



Ecole Doctorale 446
Biologie Santé de Lille



Année universitaire 2021-2022

Thèse de DOCTORAT

Caractérisation de la multi-exposition environnementale : exploration des potentiels de résilience et de vulnérabilité face aux risques sanitaires

Par Delphine BROUSMICHE

En vue de l'obtention du grade de
Docteur de l'Université de Lille

Discipline : Recherche clinique, Innovation technologique, Santé publique
Spécialité : Santé publique

Présentée publiquement le 21/09/2022

Thèse dirigée par :	Damien Cuny, PU Caroline Lanier, MCU	Université de Lille Université de Lille
Président de jury :	Samuel Rufat, MCU	Université de Cergy Pontoise
Rapporteurs :	Bruno Barroca, PU Anne Peggy Hellequin, PU	Université Gustave Eiffel Université de Paris Nanterre
Examineurs :	Joséphine Bryère, MCU Annabelle Deram, PU François Glowacki, PU-PH	Université de Caen Université de Lille Université de Lille
Membre invité :	Michaël Genin, MCU-PH	Université de Lille

Remerciements

Les études en santé environnement ont la particularité et la nécessité d'être pluridisciplinaires ce qui les rend d'autant plus intéressantes. Ce sujet de recherche m'a permis de rencontrer un grand nombre de personnes de divers horizons qui ont apporté leur contribution à ces travaux (enseignants chercheurs, géographes, médecins, statisticiens, gestionnaires de collectivité territoriales, ...).

La thèse est un moment de rendez-vous avec soi-même... Toutefois, ces travaux n'auraient pas pu aboutir sans le soutien de nombreuses personnes que je tiens à remercier ici.

Je souhaite tout d'abord remercier la région des Hauts-de-France, l'Agence Régionale de Santé des Hauts-de-France ainsi que la Communauté Urbaine de Dunkerque d'avoir financé les différents projets de recherche auxquels j'ai été amenée à participer. Je tiens également à remercier l'Ecole Doctorale Biologie Santé de Lille et l'ensemble des membres de l'ED, pour l'attribution d'un financement institutionnel en premier lieu mais aussi et surtout pour la disponibilité et l'implication dont ils font preuve auprès des doctorant-e-s.

Je tiens à exprimer mes remerciements les plus sincères à Damien Cuny et Caroline Lanier. Selon *l'Office Research Integrity*, les bons mentors en recherche doivent être respectueux, encourageants, disponibles, préparés à assumer leur rôle et honnêtes. A tout point de vue, en toute circonstance et ce même en dépit de vos agendas bien remplis, j'ai pu bénéficier d'un encadrement de haute volée par d'EXCELLENTS mentors. En disant cela, je pèse bien mes mots et ma gratitude pour ce que vous êtes, pour votre aide, pour votre soutien, va bien au-delà de tout ce que je pourrais écrire ici. Un immense merci pour votre très grande bienveillance à mon égard tout au long de la thèse, et je dirai même avant la thèse. Les formations doctorales m'ont confrontée à d'autres réalités d'encadrement qui m'ont donné l'opportunité de mesurer, si besoin était, la chance que j'ai eue d'avoir été guidée et conseillée par vous tout au long de cette grande et formidable aventure qu'est la thèse. Caroline, merci d'avoir été la première à me proposer un stage en recherche. Ce manuscrit est une belle manière, je trouve, de boucler la boucle.

J'aimerais remercier Anne Peggy Hellequin de l'Université de Paris Nanterre et Bruno Barroca de l'Université de Paris Est Créteil de m'avoir fait l'honneur d'être les rapporteurs de ce manuscrit et d'avoir probablement dû l'emmener en vacances. Anne Peggy, merci également pour votre accompagnement dans mon comité de suivi individuel, les échanges que nous avons eus ont été réellement très enrichissants.

Je tiens également à remercier Joséphine Bryère de l'Université de Caen, Michaël Genin et François Glowacki de l'Université de Lille ainsi que Samuel Rufat de l'Université de Cergy Pontoise pour leur participation au jury de soutenance.

Un remerciement particulier à Annabelle Deram, qui même si elle n'a pas été la première à me faire confiance pour un stage en recherche, m'a offert une réelle et belle opportunité de revenir au LSVF pour collaborer avec Audrey Hayet. Ces deux très intéressants stages avec vous Annabelle et Audrey m'ont indéniablement et irréversiblement conduit sur la voie d'un M2 recherche et de la poursuite en thèse. Merci Annabelle pour tes grandes qualités d'écoute, pour tes retours et interrogations lors des CSI et plus largement merci d'avoir été ce presque-premier guide scientifique vers cette aventure passionnante qu'est le doctorat.

Un remerciement également à Aline Meirhaeghe et Cécile Bogucki pour leur présence et leur implication dans mes comités de suivi de thèse. Aline, j'espère que nous aurons l'occasion de

poursuivre nos collaborations, que ce soit dans le cadre de MONICA mais également des AVC. Des remerciements également à Mohamed, Samuel ainsi qu'à mes stagiaires pour leur contribution à ces travaux.

Je tiens également à remercier un peu plus particulièrement Florent. C'est vrai que c'est génial, passionnant, enthousiasmant et exaltant de jouer aux cartes. Merci d'avoir rendu cela possible, et merci également pour tout le reste. Plus largement, merci à l'ensemble des personnes travaillant au LSVF (Nicole, Ludivine, Stéphane, Régis, Yolène, Sébastien, Malika, Sylvain, Pierre-Arthur) pour tous ces bons moments passés ensemble. Merci également aux stagiaires qui sont passés (entre autres Camille et Vincent). Stéphane, quand tu veux pour discuter de nos lectures communes. Un VRAI GRAND merci à Nicole d'avoir initié les pauses-déjeuner au soleil (si si même dans le Nord c'est possible), c'était vraiment sympa... A ton image en somme. Les party « dépistage covid », certes un peu moins, mais bon... Avant nous prenions le café ensemble, pendant un temps on a fait tourner les tests de dépistage (j'aurais préféré les serviettes). Je dois dire que malgré tout c'est quand même plus « convivial » que les cafés virtuels chacun chez soi (je ne suis toutefois pas sûre que « convivial » soit le terme le plus approprié pour une autotest-party XD). Qu'il a fait bon travailler à vos côtés...

Je n'oublie pas les autres doctorants du labo : Laakri, Martin, Lara-Maria, mais également Klervi la tout-bientôt doctorante qui aura fait un passage chez nous. Un jour on en voit le bout... Promis... Bon... Peut être avec quelques plumes en moins et quelques cheveux blancs en plus mais tout de même ! Je n'ai qu'un regret et non des moindres... Que le Covid (oui « le » et pas « la ») soit passé par là, et que nous n'ayons pas pu être plus ensemble. Heureusement que les campagnes de terrain étaient là, je n'ai qu'une chose à dire (bien que ce qui se passe en campagne reste en campagne) : vive la bèche-danse !

Merci à vous aussi les zami-e-s (Isabelle, Lucy, Fred, Kris pour ne citer que vous mais dans mon cœur vous êtes tellement plus nombreux !!). Le hasard a fait de nous des collègues, mais la vie nous a offert la chance de devenir des Ami-e-s. Merci d'avoir toujours été là pour toutes mes aventures professionnelles et personnelles, je conviens qu'il n'est pas toujours aisé de me suivre. Essayez d'être sages à l'issue de cette soutenance, j'ai toute pas-confiance en vous pour mettre en lumière ma très grande sagesse et timidité.

La familia... Il est inévitable car totalement impensable de ne pas penser à vous tou-te-s dans des moments comme ça... Ami-e-s, confident-e-s, complices, soutiens de toujours et de tous les jours. Quelques soient les kilomètres qui nous séparent, nous sommes toutes et tous toujours ensemble. Une pensée émue à ma mamie qui aurait été drôlement fière je pense. La citation de Jean d'Ormesson (que tu adorais) est tellement vraie et résume tellement bien ce que je pense « Il y a quelque chose de plus fort que la mort, c'est la présence des absents dans la mémoire des vivants », quoi que nous fassions tu es et resteras présente dans nos cœurs.

Une pensée toute particulière pour nos 4 filles, Louane, Sarah, Zoé, Olivia, premières spectatrices de l'investissement de temps, d'énergie et de l'absence totale de gestion correcte du stress en lien avec cette (folle ?) reprise d'études dans un premier temps et avec la réalisation du doctorat dans un second temps. Promis les filles, un jour on arrêtera de courir. En attendant, n'ayez pas peur de rêver grand : Oscar Wilde disait « Il faut toujours viser la lune, car même en cas d'échec, on atterrit dans les étoiles ». Rien n'est jamais simple dans la vie, parfois on se trompe, souvent on échoue, quelquefois cela peut être difficile voire sembler impossible de continuer à avancer... Mais de ce que je peux en voir, toujours on arrive. Ayez confiance en vous, suivez votre cœur il vous donnera la bonne route.

Un dernier merci pour toi (en dernier... je sais cela te fait enrager... j'ai même sauté une page c'est dire à quel point je veux t'embêter). Ne dit-on pas que l'on garde le meilleur pour la fin ? (OK comme diraient les filles « c'est une disquette » ^-^). Que te dire de plus que ce que je te dis tous les jours ? La vie ne serait pas la même sans ton soutien et ta présence à mes côtés. Merci d'avoir répondu à mes sempiternelles questions pour ce manuscrit, de l'avoir relu, d'avoir corrigé les fautes, merci d'avoir dû parfois (souvent ?) répéter ce que tu disais parce que j'étais « ailleurs » et que je n'avais pas entendu. Pardon d'avoir eu ma check-list mentale qui m'a fait oublier plein de choses et d'avoir été au bord de la crise d'angoisse beaucoup trop souvent ces dernières semaines. « Au fond d'ma bulle, je te fais une place¹, et si j't'agace, si j'suis trop nul(le) je deviendrai toute folle², tout(e) clown³, gentil(le) pour qu'tu souris. Une petite place, ici maintenant car le temps passe à pas d'géant ».

Finalement, remercier la personne qui vit avec un doctorant⁴, c'est peut-être l'exercice de remerciement le plus difficile parce que s'il faut du courage et de l'abnégation au doctorant pour arriver au bout de ce rendez-vous, il en faut également (et peut-être même plus) à son entourage proche...

Surtout quand le doctorant en question c'est moi.

Et que déjà en temps normal j'ai un niveau d'insupportabilité assez important...

Et que cette insupportabilité a une légère propension à s'accroître en situation de stress...

Merci de m'offrir des capsules de bonheur, hier, aujourd'hui, demain et d'être encore au côté de l'autiste folle que je suis. Merci simplement d'être toi et de l'être toujours.

¹ Si si promis j'essaie

² Ça c'est franchement pas trop dur

³ Ça non plus c'est pas trop un défi

⁴ Comprendre : remercier la personne qui **SUBIT** le doctorant tous les jours avant qu'il parte au labo, le soir quand il revient du labo, le week-end, les vacances, quand il mange et qu'il pense aux phrases qu'il a écrites et que ce serait bien d'en modifier la tournure, quand la nuit il fait tout sauf dormir

La science d'aujourd'hui détruit l'ignorance d'hier et elle fera figure d'ignorance au regard de la science de demain. Dans le cœur des hommes il y a un élan vers autre chose dont la clé secrète est ailleurs

(Jean d'Ormesson, « C'est une chose étrange à la fin que le monde »)

Résumé

Indépendamment des facteurs individuels, l'apparition d'une maladie est déterminée par un ensemble complexe de déterminants interconnectés et évoluant dans l'espace. Cette variabilité des phénomènes observés nécessite la validation d'approches multidimensionnelles. Une des méthodes reconnues repose sur la création d'indices composites qui regroupent différents indicateurs mesurant les multiples facettes d'un phénomène. De nombreuses études traitent des questions de santé environnementale où l'évaluation des inégalités territoriales se concentre principalement sur la vulnérabilité (*i.e.* la somme des contraintes et nuisances péjoratives auxquelles sont soumises les populations). La résilience, définie comme la capacité d'un territoire à faire face à ces inégalités de santé, vient s'intégrer à l'analyse menée au cours de nos travaux afin d'apporter une description plus précise des territoires. Cependant, le design des indices composites en santé environnementale doit être optimisé et la variabilité induite doit être maîtrisée.

