Exploratory data analysis in environmental health

Stéphane Joost & Mayssam Nehme

Introduction to spatial epidemiology

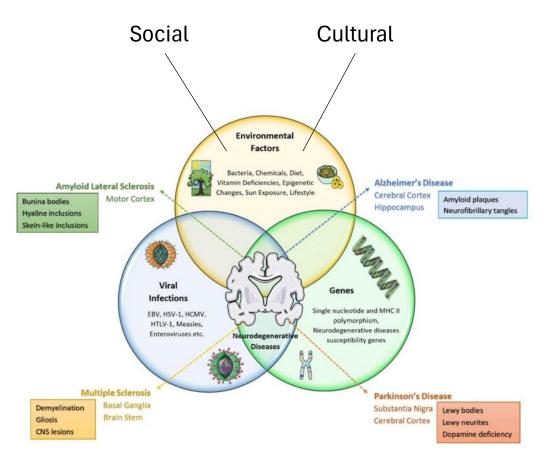


Spatial epidemiology (cf lecture #3)

- Epidemiology an approach used to find the causes of health outcomes and diseases
- Spatial epidemiology uses **epidemiologic study designs** that involve georeferenced study subjects, health facilities, or sources of exposure
- Its primary focus is on **populations** Investigates demographic, environmental, behavioral, socioeconomic, genetic, and infectious risk factors
- Four types of spatial analyses in epidemiology:
 - 1) disease mapping, 2) geographical correlation studies,
 - 3) risk assessment, 4) cluster detection and disease clustering

The multifactorial nature of disease

- The task of epidemiology to disentangle the relationships between health and various risk factors (like e.g. detrimental environmental conditions) is very difficult
- The nature of disease is multifactorial
- There is a diversity of environmental aggressors
 - Biologic, physical, social, and cultural factors
 - Combined with genetic susceptibility
- This suggests the need to implement multidisciplinary epidemiologic investigations

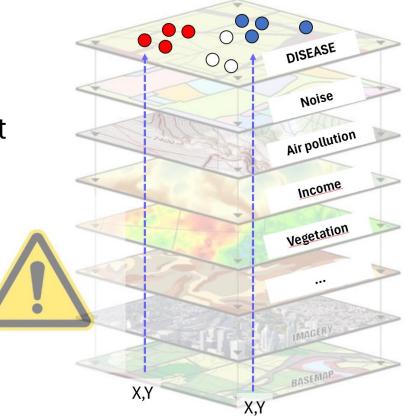


Onisoforou et al. 2020 - CC BY-NC-ND 4.0

GIS for interdisciplinary collaboration

- There is a need to utilize tools like GIS and other geospatial methods able to integrate multilevel*, spatial, and temporal factors
- The multifactorial nature of disease and availability of geospatial tools favors collaborations and creativity in the field of environmental epidemiology
- Moreover recent <u>technology</u> advances in GIS, mobile & smart phones, sensors, diffusion of innovation (web, social networks) offer research opportunities
- The effect of technology → large emphasis on molecular epidemiology and biology of the disease (<u>microlevel</u>) while the <u>macrolevel</u> picture involving the individual in a social, cultural, and physical setting may be missed
- Important to keep in mind in the context of applications in population health and in public health

*Multilevel analysis is an analytical strategy that examines simultaneously factors at the group level and at the individual level



Population health and Public health

- <u>Public health</u> focuses on improving the health outcomes and overall well-being of the public at large rather than individual patients
- Used primarily to describe the health practices and **policies** of local authorities (governments), public health encompasses a wide range of duties, responsibilities, and jobs that all work to ensure positive health outcomes for the public
- Population health is a subset of public health focusing on improving the well-being and health outcomes of a specific group of people or community within the public at large
- Professionals working in population health usually focus their efforts on very specific groups defined by common demographic factors like geographic location, ethnicity, age, or a shared disability

Source: Global Master of Public Health, Imperial College London

Integrating contributions of several disciplines

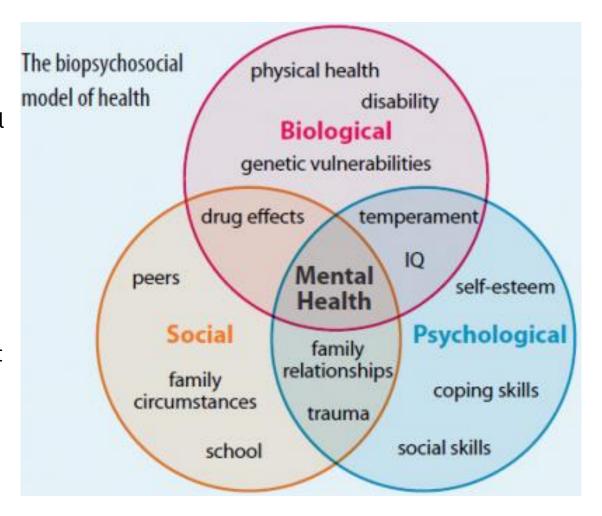
- With interdisciplinary approaches and integration, there are more and more disciplines involved, and the perspectives leading to <u>understanding</u> increase
- Use of GIS and spatial approaches confers key advantages, e.g. the ability to address population-level disease determinants (population health), which were ignored in standard nonspatial individual-level epidemiologic studies leading to simplistic formulations of risk factors (Fielding 1999)
- But we lack an **integrating process** enabling a rigorous new description of risk and disease (Lawson et al. 2016): the way this integrative approach is implemented must be theoretically described
- We need a specific model of disease taking into account the current challenges for public health

Challenges for public health

- Today public health faces unprecedented challenges
 - Important global population growth
 - Aging population
 - Possible irreversible changes in environmental health determinants (globalization and climate change)
- In the case of globalization, negative health effects implied by neoliberal model of global market are: i) spread of disease (trade and migrations), ii) loss of government policy impact, iii) increased labor insecurity (stress, deprivation)
- Climate change: major effects in urban areas where 55% of the world's population lives (68% in 2050; 74% in Switzerland) - inadequate housing, no proper sanitation, bad waste management, bad air quality, noise, water and soil contamination, urban heat islands, lack of space for walking, cycling and active living (cities are epicentres of an epidemic of non-communicable diseases (WHO, 2024)

