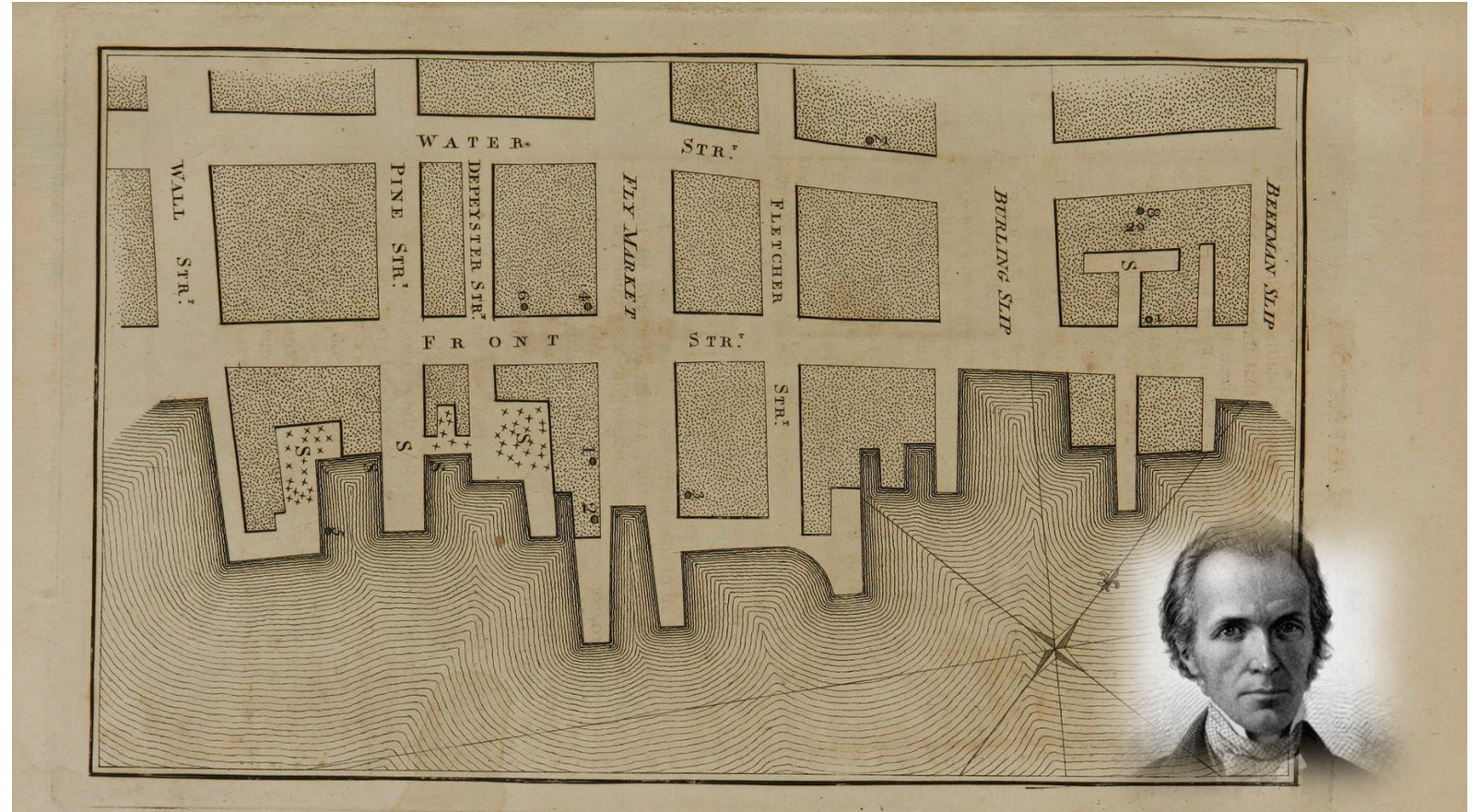
Industrial revolution

- Industrial revolution end of 18th Century: many public health issues
- Need to inform about risk areas where tuberculosis, cholera or yellow fever appear most frequently
- Emergence of medical cartography

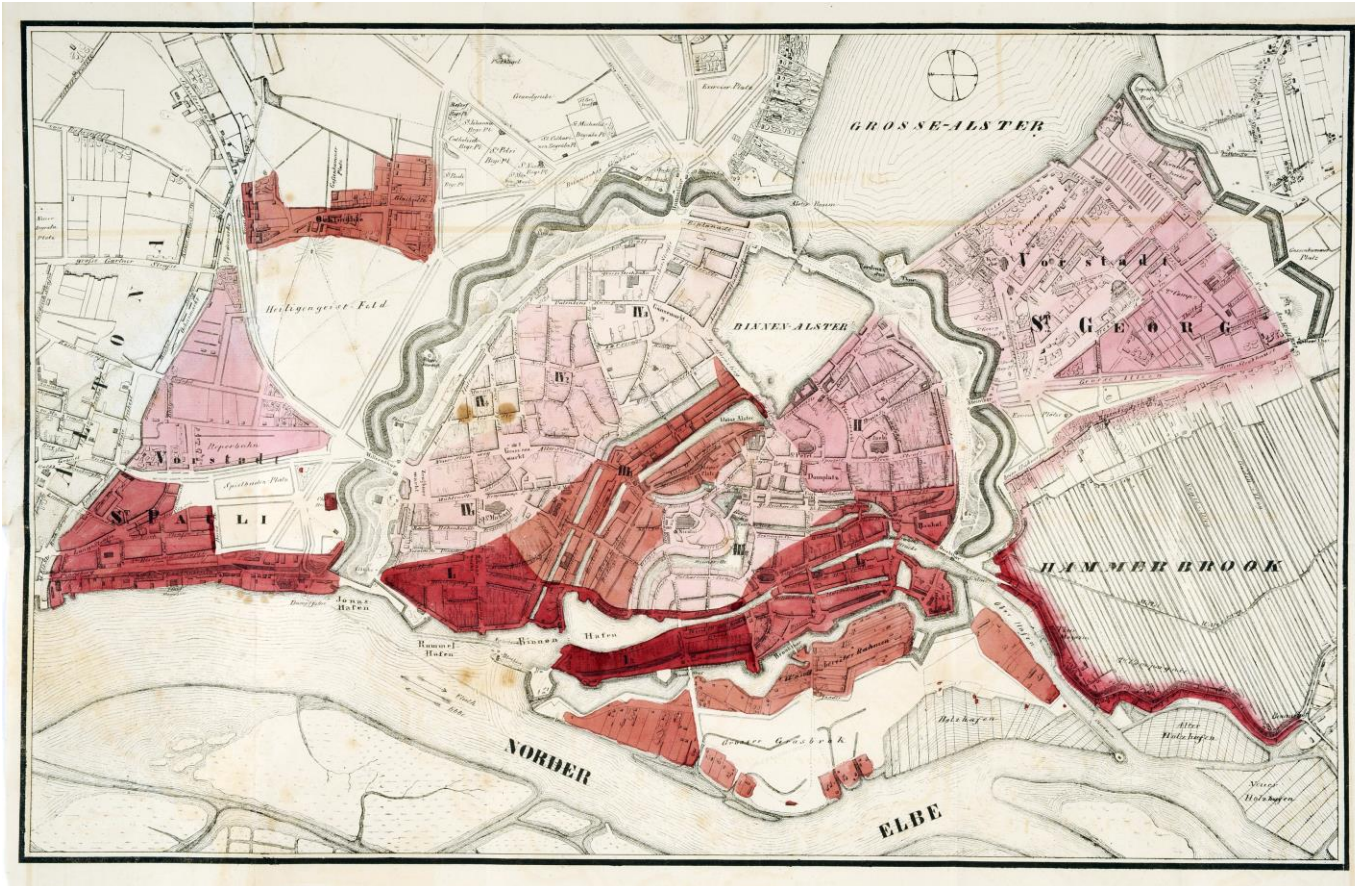


Yellow fever

- 1798 - Valentine Seaman, a surgeon at New York Hospital, investigated the yellow fever outbreak in his city
- Seaman argued the origin of the city's disease outbreak was the smell that arose from the garbage and sewage that accumulated in the harbor area
- “In the city there appears to be an intimate and inseparable connection between the prevalence of yellow fever and the existence of putrid effluvia”

