International Journal of Health Geographics

RESEARCH

Open Access

Beyond the map: evidencing the spatial dimension of health inequalities

Fayet Yohan[1,2*], Praud Delphine[3,4], Fervers Béatrice[3,4], Ray-Coquard Isabelle[1,2], Blay Jean-Yves[5], Ducimetiere Françoise[1], Fagherazzi Guy[6,7] and Faure Elodie[7,8]

Abstract

Background: Spatial inequalities in health result from different exposures to health risk factors according to the features of geographical contexts, in terms of physical environment, social deprivation, and health care accessibility. Using a common geographical referential, which combines indices measuring these contextual features, could improve the comparability of studies and the understanding of the spatial dimension of health inequalities.

Methods: We developed the Geographical Classification for Health studies (GeoClasH) to distinguish French municipalities according to their ability to influence health outcomes. Ten contextual scores measuring physical and social environment as well as spatial accessibility of health care have been computed and combined to classify French municipalities through a K-means clustering. Age-standardized mortality rates according to the clusters of this classification have been calculated to assess its effectiveness.

Results: Significant lower mortality rates compared to the mainland France population were found in the Wealthy Metropolitan Areas (SMR = 0.868, 95% CI 0.863–0.873) and in the Residential Outskirts (SMR = 0.971, 95% CI 0.964–0.978), while significant excess mortality were found for Precarious Population Districts (SMR = 1.037, 95% CI 1.035–1.039), Agricultural and Industrial Plains (SMR = 1.066, 95% CI 1.063–1.070) and Rural Margins (SMR = 1.042, 95% CI 1.037–1.047).

Conclusions: Our results evidence the comprehensive contribution of the geographical context in the constitution of health inequalities. To our knowledge, GeoClasH is the first nationwide classification that combines social, environmental and health care access scores at the municipality scale. It can therefore be used as a proxy to assess the geographical context of the individuals in public health studies.

Keywords: Health inequalities, Environment, Social deprivation, Health care access, Geography, Public health, France, GIS

Background

Health inequalities have emerged as a political challenge in many developed countries [1] but few interventions have been sufficient to address this issue, the scope of which extends well beyond public health [2]. While social inequalities have been explored comparing health outcomes according to the social characteristics of the individuals [3], maps [4] and contextual studies [5] have raised the question of the spatial dimension of health inequalities. For example, standardized mortality rates in France were two times higher in Northern counties compared to the South-East in the early 2000's [6]. At a finer spatial scale, maps have also reported better health outcomes in urban areas than in the surrounding rural areas as well as other inequalities according to neighborhoods within urban areas [7]. The breadth of spatial inequalities in health and the continuity of trends across

*Correspondence: yohan.fayet@lyon.unicancer.fr
[1] Equipe EMS – Département de Sciences Humaines et Sociales, Centre Léon Bérard, 28 rue Laennec, 69008 Lyon, France
Full list of author information is available at the end of the article

time show that these inequalities do not occur by chance. We assume they reflect the unequal distribution of health risk according to the characteristics of the geographical contexts.

Combined effects of local health care accessibility, environmental and social characteristics lead to spatial inequalities in health

Three types of contextual risk factors involved in a combined manner in the constitution of spatial inequalities in health can be outlined from the literature:

- Physical environments: Studies considering physical environment generally focused on exposure to a single environmental exposure such as air pollution [8], noise [9], water contaminants [10], ultraviolet radiation [11] or green spaces [12]. Moreover, environment contributes also to healthy behaviours, raising the question of urban planning [13] and food environment [14].
- Social characteristics of the population: This dimension is well-known thanks to deprivation indices developed at a small area level by social epidemiology, such as in the United Kingdom [15, 16] or in France [17, 18]. Multilevel analyses in health research in the 1990s found deprivation to impact on health outcomes, whilst taking into account individual characteristics [19]. This demonstrates that social context must be also considered in the analysis of spatial inequalities in health.
- Spatial accessibility to health care: Studies demonstrated a lower health care use in lower medical density areas [20] and for remote patients [21]. Health care access may improve prevention and early diagnosis of diseases but also follow-up and management [22]. Assessment of health care accessibility must also take into account the quality of care according to the facilities, their size and level of expertise. In the case of cancer care for example, remote patients having a lower accessibility to expert centers, suffer from poorer health care and consequently poorer survival rates [23].

These determinants of health suggest the geographic contexts' potential to influence health status and to produce health inequalities. However, estimating the impact of the geographic context on health inequalities remains difficult, partly because epidemiological studies aiming to identify associations between spatial characteristics and health outcomes most often investigate one spatial factor at a time, according to the objectives of the study. As a result, geographic contexts may be variously measured, in terms of characteristics and spatial scale. In addition, spatial indices may be combined differently across studies, due to methodological choices and study objectives. Consequently, studies analyzing geographical inequalities in health are based on different geographical frameworks limiting the comparability of the results [24]. Furthermore, the separated analysis of risk factors in epidemiological studies mostly impedes a comprehensive review of all the vulnerabilities related to the place of residence. Some epidemiological studies using a social deprivation index took this limit into account, investigating a potential difference in their analysis between rural and urban deprived areas [25]. Considering this challenge, developing a geographical classification is needed to assess these combined effects of geographical determinants on health and to develop a common geographical frame of reference in health studies. This paper aims to demonstrate the value of this new comprehensive classification approach to address the geographical context's contribution into health inequalities. The "Geographical Classification for Health studies" (GeoClasH) was computed from spatial data measuring physical environment, social deprivation and health care accessibility, using the example of mainland France.

Methods

We used spatial data available for the 35,798 mainland France municipalities regarding environmental factors, social environment and spatial accessibility of the health care. Ten geographical scores have been computed from these data and combined to classify French municipalities through K-means clustering. As a gold standard to study health inequalities, age-standardized all-causes Mortality Rates (SMR) have been calculated to compare the general health status according to the clusters of the GeoClasH classification.

Geographical data and scores calculation

The year of reference of the geographical data used in the classification was 2014 corresponding to the date of the last detailed French census data, provided by the National Institute for Statistics and Economic Studies (INSEE). Our study deals only with mainland France because little data were available for French overseas territories. We used the GeoFLa municipalities basemap from the National Geographic Institute (IGN) for 2016, which comprises 35,798 municipalities. For all spatial data, the projection used was the Lambert 93 being the current projection system in effect in France. As this study aims to evidence the spatial dimension of health inequalities at a fine scale, only variables measuring the contextual features and exhaustively available at the municipality scale in mainland France were selected (Table 1).