Models of disease

- Since the late 1970s, dominant model of disease is the biopsychosocial model proposed by George Engel in 1977
- To understand a person's medical condition it is not simply the biological factors to consider, but also the psychological and social factors (Gatchel et al. 2007)
- Environmental factors are not in the picture
- Environmental-social cause model: ecoepidemiology (Schwartz et al. 1999)
- Integrated Socio-Environmental Model of Health and Well-Being – ISEM (Alvarez et al. 2018)
- How social and environmental factors combine and interact to affect health and well-being over the life span
- → understand how the underlying environmental (including social, demographic, psychological) and genetic factors produce risks and how these translate into health, disease, and quality of life (Fielding 1999)
- Key role for spatial approaches

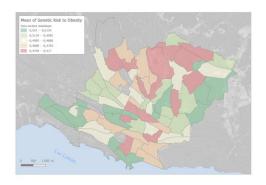


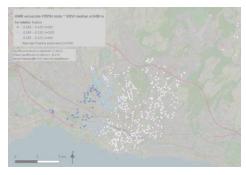
Exposome

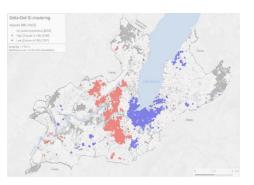
- Description and studies of the "exposome" (Wild 2005, cf lecture #3)
- Concept of <u>cumulative risk assessment</u> (chemicals + community-based and population-based stressors)
- Patel and Ioannidis (2014): model to discover environmental exposures associated with disease = "environment-wide association studies" (EWAS)
- Consensus around a framework supporting risk assessment for decision making in public health
- "Multidisciplinary epidemiology" utilizing new and innovative approaches and data resources of GIS and geospatial methods from multiple epidemiologic perspectives (Lawson et al. 2016)
- Compatible with the current public health challenges
- Can provide an integrated process to better understand disease

A place for the traditional geospatial perspective...

- 1. Disease mapping
- 2. Geographical correlation studies
- 3. Risk assessment
- 4. Cluster detection and disease clustering





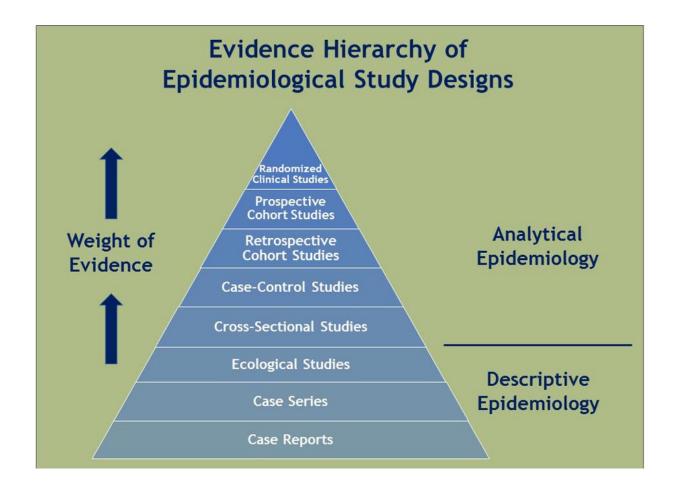


... into the multidisciplinary disease model

Geospatial perspective model	Geospatial methods	Issues faced	NIH model (US National Institute of Health)	
Traditional	Disease mapping Disease clustering Ecological analysis Exposure assessment	Emphasis on multiple risk factors	Benchside (refers to basic laboratory research, experimental)	
Acute event (crisis)	Rapid event assessments Environmental toxin spread Future disaster prevention	Unprecedented challenges Global population growth Aging	Bedside (focuses on clinical practice, diagnosis, and management of disease in real-world settings)	
	COVID-19 cluster detection	Irreversible changes in environmental health determinants		
Community	Social structure Community-derived data Maximizing public health impact via targeted interventions	Need for methods to examine historical and social forces that determine population disease risk	Community (engaging communities in research, healthcare, and public health interventions)	

Relationship between proposed geospatial perspectives and theoretical comparisons (Lawson et al. 2016)

Ecological analysis



• Ecological analyses provide **etiologic cues** about the relationship between spatial disease distributions and explanatory factors

Etiology

- Etiology is the study of the cause or origin of a disease or a condition
- In medicine and epidemiology it refers to understanding what factors lead to the development of a disease, which can include biological, environmental, genetic, or behavioral influences
- There are different types of causes identified in etiology:
 - **Primary cause**: The main factor directly responsible for the disease. For example, the bacterium *Mycobacterium tuberculosis* is the primary cause of tuberculosis
 - **Contributory factors**: Conditions or behaviors that increase the likelihood of developing a disease but are not the direct cause. For instance, smoking is a contributory factor to lung cancer, though not every smoker will necessarily develop the disease
 - Idiopathic: When the cause of a disease is unknown, it is described as having an idiopathic etiology.
- Understanding etiology is crucial for disease prevention, diagnosis, and treatment, as it helps healthcare professionals identify risk factors and develop strategies to address or mitigate the root causes of diseases

Ecological analysis - geographic correlation

- Geographic correlation studies investigate geographic variations of environmental exposures, sociodemographic indicators, and lifestyle factors in relation to health outcomes, most often at the aggregate level (Elliott and Wartenberg 2004)
- Ecologic studies take advantage of specialized spatial statistical procedures to test statistical associations between exposures and disease at an aggregate level or not
- Allow one to develop and test spatial hypotheses of disease etiology (but cf ecological fallacy or MAUP problem)
- Ecological analysis may be considered as an end product of spatial modeling, a mechanisms by which geospatial "bench discoveries" are made
- This type of analysis is able to provide etiologic cues at an ecologic level, but they can also be used at the individual level if (geo)data are available, and if ethical aspects can be addressed

Case study

Combining disease clustering and ecological analysis

- 1. First investigate the spatial dependence of a health variable. Population dataset with individual coordinates (X, Y and attributes)
- 2. Use spatial statistics like Local Moran's I, Getis-Ord Gi, local Geary's C
- 3. Identify clusters of different behaviors (distinct spatial autocorrelation classes)
- 4. Overlay environmental information characterizing the territory (often raster)
- 5. Transfer the value of environmental variables as additional attributes in the population dataset
- 6. Variance analysis to check if the mean (or another statistic) for a given environmental variable is significantly different between the classes

Spatial clusters of daytime sleepiness in Lausanne



Le PS lausannois lance l'idée d'un centre-ville à 30 km/h

Bruit La conseillère Anne-Françoise Decollogny suggère la mesure pour lutter contre le bruit routier excessif.