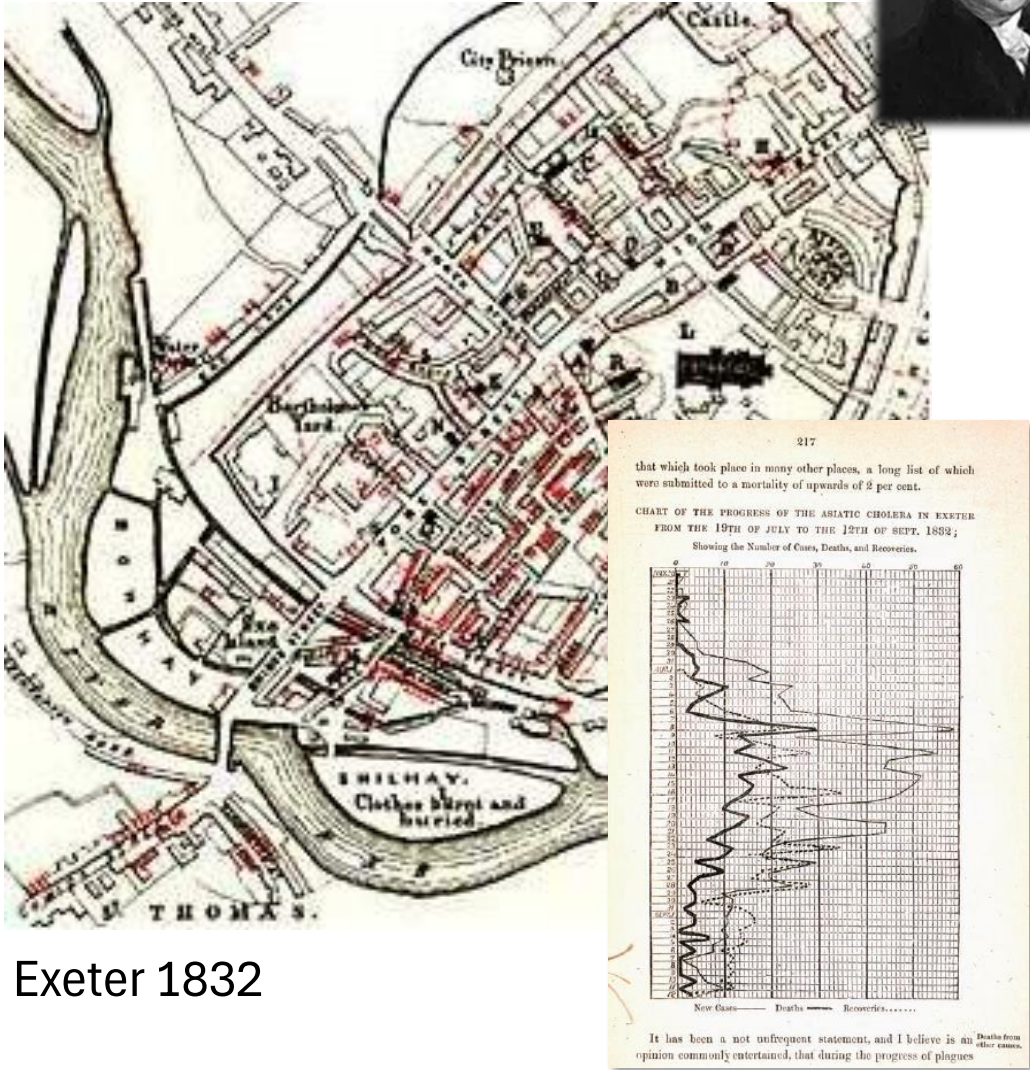


Cholera



- Map of the 1832 cholera outbreak in Hamburg, Germany
- Choropleth map to show the relative aggregated incidence of the disease in different urban districts