Table 1 Contextual features involved in spatial inequalities in health and variables included in the design of the GeoClasH classification

Contextual features	Variables	Time	Data sources
Physical environment	Population density	2014	IGN-INSEE
	Air quality score	2009–2013	INERIS
	Pesticides' risk exposure	2012	Corine Land Cover
Social characteristics of the population	Median income per household people	2014	INSEE
	Unemployment rate	2014	INSEE
	Percentage of labourers and employees	2014	INSEE
	Percentage of high school graduates	2014	INSEE
	Percentage of unattached individuals	2014	INSEE
Spatial accessibility to health care	Average journey time to hospital	2014	SAE-Odomatrix
	Spatial accessibility to primary care practitioners	2013	DREES

Every variables included in the calculation of the GeoClasH classification was produced at the municipality scale

Variables considered but unavailable at the municipality scale: water quality; noise; ultraviolet radiation; green spaces; food, alcohol and tobacco retail environments

For confidentiality reasons, because of their low population some data were missing in some low-populated municipalities. Low population may also trigger extreme values (in the rates calculation for example) which may lead to a bias in the multivariate analysis used to design the classification process if there were used in the multivariate analysis. To overcome this issue, variables with missing values at the municipality scale were also calculated at the "canton" scale. A canton in France is a continuous electoral district essentially defined on demographic bases and its scale is slightly larger than the Commune scale. There were 1989 cantons across the mainland France and 35,789 municipalities in 2014. The canton value was assigned to municipalities with missing data or with a reference population of less than 50 inhabitants.

Environmental scores

Population density The population density was calculated for 2014 using the surface area in km^2 of municipalities from the IGN's GEOFLA database, and the 2014 number of inhabitants per municipalities from the population censuses provided by INSEE. Thus, for each municipality, a population density per km^2 has been computed.

Air quality Data provided by the French National Institute for Industrial Environment and Risks (INERIS) on particulate matters (PM_{10} and $PM_{2.5}$) and nitrogen dioxide (NO_2) were used to estimate the average exposure to each pollutant for each municipality between 2009 and 2013, in accordance with methods reported by Pascal et al. [26]. To elaborate the synthetic air quality score over the period 2009–2013, we considered the different spatial distribution of particulate matters in comparison of NO_2 which are more related to road traffic (Additional file 2: Fig. S2). We first computed for each municipality the mean value of $PM_{2.5}$ and PM_{10} concentration in $\mu g/m^3$ as $PM = mean(PM_{2.5};PM_{10})$. Then, the air quality score was calculated, for each municipality, as the mean of PM and NO_2 as follows: $Air\ quality\ score = mean(NO_2;PM)$.

Risk of pesticides' exposure Due to the lack of data on pesticides applications and types used in France, we used the 2012 Corine Land Cover nomenclature to assess pesticide exposure level at the municipality scale, using the share of the municipalities' area occupied by crops using pesticides (on-irrigated arable lands, vineyards and orchards) as a proxy for exposure [27].

Social environment scores

We considered socio-economic variables used in the most sensitive deprivation index developed for France [28] and published by Rey et al. [18]:

- the median income per household people
- the unemployment rate in the population aged 15 to 64 years
- the percentage of labourers and employees in the population aged 15 to 64 years
- the percentage of high school graduates in the population aged 15 years and older
- the percentage of unattached individuals in the population aged 15 years and older.

Health care access scores

Spatial accessibility to primary care practitioners We used the reference indices to assess the spatial accessibility of healthcare practitioners in France [29, 30]: the APL

(*Accessibilité Potentielle Localisée*, localized potential accessibility) indicators measuring spatial accessibility to primary care practitioners (general practitioners, nurses, chemists, gynecologists, midwives, physiotherapists, dentists, ophthalmologists, pediatricians and psychologists). We computed a synthetic score of the primary care accessibility at the municipality scale. Considering the importance of general practitioners in the primary care system, their APL indices count for half in the calculation of the final score, measuring the accessibility to primary care. We used principal component analysis (PCA) to summarize the other primary care practitioners' indices and selected the first component (accounting for 45% of the variance in the dataset) of the PCA as the score of these practitioners' accessibility. The final score was the mean value of both accessibility scores of general practitioners and other primary care practitioners.

Spatial accessibility to hospitals French municipalities hosting a hospital (n = 536 municipalities) were identified using the Annual Statistics of the Facilities (SAE) database (year 2014), which is managed by the French ministry of health. To measure the municipalities' remoteness to hospitals, journey times by car to the closest municipality hosting a hospital were estimated using the Odomatrix® software (Source: ODOMATRIX, INRA UMR1041 CESAER, from IGN Route500®).

Preliminary analysis of the dataset and clustering process of the French municipalities

A correlation matrix was performed using the R software (version 3.5, package "Corrplot") to study potential correlation between the selected variables and to justify the relevance of each variable to be included in the clustering process. Several methodologies are used to develop multivariate classification of areas [31]. Hierarchical clustering and partitioning methods have their limitations but hierarchical clustering is known to be more computationally demanding and less efficient than partitioning methods with large datasets [32]. Considering our large dataset (ten variables, 35,798 municipalities), we selected the K-means partitioning algorithm using the R software (version 3.5, packages "NbClust") to perform the classification. Geographical variables were standardized using the R's "scale" function, before incorporating them into the clustering process. We performed several K-means partitioning from the same dataset in order to ensure the clustering process' stability and the obtained classification was always the same. While the classification aims to allow health outcomes' comparison across geographical contexts, we were careful not to deal with too heterogeneous clusters considering their population size in order to be statistically comparable. We chose therefore the five-cluster partitioning because the six-cluster partitioning splits the RM (Rural Margins) cluster in two parts whereas the population size of this cluster was already the lowest.

Calculation of the age-standardized mortality rates according to the GeoClasH clusters

In order to assess the relevance of the classification to display health inequalities, the age-standardized mortality ratio (SMR) and their 95% confidence interval (CIs) was calculated in each cluster, as the ratio of observed deaths and the expected deaths in each cluster based on the distribution of mainland France, using the indirect method. The expected deaths were based on age-specific deaths in the mainland France. We obtained data on overall population and mortality from the INSEE databases, for the mainland France at the municipality scale, in 5-year age groups (i.e. 0–4, 5–9… 80–84, and ≥ 85 years) from 2011 to 2015 [33]. Data were not available at the municipality scale to standardize also on sex distribution.

We interpreted the SMR using their CI; we considered that observed mortality rates were significantly different to expected mortality rates (based on mainland France rates) when SMRs were significantly greater than 1. We considered SMRs as significantly different between each other when their CIs did not overlap.