Anne-Françoise Decollogny se félicite des tests nocturnes menés à Vinet-Beaulieu et propose désormais d'aller plus loin.

Image: Florian Cella

Un logiciel pour prédire le bruit qu'il fera

Dans le cadre de la lutte contre le bruit à Lausanne, la Municipalité veut acheter un logiciel (CadnaA) spécialisé dans sa prédiction. En simple: un outil informatique pour prévoir le bruit qu'il va faire. Il est présenté par ses concepteurs comme la solution «la plus performante» pour le calcul, l'évaluation, la prévision et la représentation de l'impact du bruit dans l'environnement. Son coût, qui figure dans la première série de crédits supplémentaires pour le budget 2018 de la Ville de Lausanne: 12'000 francs.



Spatial clusters of daytime sleepiness in Lausanne

- Colaus-PsyColaus study, a longitudinal study in the population of Lausanne
- 6700 participants recruited between 2001 and 2003 (Baseline)
- Follow-up medical visits: FU1:2009–2012; FU2: 2014–2017; FU3: 2018–2021
- Hypnolaus substudy: 3700 subjects from the baseline sample, georeferenced at their place of residence
- Questionnaires on daytime sleepiness
- Sleepiness measured using the Epworth Sleepiness Scale (ESS)

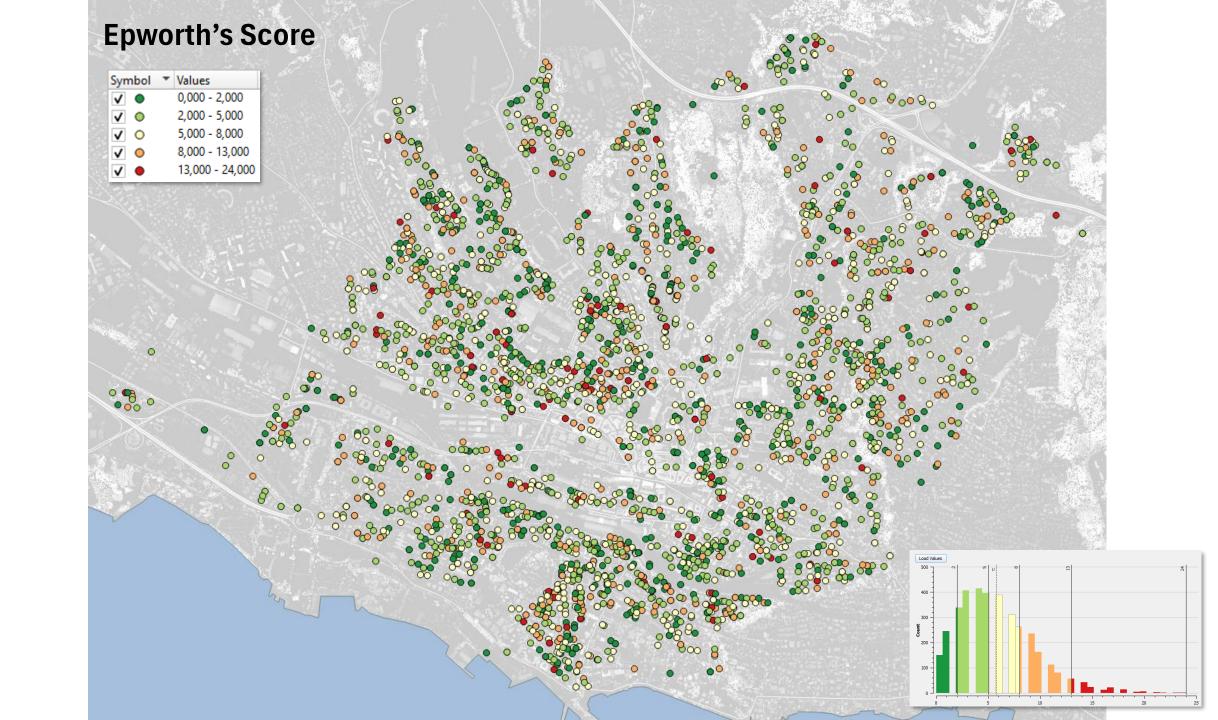
Epworth Sleepiness Scale (ESS)