Thomas Shapter



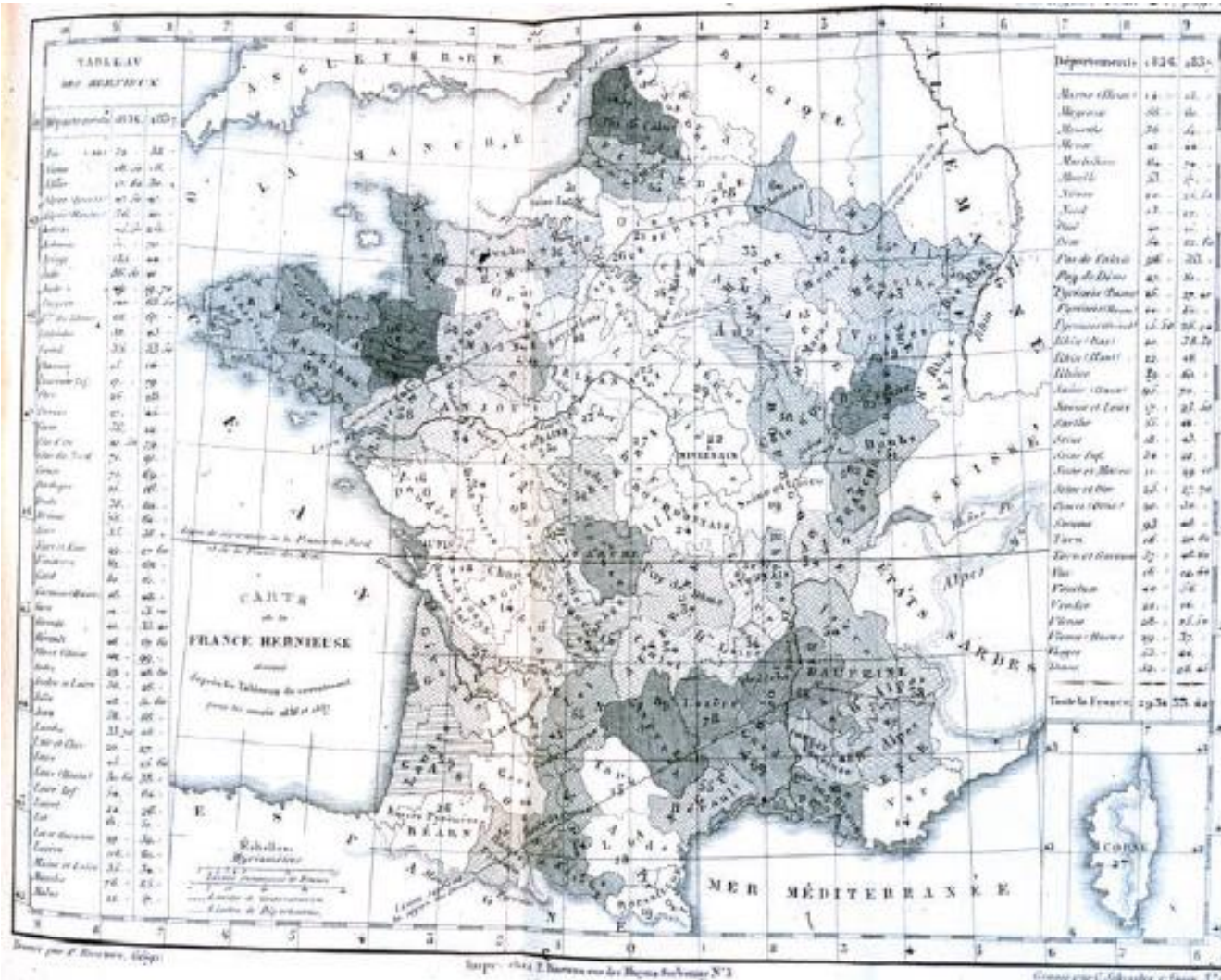
Exeter 1832

John Snow



London, 1854

Hernia



- 1840 - Joseph-François Malgaigne's map of the incidence of hernia in France
- Early example of the "Paris School" of medical statistics
- The map was used to test various hypotheses about the potential causes of hernias

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)

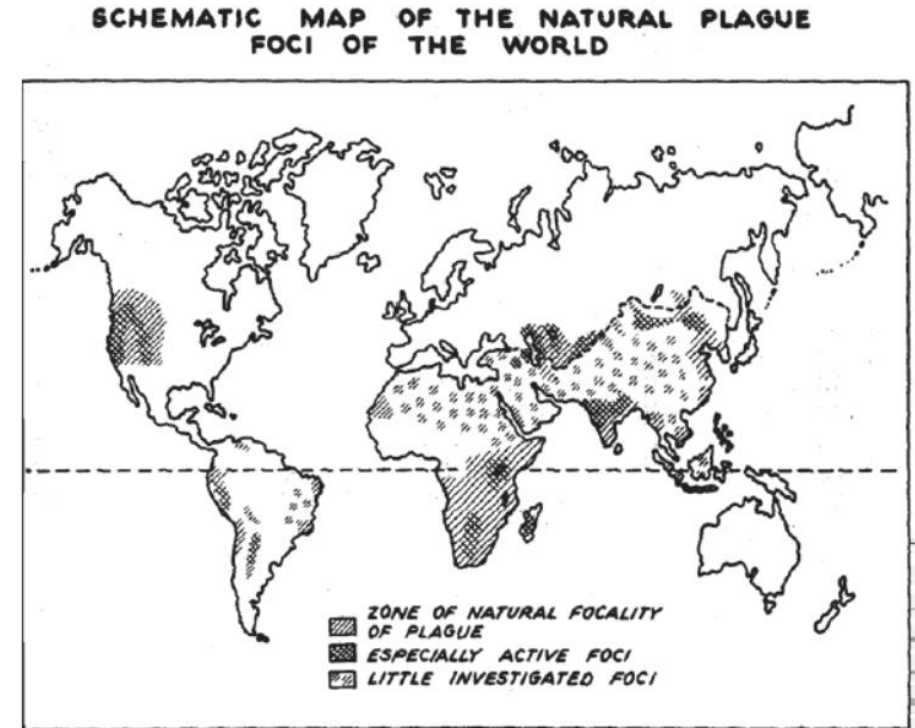


Figure 9.2| May's schematic map of plague foci, distinguishing between environments where plague was known, recurrent, and in some cases, epidemic. Source: May, American Geographical Society (1961).

Health and place: which discipline ?

Medical Geography

Health Geography

Spatial epidemiology

Environmental epidemiology

Environmental health

Medical geography versus spatial epidemiology

- Medical geography (health geography) refers to research applying geographic analytical methods to health, disease, or health care issues
- Its primary focus is on spatial patterns and context (Dummer 2008)
- Spatial epidemiology uses epidemiologic study designs that involve georeferenced study subjects, health facilities, or sources of exposure
- Its primary focus is on populations (Kirby et al. 2017; Szklo & Nieto 2014)

Epidemiology

- Epidemiology is the study of the distribution of the determinants of health (-related states or events) in specified populations
- Application of this study to control health problems
- Epidemiology – an approach used to find the causes of health outcomes and diseases in populations

*Principles of Epidemiology in Public Health Practice,
Centers for Disease Control and Prevention
<https://www.cdc.gov/>*



Image Credit: MarcoVector/Shutterstock.com

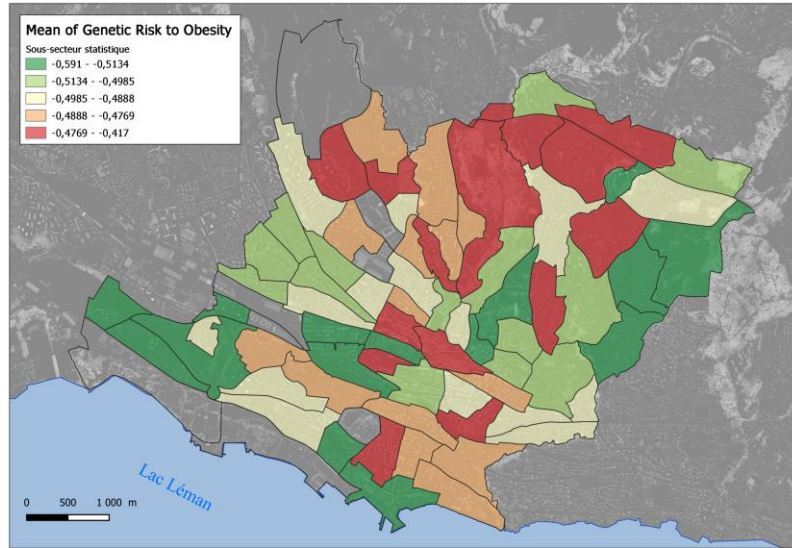
Incidence vs prevalence

- Spatial epidemiology “concerns the analysis of the spatial/geographical distribution of the incidence of disease” by means of the use of population health data (Lawson 2013)
- Incidence refers to the number of new cases (e.g. disease) over a given period (usually a year)
- Prevalence refers to the proportion of existing cases in a defined population at a given time
- It is important to have incidence and prevalence figures differentiated according to sub-populations

Types of analyses in spatial epidemiology

- Elliott et al. (2000) identified four types of spatial analyses in epidemiology:
 - (1) disease mapping
 - (2) geographical correlation studies
 - (3) risk assessment in relation to point or line sources (e.g. along rivers)
 - (4) cluster detection and disease clustering