Results

The maps of the ten geographical scores included in the GeoClasH classification are provided in Additional files to describe and improve the understanding of these scores, giving an overview of spatial inequalities in France (see Additional files 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10). The correlation matrix (Fig. 1) reports some moderated correlations for population density with air pollution (cor = 0.7) and access to practitioners (cor = 0.66), as well as between income and people without higher education diploma (cor = 0.63). But we consider these correlations are not strong enough to exclude one of these variables, which each have their own relevance and interest with regards to literature.

Average values (Table 2) and z-score means (Fig. 2) for each of the geographical variables used in the classification process have been calculated and compared to characterize the clusters. The geographical distribution of the clusters (Fig. 3) was considered and used to name them.

Wealthy Metropolitan Areas (WMA) stand out from the other clusters for their higher population density,

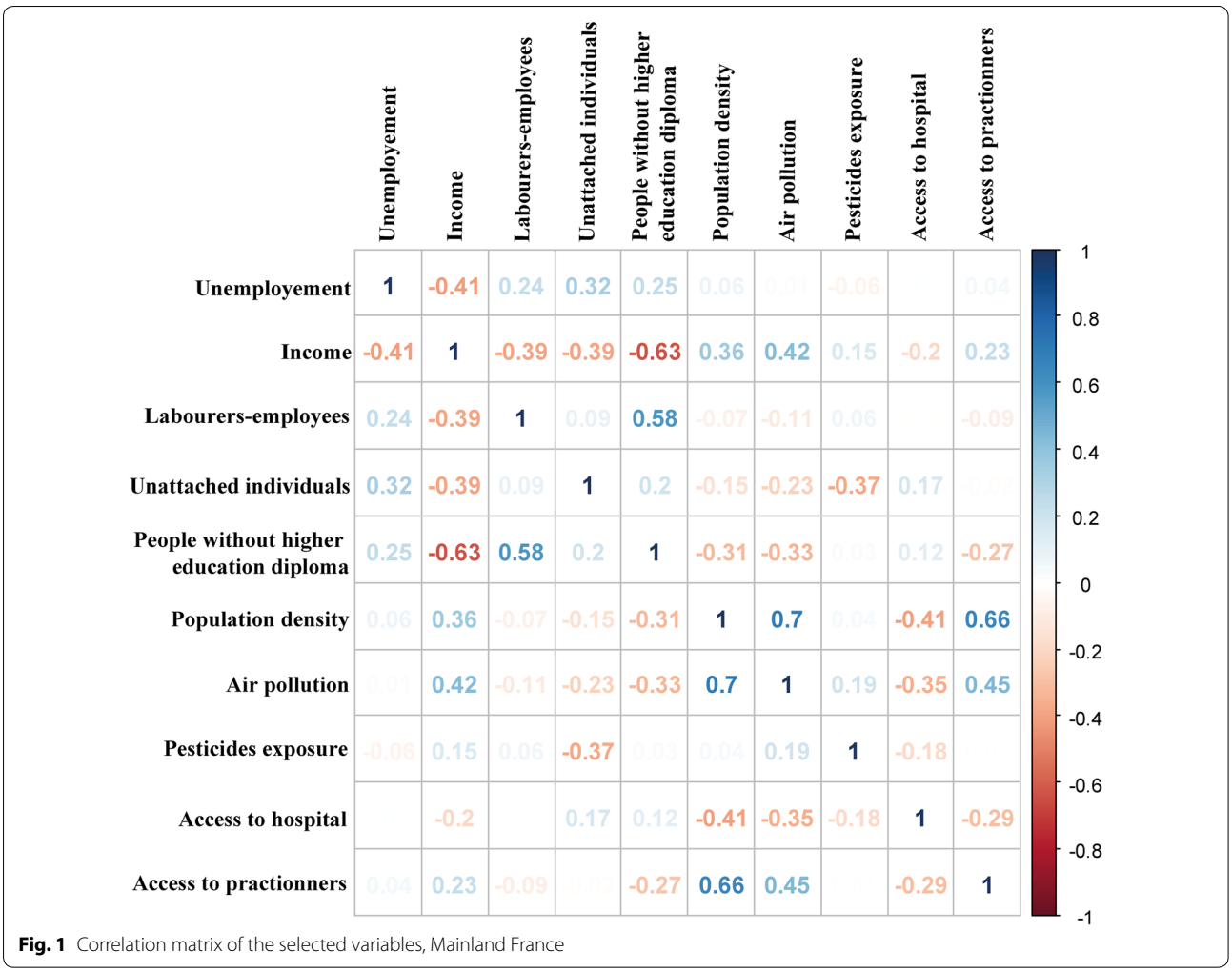

Fig. 1 Correlation matrix of the selected variables, Mainland France

very low social deprivation and good spatial accessibility to health care. This cluster has the most pronounced profile (Fig. 2) and includes municipalities hosting the wealthiest populations of major French urban areas, who benefited from their economic dynamism but are also exposed to their drawbacks (air pollution). Depending on local configurations, these municipalities may not be located in the center of urban areas but rather in the near periphery. Precarious Population Districts (PPD) are densely populated municipalities with the best spatial accessibility to health care. Contrary to the WMA, most of the social deprivation indices report the social vulnerabilities of the PPD population. Lower access to practitioners in the WMA in comparison to PPD can be linked to the higher cost of life and housing in these areas that can dissuade the establishments of practitioners. The PPD cluster is the biggest cluster in terms of population (more than 32 million inhabitants) and encompasses some downtowns or working-class suburbs of the biggest cities, previous and current industrial districts and most of the little and medium-sized cities.

Residential Outskirts (RO) are situated around the urban areas and have a rather intermediate profile with regard to spatial accessibility to health care and to the environment, except for exposure to pesticide risk, probably due to the proximity of countryside and agricultural activities. This population is rather privileged compared to the national average. Agricultural and Industrial Plains (AIP) are rural spaces strongly marked by productive activities in terms of population (63% of laborer or employee, 83% of people without higher education diploma) and environmental exposures (risk of pesticides exposure). This cluster includes most of the countryside municipalities in the North of France and some valleys in the South. Rural Margins (RM) are remote spaces far