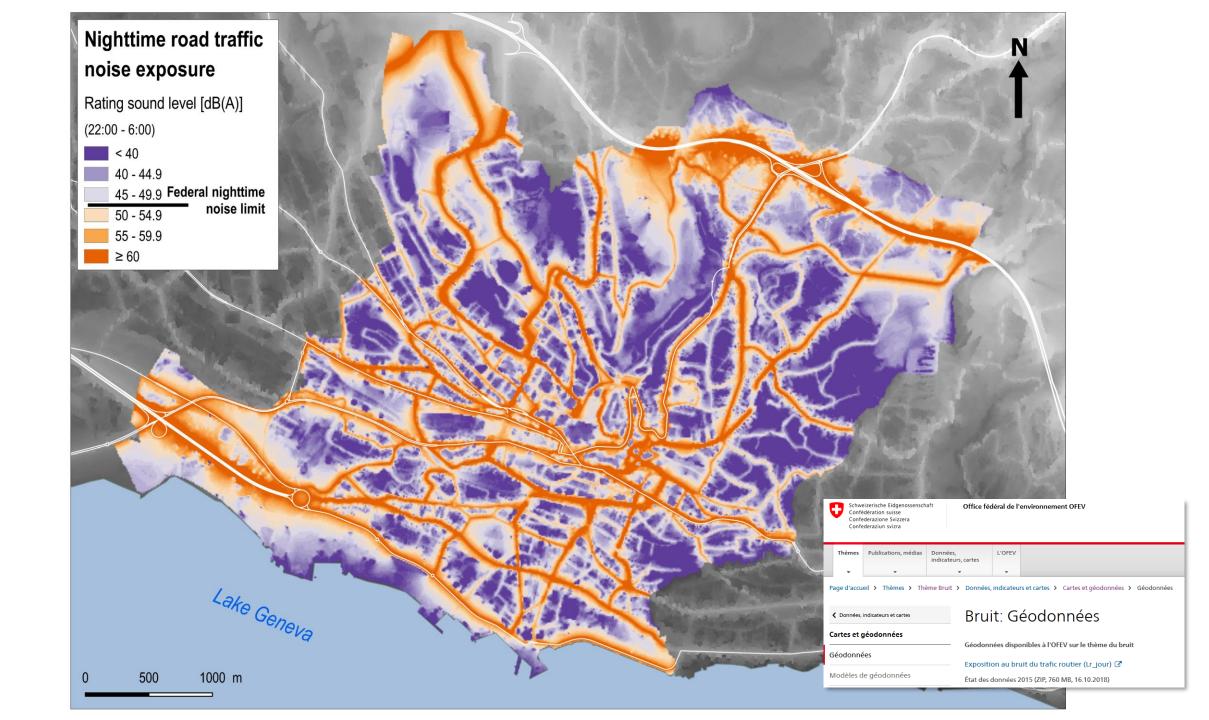
- The 3,700 participants in the GeoHypnolaus study completed a short questionnaire to quantify daytime sleepiness
- The Epworth Sleepiness Scale estimates the probability of falling asleep in everyday situations
- The higher the score, the greater the daytime sleepiness, and the more likely it is that the person is experiencing night-time sleep problems, including sleep disturbed by environmental factors

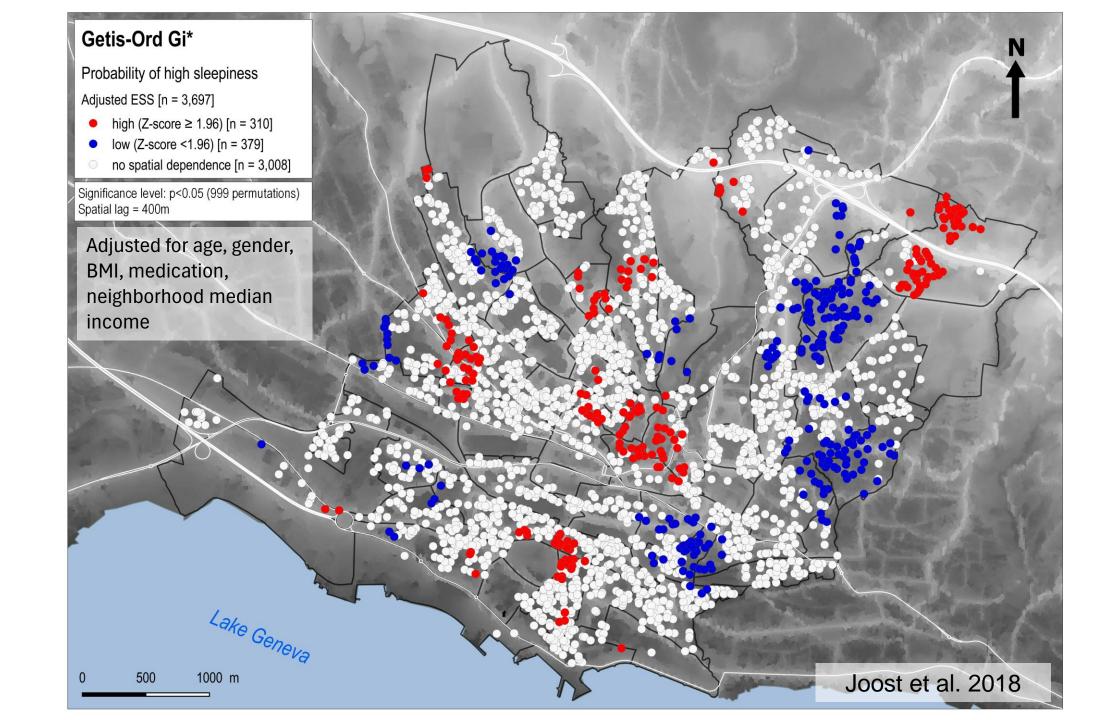
Test d'Epworth

Afin de pouvoir évaluer chez vous une éventuelle somnolence dans la journée, vous pouvez répondre aux questions de l'échelle de somnolence d'Epworth. Voici quelques situations relativement usuelles où nous vous demandons d'évaluer le risque de vous assoupir. Si vous n'avez pas été récemment dans l'une de ces situations, essayez d'imaginer comment cette situation pourrait vous affecter. Pour répondre, utilisez l'échelle suivante en choisissant le chiffre le plus approprié pour chaque situation : 0 = aucun risque de somnoler ou de m'endormir 1 = faible risque 2 = risque moyen 3 = fort risque

	0	1	2	3
Pendant que vous êtes occupé à lire un document	0	0	0	0
Devant la télévision ou au cinéma	0	0	0	0
Assis inactif dans un lieu public(salle d'attente, théâtre,)	0	0	0	0
Passager, depuis au moins une heure sans interruptions, d'une voiture ou d'un transport en commun(train, bus, avion,)	0	0	0	0
Allongé pour une sieste, lorsque les circonstances le permettent	0	0	0	0
En position assise au cours d'une conversation (ou au téléphone) avec un proche	0	0	0	0
Tranquillement assis à table à la fin d'un repas sans alcool	0	0	0	0
Au volant d'une voiture immobilisée depuis quelques minutes dans un embouteillage	0	0	0	0