Le premier objectif de ce projet de thèse est la définition et l'analyse spatiale des typologies de territoires au regard de la multi-dimensionnalité des phénomènes impliqués en santé environnement. Une analyse de la littérature a permis d'identifier une grande diversité des terminologies et des méthodologies mais également sur des variables incluses dans les indices composites spatialisés en santé environnement. Sur la base des conclusions de l'analyse, un cadre méthodologique de construction d'indices composites spatialisés permettant de prendre en compte une multi-exposition environnementale, qu'elle soit bénéfique ou néfaste pour la santé humaine, a été proposé. Celle-ci a été appliquée aux communes des Hauts-de-France et a permis d'obtenir une représentation holistique des niveaux de vulnérabilité et de résilience des communes. Une réflexion a également été menée sur l'adaptabilité de cette méthodologie à très fine échelle sur la Communauté Urbaine de Dunkerque.

Le second objectif est l'analyse des relations entre les potentiels de vulnérabilité et de résilience et la survenue de maladies chroniques prégnantes sur certains territoires. La collaboration avec des registres de santé nous a permis de pouvoir garantir l'exhaustivité des cas pour une pathologie donnée sur un territoire donné. Dans un premier temps, l'influence de ces caractéristiques territoriales sur la variabilité résiduelle de l'insuffisance rénale chronique terminale (IRCT) a été menée à l'échelle des communes du Nord et du Pas de Calais sur la période 2005-2018. Après ajustement sur des facteurs de confusion, des régressions écologiques ont été menées. Les résultats ont mis en évidence l'influence et la prépondérance des facteurs bénéfiques sur l'incidence de l'IRCT. Dans un second temps, des indices de vulnérabilité et de résilience ont été développés à l'échelle des IRIS de la Métropole Européenne de Lille pour évaluer leur influence sur la survenue des maladies coronariennes sur la période 2012-2016. Bien que les méthodologies mises en place diffèrent, les résultats rejoignent ceux obtenus dans le cadre de l'insuffisance rénale chronique terminale, avec une influence plus importante des facteurs bénéfiques sur la survenue des maladies coronariennes. Afin de poursuivre l'analyse des interactions entre les déterminants en lien avec la résilience et la vulnérabilité et les maladies coronariennes, des profils complexes de territoire ont été développés. Leur caractérisation par l'incidence des maladies coronariennes a souligné que la sur-morbidité n'était pas la seule résultante d'une accumulation de déterminants en lien avec la vulnérabilité. L'enjeu de ce travail était d'identifier des facteurs de risques territoriaux qui permettront de décrire de nouvelles hypothèses étiologiques dans le cadre de futures études épidémiologiques.

Mots-clés : Santé environnementale ; indice composite ; statistiques spatiales ; maladies coronariennes ; insuffisance rénale chronique terminale ; inégalités territoriales en santé environnement

Abstract

Regardless of individual factors, the development of disease is determined by a complex set of interconnected and spatially evolving determinants. This variability of observed phenomena requires the validation of multidimensional approaches. One of the recognized methods is based on the development of composite indices that aggregate together different indicators describing the multiple facets of a phenomenon. Many studies deal with environmental health issues where the evaluation of territorial inequalities focuses mainly on vulnerability (*i.e.* the sum of the pejorative constraints and nuisances to which populations are subjected). Resilience, defined as the capacity of a territory to cope with these health inequalities, is integrated into the analysis carried out during our work in order to provide a more precise description of the territories. However, the design of composite environmental health indices must be optimized and the resulting variability must be mastered.

The first objective of this thesis project is the definition and spatial analysis of the typologies of territories in view of the multi-dimensionality of the phenomena involved in environmental health. An analysis of the literature made it possible to identify a great diversity of terminologies and methodologies but also of the variables included in the spatialized composite indices in environmental health. Based on the conclusions of the analysis, a methodological framework for building up spatialized composite indices to consider environmental multi-exposure, whether beneficial or harmful to human health, has been proposed. This was applied to the municipalities of Hauts-de-France and made it possible to obtain a holistic representation of the levels of vulnerability and resilience of the municipalities. A reflection was also carried out on the adaptability of this methodology at a very fine scale on the Dunkerque Urban Community perimeter.

The second objective is to analyze the relationships between the potential for vulnerability and resilience and the incidence of significant chronic diseases in certain territories. Collaboration with health registries has enabled us to guarantee the exhaustiveness of cases for a given pathology in a given territory. Firstly, the influence of these territorial characteristics on the residual variability of end-stage renal disease (ESRD) was assessed at the level of the municipalities of Nord and Pas de Calais over the period 2005-2018. After adjusting for confounding factors, ecological regressions were carried out. The results highlighted the influence and the preponderance of beneficial factors on the incidence of ESRD. Secondly, vulnerability and resilience indices were developed at the scale of the neighborhood of the European Lille Metropolis to assess their influence on the incidence of coronary heart disease over the period 2012-2016. Although the methodologies implemented differ, the results are similar to those obtained in the context of end-stage chronic renal disease and highlight a greater influence of beneficial factors on the occurrence of coronary heart disease. In order to continue the analysis of the interactions between the determinants related to resilience and vulnerability and coronary heart disease, complex territory profiles have been developed. Their characterization by the incidence of coronary heart disease highlighted that excess morbidity was not the only result of an accumulation of determinants linked to vulnerability. The challenge of this work was to identify potential territorial risk factors that will make it possible to describe new etiological hypotheses in the context of future epidemiological studies.

Keywords: Environmental health; composite index; spatial statistics; coronary heart diseases; end-stage renal disease; territorial inequalities in environmental health

Table des matières

Résumé	7
Abstract	8
Liste des acronymes	11
Liste des figures	13
Liste des tableaux	15
Introduction générale.....	16
Les objectifs	21
Chapitre I : Evaluation des caractéristiques du territoire	22
Introduction.....	23
Article 1	25
Article 2	58
Article 3	77
Article 4	96
Discussion Chapitre I	112
Chapitre II : Caractéristiques du territoire et maladies chroniques.....	122
Introduction.....	123
Partie II-1 : l'Insuffisance Rénale Chronique Terminale	125
Introduction.....	125
Matériel et méthodes.....	126
Résultats	131
Discussion	142
Partie II.2 : les maladies coronariennes	152
Introduction.....	152
Matériel et méthodes.....	153
Résultats – Discussion	154
Article 5	161
Discussion Partie II.2 : profil territorial complexe et maladies coronariennes	178
Discussion générale	182
Valorisations en lien avec les travaux de thèse.....	193
Activités d'encadrement	195
Activités d'enseignement	196
Implication dans des sociétés savantes.....	197
Organisations de colloques	198
Bibliographie.....	199
Annexes	227

Liste des acronymes

ACP :	Analyse en Composantes Principales
ALD :	Affection de Longue Durée
APL :	Accessibilité Potentielle Localisée
BPCO :	Broncho-Pneumopathie Chronique Obstructive
CAH :	Classification Ascendante Hiérarchique
CUD :	Communauté Urbaine de Dunkerque
DFG :	Débit de Filtration Glomérulaire
EDCH :	Eau Destinée à la Consommation Humaine
EDI :	<i>European Deprivation Index</i>
EPCI :	Etablissement Public de Coopération Intercommunale
ETM :	Elément Trace Métallique
FDep :	<i>French Deprivation index</i>
HAP :	Hydrocarbure Aromatique Polycyclique
HDF :	Hauts-de-France
IDH :	Indice de Développement Humain
ISDH :	Indice Soutenable de Développement Humain
IMC :	Indice de Masse Corporelle
IPH :	Indice de Pauvreté Humaine
IRCT :	Insuffisance Rénale Chronique Terminale
IRIS :	Îlot Regroupé pour l'Information Statistique
MAUP :	<i>Modifiable Areal Unit Problem</i>
MCV :	Maladie Cardio-Vasculaire
MEL :	Métropole Européenne de Lille
MNT :	Maladie Non Transmissible
MRC :	Maladie Rénale Chronique
NDVI :	<i>Normalized Difference Vegetation Index</i>
NEAP :	<i>Neighborhood Effect Averaging Problem</i>
OMS :	Organisation Mondiale de la Santé
PIB :	Produit Intérieur Brut
PLU :	Plan Local d'Urbanisme

pmh :	Par million d'habitants
PNSE :	Plan National Santé Environnement
PNUD :	Programme des Nations Unies pour le Développement
PPP :	Produit Phyto-Pharmaceutique
PRSE :	Plan Régional Santé Environnement
PRASE :	Plan Régional d'Action en Santé Environnement
REACH :	Registration Evaluation Authorization and restriction of Chemicals
REIN :	Réseau Epidémiologique et Information en Néphrologie
RR :	<i>Risk Ratio</i>
REM :	Ratio d'Emission Médian
SIG :	Systèmes d'Information Géographique
SIR :	Ratio d'Incidence Standardisé
UAD :	Unité d'Auto-Dialyse
UDM :	Unité de Dialyse Médicalisée
UGCoP :	<i>Uncertain Geographic Context Problem</i>

Liste des figures

Figure 1 : Article 1 – Figure 1: Results of the article search and outline of the entire article searching process for this systematic review	29
Figure 2 : Article 1 – Figure 2: Composite indices in environmental health, scientific vs institutional bodies, published since 2001	31
Figure 3 : Article 1 – Figure 3: Representativeness of the number of variables within dimensions.....	32
Figure 4 : Article 1 – Figure 4: Good practice for developing composite indices (inspired by European Commission et al. (2008)).....	53
Figure 5 : Article 3 – Figure 1: A comprehensive methodology for the development of composite indices to assess environmental health inequalities. The general methodological framework presents the main steps to be followed for the development of composite indices, while the detailed methodological framework explains the different methods proposed for each step (step (i): analysis of data, step (ii): calculation of the index, step (iii): evaluation of the spatial distribution, step (iv): discretization and step (v): joint analysis of resilience and vulnerability indices).....	81
Figure 6 : Article 3 – Figure 2: Diagram illustrating the orientation of variables towards resilience or vulnerability index, according to the effect of an increase in the variable, for use in joint analysis....	83
Figure 7 : Article 3 – Figure 3: Flowchart summarizing the variable selection steps, from the initial dataset (50 variables) to the Resilience Index (21 variables) and the Vulnerability Index (13 variables)	86
Figure 8 : Article 3 – Figure 4: (a) The Vulnerability Index for the municipalities in the Hauts-de-France region. Three Jenks’s natural breaks were calculated (0.143, 0.205, 0.461) to obtain four classes of vulnerability (very low, low, high, very high).	87
Figure 9 : Article 3 – Figure 5: Diagram showing the distribution of the number of municipalities among each typology (To monitor, Resilient, Have Resources, Territorial blackspot) after the cross-classification of resilience (RI) and vulnerability indices (VI) in the Hauts-de-France region	88
Figure 10 : Article 3 – Figure 6: Spatial distribution of the joint analysis of resilience and vulnerability indices for the 3,817 municipalities of the Hauts-de-France region. The “Territorial blackspots” are characterized by a high (or very high) vulnerability, paired with a low (or very low) resilience. The municipalities classified in the “Resilient” category are characterized by a low (or very low) vulnerability and a high (or very high) resilience. The municipalities classified in the “Have resources” category are characterized by a high (or very high) vulnerability and a high (or very high) resilience. The municipalities classified in “To monitor” are characterized by a low (or very low) vulnerability and a low (or very low) resilience	89
Figure 11 : Article 4 – Figure 1: Dunkerque Urban Community, located in the Hauts-de-France region (on the coast in the north of France)	100
Figure 12 : Article 4 – Figure 2: Methodological framework used to develop composite indices (Brousmitche et al., 2020a)	101
Figure 13 : Article 4 – Figure 3: Synthesis of cases examined to assess the influence of the spatial unit on the development and results of composite indices that are used to study resilience and vulnerability in environmental health inequalities (RI: Resilience Index; VI: Vulnerability Index; IRIS: smallest aggregated unit for statistical information).....	102
Figure 14 : Article 4 – Figure 4: Composite indices based on 34 variables – focus on the municipalities of the Dunkerque Urban Community (Brousmitche et al., 2020a); case 1: Vulnerability Index (a), Resilience Index (b), Joint analysis of Resilience and Vulnerability (c); Case2: Vulnerability Index (d), Resilience Index (e), Joint analysis of Resilience and Vulnerability (f)	106