Table 2 Main characteristics of the five clusters, GeoClasH, Mainland France

Characteristics, mean (SD)	Wealthy Metropolitan Areas (WMA)	Precarious Population Districts (PPD)	Residential Outskirts (RO)	Agricultural and Industrial Plains (AIP)	Rural Margins (RM)	France
Population density (pop/km^2)	813 (2793)	701 (1496)	77 (74)	43 (36)	19 (21)	178 (973)
Exposure to NO_2 (µg/m^3)	17.0 (6.0)	14.2 (5.4)	10.9 (2.8)	9.5 (2.3)	6.5 (2.2)	10.3 (4.4)
Exposure to PM (µg/m^3)	19.4 (1.9)	18.6 (1.7)	18.2 (1.2)	18.0 (1.2)	17.1 (0.9)	18.0 (1.4)
Share of the area occupied by crops using pesticides (%)	34.0 (28.4)	23.3 (22.9)	45.7 (29.5)	52.0 (26.0)	10.0 (15.5)	35.6 (30.0)
Unemployment rate (%)	6.2 (1.7)	11.2 (2.9)	6.3 (1.9)	8.8 (2.6)	8.5 (3.3)	8.0 (3.0)
Annual income per household (€)	26,050 (3372)	19,162 (1750)	21,733 (1794)	19,270 (1376)	18,543 (1722)	20,335 (2808)
Labourers-employees (%)	38.4 (8.6)	59.1 (7.5)	50.0 (9.0)	63.2 (8.5)	52.5 (10.8)	54.4 (11.7)
Unattached individuals (%)	11.7 (4.3)	18.2 (5.1)	11.1 (3.0)	13.7 (3.5)	18.5 (4.7)	14.4 (5.3)
People without higher education diploma (%)	63.0 (8.1)	79.9 (5.8)	74.7 (5.3)	83.4 (4.3)	80.2 (6.2)	78.1 (7.9)
Journey time to hospital (min)	18.5 (8.5)	16.5 (10.4)	21.6 (8.3)	22.4 (8.2)	33.5 (14.2)	23.9 (11.7)
APL score to general practitioners	57.7 (22.9)	84.4 (33.4)	45.7 (26.2)	41.7 (24.7)	31.5 (31.9)	46.1 (31.6)
Number of municipalities	2920	3742	10159	10330	8647	35798
Overall population	14,435,670	32,023,919	8,717,397	5,902,766	2,948,094	64,027,846

SD: Standard deviation, NO_2: Nitrogen dioxide, PM: particulate matter, APL: localized potential accessibility

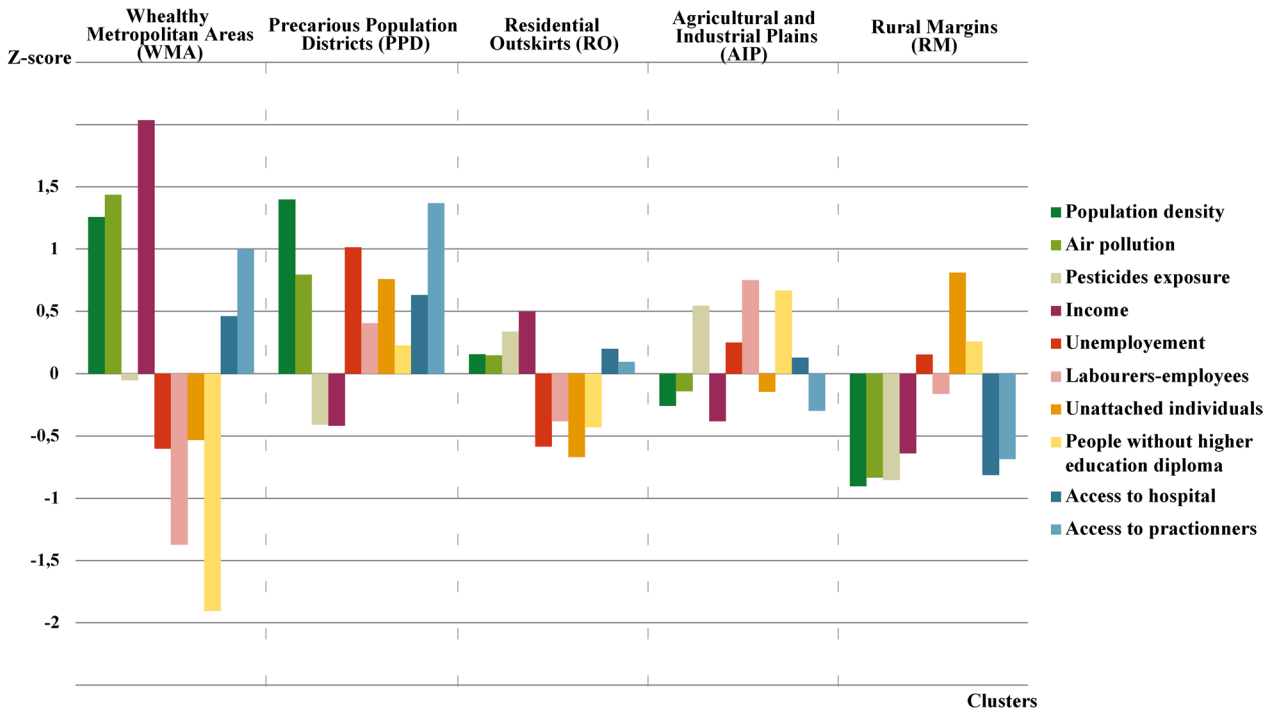

Fig. 2 Mean z-score for each of the variables within the clusters, GeoClasH, Mainland France

from urbans dynamics and stresses. The population of this cluster benefits from the quality of its environment (fewest exposures to air pollution and pesticides) but experiences social deprivation (lowest average income and highest rate of unattached individuals) and difficulties to get access to health care services (worst indices regarding the health care access). Most of these municipalities are situated in relief or mountainous areas.