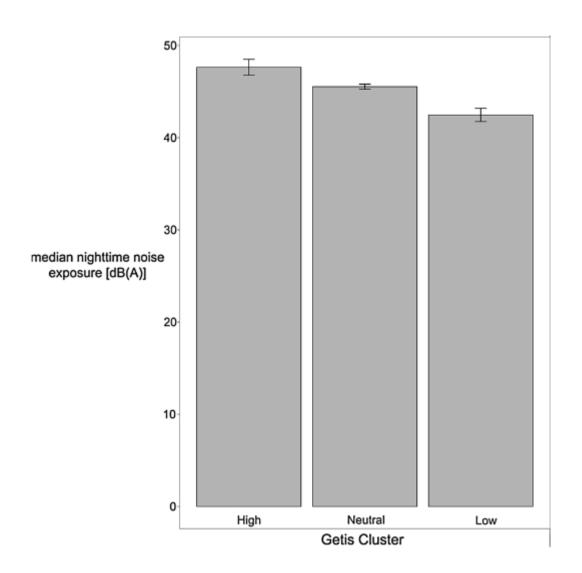






Results

- Variance analysis on cluster classes (Tukey's HSD)
- We observe a dose-response effect
- Getis-Ord: median nighttime noise in the high clusters was 2.10 dB(A) higher than in the neutral class (p < 0.001) and 5.16 dB(A) higher than in low clusters (p < 0.001)
- Local Moran's I shows a dose-response effect too
- The difference between the nighttime noise levels in the high-high and low-low clusters was 4.49 dB(A) (p < 0.001)
- The difference is of 1.80 dB(A) between high-high and the neutral class (p < 0.05).



Use of results by the city of lausanne

International Journal of Hygiene and Environmental Health 221 (2018) 951–957



Contents lists available at ScienceDirect

International Journal of Hygiene and Environmental Health

journal homepage: www.elsevier.com/locate/ijheh

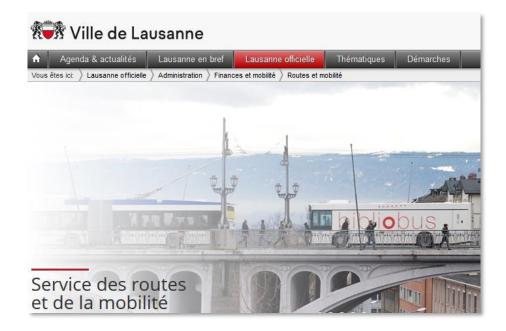


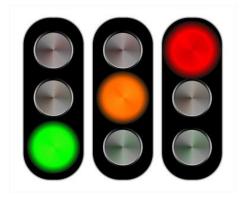
Spatial clusters of daytime sleepiness and association with nighttime noise levels in a Swiss general population (GeoHypnoLaus)



Stéphane Joost^{a,b,c,1}, José Haba-Rubio^{d,1}, Rebecca Himsl^{a,b,c}, Peter Vollenweider^e, Martin Preisig^g, Gérard Waeber^e, Pedro Marques-Vidal^{c,e}, Raphaël Heinzer^{d,2}, Idris Guessous^{b,c,f,2,*}

- a Laboratory of Geographic Information Systems (LASIG), School of Architecture, Civil and Environmental Engineering (ENAC), Ecole Polytechnique Fédérale de Lausanne (FPFL), Lausanne Switzerland
- b Unit of Population Epidemiology, Division of Primary Care Medicine, Department of Community Medicine, Primary Care and Emergency Medicine, Geneva University Hospitals, Geneva, Switzerland
- c GIRAPH Lab (Geographic information for research and analyses in public health), Switzerland
- d Center for Investigation and Research in Sleep, Lausanne University Hospital (CHUV) and Lausanne University, Lausanne, Switzerland
- ^c Department of Medicine, Internal Medicine, Lausanne University Hospital (CHUV) and Lausanne University, Lausanne, Switzerland
- f Department for Ambulatory Care and Community Medicine, University of Lausanne, Lausanne, Switzerland
- ⁸ Department of Psychiatry, Lausanne University Hospital and Lausanne University, Lausanne, Switzerland





Ici on teste
30km/h
de nuit

Sur les avenues de Beaulieu et Vinet

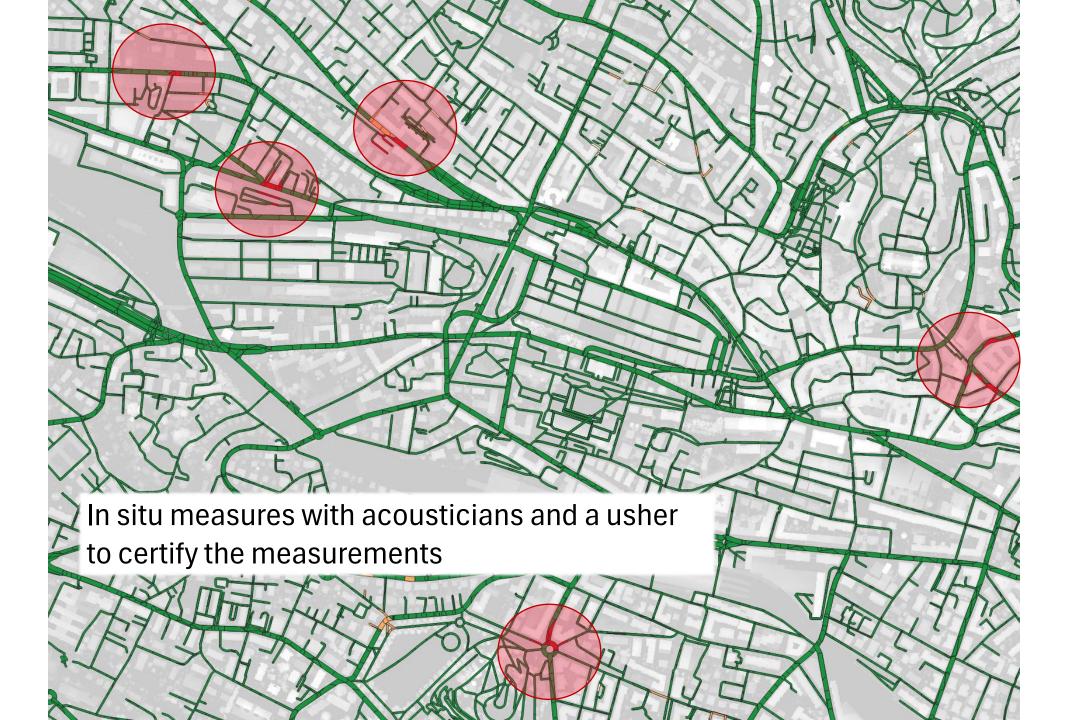
22h - 06h



Setting priorities between neighborhoods

- Where are the more exposed neighborhood: hotspots with the highest dB(a) mean values
- Use Local Moran's I classes