Figure 15 : Article 4 – Figure 5: Development of composite indices based on 31 variables carried out for the 93 IRISs of Dunkerque Urban Community. (a): Vulnerability Index; (b): Resilience Index; (c): Joint analysis of resilience and vulnerability indices	108
Figure 16 : Analyse de sensibilité vis-à-vis des données manquantes, comparaison de deux méthodes de développement d'indices composites.....	119
Figure 17 : Taux d'incidence brute d'IRCT sur 2005-2018, selon l'âge et le sexe	131
Figure 18 : Clusters d'IRCT obtenus après ajustement sur l'âge, le sexe et la défaveur socio-économique et le temps de trajet par la route en minutes vers le centre lourd le plus proche	132
Figure 19 : Distribution spatiale de la moyenne pondérée du DFG à l'initiation des patients atteints d'IRCT (2005-2018).....	133
Figure 20 : Distribution spatiale du FDep sur le Nord-Pas de Calais, moyenne sur la période couverte par le registre (2005-2018).....	134
Figure 21 : Distribution spatiale de l'indice de résilience sur les 1 539 communes du Nord-Pas de Calais	135
Figure 22 : Distribution spatiale de l'indice de vulnérabilité sur les 1 539 communes du Nord-Pas de Calais.....	136
Figure 23 : RR d'IRCT après ajustement sur l'âge et le sexe sur 2005-2018 (modèle 1).....	137
Figure 24 : RR d'IRCT après ajustement sur l'âge et le sexe et les facteurs de confusion liés à la santé sur 2005-2018 (modèle 2)	138
Figure 25 : RR d'IRCT après ajustement sur l'âge, le sexe, les facteurs de confusion liés à la santé et le FDep sur 2005-2018 (modèle 3).....	139
Figure 26 : RR d'IRCT après ajustement sur l'âge, le sexe, les facteurs de confusion liés à la santé, le FDep et l'indice de résilience sur 2005-2018 (modèle 4).....	140
Figure 27 : Incidence brute des maladies coronariennes, pour 100 000 habitants (2012-2016)	154
Figure 28 : Ratios d'Incidence Standardisé sur l'âge et le sexe pour les maladies coronariennes sur 2012-2016, après lissage spatial (cette figure sera publiée dans l'article 5)	155
Figure 29 : Distribution spatiale de l'indice de vulnérabilité vis-à-vis de la santé cardiovasculaire sur les 472 IRIS de la MEL	157
Figure 30 : Distribution spatiale de l'indice de résilience sur les 472 IRIS de la MEL.....	159
Figure 31 : Distribution des SIRs lissés en fonction des classes (a) de vulnérabilité et (b) de résilience, la ligne rouge représentant la médiane du SIR lissé	160
Figure 32 : Article 5 – Figure 1: European Lille Metropolis (ELM), located in the north of France.....	164
Figure 33 : Article 5 – Figure 2: Age and gender smoothed Standardized Incidence Ratios for CHD (2012-2016).....	168
Figure 34 : Article 5 – Figure 3: Spatial distribution of classes obtained after PCA-HCPC based on the 24 selected variables	169
Figure 35 : Article 5 – Figure 4: Averaged age and gender smoothed SIRs for CHD according to class obtained after PCA and HCPC; line in red represents the overall mean of SIRs on ELM	170

Liste des tableaux

Tableau I : Article 1 - Table I: Synthesis of studied indices.....	33
Tableau II : Article 1 – Table II: Characterization of sub-dimensions and distribution of variables within dimensions	39
Tableau III : Article 2 – Table I: Specifications Table	60
Tableau IV : Article 2 – Table II: Metadata related to the “Economics” dimension.....	62
Tableau V : Article 2 – Table III: Metadata related to the “Environment” dimension	64
Tableau VI : Article 2 – Table IV: Metadata related to the “Policy” dimension	65
Tableau VII : Article 2 – Table V: Metadata related to the “Health” dimension	68
Tableau VIII : Article 2 – Table VI: Metadata related to the “Social” dimension	69
Tableau IX : Article 2 – Table VII: Metadata related to the “Services” dimension.....	71
Tableau X : Article 3 – Table I: Presentation of sub-dimensions that aim to describe multiple facets for each of the six dimensions	82
Tableau XI : Article 3 – Table II: Correlation matrix between the Vulnerability Index, the Resilience Index and other indicators (HDI4, French EDI, FDep). Results of the significant Spearman’s correlation tests (i.e. $p < 0.001$) are marked with an asterisk “*”	88
Tableau XII : Article 4 – Table 1: Correlation matrix between the Vulnerability Index, the Resilience Index and other indicators (HDI4, French EDI, FDep). Results of the significant Spearman’s correlation tests (i.e. $p < 0.001$) are marked with an asterisk	104
Tableau XIII : Article 4 – Table 2: For each of the three cases, results of the variable selection step to build up resilience and vulnerability composite indices	104
Tableau XIV : Article 4 – Table 3: For each of the three cases, (i) distribution of spatial units among the Very Low, Low, High, Very High classes for the resilience and vulnerability indices, (ii) results of the Spearman’s correlation test between RI and VI, significant Spearman’s correlation tests (i.e. $p < 0.001$) are marked with an asterisk “*” , and (iii) distribution municipalities and IRIS within the defined typologies	105
Tableau XV : Liste des données qui ont été retenues pour le développement des indices composites de résilience et de vulnérabilité	129
Tableau XVI : Influence du FDep sur le RR de l’IRCT sur 2005-2018	134
Tableau XVII : RR, intervalle de confiance à 95 % et part de la variation spatiale observée expliquée par les covariables pour les différents modèles	141
Tableau XVIII : RR et IC de l’incidence d’IRCT ajustée sur l’ensemble des facteurs de confusion de l’ensemble des variables intégrées à l’indice de résilience.....	142
Tableau XIX : Contribution des 13 variables en lien avec la vulnérabilité aux quatre composantes retenues après la rotation VARIMAX	156
Tableau XX : Contribution des 11 variables en lien avec la résilience aux quatre composantes retenues pour la rotation VARIMAX	158
Tableau XXI : Article 5 – Table XXII: Distribution by dimension and sub-dimension of the selected variables in the dataset, each had been assigned to resilience or vulnerability	167
Tableau XXIII : Article 5 – Table II: Mean (standard deviation) of each of the 24 variables within the complex territorial profiles obtained after PCA and HCPC. A mean (sd) or a variable in grey was non-significant while significant means (sd) in class are marked with an asterisk “*” (i.e. $p < 0.05$)	171

Introduction générale

Dans le préambule de sa constitution entrée en vigueur le 7 avril 1948, l'Organisation Mondiale de la Santé (OMS) a défini la santé comme « un état complet de bien-être physique, mental et social et ne consiste pas seulement en une absence de maladie ou d'infirmité ». La charte d'Ottawa de 1986 a développé cette définition : « la promotion de la santé est le processus qui permet aux individus de mieux contrôler et améliorer leur santé. Pour atteindre un état de bien-être physique, mental et social complet, un individu ou un groupe doit être capable d'identifier et de réaliser des aspirations, de satisfaire des besoins et de changer ou de faire face à son environnement. La santé est donc considérée comme une ressource de la vie quotidienne et non comme un objectif de vie. La santé est un concept positif qui met l'accent sur les ressources sociales et personnelles, ainsi que sur les capacités physiques. Par conséquent, la promotion de la santé n'est pas seulement de la responsabilité du secteur de la santé, mais va au-delà des modes de vie sains pour le bien-être ». Ces définitions soulignent la nature positive de la santé et les aspects multidimensionnels qui la composent (McCartney *et al.*, 2019).

L'OMS lors de la conférence d'Helsinki de 1994 a défini une nouvelle notion intégrant l'environnement, « la santé environnementale comprend les aspects de la santé humaine, y compris la qualité de la vie, qui sont déterminés par les facteurs physiques, chimiques, biologiques, psychosociaux et esthétiques de notre environnement. [La santé environnementale] concerne également la politique et les pratiques de gestion, de résorption, de contrôle et de prévention des facteurs environnementaux susceptibles d'affecter la santé des générations actuelles et futures ». Selon l'OMS, le meilleur état de santé pour tout être humain, quelles que soient sa race, sa religion, ses opinions politiques, sa condition socio-économique, constitue un droit fondamental bien que les inégalités en termes de santé entre les pays y portent préjudice. Selon la définition de Arcaya *et al.* (2015), les inégalités de santé caractérisent les variations de tout aspect mesurable de la santé d'un individu à un autre, d'un groupe socialement pertinent à un autre. Cette définition des inégalités de santé peut porter à confusion avec celles des disparités de santé, traduisant l'observation de simples différences entre des groupes ou des différences subsistant après avoir pris en compte d'autres explications (McCartney *et al.*, 2019). Arcaya *et al.* (2015) distinguent les iniquités de santé, qu'ils considèrent comme un type spécifique d'inégalités de santé : les inégalités injustes et évitables. Le terme d'iniquité (« inequity » en anglais) est linguistiquement le plus pertinent pour traduire l'injustice des différences en matière de santé car il comporte une dimension morale et éthique (McCartney *et al.*, 2019; Whitehead, 1992). Toutefois, en Europe « health inequities » est peu utilisée et l'expression principalement utilisée est « health inequalities ». L'OMS quant à elle définit les inégalités de santé comme des différences de santé systématiques socialement produites et de fait modifiables et injustes (Whitehead and Dahlgren, 2006). D'après Marmot *et al.* (2008), ces inégalités surviennent parce que les individus grandissent, vivent, travaillent et vieillissent dans des circonstances différentes, modelées par des forces politiques, sociales et économiques. Le lien entre la santé et la position sociale de l'individu est l'exemple des différences systématiques de santé le plus notable : ce gradient social de la santé est universel et une relation statistiquement significative existe entre le statut socio-économique des quartiers de résidence et la mortalité, avec une association plus forte pour les hommes et les jeunes (Meijer *et al.*, 2012; Potvin *et al.*, 2010; Whitehead and Dahlgren, 2006). Les inégalités sociales de santé font l'objet de nombreux travaux depuis plusieurs décennies au regard de l'ampleur et de la persistance de ces phénomènes (Menvielle and Lang, 2021). Les déterminants sociaux de la santé peuvent permettre de plus de prédire en partie les inégalités de santé des individus et des populations (Alla, 2016). Toutefois, les inégalités de santé peuvent également être la résultante d'autres déterminants, tels que les

conditions générales de vie et de l'environnement dans lesquels s'inscrivent les populations (Menvielle and Lang, 2021; Potvin *et al.*, 2010). Le concept des inégalités environnementales et sociales de santé a dès lors émergé afin de pouvoir considérer ces déterminants dans les sources d'inégalités de santé. Pour caractériser de manière étendue les multiples facettes des déterminants impliqués dans la santé des populations et qui ne relèvent pas des seuls aspects socioéconomiques et environnementaux, ces travaux de thèse portent sur l'analyse des inégalités territoriales en santé environnement.

L'OMS souligne le paradoxe français : l'espérance de vie en France figure parmi les plus longues d'Europe tandis que les inégalités de santé y sont parmi les plus importantes (Bourret, 2016; INERIS, 2014; Lang *et al.*, 2017). Ces inégalités de santé au sein de l'Europe étaient déjà soulignées dans les travaux de Whitehead (1992), celles-ci restent très marquées et continuent à s'accroître en dépit de l'amélioration des conditions environnementales qui peut être observée dans la plupart des pays (Danet, 2010; Maresca and Helmi, 2014; Rican *et al.*, 2011; WHO, 2019). La mortalité prématurée en France (survenant avant 65 ans) est supérieure à la moyenne européenne et de grandes disparités selon les régions peuvent de plus être observées (Leroux et Riem, 2020). Toute la population française est concernée par la question des inégalités de santé qui suivent un gradient social, intervenant dès l'enfance et pouvant perdurer par la suite (Menvielle and Lang, 2021; Santé Publique France, 2021). Ces inégalités persistent et sont particulièrement préoccupantes pour les populations âgées de moins de 65 ans (DREES and Santé Publique France, 2017). La région Hauts-de-France, la deuxième région la plus densément peuplée de la métropole, présente une surmortalité avant 65 ans de 33 % supérieure la moyenne nationale pour les hommes et de 26 % pour les femmes (OR2S, 2019). Les espérances de vie à la naissance y sont également les plus faibles de France, les différentiels avec les moyennes nationales étant toutefois moins élevés en zones rurales (Leroux and Riem, 2020; MTES, 2019; OR2S, 2019). Au sein des départements des Hauts-de-France, les inégalités de santé se renforcent et certains territoires sont en surmortalité (Rollet et Thomas, 2021; Vigneron, 2013). Les cancers et les maladies cardiovasculaires sont les principales causes de mortalité au sein de la région de manière constante depuis une trentaine d'années (respectivement 28,1 % et 24,5 % sur 2011-2013), la mortalité en lien avec les cancers dans le Pas-de-Calais étant la plus élevée de la région (ARS, 2018). De la même manière, les taux d'admission en Affection Longue Durée (ALD) pour le diabète de type 2 sont supérieurs dans la région Hauts-de-France (OR2S, 2019). Ce diabète est un facteur aggravant des maladies cardiovasculaires et constitue la première cause d'Insuffisance Rénale Chronique Terminale (IRCT) en France (Pillon *et al.*, 2014).