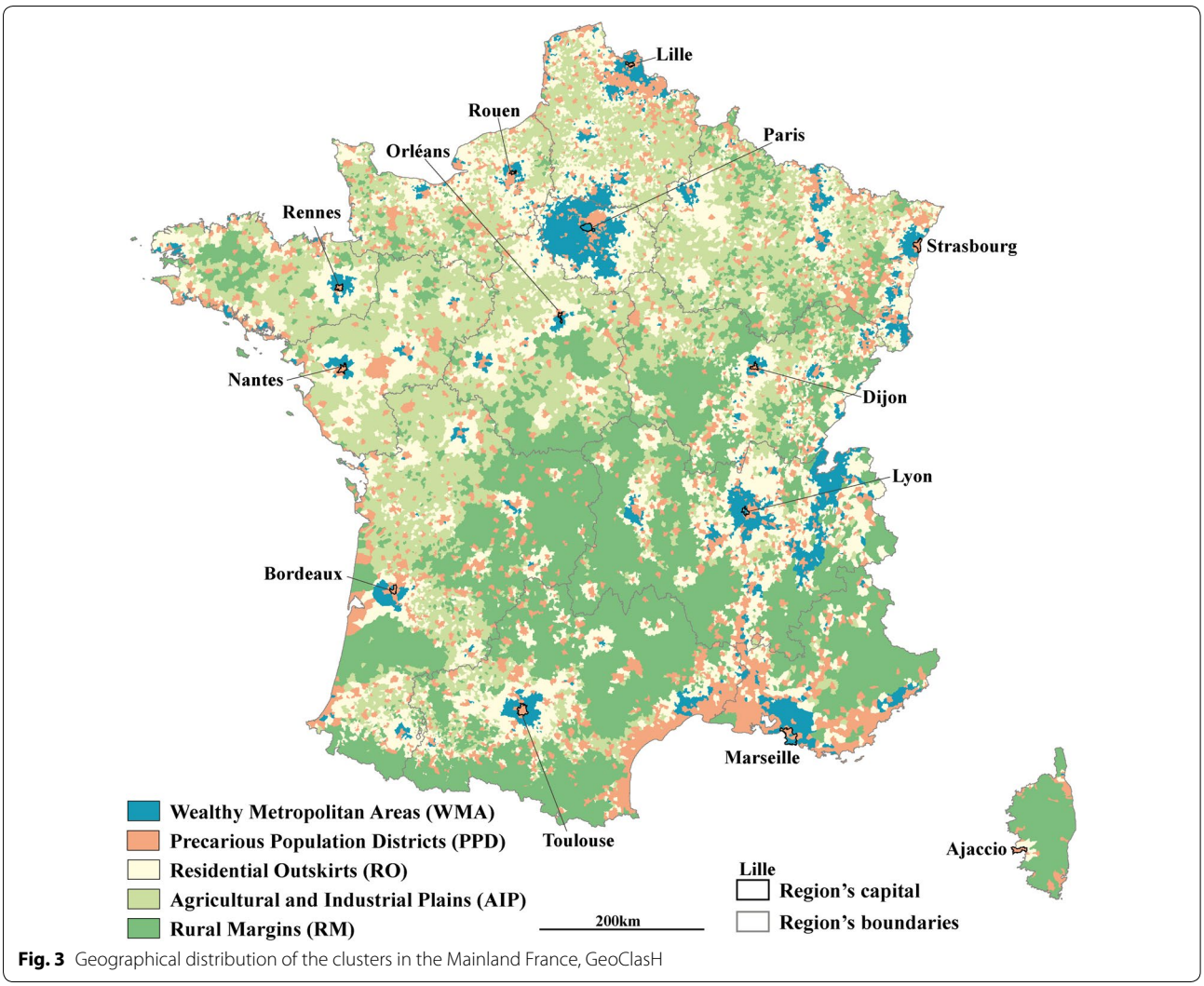

Fig. 3 Geographical distribution of the clusters in the Mainland France, GeoClasH

Table 3 Age-standardized mortality ratios (SMR) according to the GeoClasH clusters, France, 2011–2015

	Population	Observed deaths (rate)	Expected deaths (rate)	SMR (95% CI)
Wealthy Metropolitan Areas (WMA)	14,436,150	95,440 (0.66%)	109,929 (0.76%)	0.868 (0.863–0.873)
Residential Outskirts (RO)	8,716,474	63,206 (0.73%)	65,106 (0.75%)	0.971 (0.964–0.978)
Precarious Population Districts (PPD)	32,024,097	300,523 (0.94%)	289,795 (0.90%)	1.037 (1.033–1.041)
Rural Margins (RM)	2,947,557	35,969 (1.22%)	34,525 (1.17%)	1.042 (1.031–1.053)
Agricultural and Industrial Plains (AIP)	5,903,512	59,044 (1.00%)	55,365 (0.94%)	1.066 (1.058–1.075)

Table 3 gives observed and expected numbers of deaths and the corresponding SMRs with their 95% confidence intervals for each cluster. The number of deaths was significantly lower than expected in the WMA (95,440 vs 109,929, SMR = 0.868, 95% CI 0.863–0.873) and in the RO (63,206 vs 65,106, SMR = 0.971, 95% CI 0.964–0.978) compared to the mainland France population. Significant excess mortality were found for PPD (SMR = 1.037, 95% CI 1.035–1.039),

RM (SMR = 1.042, 95% CI 1.037–1.047) and AIP (SMR = 1.066, 95% CI 1.063–1.070) compared to the mainland France population.

Discussion

This paper shows the potential of geographical context to influence health outcomes and its significant involvement in the constitution of health inequalities. As the Marmot Review recommended to fully integrate planning, transport, housing, environmental and health systems to address the social determinants of health in each locality [1], the GeoClasH classification supports a comprehensive understanding of the spatial dimension of health inequalities. Based on ten scores measuring local health care accessibility, environmental and social characteristics, the GeoClasH classification is, to our knowledge, the first nationwide geographical classification designed for health studies at the municipality scale combining these different contextual features impacting health outcomes. Indeed, geographical indicators combining several features are usually calculated as a continuous variable [34] and often include health outcomes variables in their calculation. As an example, the Index of Multiple Deprivation (IMD) includes an "Health Deprivation and Disability Domain", based on morbidity, disability and premature mortality measures, into its score calculation [24, 35]. Combining health outcomes and spatial data in the calculation of the index can aid policy-makers identify the most deprived areas in need of priority public health interventions [36, 37] while the experimental GeoClash classification has the specific objective to assess the contribution of geographical context in the constitution of health inequalities. The classification allows the comparison of populations that are heterogeneously exposed to health risks because of their geographical context.

Using the clustering process is relevant to distinguish and address areas of distinct characters, as specific combination of geographical features, since previous studies have already shown it [38–40]. This multidimensional and combined approach is expected to have a strong beneficial impact on future epidemiological studies to estimate the environmental context of the individuals in the larger sense of the term. Completing social deprivation indices, the GeoClasH classification could therefore be used to provide a simple and comprehensive assessment of the geographical context of the studied subjects and to document health inequalities. It can also be useful for multi-level analyses, for adjusting regression models or leading subgroup analyses. The clustering process used avoids the current limitations of spatial analysis at a fine geographic scale, like the modifiable area unit problem [41] since municipalities are aggregated according to their own characteristics and not according to their surface area or proximity. The classification can be used for recent years but can also be valid for last decades due to small changes and constant trends in the French geographical area regarding variables included in the classification. Our methodology can be potentially reproducible in other countries and scalable with different number of clusters, in accordance with the dataset. As an example, air pollution data used in this study can be obtained in other European countries as well as the data used to estimate the probability of pesticide use [42].