Disease Mapping

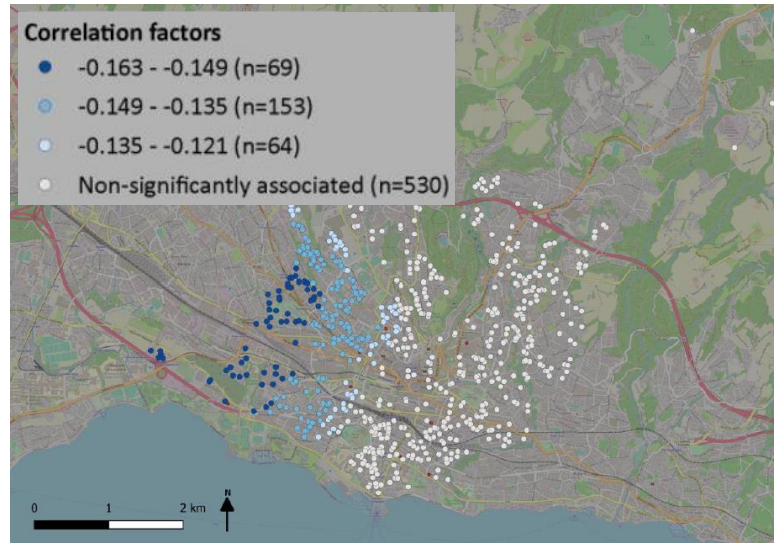


Spatial distribution of the GRS score for genetic susceptibility to obesity by statistical sub-sector in Lausanne

(UK Biobank & [ColaUS-PsycholaUS](#) studies)

Tyrell et al. (2017)

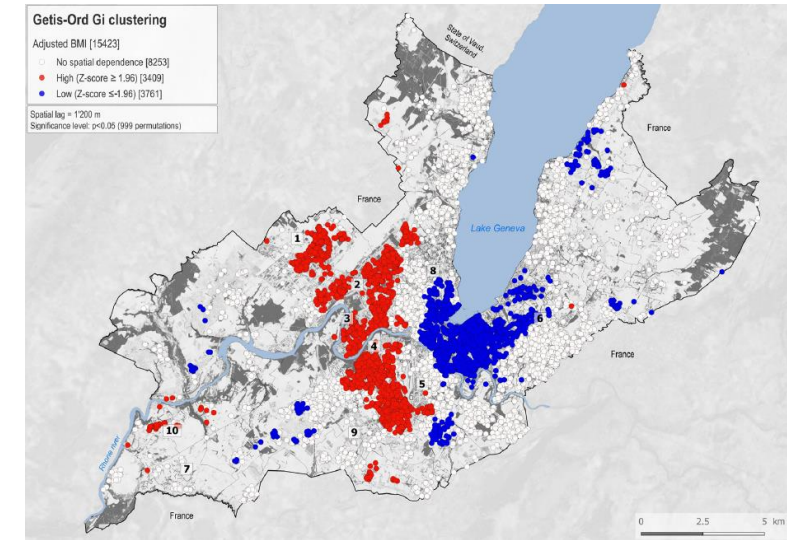
Geographic Correlation



Geographically weighted regression (GWR) between the anxiety indicator State-Trait Anxiety Inventory (STAI) and a vegetation index (NDVI) in Lausanne ([ColaUS-PsycholaUS](#) study)

Vieira (2020) Master thesis SIE/EPFL & LREN/UNIL & CHUV

Disease Clustering



Getis-Ord's Gi illustrating the spatial dependence of body mass index ([Bus santé](#) study, HUG, Genève)

Joost et al. (2019)

Geographic variations of diseases in populations

- In summary, **spatial epidemiology** is the description and analysis of the geographic variations of diseases in populations
 - Investigates demographic, environmental, behavioral, socioeconomic, genetic, and infectious risk factors (Elliott & Wartenberg 2004)
 - The discipline has made significant progress over the past 30 years thanks to:
 - The advent of geographic information systems (GIS)
 - New statistical methods
 - High-resolution georeferenced health and environmental data
- The discipline made important contributions to **environmental epidemiology** and **environmental health**

Epidémiologie environnementale

- Analysis of the impact of environmental exposures on human health
- Identification of natural or artificial risk factors in the environment
- Aims to understand the effects of external risk factors that predispose to disease, developmental abnormalities or death



Santé environnementale

- A branch of public health
- Studies the aspects of human health and quality of life determined by the physical, chemical, biological, social, psychosocial and aesthetic factors of our environment (WHO 1994)
- Encompasses prevention, risk management and public policy



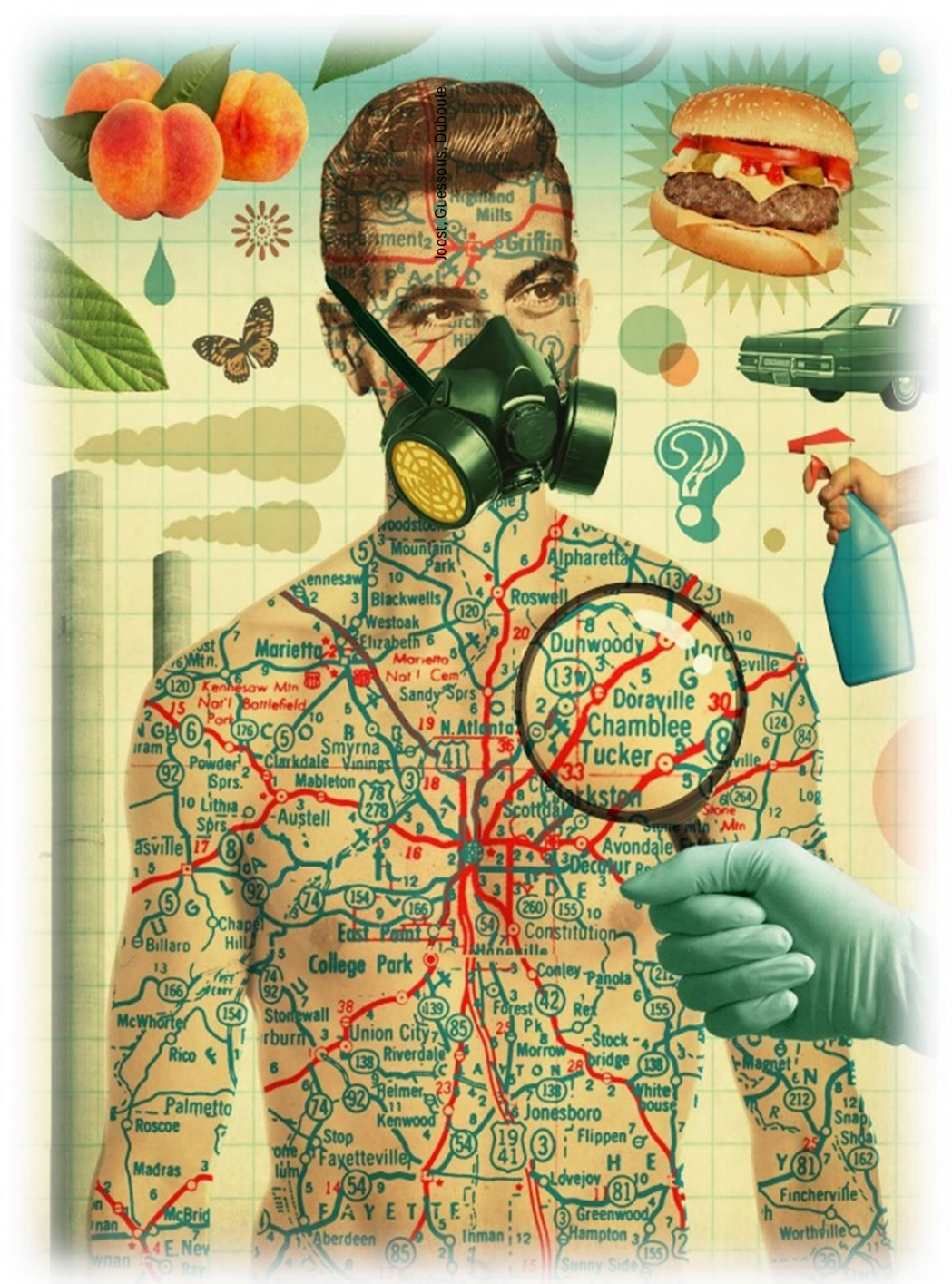
Environmental epidemiology vs environmental health