Case study 2: sugar-sweetened beverage and BMI in Geneva

- Obesity and obesity-related diseases represent a major public health concern (diabetes type 2, hypertension, high cholesterol, coronary artery disease, etc.)
- Studies have highlighted the role of sugar-sweetened beverages (SSBs) consumption in the development of these diseases
- The fine identification of populations and areas in need for public health intervention remains challenging
- This study investigates the existence of spatial clustering of SSB intake frequency (SSB-IF) and body mass index (BMI), and **their potential spatial overlap** in a population of adults of the state of Geneva using a fine-scale geospatial approach

Data

- Data on adults collected using the Bus Santé study
- A cross-sectional population-based study that collects information on cardiovascular risk factors
- Since 1995, every year, a stratified sample of 500 men and 500 women representative of the State of Geneva's general population aged 35–74 (20–74 since 2011) is recruited and studied
- For this analysis, data from surveys **1995 to 2014** were used, corresponding to 15,423 participants
- The average participation rate for 1995–2014 was 61% (range: 53–69%)
- Medical visit + questionnaires with a health professional in the Bus santé

Variables analyzed

- Body mass index (BMI) is calculated as weight (kg)/height (m2)
- Sugar-sweetened beverages intake frequency (SSB-IF) is assessed for every participant using a food frequency questionnaire (FFQ)
- This FFQ assesses the dietary intake of the previous four weeks
- It consists of 97 different food and beverage items, including SSB (colas, sodas, lemonades, syrups)
- For each item, consumption frequencies ranges from "less than once during the last four weeks" to "2 or more times per day"
- Daily SSB-IF was recoded as follows: from 0 for "less than once during the last 4 weeks" to 2.5 for "2 or more times per day"

Covariates

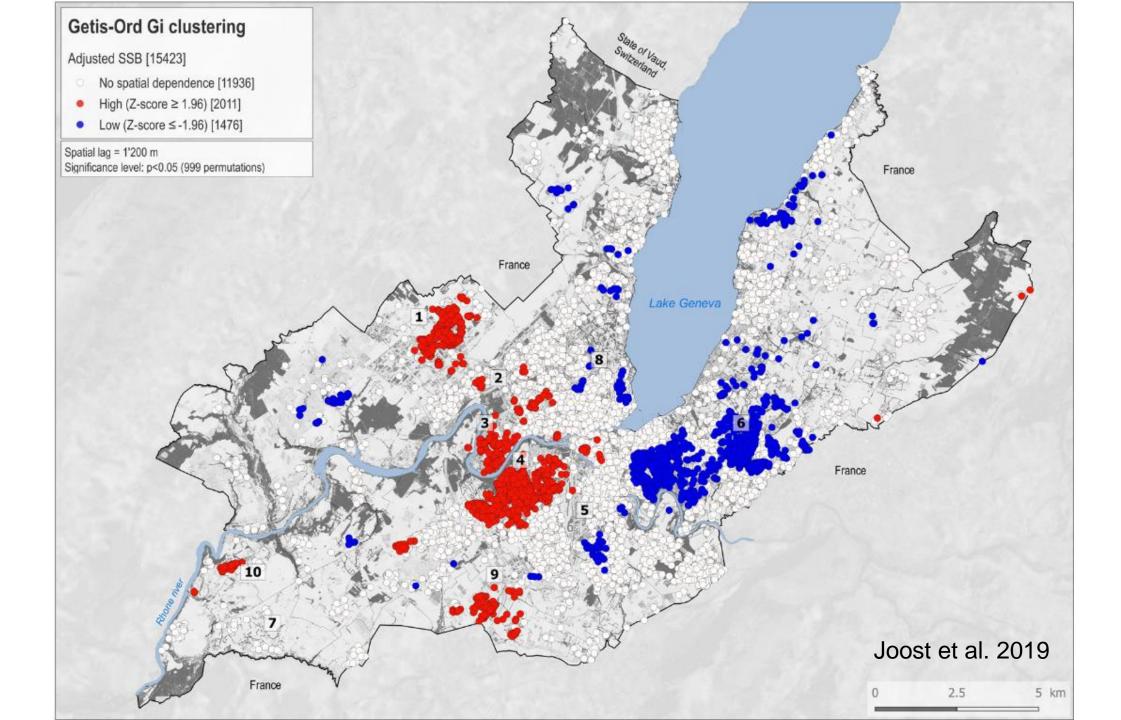
- Education level, sex, age, nationality, and the neighborhood level median income of the area
- Education level was dichotomized as having tertiary education or not
- Age was defined as a continuous variable
- Nationality was dichotomized as having Swiss nationality or not
- We used income data characterizing the 475 Geneva statistical sectors in 2009 (Statistique Genève)
- The yearly income value was attributed to Bus santé participants based on their postal address within the corresponding statistical sector