Our results report significant health inequalities according to the GeoClasH clusters and especially the positive situation of Wealthy Metropolitan Areas (WMA) in comparison to mainland France, with a 15.1% lower mortality. These results are particularly remarkable because the GeoClasH classification was only designed on theoretical basis to demonstrate the complex constitution of spatial inequalities in health which result from the combining effects of local health care accessibility, environmental and social characteristics. Indeed, the ability of the GeoClasH classification to evidence the real extents of spatial inequalities in mortality in France could have been much better if we could have overtaken some limitations. First, the GeoClasH classification considers each geographical variable equally while each one probably does not have the same impact on health outcomes. Weighting the contribution of each variable in the clustering process according to their real influence on mortality can help to improve the quality of the classification for example. But this information is currently not available in the literature. Studies that would split the contribution of each major geographical variable may be helpful to improve our understanding of the contextual potential to influence health and the precision of geographical classification used in health studies. Considering the geographical scale, we may have chosen the IRIS (an acronym of 'aggregated units for statistical information') scale, the infra-municipal scale produced by the INSEE. But only social data are available at this scale. An IRIS scale classification would perhaps have been more sensitive to social disparities inside urban areas, and deprivation indices in France are already developed to precisely study social inequalities inside the urban areas [17, 18]. Finally, protection of individuals' privacy supported restrains the use of IRIS scale in health studies because collection of identifying data as patients' address is mostly limited by the independent data protection authority in France ("Commission nationale de l'informatique et des libertés", https://www.cnil.fr/en/home). We have therefore opted for the municipal level to protect the balance of the GeoClasH classification and promote its compatibility and its implementation with most health studies.

Among the variables considered, only data with good temporal consistency and with a sufficiently precise scale has been used. Therefore, water quality data have not been integrated because of the low number of sampling in sparsely populated areas. For example, weekly biological measurements were made in some densely populated areas while just one value in 5 years was available for sparsely populated areas. The low number of samplings raises question about the representativeness of these few measurements and could also introduce a potential bias between urban and countryside municipalities. We only had annual data on PM and NO_2 concentrations and it would be interesting to use hourly data to consider the number of threshold exceedances that are known to have short term effects on health [43]. Corine Land Cover data used to estimate environmental exposures to pesticide were not initially designed to this purpose even if several studies recognized Corine Land Cover data as the most suitable database available for national scale studies [27, 44]. Crop density by municipality has been used in ecological studies as a surrogate for residential exposure to agricultural pesticides [45]. Each plot (25 ha) has been characterized by its most extensive crop that may lead to not considering small specific acreage within the plot. Data to provide a nationwide measure of food, alcohol and tobacco retail environments are not currently available in France [46]. Similarly, we do not currently have consistent, reliable and comprehensive data in France to measure spatial coverage dedicated to greenspaces [12, 13]. As developing countries can lack from freely available data, the use of data from satellite images can be considered as interesting option to measure the physical environments. For example, nighttime lights can be considered as a relevant surrogate of the level of urbanization.

Social deprivation scores are generated from national census data but it does not take into account some local initiatives led by associations, foundations, local policies or management networks dedicated to specific diseases that have been implemented in some deprived areas (especially in urban areas) to improve behaviors and health care accessibility in its multiple dimensions (spatial, cultural, financial, etc.). Other dimensions of the health care spatial accessibility may have been taken into account including the different level of expertise of technical equipment according to the facilities [47]. While telemedicine has been recently implemented, its use and potential to improve spatial accessibility of health care access could be taken into account and regularly updating the GeoClasH model over time will help to evaluate the impact of such major changes in healthcare organization.

Despite these limitations and although the classification was designed a priori, significant mortality inequalities between the clusters have been emphasized. These results highlight the potential of the geographical context to promote health outcomes and its relevance in the study of health inequalities. It also evidences the extent of health inequalities according to geographical contexts, which due to their characteristics, are more or less favorable to health. Despite some environmental exposures whose impact on health outcomes is clearly demonstrated, WMA appear to be the most favorable context to health. This finding underscores the importance of making a strong distinction between studies on health inequalities and determinants to avoid confusion and design specific policies [48]. Studies in health determinants are seeking to evaluate the specific impact of each determinant on health at the individual scale. Identifying the causes of significant differences in health outcomes at the population level is required to analyze the constitution of health inequalities. As epidemiological research and general interest are increasing on environmental exposures and their impacts on health, our results confirm that spatial inequalities in health express a more complex process through the combined effects of many contextual features.

Variable availability of spatial data in different countries can currently limit the identical reproduction of this nationwide classification of the municipalities, even if more data can be available in urban areas. Moreover, this study considering the combined effects of local health care accessibility, environmental and social characteristics promotes a comprehensive approach which can be implemented everywhere to better understand the constitution of these inequalities and to support the design of relevant actions to efficiently reduce it. Nationwide exhaustive data as well as compatible and cross-sectional analyses are needed to improve knowledge of geographical inequalities in health and the understanding of its underlying mechanisms, in order to support evidence-based public policies addressing these major public health issues. Using the same geographical referential in health studies will make possible to compare gaps between health outcomes, and to define which outcomes are more determined by the geographical context. Moreover, it can contribute to identify and target specific public health issues to each geographical context. The GeoClasH classification can be used as a general contextual proxy that will be useful and relevant to assess the geographical context of individuals in public health studies and it support also dedicated research to efficiently address the spatial inequalities in public health.

Conclusions

The Geographical Classification for Health studies (GeoClasH) was developed to demonstrate the value of a new comprehensive classification approach to evidence the spatial dimension of health inequalities. To our knowledge, this is the first nationwide classification that combines social, environmental and health care access scores at the municipality scale. Our results report significant inequalities in age-standardized mortality according to the GeoClasH clusters and especially the positive situation of Wealthy Metropolitan Areas (WMA) in comparison to mainland France. Using the same geographical referential in health studies will make possible to compare inequalities according to the health outcomes and to identify specific public health issues in each type of geographical context. The GeoClash classification can also be used as a proxy to assess the geographical context of the individuals in public health studies.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12942-020-00242-0.

Additional file 1. Population density in the mainland France municipalities.

Additional file 2. Air pollution in the mainland France municipalities.

Additional file 3. Risk of pesticides exposure in the mainland France municipalities.

Additional file 4. Average income in the mainland France municipalities.

Additional file 5. Unemployment in the mainland France municipalities.

Additional file 6. Lower occupations in the mainland France municipalities.

Additional file 7. Unattached individuals in the mainland France municipalities.

Additional file 8. Level of education in the mainland France municipalities.

Additional file 9. Spatial accessibility to the hospitals in the mainland France municipalities.

Additional file 10. Spatial accessibility to the primary care practitioners in the mainland France municipalities.