- Environmental epidemiology focuses on researching and analyzing the links between environmental factors and health, often through statistical studies
- Environmental health covers a broader field, integrating management, prevention and the implementation of policies to reduce environmental impacts on human health

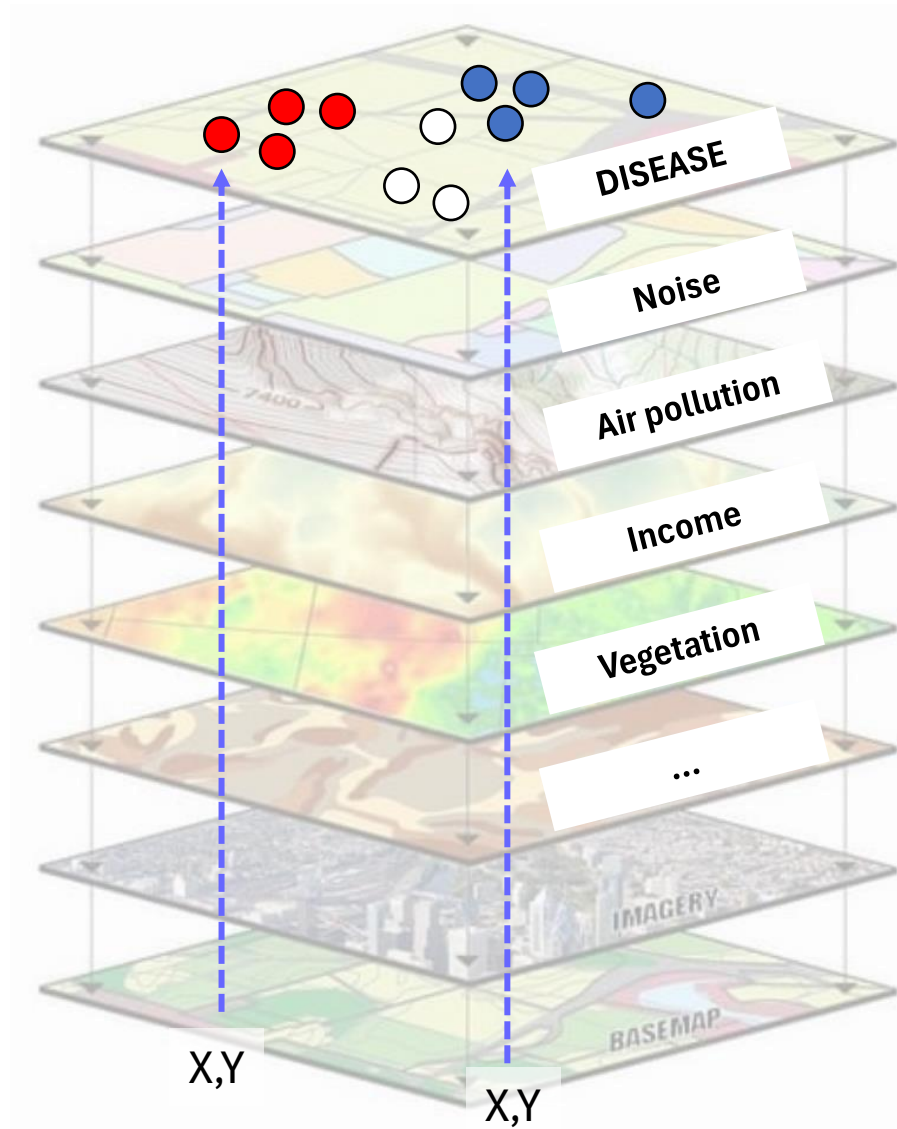
The two disciplines are complementary: environmental epidemiology often provides data and evidence to guide environmental health actions

Exposome

- Totality of exposure to environmental factors
- Key concept in environmental epidemiology & environmental health
- The exposome is described by measurements of environmental conditions: air pollution, temperature, noise, socio-economic conditions (income statistics, education, etc.)
- Geographical information is used to relate the subjects studied and their biological variables/diseases to the descriptors of the exposome