La région des Hauts-de-France présente également de nombreux indicateurs socio-économiques défavorables par rapport aux moyennes nationales, *e.g.* revenu annuel médian plus faible ou taux de pauvreté des jeunes plus élevé (Rollet et Thomas, 2021). De plus, les jeunes y présentent de plus grandes difficultés d'insertion et l'illettrisme y est le plus élevé (11,5 % dans les Hauts-de-France contre 7 % de moyenne métropolitaine) (MTES, 2019). De fortes disparités socio-économiques existent également entre les différents départements de la région, *e.g.* taux de chômage plus élevé dans le département du Nord, déficits de certains professionnels de santé libéraux dans l'Oise, proportion d'ouvriers plus importantes dans la Somme (ARS, 2018). D'autres déterminants des inégalités territoriales en santé environnement présentent de la même manière des résultats défavorables au sein de la région. L'augmentation de l'artificialisation y figure parmi les plus élevées avec des conséquences sur la fragmentation des milieux naturels engendrant des menaces sur la biodiversité (MTES, 2019). L'ancienne région du Nord-Pas de Calais présente des dépassements des valeurs limites et/ou cibles normes en matière de pollution atmosphérique, les oxydes d'azote et les particules en étant les principaux responsables. Depuis 2008, la quasi-totalité de ces deux départements sont concernés par des dépassements de seuils pour les particules, et notamment les PM10 (DREAL, 2020).

Face à ces problématiques d'inégalités territoriales en santé environnement, la France figure parmi les pays d'Europe les plus engagés sur ces problématiques de santé environnementale et élabore tous les cinq ans des Plans Nationaux de Santé Environnement (PNSE). Ils sont inscrits au Code de la Santé Publique et ont pour but d'approfondir les connaissances sur les multiples facteurs environnementaux (qualité des milieux caractérisée par des nuisances) et la santé. Les PNSE ont fait l'objet de déclinaisons régionales au travers des Plans Régionaux Santé Environnement (PRSE) en concertation avec les territoires et les acteurs locaux afin de tendre vers une territorialisation des politiques dans tous les domaines en lien avec la santé environnement qui prend en compte de ce fait les préoccupations locales. L'ancienne région du Nord-Pas de Calais a été la première à mettre en place un Plan Régional d'Action en Santé Environnement (PRASE) en 2000 afin de pouvoir approfondir les connaissances en lien avec les impacts sanitaires des sites et sols pollués, du monoxyde de carbone, du plomb dans les logements et de l'insalubrité et/ou de l'indécence des logements. Le PNSE1 pour la période de 2004-2008 reposait sur une approche intégrée et globale de l'ensemble des polluants et des milieux de vie. Il a contribué à réduire les impacts de l'environnement sur la santé et à améliorer à toutes les échelles du territoire la prise en compte de la santé environnementale. Suite aux conclusions du groupe de travail composé d'élus, de représentants d'associations, de salariés, d'entreprises, d'experts et de l'Etat, le PNSE2 (2010-2014) a eu comme principaux objectifs de réduire (i) les expositions responsables de pathologies à fort impact sanitaire (réduction des substances toxiques, réduction des substances identifiées dans le cadre de REACH, qualité de l'air intérieur, plan particules) et (ii) les inégalités environnementales (micropolluants et résidus médicamenteux dans le compartiment aquatique, plan radon, perturbateurs endocriniens, caractérisation des expositions faibles doses de la population par la biosurveillance, lutte contre les points noirs environnementaux, réduction de l'exposition aux champs électromagnétiques), ces objectifs devant être en harmonie avec d'autres plans (e.g. plan cancer, plan santé travail). Le PNSE3 (2015-2019) portait sur la consolidation des progrès obtenus et la proposition d'une nouvelle approche de la santé environnement, celle-ci devant être plus positive et ancrée sur les territoires. Cette approche doit également intégrer le nouveau paradigme de l'exposome qui nécessite de prendre en considération toutes les sources de pollution et de nuisances (et leurs possibles interactions) susceptibles de nuire à la santé des individus pendant toute la durée de leur vie (Fang *et al.*, 2021). Le PNSE4 (2021-2025) s'inscrit totalement dans la démarche « One health » de l'OMS qui a émergé de la nécessité de décloisonner les instances en charge de la santé publique, de la santé animale (75 % des nouvelles maladies apparues depuis 1970 sont d'origine animale) et de la préservation de l'environnement pour tendre vers une santé globale (Parodi, 2021). Le PNSE4 vise ainsi à mieux comprendre et réduire les risques liés aux substances chimiques, aux agents physiques (bruit, onde) ou aux agents infectieux en lien avec les pathologies pouvant se transmettre de l'animal à l'Homme (*i.e.* les zoonoses). Parmi les quatre actions prioritaires du PNSE4, le développement de la recherche en santé environnement permettra d'améliorer les connaissances sur les liens entre exposition et effets sur la santé humaine, animale et des écosystèmes. La réduction des expositions environnementales reste toutefois au cœur des préoccupations des PNSE au regard de l'importance et de l'accroissement des pathologies humaines et animales en lien avec l'environnement.

La recherche et l'action en santé publique reposent sur la compréhension de l'influence des déterminants qui ont un impact sur la santé des populations, mais également sur les inégalités de santé qui en est son concept connexe (INSERM, 2021; McCartney *et al.*, 2019). Un déterminant de santé est un « facteur, qu'il s'agisse d'un évènement, d'une caractéristique ou de toute autre entité définissable, qui provoque un changement de l'état de santé ou d'une autre caractéristique définie » (Last and International Epidemiological Association, 2001). La notion de facteur de risque est définie comme « une caractéristique individuelle ou collective dont la présence est associée à une modification de la

probabilité de survenue d'un évènement de santé ». Les différences entre ces notions souvent employées l'une pour l'autre tendent à disparaître, le terme de déterminant étant dorénavant utilisé de manière plus consensuelle (Alla, 2016). Ce concept de déterminant de santé est un élément clé en santé publique car une action sur un déterminant a une influence sur l'état de santé des populations (Alla, 2016).

Parce qu'elle combine les effets et les interactions de nombreux déterminants de différentes natures et qui ont de nombreuses ramifications, la santé est considérée comme un phénomène complexe (Alla, 2016; Anctil *et al.*, 2012; Moleux *et al.*, 2011). Les déterminants de santé peuvent être en lien avec les caractéristiques individuelles (*e.g.* âge, sexe, patrimoines génétique et biologique), les habitudes de vie, le milieu de vie (familial, scolaire, professionnel, communautaire, hébergement), les systèmes d'administration publique (systèmes éducatifs, systèmes de santé, services sociaux, aménagement du territoire) ou encore le contexte global comprenant des déterminants plutôt macroscopiques qui influencent le cadre de vie des populations (environnement, contexte démographique, social, économique, culturel et politique) (Anctil *et al.*, 2012; Dahlgren and Whitehead, 1991). Réfléter la pluralité et la complexité de ces déterminants s'avère primordial dans une approche globale de la santé visant à implémenter des actions préventives en santé publique (Alla, 2016).

L'espace et le lieu dans lesquels s'inscrivent les populations jouent également un rôle important car ils permettent d'avancer dans la compréhension de l'influence du territoire sur la santé (Arcaya *et al.*, 2016). L'espace traite plutôt de déterminants spatialement répartis qui évoluent en fonction de l'emplacement des populations. Ces déterminants ont plutôt trait à l'accessibilité, la proximité ou encore l'exposition à un facteur de risque (*e.g.* la pollution atmosphérique) ou à une aménité (*e.g.* proximité à un espace vert). Le lieu se réfère plutôt à l'appartenance à une unité administrative qui a des spécificités propres uniformément réparties au sein de l'unité spatiale (*e.g.* politique d'aide pour une population précise au sein d'une unité administrative) (Arcaya *et al.*, 2012). Cette notion de lieu renvoie à la définition du territoire « prescrit » proposée par Raulet-Croset (2014) qui identifie un territoire selon un découpage géographique issu d'une logique descendante et est en lien avec des actions et politiques publiques. Les disparités de santé correspondent aux variations géographiques de la santé, ces disparités étant influencées par des déterminants en lien avec l'espace, le lieu ou les deux simultanément (Rican *et al.*, 2014; Vallée, 2015). Des indicateurs de santé en lien avec des habitudes de vie à l'échelle des quartiers et plus largement les effets de l'espace ont été souvent considérés en épidémiologie ou en géographie de la santé, néanmoins le territoire devrait être considéré à une échelle plus large pour l'analyse des inégalités de santé (Vallée, 2015). Prendre en compte l'espace et le lieu dans un diagnostic permettrait de mieux contextualiser les programmes publics visant à réduire ces disparités de santé (Arcaya *et al.*, 2016). De plus, le rapport de la Lancet Commission a mis en évidence le besoin d'une approche systémique complexe pour caractériser la santé des population qui est influencée par les interactions et les interconnexions entre les déterminants de santé qui évoluent dans le temps et l'espace (Rydin *et al.*, 2012; Sarkar and Webster, 2017). Ces travaux de thèse se focalisent sur l'influence spatiale des déterminants intégrés dans l'analyse de la balance résilience vulnérabilité pour caractériser de manière holistique le territoire.

Lors de l'analyse des inégalités de santé, l'accent a longtemps été mis sur les conditions sociales, économiques et culturelles des individus, la position socio-économique des populations ayant largement été considérée et privilégiée comme un déterminant clé de la santé (Gkiouleka *et al.*, 2018; Rican *et al.*, 2011). Néanmoins pour Rican *et al.* (2011), les contextes sociaux, politiques et environnementaux dans lesquels les populations évoluent devraient également être inclus. L'environnement est en effet une des causes d'apparition et/ou d'aggravation de pathologies : près d'un quart des décès et des pathologies chroniques, généralement plurifactorielles (*e.g.* pathologies

respiratoires, maladies cardiovasculaires) dans le monde lui sont imputables (PNSE4, 2021; Prüss-Üstün *et al.*, 2016). Les connaissances restent toutefois encore parcellaires sur le poids de certains facteurs environnementaux dans la survenue de maladies chroniques (Kamel et El-Minshawy, 2010). La recherche en santé publique vise de plus en plus à comprendre l'impact des polluants anthropiques environnementaux sur la santé (Chang, 2016). Dans de nombreuses études traitant des questions de santé environnementale, l'évaluation des inégalités territoriales en santé environnement s'est principalement concentrée sur les aspects en lien avec la vulnérabilité (*i.e.* la somme des contraintes et nuisances péjoratives auxquelles sont soumises les populations) (Lanier *et al.*, 2020). Toutefois, de nouveaux axes de recherche ont émergé et souligné l'influence positive sur la santé de certains déterminants qui pourrait avoir été sous-estimée, les travaux de Sirven and Debrand (2012) portent sur les liens entre capital social et résultats de santé individuels tandis que Twohig-Bennett and Jones (2018) analysent les liens entre exposition aux espaces verts et leurs impacts bénéfiques sur la santé. La résilience en santé environnement peut alors être définie comme la capacité d'un territoire à faire face aux inégalités de santé (Lanier *et al.*, 2020). Bien que ces notions de vulnérabilité et de résilience soient étroitement entrelacées, la résilience n'est pas le simple antonyme de la vulnérabilité (Fekete, 2018). En effet une faible vulnérabilité ne signifie pas nécessairement qu'un territoire possède une ou des ressources à mettre en œuvre pour pallier cette vulnérabilité si le contexte territorial est dégradé ou se dégrade. Le challenge est donc de proposer une analyse des inégalités territoriales de santé la plus large possible et qui considèrerait les déterminants qui ont un impact négatif et bénéfique sur la santé. De plus, il s'avère que les évaluations intégrées et collectives apparaissent comme plus pertinentes que des approches fragmentaires pour relever le défi des problématiques de santé environnementale soulevé par les décideurs et gestionnaires territoriaux qui cherchent à réduire voire éliminer les inégalités territoriales en santé environnement (Arcaya *et al.*, 2016; Knol *et al.*, 2010).