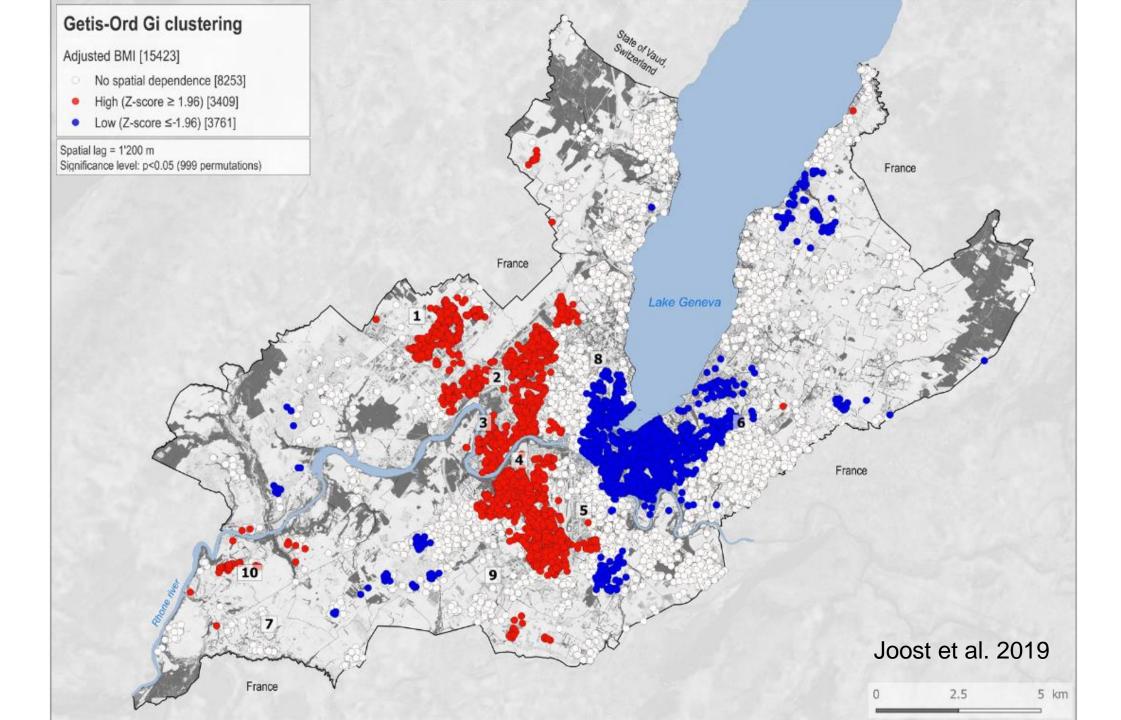
Table 1 Summary characteristics, 1995–2014 Bus Santé study participants (n = 15,423)

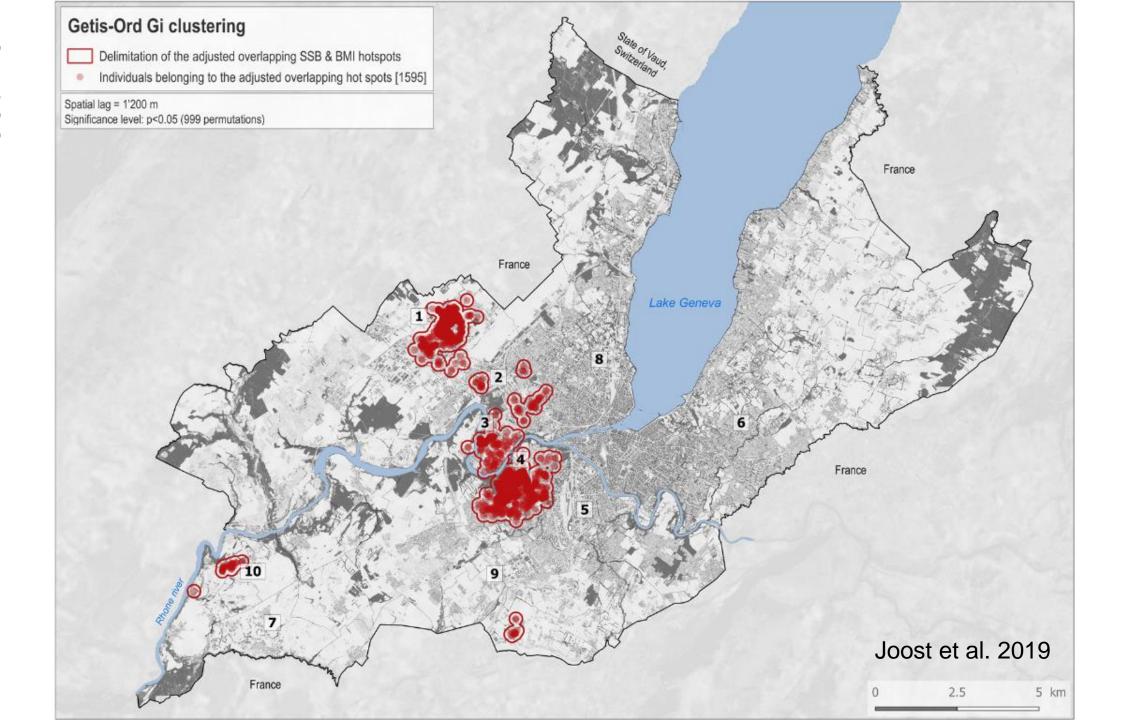
Variable	n (%)	Mean (SD)
Gender		
Men	7713 (50)	-
Women	7710 (50)	_
Age (years)	15,423 (100)	52.3 (11.0)
Neighborhood-level median income (CHF)		(19961.4)
Education		
Tertiary	5820 (37.8)	-
Others	9603 (62.2)	_
Nationality		
Swiss	10,883 (70.5)	_
Others	4540 (29.5)	-
Body mass index (kg/m²)	15,423 (100)	24.9 (4.0)
Sugar-sweetened beverage intake (SSB per day)	15,423 (100)	0.2 (0.5)

Methods

- Using the geographical coordinates of the place of residence, we used the Getis-Ord Gi statistic to investigate whether SSB-IF and BMI were spatially dependent
- To determine whether SSB-IF, BMI, and their spatial dependence were stable during the 1995–2014 period, the dataset was divided into 3 subperiods
- Subperiod 1 (P1) = 1995–2001 (n = 5511), subperiod 2 (P2) = 2002–2008 (n = 4714), and subperiod 3 (P3) = 2009–2014 (n = 5357)
- We then conducted a Tukey's multiple comparison analysis to ensure that the mean of the SSB-IF and the BMI had not increased or decreased sharply between the three subperiods
- Finally, we calculated global Moran's I statistics to verify that there was no difference in global spatial autocorrelation between the three subperiods