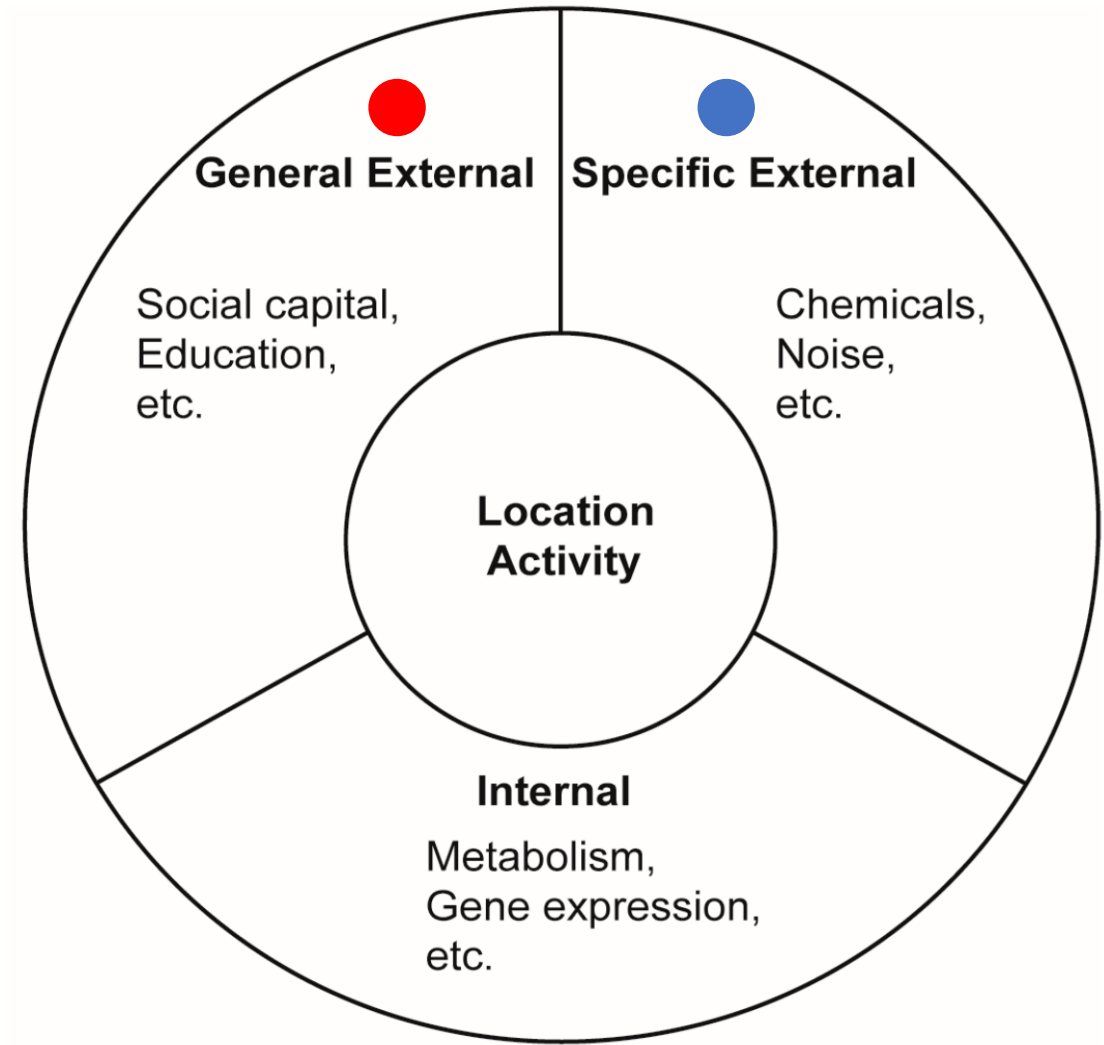


Geographic information as a connector



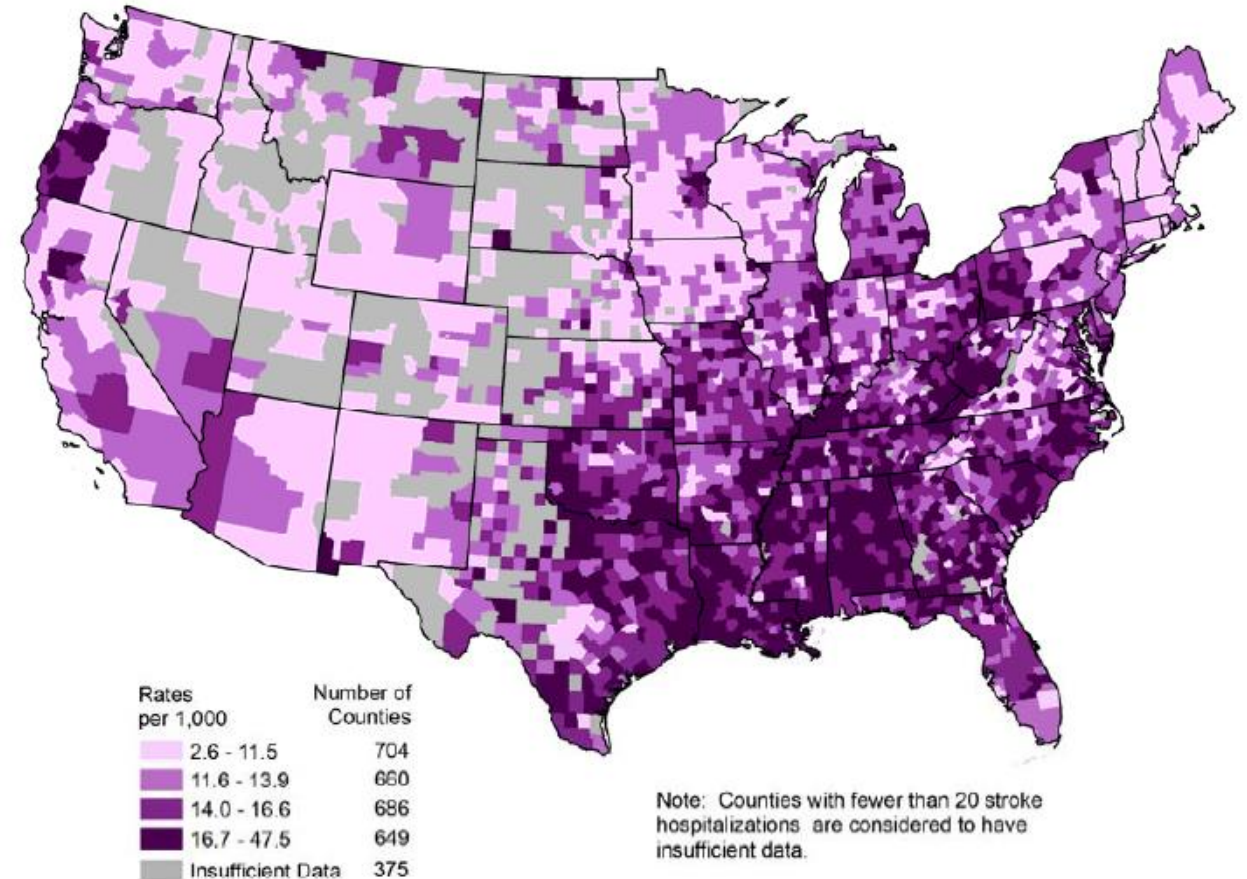
Exposome

- Total exposure to environmental, social and cultural factors (Wild 2012)
- **Specific external exposome**
 - Diet, occupational & environmental exposures
- **General external exposome**
 - Home location
 - Education level
 - Socioeconomic status
- **Internal exposome**
 - Physiology, Age
 - Body morphology, Metabolism
- Need to be described at the correct scale