Les objectifs

Les déterminants de la santé et des inégalités environnementales et sociales de santé sont nombreux et multidimensionnels (Commission on Social Determinants of Health, 2008). En France il existe une grande hétérogénéité de la distribution spatiale de ces déterminants de santé, et celles-ci ne se superposent pas à celles des inégalités de santé qui peuvent être observées (Vigneron, 2013). En raison de cette multidimensionnalité et des possibles interactions entre les déterminants, les inégalités territoriales de santé sont difficiles à caractériser.

La première problématique de cette thèse porte sur la caractérisation des déterminants pertinents pour une évaluation spatialisée des inégalités de santé. Le premier objectif est de proposer une méthodologie d'évaluation des inégalités territoriales de santé qui permet de prendre en compte les multiples facettes des déterminants impliqués en santé environnement. Il s'agit également d'explorer l'influence de la résilience qui capitalise les capacités d'un territoire à faire face aux inégalités de santé et l'influence de la vulnérabilité qui reflète le cumul des nuisances et contraintes auxquelles sont soumises les populations. L'analyse de cette balance résilience / vulnérabilité permettra de définir des typologies de territoires et contribuera à identifier et hiérarchiser des territoires appelant des actions spécifiques en santé environnement.

- ❖ Le développement de cette méthodologie sur différents périmètres (*i.e.* la région des Hauts-de-France, la Communauté Urbaine de Dunkerque) à différentes échelles spatiales (*i.e.* la commune, l'Îlot Regroupé pour l'Information Statistique) permet d'en évaluer la transposabilité. Les résultats de ces travaux sont présentés dans le chapitre I.

Le second objectif de la thèse porte sur la caractérisation de l'influence de l'environnement dans l'étiologie de deux maladies chroniques et multifactorielles : les maladies coronariennes et l'Insuffisance Rénale Chronique Terminale (IRCT). Des études écologiques sont menées sur différents périmètres afin de pouvoir identifier les déterminants environnementaux générateurs d'effets négatifs ou positifs associés à ces maladies sur la santé des populations. Pour ce faire, l'*open data* et les outils SIG seront mis en œuvre pour optimiser la caractérisation des territoires et formuler des hypothèses étiologiques.

- ❖ Le Chapitre II présente l'analyse du lien entre les caractéristiques territoriales et (i) l'incidence de l'insuffisance rénale chronique terminale à l'échelle des communes du Nord et du Pas de Calais et (ii) la survenue des maladies coronariennes à l'échelle des IRIS de la Métropole Européenne de Lille. Les contributions territoriales seront retranscrites par des indices de résilience et de vulnérabilité.

Chapitre I : Evaluation des caractéristiques du territoire

Introduction

L'évaluation des inégalités territoriales de santé est une problématique complexe au regard de la multitude des déterminants à prendre en compte qui varient géographiquement (Ministère des Solidarités et de la Santé, 2022; Watts *et al.*, 2018). Plusieurs méthodologies peuvent être mises en place pour étudier un concept multidimensionnel. La recherche en santé publique à l'aide de l'épidémiologie analyse l'influence des déterminants de santé (biologiques, économiques, environnementaux) sur la santé afin de réduire les inégalités de santé et plus largement pour améliorer la santé des populations. L'épidémiologie descriptive s'attache plutôt à étudier la répartition des pathologies ainsi qu'à identifier le risque d'être malade, ou de le devenir, dans une population donnée (Raimondeau and Carmona, 2020). L'épidémiologie analytique permet quant à elle de formuler des hypothèses étiologiques et d'identifier des facteurs de risque en lien avec la survenue d'un problème de santé. Les principales études d'épidémiologie analytique sont les études de cohorte (*i.e.* suivi de sujets) et les études cas-témoins (*i.e.* identification de facteur(s) de risque par comparaison entre un groupe de malades atteints d'une pathologie et un groupe de non-malades) (Goldberg *et al.*, 2008). Toutefois ces études présentent certaines limites en lien avec des problèmes méthodologiques (*e.g.* biais de sélection, biais de mesure) (Goldberg *et al.*, 2008). Afin de proposer une représentation détaillée de chacune des dimensions et sous-dimensions impliquées dans les inégalités de santé, des tableaux de bord peuvent également être proposés. Néanmoins, aussi précise et circonstanciée que soit la caractérisation obtenue par tableau de bord, cette méthodologie ne permet pas d'obtenir une synthèse des informations incluses (Ceccarelli *et al.*, 2020).

Le développement d'indices composites constitue une des méthodes qui permettent une approche multidimensionnelle par agrégation d'indicateurs individuels (Freudenberg, 2003). Différents travaux ont souligné la pertinence de cette approche pour évaluer un concept multidimensionnel, car en une valeur unique, les indices composites peuvent permettre de retranscrire les multiples facettes des phénomènes (Dialga, 2016; Freudenberg, 2003; Mauro *et al.*, 2021). Cette méthodologie intégrant de multiples paramètres constitue de plus pour les politiques publiques un outil de diagnostic pouvant être décliné aux échelles spatiales les plus fines (Saib, 2015). Les indices composites ont émergé depuis les années 1980 et sont à la base de nombreux travaux qui ont d'abord mis l'accent sur des aspects économiques (HCSP, 2021). Parmi les plus connus, le Produit Intérieur Brut (PIB) vise à mesurer le développement humain au travers de la richesse d'un pays (Daoui, 2018). Néanmoins, cette mesure du développement humain par le PIB porte sur la seule dimension économique et ne permet pas de rendre compte des inégalités sociales et économiques qui peuvent exister entre les différents individus d'une même population (Gadrey et Jany-Catrice, 2016). Le Programme des Nations Unies pour le Développement (PNUD) a initié à partir de 1990 des travaux sur le développement humain qui ont suscité un grand intérêt de la communauté scientifique internationale mais également des responsables politiques (Klugman *et al.*, 2011a). L'Indice de Développement Humain (IDH) développé par le PNUD propose une mesure du développement humain et du bien-être des pays en allant au-delà de la seule croissance économique (Dialga, 2016). L'IDH est mis à jour annuellement et agrège trois indicateurs apportant un élément de caractérisation de la dimension économique (revenus), sanitaire (espérance de vie) et éducative (scolarisation). Des dimensions particulières du développement sont complétées par des indices plus spécifiques que l'IDH (*e.g.* l'Indice de Pauvreté Humaine IPH), ce qui concourt à améliorer la caractérisation des inégalités de développement principalement à l'échelle mondiale (Daoui, 2018).

Afin de pouvoir proposer cet indicateur à plus fine échelle, des déclinaisons régionales (IDH-2) ou communales (IDH-4) ont émergé, les variables à agréger étant adaptées pour être plus spécifiques de

certaines problématiques locales (Marlier *et al.*, 2015). Toutefois, les principaux défauts de l'IDH et de ses déclinaisons à plus fine échelle est (i) qu'il apparaît très synthétique pour pouvoir retranscrire les multiples facettes des déterminants de santé et (ii) qu'il ne prend pas en compte l'environnement (Nourry, 2008). L'Indice Soutenable de Développement Humain (ISDH) a donc été proposé afin de pouvoir amender et compléter l'indice composite du PNUD en intégrant l'environnement. Toutefois, les ISDH et IDH s'avèrent fortement corrélés, les émissions de carbone par habitant utilisées comme indicateur environnemental dans le ISDH étant positivement corrélées aux revenus (Bravo, 2014). En dépit des nombreux travaux sur l'IDH et la reconnaissance de cet indicateur, sa simplicité de construction suscite des interrogations sur sa réelle capacité à retranscrire le bien-être, d'autres dimensions comme par exemple l'équité ou encore la durabilité en étant également exclues (Klugman *et al.*, 2011a).

En santé environnement, les inégalités de santé sont en lien avec les conditions dans lesquelles les populations vivent et travaillent et qui sont elles-mêmes façonnées par des forces sociales, économiques et politiques auxquelles peuvent s'ajouter les conditions environnementales et culturelles (Commission on Social Determinants of Health, 2008; Marmot *et al.*, 2008). De plus, la question des inégalités de santé constitue un défi mondial de santé publique mais également une priorité de politique publique (Regmi et Mudyarabikwa, 2020). Le rapport de 2012 de la Commission Lancet a mis en évidence que des approches systémiques complexes étaient nécessaires pour caractériser les processus multi-niveaux qui influencent la santé des populations (Rydin *et al.*, 2012). Les indices composites en santé environnement peuvent permettre d'agrèger des données en lien avec la multiplicité des expositions ou des effets sanitaires et aider les décideurs dans (i) la hiérarchisation et la priorisation des unités spatiales les unes par rapport aux autres pour la mise en place de politiques spécifiques et (ii) la prise de décisions vis-à-vis d'actions contextualisées de planification urbaine de santé publique (Hat and Stöglehner, 2019 ; HCSP, 2021). Depuis la fin des années 1970, de nombreuses recherches visent à élaborer des indicateurs de défaveur (Rican *et al.*, 2011) pour (i) caractériser l'hétérogénéité spatiale de la santé (Rey *et al.*, 2009), (ii) obtenir une population d'étude avec des caractéristiques socio-économiques particulières (Macintyre and Ellaway, 2000), (iii) être mis en rapport avec des résultats épidémiologiques ou être considérés comme facteur confondant d'un évènement de santé (Saib, 2015).

Une revue exhaustive de la littérature a été menée afin d'apporter une analyse critique de l'utilisation des indices composites dans l'évaluation de l'hétérogénéité spatiale des ressources et des nuisances environnementales qui ont un impact sur la santé. Cette analyse s'est principalement focalisée sur les premières étapes du développement des indices composites afin de pouvoir identifier les « descripteurs » classiquement ou ponctuellement mentionnés dans la littérature pour décrire les différentes dimensions. Dans un second temps, cette analyse a permis d'identifier s'il existe un consensus sur les méthodologies de développement des indices composites car les résultats peuvent être relativement sensibles à ces dernières (HAS, 2010a). Les résultats de cette analyse de la littérature sont présentés dans l'Article 1.

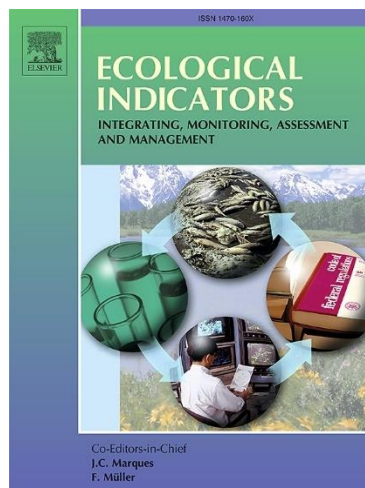
Article 1

Spatialized composite indices to evaluate environmental health inequalities: meeting the challenge of selecting relevant variables

Delphine Brousmiche, Florent Occelli, Michaël Genin, Damien Cuny, Annabelle Deram, Caroline Lanier

Ecological Indicators 111, 106023

<https://doi.org/10.1016/j.ecolind.2019.106023>



Spatialized composite indices to evaluate environmental health inequalities: meeting the challenge of selecting relevant variables

Delphine Brousmiche^a, Florent Occelli^a, Michaël Genin^b, Damien Cuny^a, Annabelle Deram^a, Caroline Lanier^{*a}

^a Univ. Lille, CHU Lille, Institut Pasteur de Lille, EA 4483 IMPECS (IMPact of Environmental Chemicals on human health), F-59000 Lille, France

^b Univ. Lille, EA 2694, Santé Publique épidémiologie et qualité des soins, Centre d'Etudes et Recherche en Informatique Médicale (CERIM), F-59000 Lille, France

*Corresponding author: caroline.lanier@univ-lille.fr

Abstract

The wide range of factors involved in environmental health and the complexity of interactions between all environmental determinants require the validation of multidimensional approaches. While the development of composite indices is receiving growing attention by scientists and public authorities, the concept continues to lack transposability and robustness partly due to varying conceptualizations and/or methodologies. This review aims to promote harmonizing practices governing the first step of development of composite index, namely identification and characterization of the dimensions and variables that are included in environmental health indices. A review of available literature (more than 1,500 studies) was conducted to identify the composite indices developed to assess territorial determinants from an environmental health perspective. This process made it possible to identify 23 spatialized composite indices and to assess a total of 329 variables. This diversity highlights that the absence of a common framework can lead to a strong subjectivity and limit comparisons between different environmental health indices. The specificity and the availability of certain variables would limit the transposability of indices. In light of current knowledge, this review proposes a consolidated methodological framework based on a categorization of variables into dimensions and sub-dimensions related to health, environment, social, economics, services and policy. To characterize the sub-dimensions, several variables are possible and can be chosen according to the availability and/or accessibility of the data. The adaptation of a composite index to a specific territory or to a specific issue would then be effective through the included variables. This also aims to be transposable to any spatial unit (country, region, census tract). This work is a first step towards a proposal of guidelines designed to provide a consensual framework that could facilitate the exploitation of environmental health indices. This transparency could also increase the understanding and adoption of these tools by public authorities and general public.