Abbreviations

AIP: Agricultural and Industrial Plains; APL: Accessibilité Potentielle Localisée (localized potential accessibility; GeoClasH: Geographical Classification for Health studies; IGN: Institut National Géographique (National Geographic Institute); INERIS: Institut National de l'Environnement industriel et des RISques (National Institute for Industrial Environment and Risks); INSEE: Institut National de la Statistique et des Etudes Economiques (National Institute of Statistics and Economic Studies); IRIS: Îlots Regroupés pour l'Information Statistique (aggregated units for statistical information); NO_2: Nitrogen Dioxide; PCA: Principal Component Analysis; PM: Particulate Matters; PPD: Precarious Population Districts; RM: Rural Margins; RO: Residential Outskirts; SAE: Statistique Annuelle des Etablissements (Annual Statistics of the Facilities); SMR: Standardized all-causes Mortality Rates; WMA: Wealthy Metropolitan Areas.

Acknowledgements

We would like to thank Laure Malherbe and Laurent Letinois (INERIS) for its expertise on the CYSCLAD project and the data processing, Veronique Lucas-Gabrielli and Charlène Le Neindre (IRDES) for their advice on the measure of health care accessibility in France, Thomas Coudon and Lény Grassot (Centre Léon Bérard) for their advice on air quality measure and statistical analysis, Conor Macdonald (Center of Epidemiology and Population Health) for the English review, as well as Marc Billaud (Centre Léon Bérard) for his support.

Authors' contributions

YF conceptualized the study. He led the project administration and the funding acquisition with the support of IRC and FD. YF and EF worked on the data curation. YF, DP and EF designed the methodology and analyzed data. BF, IRC, JYB and FD provided resources required to perform the study and supervised it. YF wrote the first draft of the paper with the support of DP, GF and EF. All authors contributed to the interpretation of the results and critical revision of the manuscript. All authors read and approved the final manuscript.

Fundings

This research was supported by the ARC Fundation (Grant number: PGA1*20160203865), the French National Cancer Institute (Grant number: SHSESP16-063) and the SIRIC LYriCAN (Grant INCa-DGOS-Inserm_12563).

Availability of data and materials

All data of the geographical scores and the GeoClasH classification presented in this study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1] Equipe EMS – Département de Sciences Humaines et Sociales, Centre Léon Bérard, 28 rue Laennec, 69008 Lyon, France. [2] EA 7425 Health Services and Performance Research, Université de Lyon, Lyon, France. [3] Department Prevention Cancer Environment, Centre Léon Bérard, Lyon, France. [4] Inserm UA 08: Radiations, Défense, Santé, Environnement, Centre Léon Bérard, Lyon, France. [5] Department of Medical Oncology, Centre Léon Bérard, Université Claude Bernard, Lyon, France. [6] Digital Epidemiology and e-Health Research Hub, Department of Population Health, Luxembourg Institute of Health, Strassen, Luxembourg. [7] Center of Epidemiology and Population Health, UMR 1018, Inserm, Paris South, Paris Saclay University, Villejuif, France. [8] Gustave Roussy Institute, Villejuif, France.

Received: 2 July 2020 Accepted: 29 October 2020
Published online: 09 November 2020

References

1. The Marmot review. Fair society, healthy lives: the Marmot review. London, Royaume-Uni de Grande-Bretagne et d'Irlande du Nord; 2010, p. 242.
2. Thomson K, Hillier-Brown F, Todd A, McNamara C, Huijts T, Bambra C. The effects of public health policies on health inequalities in high-income countries: an umbrella review. BMC Public Health. 2018;18(1):869.
3. Marmot M. Social determinants of health inequalities. Lancet. 2005;365(9464):1099–104.
4. Santana P, Costa C, Freitas Â, Stefanik I, Quintal C, BanaCosta C, et al. Atlas of population health in European Union regions. Coimbra: Imprensa da Universidade de Coimbra; 2017. p. 263.
5. Arcaya MC, Tucker-Seeley RD, Kim R, Schnake-Mahl A, So M, Subramanian SV. Research on neighborhood effects on health in the United States: a systematic review of study characteristics. Soc Sci Med. 2016;168:16–29.
6. Trugeon A, Thomas N, Michelot F, Fédération nationale des observatoires régionaux de santé (France). Inégalités socio-sanitaires en France : de la