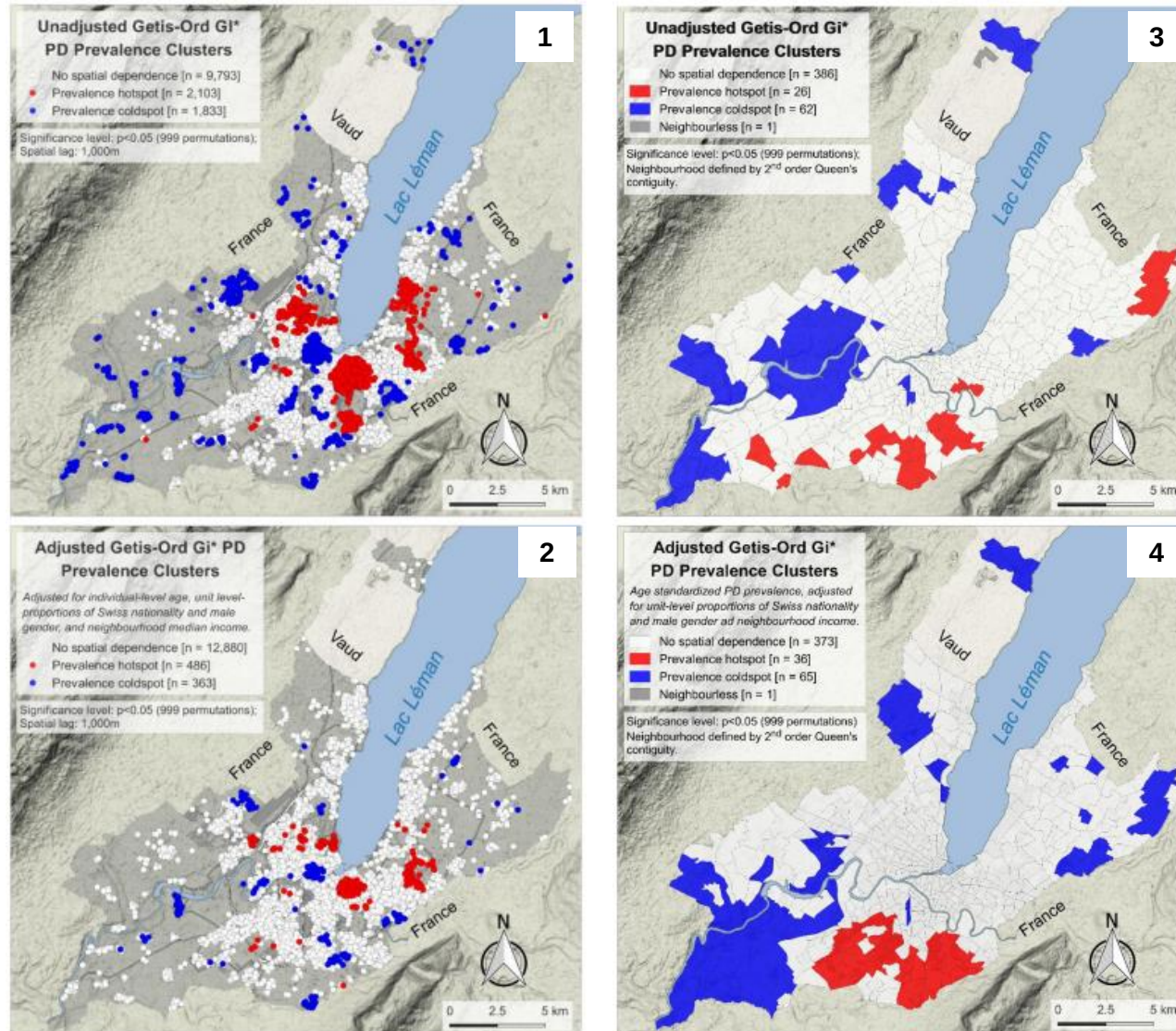


Scale of analysis in spatial epidemiology

- In spatial epidemiology, geolocated data are usually aggregated within administrative units
- Here: stroke hospitalization rates of medicare beneficiaries aged 65 or older in 2005-2006 at the county level
- Exposome data are aggregated at the corresponding level
- Poses methodological problems like ecological fallacy or Modifiable Areal Unit Problem (MAUP) (Openshaw 1984)
- For what use and effectiveness in public health, in the field of prevention in particular?



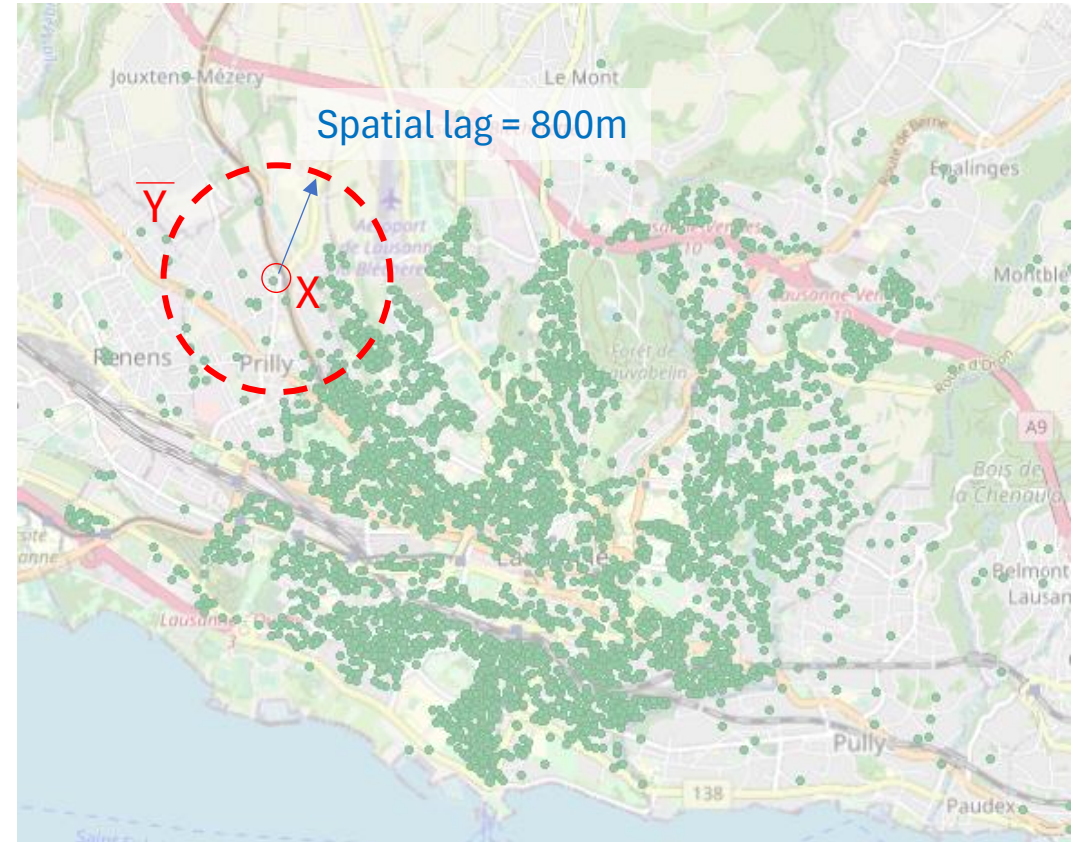
Ecological fallacy



- Ecological fallacy illustrated with an example on Parkinson's disease in the state of Geneva, Switzerland
- Map 1 and 2 show data with points representing subjects for which a probability to get the disease was calculated and then clustered using the Getis-Ord spatial statistics
- Map 1 shows the situation with raw data and map 2 shows the situation after having adjusted the output variable for possible confounders
- Maps 3 and 4 display the same information, but the data were aggregated at the level of Geneva statistical subsectors