Keywords

combined environmental pressure ; indicators ; health determinants ; overlapping inequalities ; interterritorial comparisons

Introduction

The WHO first tackled the environmental health issue in 1994 by defining the concept as follows: “Environmental health comprises those aspects of human health, including quality of life, that are determined by physical, chemical, biological, social, psychosocial and aesthetic factors in the environment”. In its definition, WHO considers any health problems that are not caused by genetic factors or due to individual choices. The idea of the environment therefore reflects the different neighborhoods of populations and excludes the effect of individual behaviour and voluntary exposition to factors such as alcohol consumption, smoking or unhealthy dietary habits. Although these behavioural and individual factors are socially distributed, many studies suggest that other factors could also contribute to explaining observed health inequalities (Saib, 2015). The health status of populations is thus strongly influenced by complex interactions between all environmental determinants, which are interconnected and evolve over time and space (Sarkar and Webster, 2017). The environment, considered in this context to be the outdoor living environment, is characterised by a number of determinants such as environmental quality (water, air and ground) and living conditions (e.g. surroundings, access to healthcare) and by socio-economic parameters such as local economic drive, social disadvantage and education. These different environmental determinants can influence the health status of the population in a combined or sequential way. The environment is currently the subject of increasing apprehension and concern (Scarwell *et al.*, 2013), with a particular emphasis on the possible impacts of environmental quality on health. These consequences are difficult to characterize due to the many aspects that can be harmful or beneficial.

The wide range of factors involved in environmental health, the complexity of interactions and their spatio-temporal variabilities all require the validation of multidimensional approaches that would provide more than a simple aggregation of health determinants. One of the recognized methods to explain a complex research problem and follow the evolution of a given phenomenon is based on the creation of composite indices. A composite index is a mathematical combination of variables reflecting one or more selected dimensions that are usually evaluated separately (Nardo *et al.*, 2005; Nascimento and Carrage, 2007). It derived from quantitative or qualitative measurements obtained from field observations (Freudenberg, 2003). In light of the wide range of methodological approaches, the OECD has developed a guidance manual for evaluators, non-specific to environmental health issues (European Commission *et al.*, 2008). This manual confirms the interest of grouping together different variables that measure the multiple facets of the phenomenon and provide a picture of it. A spatial approach, that is the calculation of the composite index for each spatial unit, can be added to provide a further level of information (Beale *et al.*, 2008; Flacke, 2015). A wider availability of spatialized data for the environment, health and population has led to an increase in the number of spatialized epidemiological studies. Composite indices are increasingly recognized as a useful measurement tool for not only etiological approaches but also strategy and policy development, and public communication by institutions (Cutter *et al.*, 2010; Saib *et al.*, 2015; Saisana and Cartwright, 2007).

Although an increasing number of studies in the literature show interest in composite indices and report on the growing demand for environmental health variables, a number of questions remain to be answered, particularly with respect to methodological issues (WHO, 1999). Informative reviews have been published on specific aspects of methodological constructions and concentrate on mathematical aspects. The findings of these reviews provide guidelines for future environmental composite indices, mainly in terms of the overall uncertainties linked to each step of the methodology. Many focused on weighting and/or aggregation (Becker *et al.*, 2017; Gan *et al.*, 2017; Greco *et al.*, 2018; Wiréhn *et al.*, 2015). Becker *et al.* (2017) highlighted the complexity of including weighting in composite indices and presented tools to help developers investigate the effects of weights. The main

aim of these works was to propose a methodological framework or a means to evaluate the qualitative aspects of the methodology used to develop composite indices. However, no studies to date have considered the potential inclusion of data, as a carrier of information, for use in composite environmental health indices. There is no recognized standard or internationally recognized rule to determine the number and type of variables to be included in the quantification of environmental health (He *et al.*, 2018). The design of composite indices in environmental health has to be optimized and induced variability must be mastered (Burgass *et al.*, 2017).

To this end, this study analyses the literature to identify and characterize the dimensions and variables that are included in environmental health indices. Our analysis will highlight (i) the diversity of the published composite indices, but also (ii) the diversity of variables used in these composite indices. For this purpose, a literature review on the 23 main papers dealing with composite indices in environmental health was conducted. This work is a first step towards a proposal of guidelines designed to provide a consensual framework that facilitates the exploitation of environmental health indices. Their transposability and comparability regardless of the spatial scale were considered as mandatory criteria. The availability or accessibility of data have also been given a special attention.

Materials and methods

Methodology for the selection of publications

An analytical framework was developed to ensure the exhaustive nature of this review. The article selection methodology took place in two stages (Figure 1). First, a search with the keywords “environmental health” and “composite indices” was conducted on scientific databases (Wiley, ScienceDirect, Google Scholar, Web of Science). In a second step, non-academic searches were carried out on the Google search engine to find methodologies proposed by public authorities, which were not indexed in the scientific databases. The present review summarizes the data published before December 2018. Special attention has been given to the use of composite indices to assess the possible spatial heterogeneity of exposure to the beneficial and harmful aspects of the environment (Figure 1).

At the time this paper was written, a meticulous screening of the 1,551 initial studies was carried out. The exclusion criteria were (i) studies with an abstract but no provision of the full article, (ii) studies that developed non-spatialized indices, (iii) composite indices that do not include both health and environmental data and (iv) composite indices that included behavioural variables.

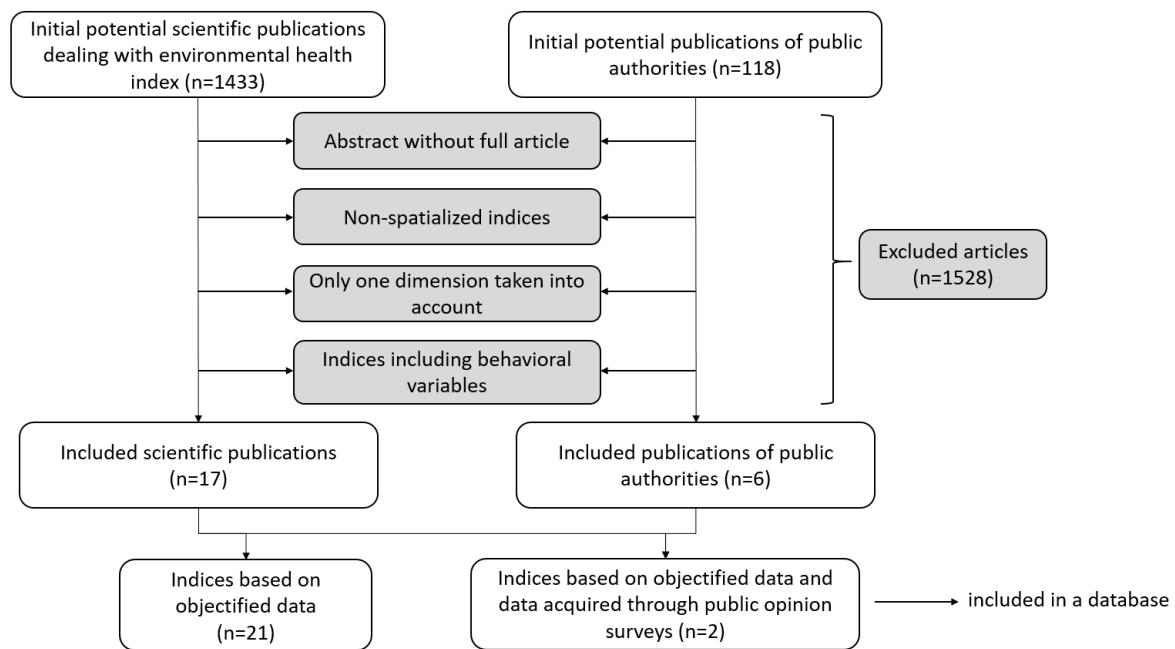


Figure 1 : Article 1 – Figure 1: Results of the article search and outline of the entire article searching process for this systematic review

23 spatialized composite indices tested on different geographical perimeters or specific countries met the inclusion criteria for our analysis. 73.9 % of composite indices were originating from scientific teams, and 26.1 % from public authorities. 91.3 % were constructed using objectified data from scientific measurements, 8.7 % based on objectified data and from public opinion data collected through questionnaire surveys.

Classification of variables within dimensions

A variable, which can also refer to an indicator in some studies although these two terms do not overlap (Maggino, 2017), is linked to an observable phenomenon to which different measured values can be attributed. A variable is a basic element permitting the understanding of a dimension. Several variables may be needed to obtain a complete characterization of a dimension.

The dimensions partly reflect the determinants of health presented by Dahlgren and Whitehead (1991), namely environmental, social, economic, housing and educational conditions, all recognized as playing a role in health inequalities. These dimensions are not all systematically found within the composite indices. The authors generally present a classification of variables within each dimension before building the composite index. We observed that a given variable may be classified in different dimensions by different authors, because of a possible overlapping between dimensions.

In order to provide an unambiguous reading of the distribution of variables within composite index dimensions, an iterative process was used to categorize variables, based on the principle of greater proximity between the information provided. Thus, the variables that are thematically related have been grouped together within the same dimension and the analysis led to a six-dimensional organization. The dimensions names (health, environment, social determinants, economics, services and policy) are the most frequently mentioned in the literature and are in accordance with the work of (i) Dahlgren and Whitehead (1991) on the determinants of health and (ii) Stiglitz *et al.* (2009) on the quality of life.

The health dimension will include all variables related to health in general, but will also incorporate information about resources that can be used to solve public health events in the study area. The environment dimension will group together all aspects of environmental quality, the natural characteristics of the spatial unit in question or the environmental hazards to which the populations may be exposed. The social factors dimension will include any variable that characterizes the population in terms of its integration, structure, sensitivity and social network, but will also consider educational variables, as education is important for the quality of life regardless of its effects on income (Stiglitz *et al.*, 2009). The economics dimension contains any variable that characterizes employability, economic drive and influence, and also the incomes and inequalities of populations in the geographic area considered. The services dimension will include all variables that are linked to regional development, population mobility and access to basic necessities and will also incorporate variables for accommodation, a key factor in quality of life for populations (Lejeune *et al.*, 2012). Finally, the policy dimension will group together any measures concerning the civil rights and liberties, environmental governance and investments made in the spatial units considered in these studies.

Overview of the spatialized composite indices in environmental health

Temporal dynamics

The use of spatialized composite indices is relatively recent. The first index to be recognized was the Human Development Index (HDI) (UNDP, 1990), which was developed in order to “direct the developmental economy through policies that are based on human capital rather than an evaluation of national revenue” (WHO, 2012a). It is one of the most widely used composite indices (Biggeri and Mauro, 2018) and aims to measure the level of development of countries by considering: 1) health, measured in terms of life expectancy at birth, 2) education, evaluated with the average duration of scolarity, and 3) lifestyle, as proxied by the average gross income per capita. However, the use of this index is considered to be questionable for the management of natural resources and insufficient for human needs (Lasso de la Vega and Urrutia, 2001).

Environmental concerns progressively became a central element of socio-economic conditions, leading to suggest that different environmental variables such as pollution or energy consumption should be added to the HDI in order to improve it with a dimension for sustainable development (Lasso de la Vega and Urrutia, 2001; Togtokh and Gaffney, 2010). In environmental health, interest began in 2001 and the vast majority of composite indices which took both health and environment into account were developed after 2009. One or more indices are repeatedly observed in the literature every year. On how the harmful and beneficial impacts are taking into account, two types of index construction are revealed: (i) unique indices that combine variables of negative impact and environmental benefits (Braconnier *et al.*, 2011; Kerényi, 2011; Pearce *et al.*, 2010; Pearson *et al.*, 2013; Stewart *et al.*, 2014; Zuideau and Lipovac, 2009), or (ii) two independently constructed sub-indices that permit the comparison of beneficial or harmful environmental issues (Estoque and Murayama, 2014) or more generally any determinants that have a positive or negative impact on territorial inequalities (Miller *et al.*, 2017). Figure 2 shows the dynamics of publications based on environmental health composite indices.