région au canton. Issy-les-Moulineaux : Masson. 2010. (Abrégés (Paris. 1971), ISSN 0768-1992).
7. Vigneron E, Cartier N (1975). Les inégalités de santé dans les territoires français : état des lieux et voies de progrès. Issy-les-Moulineaux: Elsevier Masson; 2011.
8. Ambient (outdoor) air quality and health, WHO, 2018, available https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health. Accessed 12 Sep 2019.
9. Weltgesundheitsorganisation, Regionalbüro für Europa. Environmental noise guidelines for the European Region. 2018.http://www.euro.who.int/en/publications/abstracts/environmental-noise-guidelines-for-the-european-region-2018. Accessed 13 Mar 2019.
10. Roh T, Lynch CF, Weyer P, Wang K, Kelly KM, Ludewig G. Low-level arsenic exposure from drinking water is associated with prostate cancer in Iowa. Environ Res. 2017;159:338–43.
11. Savoye I, Olsen CM, Whiteman DC, Bijon A, Wald L, Dartois L, et al. Patterns of ultraviolet radiation exposure and skin cancer risk: the E3N-SunExp Study. J Epidemiol. 2018;28(1):27–33.
12. Rojas-Rueda D, Nieuwenhuijsen MJ, Gascon M, Perez-Leon D, Mudu P. Green spaces and mortality: a systematic review and meta-analysis of cohort studies. Lancet Planet Health. 2019;3(11):e469–77.
13. Porcherie M, Lejeune M, Gaudel M, Pommier J, Faure E, Heritage Z, et al. Urban green spaces and cancer: a protocol for a scoping review. BMJ Open. 2018;8(2):e018851.
14. Chaix B, Bean K, Daniel M, Zenk SN, Kestens Y, Charreire H, et al. Associations of supermarket characteristics with weight status and body fat: a multilevel analysis of individuals within supermarkets (RECORD study). PLoS ONE. 2012;7(4):e32908.
15. Carstairs V, Morris R. Deprivation: explaining differences in mortality between Scotland and England and Wales. BMJ. 1989;299(6704):886–9.
16. Townsend P. Deprivation. J Soc Policy. 1987;16(2):125–46.
17. Pornet C, Delpierre C, Dejardin O, Grosclaude P, Launay L, Guittet L, et al. Construction of an adaptable European transnational ecological deprivation index: the French version. J Epidemiol Commun Health. 2012;66(11):982–9.
18. Rey G, Jougla E, Fouillet A, Hémon D. Ecological association between a deprivation index and mortality in France over the period 1997–2001: variations with spatial scale, degree of urbanicity, age, gender and cause of death. BMC Public Health. 2009;22(9):33.
19. Riva M, Gauvin L, Barnett TA. Toward the next generation of research into small area effects on health: a synthesis of multilevel investigations published since July 1998. J Epidemiol Commun Health. 2007;61(10):853–61.
20. Tanke MAC, Ikkersheim DE. A new approach to the tradeoff between quality and accessibility of health care. Health Policy. 2012;105(2–3):282–7.
21. Korda RJ, Butler JR, Clements MS, Kunitz SJ. Differential impacts of health care in Australia: trend analysis of socioeconomic inequalities in avoidable mortality. Int J Epidemiol. 2007;36(1):157–65.
22. Guagliardo MF. Spatial accessibility of primary care: concepts, methods and challenges. Int J Health Geogr. 2004;3(1):3.
23. Cramb SM, Mengersen KL, Turrell G, Baade PD. Spatial inequalities in colorectal and breast cancer survival: premature deaths and associated factors. Health Place. 2012;18(6):1412–21.
24. Abel GA, Barclay ME, Payne RA. Adjusted indices of multiple deprivation to enable comparisons within and between constituent countries of the UK including an illustration using mortality rates. BMJ Open. 2016;6(11):e012750.
25. Bertin M, Chevrier C, Pelé F, Serrano-Chavez T, Cordier S, Viel J-F. Can a deprivation index be used legitimately over both urban and rural areas? Int J Health Geogr. 2014;14(13):22.
26. Pascal M, Corso M, Chanel O, Declercq C, Badaloni C, Cesaroni G, et al. Assessing the public health impacts of urban air pollution in 25 European cities: results of the Aphekom project. Sci Total Environ. 2013;1(449):390–400.
27. Camille C, Ghislaine B, Yolande E, Clément P, Lucile M, Camille P, et al. Residential proximity to agricultural land and risk of brain tumor in the general population. Environ Res. 2017;159:321–30.
28. Temam S, Varraso R, Pornet C, Sanchez M, Affret A, Jacquemin B, et al. Ability of ecological deprivation indices to measure social inequalities in a French cohort. BMC Public Health. 2017;17(1):956.
29. Lucas-Gabrielli V, Mangeney C. How can accessibility measures be improved to better target underserved areas? Rev Epidemiol Sante Publique. 2019;67(Suppl 1):S25–32.
30. Barlet, Coldefy M., Collin C., Lucas-Gabrielli V. L'Accessibilité potentielle localisée (APL) : une nouvelle mesure de l'accessibilité aux soins appliquée aux médecins généralistes libéraux en France. Institut de Recherche et Documentation en Economie de la Santé. (I.R.D.E.S.). Paris. FRA, editor. Paris : Irdes; 2012. (Irdes Working Document ; 51). http://www.irdes.fr/EspaceRecherche/DocumentsDeTravail/DT51AccessibilitePotentielleLocalisee.pdf.
31. Reibel M. Classification approaches in neighborhood research: introduction and review. Urban Geogr. 2011;32(3):305–16.
32. Rokach L, Maimon O. Clustering Methods. In: Maimon O, Rokach L, editors. Data mining and knowledge discovery handbook. Boston: Springer; 2005. p. 321–52. https://doi.org/10.1007/0-387-25465-X_15.
33. www.insee.fr.
34. Green MA, Daras K, Davies A, Barr B, Singleton A. Developing an openly accessible multi-dimensional small area index of 'Access to Healthy Assets and Hazards' for Great Britain, 2016. Health Place. 2018;1(54):11–9.
35. Noble M, Wright G, Smith G, Dibben C. Measuring multiple deprivation at the small-area level. Environ Plan A. 2006;38(1):169–85.
36. Chevillard G, Mousquès J. Accessibilité aux soins et attractivité territoriale : proposition d'une typologie des territoires de vie français. Cybergeo : European Journal of Geography. 2018. http://journals.openedition.org/cybergeo/29737. Accessed 4 Dec 2019.
37. Jones M, Huh J. Toward a multidimensional understanding of residential neighborhood: a latent profile analysis of Los Angeles neighborhoods and longitudinal adult excess weight. Health Place. 2014;27:134–41.
38. Weden MM, Bird CE, Escarce JJ, Lurie N. Neighborhood archetypes for population health research: is there no place like home? Health Place. 2011;17(1):289–99.
39. Gershoff ET, Pedersen S, Lawrence Aber J. Creating neighborhood typologies of GIS-based data in the absence of neighborhood-based sampling: a factor and cluster analytic strategy. J Prev Interv Community. 2009;37(1):35–47.
40. Arcaya M, Reardon T, Vogel J, Andrews BK, Li W, Land T. Tailoring community-based wellness initiatives with latent class analysis–Massachusetts Community Transformation Grant projects. Prev Chronic Dis. 2014;11:E21.
41. Stafford M, Duke-Williams O, Shelton N. Small area inequalities in health: are we underestimating them? Soc Sci Med. 2008;67(6):891–9.
42. Terrenoire E, Bessagnet B, Rouïl L, Tognet F, Pirovano G, Létinois L, et al. High-resolution air quality simulation over Europe with the chemistry transport model CHIMERE. Geosci Model Dev. 2015;8(1):21–42.
43. Yang Y, Ruan Z, Wang X, Yang Y, Mason TG, Lin H, et al. Short-term and long-term exposures to fine particulate matter constituents and health: a systematic review and meta-analysis. Environ Pollut. 2019;247:874–82.
44. Gómez-Barroso D, García-Pérez J, López-Abente G, Tamayo-Uria I, Morales-Piga A, Pardo Romaguera E, et al. Agricultural crop exposure and risk of childhood cancer: new findings from a case-control study in Spain. Int J Health Geogr. 2016;15(1):18.
45. Booth BJ, Ward MH, Turyk ME, Stayner LT. Agricultural crop density and risk of childhood cancer in the midwestern United States: an ecologic study. Environ Health. 2015;14:82.
46. Wami WM, Dundas R, Molaodi OR, Tranter M, Leyland AH, Katikireddi SV. Assessing the potential utility of commercial 'big data' for health research: enhancing small-area deprivation measures with Experian™ Mosaic groups. Health Place. 2019;57:238–46.
47. Bissonnette L, Wilson K, Bell S, Shah TI. Neighbourhoods and potential access to health care: the role of spatial and aspatial factors. Health Place. 2012;18(4):841–53.
48. Graham H. Social determinants and their unequal distribution: clarifying policy understandings. Milbank Q. 2004;82(1):101–24.

Publisher's Note