Towards more precision in spatial epidemiology

- 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

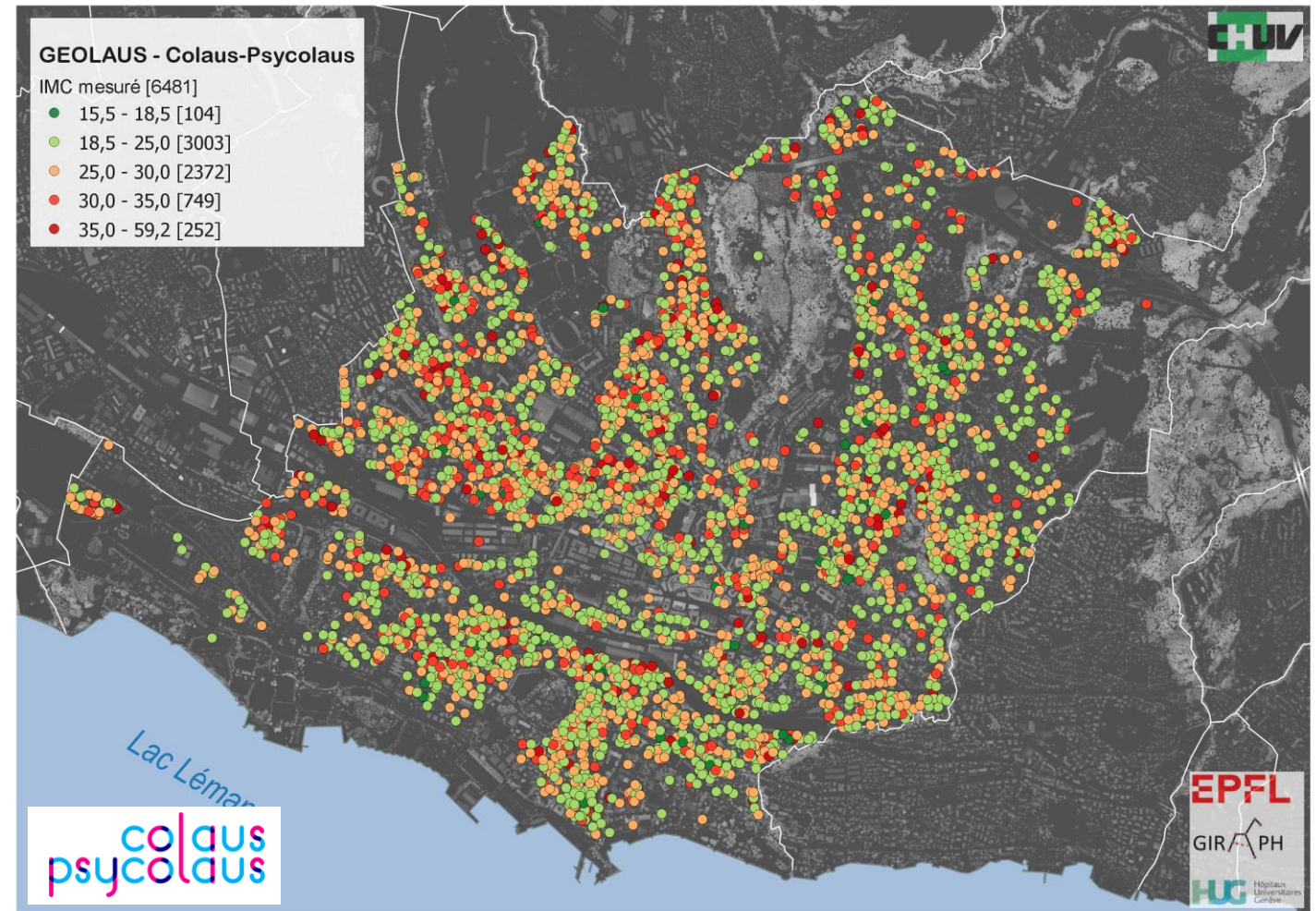


Does X look like $\bar{Y}$?

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

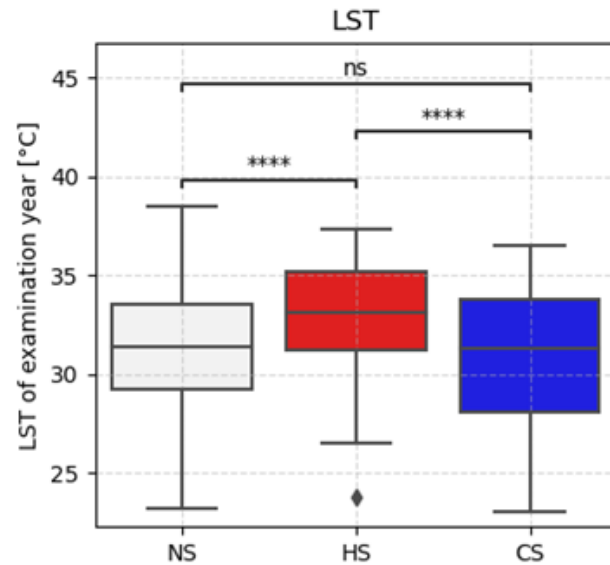


Body Mass Index, Colaus-PsyCola cohort

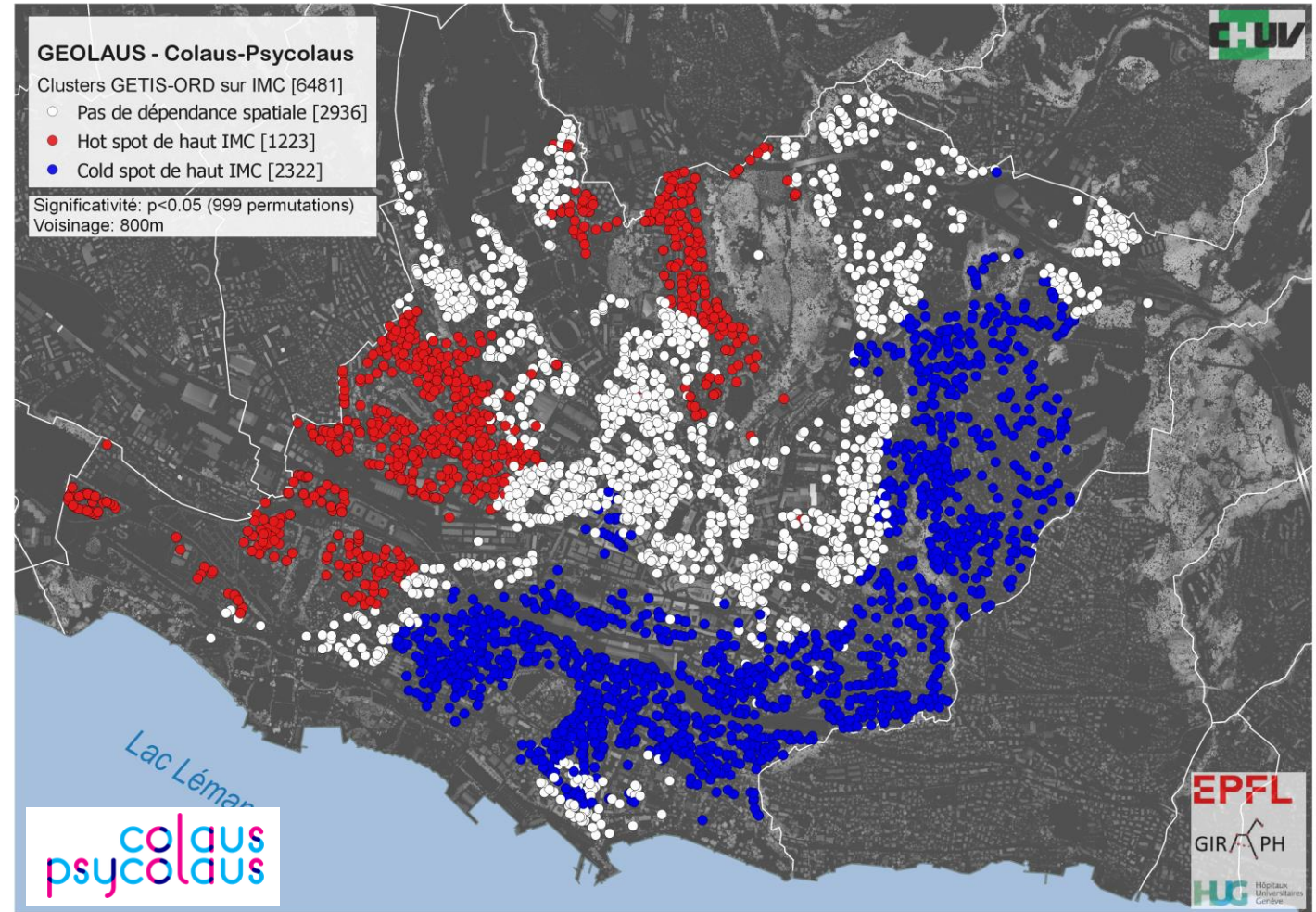
Joost et al. 2016

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



Thank you for your attention !