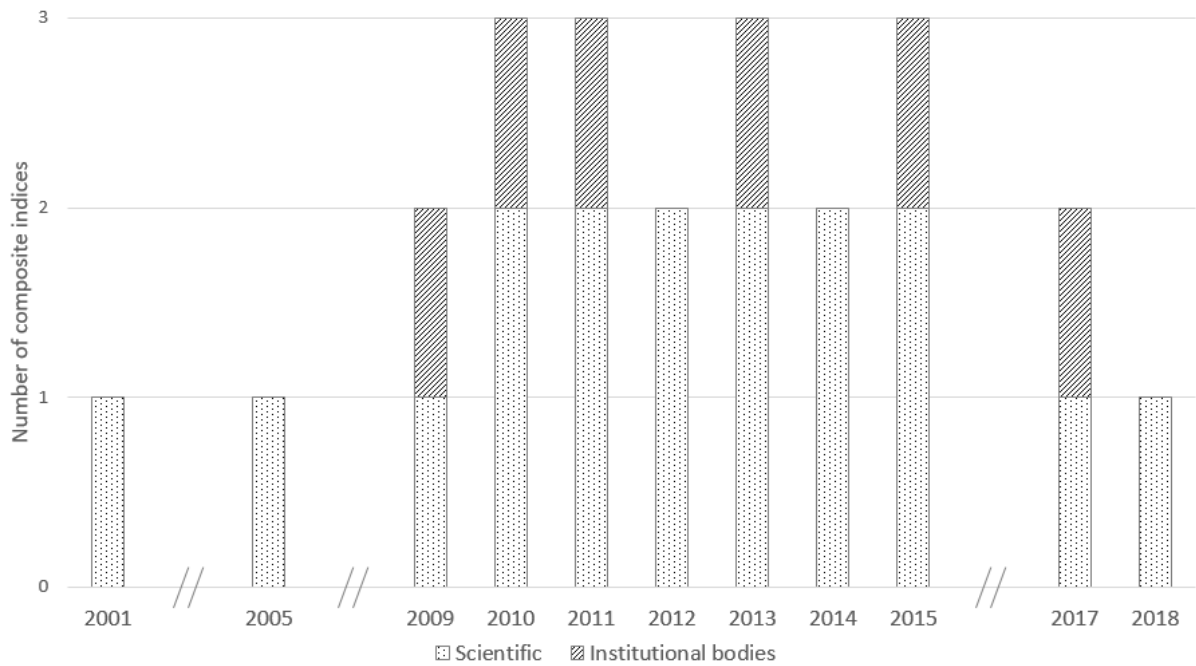


Figure 2 : Article 1 – Figure 2: Composite indices in environmental health, scientific vs institutional bodies, published since 2001

Structuration, aim and exploitability of composite indices

Since 2001, environmental health composite indices have been developed by either the scientific community or by institutional bodies, and have been improved with the addition of different variables that make it possible to consider (i) the main criticisms of the HDI, and (ii) the multiple factors that play a role in environmental health. Figure 3 illustrates this diversity of information, showing the number of variables included in each dimension, as indicated by numbers in each line on the graph. It does not take into account the effect of a possible weighting (between the variables within a dimension) in the methodologies developed by the authors.

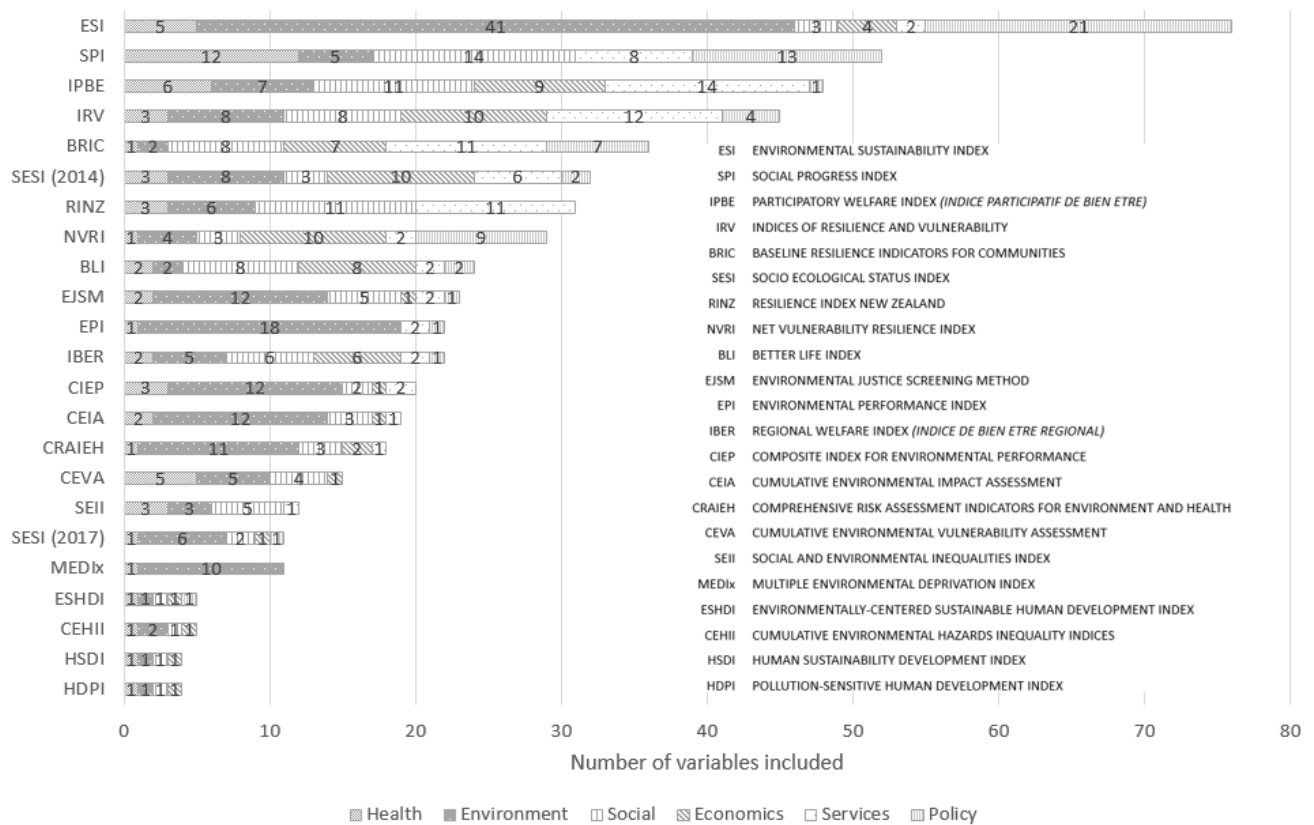


Figure 3 : Article 1 – Figure 3: Representativeness of the number of variables within dimensions

5 of the 7 indices that are composed of more than 30 variables describe all six dimensions. The RINZ does not account for economic and political variables (Pearson *et al.*, 2013), and the SPI does not include economic variables (Porter *et al.*, 2013).

The “health” and “environment” dimensions are systematically present, as required by the perimeter of this literature review. The “social” dimension, including educational data, was found in over 90 % of all the indices studied, while the “economics” or “services” dimensions that included accommodation-linked data were found in 78 and 70 % of all the indices studied, respectively. The “policy” dimension was only found in 56 % of all the indices analysed. The number of variables used to describe a given dimension differs from one study to another and the conclusions of the authors did not find this number to be related to the importance of the dimension within the index.

The variables that authors include in these dimensions are generally selected to solve the specific research problematic for which the index was designed, without the use of specific methodological framework that is reliable or recognized for the creation of composite indices for use in environmental health. Consequently, each of these indices is situation specific and is difficult to transpose and/or generalize.

Table 1 summarizes the dimensions of each of the 23 composite indices, and specifies the number of variables for each dimension concerned. The table also specifies the country where each index was created, the spatial unit concerned, the aim of the study and the exploitability of the results. The indices marked with an asterisk were developed by agencies or institutional bodies, the others were described in the scientific literature.

Tableau I : Article 1 - Table I: Synthesis of studied indices

Indices	Identified dimension	Number of variables	Country of origin-spatial unit of study	Weighting of dimensions within index	Aim Exploitability	Reference
Pollution-sensitive Human Development Index (HDPI)	Health	1	Spain. 165 countries	None	Evaluation and ranking of countries. The HDPI penalizes countries that achieved strong economic growth to the detriment of the environment.	Lasso de la Vega and Urrutia (2001)
	Environment	1				
	Social	1				
	Economics	1				
Environmental Sustainability Index (ESI)	Health	5	USA, Switzerland. 146 countries	None	Evaluation of a country's capacity to efficiently preserve its environmental resources over several decades and rank countries. A low score results from low performances for a high number of variables. A decision-making tool.	Etsy <i>et al.</i> (2005)
	Environment	41				
	Social	3				
	Economics	4				
	Services	2				
	Policy	21				
Cumulative Environmental Hazards Inequality Indices (CEHII)	Health	1	USA. Census tract (Los Angeles county)	Population weighted	Comparison of additive and multiplicative methods to characterize inequalities and cumulative environmental hazards.	Su <i>et al.</i> (2009)
	Environment	2				
	Social	1				
	Economics	1				
Indice Participatif de Bien Être (IPBE)* Participatory welfare index	Health	6	France. Three spatial units (France, Nord Pas de Calais region, agglomeration)	Budget Allocation Process (dimension with the strongest weighting: employment or economy)	Spatial comparison and positioning of the results for evaluation of well-being in the three spatial units. Possibility to monitor over time. Also designed to improve communication and alert local authorities.	Zuindeau and Lipovac (2009)
	Environment	7				
	Social	11				
	Economics	9				
	Services	14				
	Policy	1				
Human Sustainability Development Index (HSDI)*	Health	1	Mongolia, Sweden. 169 countries	None	Evaluation and ranking of countries. Like the HDPI, the HSDI penalizes countries that achieved strong environmental growth to the detriment of the environment. Strongly correlated with the HDI (Bravo, 2014).	Togtokh and Gaffney (2010)
	Environment	1				
	Social	1				
	Economics	1				
Baseline Resilience Indicators for Communities (BRIC)	Health	1	USA. 736 counties (Federal Emergency Management Agency IV)	None	Monitoring through time and comparison of the different entities. An increase in the index corresponds to a higher resistance to disasters through individual resources (resilience). Definition of priorities, measurement of progress, help in decision-making processes.	Cutter <i>et al.</i> (2010)
	Environment	2				
	Social	8				
	Economics	7				
	Services	11				
	Policy	7				
Multiple Environmental Deprivation Index (MEDix)	Health	1	United Kingdom. 10,654 census area statistics wards (United Kingdom)	Population weighted	Study of how multiple environmental disadvantages impact health inequality. The index, and thus the inequalities, increase with the north/south gradient.	Pearce <i>et al.</i> (2010)
	Environment	10				

Tableau I : Article 1 - Table I: Synthesis of studied indices (cont.)

Indices	Identified dimension	Number of variables	Country of origin – spatial unit of the study	Weighting of dimensions within index	Aim Exploitability	Reference
Indice de Bien Être Régionale (IBER) Regional welfare index	Health	2	France. 22 regions	None	Evaluation and ranking of regions. Identification of the dimensions with the weakest results. Comparison of the IBER with other indices (HDI, SHI et GDP per capita). IBER penalizes the regions with contradictory results between dimensions.	Braconnier <i>et al.</i> (2011)
	Environment	5				
	Social	6				
	Economics	6				
	Services	2				
Policy	1					
Environmental Justice Screening Method (EJSM)	Health	2	USA. Census tracts (Southern California)	None	Evaluation of the cumulative effects of social and environmental stress factors, shown by a high score; ranking of spatial units to enable decision-makers to find more efficient solutions for cumulative negative impacts.	Sadd <i>et al.</i> (2011)
	Environment	12				
	Social	5				
	Economics	1				
	Services	2				
Policy	1					
Better Life Index (BLI)*	Health	2	34 countries	None	Comparison of the performances of the different countries after adding further variables linked to well-being.	Kerényi (2011)
	Environment	2				
	Social	8				
	Economics	8				
	Services	2				
Policy	2					
Cumulative Environmental Vulnerability Assessment (CEVA)	Health	5	USA. Census block group (California's San Joaquin Valley)	None	Study of cumulative impacts designed to help authorities by concentrating on units that have a high index, <i>i.e.</i> those with the highest levels of vulnerability (in terms of environmental hazards and social vulnerability as a whole).	Huang and London (2012)
	Environment	5				
	Social	4				
	Economics	1				
Social and Environmental Inequalities Index (SEII)	Health	3	Portugal. 15 neighbourhoods (Porto)	Analytic Hierarchy Process (heaviest weighting: health)	Study of the relations between urban conditions, lifestyles and health events designed to identify the priorities for a strategy to improve the quality of urban life. A high index indicates high inequalities.	Monteiro <i>et al.</i> (2012)
	Environment	3				
	Social	5				
	Services	1				
Resilience Index New Zealand (RINZ)	Health	3	New Zealand, United Kingdom. 1,919 census area units (New Zealand)	None	Identification and quantification for each spatial unit of the characteristics that are linked to resilience, which is strongly associated with the accessibility of certain structures. Can be used by authorities to improve living conditions.	Pearson <i>et al.</i> (2013)
	Environment	6				
	Social	11				
	Services	11				
Environmental Performance Index (EPI)	Health	1	USA, Abu Dabi. 132 countries	Experts adjust according to the relative importance of certain dimensions (heaviest weighting: ecosystem vitality).	Ranking of countries according to performance variables for all categories; and the performances of a given country are measured in relation to the goals.	Hsu <i>et al.</i> (2013)
	Environment	18				
	Services	2				
	Policy	1				

Tableau I : Article 1 - Table I: Synthesis of studied indices (cont.)