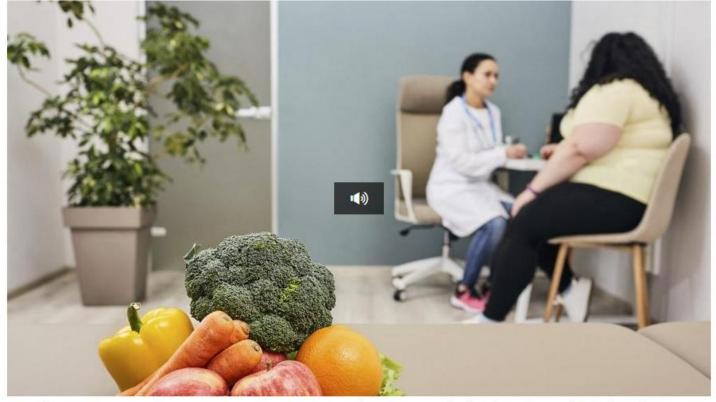
Main result and take-home message

- Spatial overlap between high SSB-IF and high BMI clusters was identified and included 11.1% (n = 1719) of the participants
- After adjustment for education level, gender, age, nationality, and the median income of the area, the overlap between high SSB-IF and high BMI clusters included 10.3% (n = 1595) of the participants
- The identification of specific areas presenting higher SSB-IF associated with higher BMI values enables local legislators and public health experts to develop targeted interventions
- It paves the way for precision public health delivery
- The allocation of resources to these populations in **high need of intervention** could **improve the efficiency of local programs** and potentially **diminish resistance against SSB taxation**.

Genève Publié le 1 février 2023 à 14:28

≪ Partager

Le Canton de Genève pourrait améliorer la prévention contre le surpoids



A Genève, la Cour des Comptes se penche sur la politique de prévention de l'obésité (1) / La Matinale / 1 min. / le 1 février 2023

Près de 42% des adultes et 15% des enfants sont en surpoids ou obèses en Suisse. Ces chiffres sont en forte augmentation par rapport au début du siècle. Un constat qui a amené la Cour des Comptes de Genève à se pencher sur la qualité de la politique de prévention du Canton.

Les autorités genevoises vont dans la bonne direction en matière de prévention, estime l'organe d'audit qui insiste toutefois sur plusieurs points d'amélioration.

Les pouvoirs publics sont par exemple appelés à mieux cibler les personnes concernées et à faciliter l'accès à une alimentation saine, ainsi qu'à une activité physique.

"La population qualifiée de 'vulnérable', c'est-à-dire les personnes en situation économique défavorisée, est la plus touchée par le surpoids et l'obésité", relève mercredi dans La Matinale Isabelle Terrier, magistrate de la Cour des comptes. "Les programmes ne parviennent que peu à atteindre ce public", constate-t-elle. Elle souligne par ailleurs l'importance de l'activité physique, un besoin mis en évidence par la pandémie.

Les conclusions de la Cour des comptes genevoise vont d'ailleurs dans le même sens que les constats de son homologue vaudoise, qui s'est penchée sur la question il y a quelques mois.

Changement de paradigme demandé

De manière générale, la problématique du surpoids et de l'obésité prend de l'ampleur et beaucoup demandent une meilleure prévention pour y faire face.

Interrogée mercredi dans La Matinale, Myriam Pasche, la responsable du département Promotion de la santé et préventions à Unisanté à Lausanne, appelle de son côté au "changement de paradigme". La méthode privilégiée en Suisse ces trois dernières décennies, informer et sensibiliser, ne suffit pas.

Selon l'infirmière de formation et licenciée de sciences sociales, la focalisation sur la responsabilité individuelle est lacunaire. La responsabilité sociétale et les environnements de vie doivent également être pris en compte.

"Social responsibility and living environments must also be taken into account" (versus the "individual responsibility" usually invoked by a category of politicians)

"Some environments are obesogenic. Because of the uneven distribution of socio-economic and environmental determinants, access to a healthy diet and sufficient physical activity differs between social groups".... and this must be taken into account!

Créer des environnements moins obésogènes

"Des environnements sont obésogènes. Des déterminants socio-économiques et environnementaux font que l'accès à une alimentation saine et à une activité physique suffisante est différent selon les groupes sociaux", explique-t-elle.

Myriam Pasche plaide donc en faveur de la création d'environnements favorables à la santé, par exemple la taxation de certains produits alimentaires ou la limitation du marketing pour des produits particulièrement sucrés ou gras. Ce sont des "mesures qui fonctionnent", assure-t-elle, ajoutant qu'elles doivent faire l'objet d'un débat en Suisse.

>> A lire aussi : Une étude menée à Genève fait le lien entre obésité et boissons sucrées

Mais elle tient aussi à "nuancer l'échec" de la politique contre le surpoids en Suisse, car celui-ci, après avoir augmenté en trois décennies, stagne depuis 2017.

References

- Diez-Roux AV. Multilevel analysis in public health research. Annu Rev Public Health. 2000;21:171-92. doi: 10.1146/annurev.publhealth.21.1.171. PMID: 10884951
- Gatchel RJ, Peng YB, Peters ML et al. (2007) The biopsychosocial approach to chronic pain: Scientific advances and future directions, Psychological Bulletin, Vol 133(4), 581-624
- Labonté R. Globalization and Health. International Encyclopedia of the Social & Behavioral Sciences. 2015:198–205. doi: 10.1016/B978-0-08-097086-8.14022-X. Epub 2015 Mar 12. PMCID: PMC7152238
- Lawson AB, Banerjee S, Haining R, Ugarte MD (2016) Handbook of Spatial Epidemiology, Ed. Chapman & Hall/CRC Press, Taylor & Francis
- Patel, C. J. and J. P. Ioannidis. (2014). Studying the elusive environment in large scale. JAMA 311(21): 2173–2174.
- Schwartz, S., E. Susser, and M. Susser. (1999). A future for epidemiology? Annu Rev Public Health 20: 15–33.