Indices	Identified dimension	Number of variables	Country of origin-spatial unit of study	Weighting of dimensions within index	Aim Exploitability	Reference
Social Progress Index (SPI)*	Health	12	USA. 50 countries	None	Measure performance, facilitate improvement, and seek similarities and differences between countries. Compared with other indices, compensation between dimensions changes the ranking. It seeks to prioritize investments and strengthen the ability to implement life improvement strategies.	Porter <i>et al.</i> (2013)
	Environment	5				
	Social	14				
	Services	8				
	Policy	13				
Socio-Ecological Status Index (SESI)	Health	3	Japan. 81 provinces (Philippines)	None	Ranking and comparison of provinces according to a socio-ecological variable that considers the interactions between humans and the environment without favouring either of these dimensions (low correlation with HDI and HSDI). Ranking of the provinces for all dimensions and sub-dimensions.	Estoque and Murayama (2014)
	Environment	8				
	Social	3				
	Economics	10				
	Services	6				
	Policy	2				
Cumulative Environmental Impact Assessment (CEIA)	Health	2	USA. 372 census tracts (Santa Clara County)	None	Evaluation of the links between social vulnerability, exposition to environmental hazards and access to amenities. A high index indicates areas of high environmental hazards associated with low access to amenities. The results can be used to reform public health policies.	Stewart <i>et al.</i> (2014)
	Environment	12				
	Social	3				
	Economics	1				
	Services	1				
Composite Index for Environmental Performance (CIEP)	Health	3	Spain. 152 countries	Criteria Importance Through Inter-Criteria Correlation (highest weighting: environment).	Ranking of countries and analysis of their performances. A high index indicates better environmental quality (influenced by growth and population density). Monitoring over time is possible. Facilitates decision-making for environmental policies.	García-Sánchez <i>et al.</i> (2015)
	Environment	12				
	Social	2				
	Economics	1				
	Services	2				
Comprehensive Risk Assessment Indicators for Environment and Health (CRAIEH)*	Health	1	China. 58 counties (Jiangsu province, Eastern China)	Multi-Criteria-Decision-Analysis (highest weighting: pressure and state indices)	Evaluation of the risks linked to health and the environment across a wide geographical area. A high index indicates high-risk zones that are affected by industrialization and population density. Facilitates the creation of environmental and public health policies.	Zhang <i>et al.</i> (2015)
	Environment	11				
	Social	3				
	Economics	2				
	Services	1				
Net Vulnerability Resilience Index (NVRI)	Health	1	France. 95 countries	None	Comparison of countries and identification of their individual strengths and weaknesses. The final index combines all the information that reveals a country's ability to create resilience. A tool to provide advice for policy needs.	Angeon and Bates (2015)
	Environment	4				
	Social	3				
	Economics	10				
	Services	2				
	Policy	9				

Tableau 1 : Article 1 - Table 1: Synthesis of studied indices (final section)

Indices	Identified dimension	Number of variables	Country of origin – spatial unit of study	Weighting of dimensions within index	Aim Exploitability	Reference
Socio-Ecological Status Index (SESI)	Health	1	Japan. 144 countries	None	Ranking and comparison of counties according to a socio-ecological variable that considers the interactions between humans and the environment. An increase in the index indicates a higher status. Applicable and adaptable on a scale from a province/region to an entire country.	Estoque and Murayama (2017)
	Environment	6				
	Social	2				
	Economics	1				
	Policy	1				
Index of Resilience and Vulnerability*	Health	3	USA. 3,077 counties	None	Identification, measurement and comparison through the use of a resilience / vulnerability quadrant. The spatial units are attributed to one of the four quadrants according to their resilience and vulnerability scores. Facilitates the elaboration of programmes to manage human and natural threats.	Miller <i>et al.</i> (2017)
	Environment	8				
	Social	8				
	Economics	10				
	Services	12				
	Policy	4				
Environmentally-centered Sustainable Human Development Index (ESHDI)	Health	1	Italy. 50 countries	None	Ranking of countries. Results are similar for homogenous countries (the countries with the highest CO ₂ emissions are the most penalized); comparison of this index with the HDI and comparison of the new construction method that penalize heterogeneity with other aggregation methods.	Biggeri and Mauro (2018)
	Environment	1				
	Social	1				
	Economics	1				
	Policy	1				

These 23 composite indices can be differentiated in several ways, thus shedding light on their main strengths and limitations.

First, the number of variables included impacts the complexity of the composite indices. By providing an additional environmental dimension to the three initial dimensions of the HDI, Biggeri and Mauro (2018), Lasso de la Vega and Urrutia (2001) and Togtokh and Gaffney (2010) developed an index construction methodology that is close to the reasoning behind the HDI, *i.e.* one variable per dimension. This “minimalism” makes it possible to (i) improve the accessibility and/or availability of variables that give opportunity of comparability across spatial units and across years, and (ii) simplify the reading and understanding of measured phenomena. Etsy *et al.* (2005), Miller *et al.* (2017), Porter *et al.* (2013) and Zuideau and Lipovac (2009) integrated more than 40 variables within their composite index. Their general objective was therefore to improve the representativeness of the index by including a maximum of variables in order to characterize each dimension as well as possible. These authors incorporated issues that were local in scope as well as those that were global in scale. However, a large number of variables increases the probability that variables will not be available for all the spatial units of the study area and the risk of strong inter-variable correlation (Miller *et al.*, 2017), which can complicate the development and interpretation of an index.

The authors’ choice of construction methodology varies according to the spatial unit that is being considered. 10 indices were created on a nationwide scale, while 13 others considered a smaller spatial unit. Of these 13, 4 considered an intermediate spatial unit (county, province, region). The 9 other composite indices were created for “small” spatial scales, *i.e.* census tracts or neighborhoods. It has been shown that the use of a composite index must provide relevant information, regardless of spatial granularity (Zuideau and Lipovac, 2009).

The assignment of weights is not systematic. In the case where each dimension is weighted equally and the number of variables is different within these dimensions, the integrated variables do not necessarily have the same importance in the final composite index (Greco *et al.*, 2018). In addition, if two colinear variables are integrated within the same dimension, this dimension will therefore have a greater weight due to the phenomenon of “double counting” (Greco *et al.*, 2018). Becker *et al.* (2017) underlined the importance of correlation phenomena on the assignment of weights, which can be optimized to reflect the desired importance of a dimension. 16 of the 23 indices did not assign weights during the aggregation process. This lack of weighting is justified in two ways: (i) this method is transparent and easy to understand for potential users, and (ii) there is no adequate scientific justification for a dimension having more weight than another (Cutter *et al.*, 2010; Estoque and Murayama, 2014; Porter *et al.*, 2013). For Burgass *et al.* (2017), this lack of knowledge concerning the importance of one dimension with regard to another, or the reticence to prioritize one dimension over another, even if the former is considered neutral, are in themselves a weighting decision.

The composite indices analysed could also be distinguished according to the authors’ recommendations as for the exploitation of their research results. In 13 of the 23 composite indices, the authors consider that the results could be used for the development of public health or environmental policies. They also underline that their indices are developed to meet the needs of decision makers in terms of tools to identify, understand and monitor the phenomena occurring in their territory, but also to allocate investments (Etsy *et al.*, 2005). Of the 13 indices, 4 were developed by agencies or institutions, of which only 2 mention a possible direct exploitation of results (Miller *et al.*, 2017; Zhang *et al.*, 2015). This low number can be explained in two ways: either the composite index developed was more an “experiment” requiring further improvement in the light of identified limitations, or the index was developed by an agency that was not in a position to implement new public and/or environmental health policies.

Identification of dimensions and associated variables

The analysis of the 23 indices made it possible to highlight the existence of sub-dimensions. Each sub-dimension aims to bring an additional level in the description of a dimension. The variables used to characterize each sub-dimension were justified by the spatial scale of the study, the level of development, the availability or the accessibility of data. Table 2 presents examples of variables from the 23 studied articles, these have been classified in the appropriate sub-dimensions.

Tableau II : Article 1 – Table II: Characterization of sub-dimensions and distribution of variables within dimensions

Dimension	Sub-dimensions	Number of identified variables	Examples of variables	References
Health	General population health	9	Life expectancy at birth, at 60 and at 65 years of age	Biggeri and Mauro (2018), Braconnier <i>et al.</i> (2011), Estoque and Murayama (2014, 2017), García-Sánchez <i>et al.</i> (2015), Kerényi (2011), Lasso de la Vega and Urrutia (2001), Miller <i>et al.</i> (2017), Porter <i>et al.</i> (2013), Togtokh and Gaffney (2010), Zuideau and Lipovac (2009)
			General death rate	Pearce <i>et al.</i> (2010), García-Sánchez <i>et al.</i> (2015)
			Children death rate	Hsu <i>et al.</i> (2013), Etsy <i>et al.</i> (2005), García-Sánchez <i>et al.</i> (2015), Porter <i>et al.</i> (2013)
			Birth rate	Zuideau and Lipovac (2009)
			Fertility rate	Etsy <i>et al.</i> (2005)
	Environmental health indicators or complementary indicators of general public health	18	Rate of low weight at birth	Estoque and Murayama (2014), Huang and London (2012), Miller <i>et al.</i> (2017), Sadd <i>et al.</i> (2011), Stewart <i>et al.</i> (2014), Su <i>et al.</i> (2009)
			Rates of obesity	Porter <i>et al.</i> (2013)
			Death due to cardiovascular disease	Porter <i>et al.</i> (2013)
			Death due to HIV	Porter <i>et al.</i> (2013)
			Chronic Obstructive Pulmonary Disease rate	Monteiro <i>et al.</i> (2012)
	Health care facilities	8	Rate of hospitalisation for asthma for 0-19 year-olds	Huang and London (2012), Estoque and Murayama (2014)
			Number of hospital beds or General Practitioners	Braconnier <i>et al.</i> (2011), Cutter <i>et al.</i> (2010), Miller <i>et al.</i> (2017), Pearson <i>et al.</i> (2013), Zhang <i>et al.</i> (2015), Zuideau and Lipovac (2009)
			Access to pharmacies	Pearson <i>et al.</i> (2013)
			Accessibility of all possible medical specialities	Estoque and Murayama (2014), Huang and London (2012), Miller <i>et al.</i> (2017), Pearson <i>et al.</i> (2013), Porter <i>et al.</i> (2013)
Accessibility of health care specialities			Angeon and Bates (2015), Estoque and Murayama (2014), Porter <i>et al.</i> (2013)	
Environment	Anthropogenic pressure	16	Greenhouse gas emissions	Angeon and Bates (2015), Biggeri and Mauro (2018), Estoque and Murayama (2017), Etsy <i>et al.</i> (2005), García-Sánchez <i>et al.</i> (2015), Hsu <i>et al.</i> (2013), Lasso de la Vega and Urrutia (2001), Porter <i>et al.</i> (2013), Togtokh and Gaffney (2010)
			Ecological footprint	Estoque and Murayama (2017), Etsy <i>et al.</i> (2005), Porter <i>et al.</i> (2013)
			Use of fertilizers and pesticides per hectare of arable land	Etsy <i>et al.</i> (2005), García-Sánchez <i>et al.</i> (2015), Huang and London (2012), Stewart <i>et al.</i> (2014)
	Natural resources	23	Fresh and/or underground water resources	Etsy <i>et al.</i> (2005), Zhang <i>et al.</i> (2015)
Closeness to green spaces, parks and open spaces			Monteiro <i>et al.</i> (2012), Pearce <i>et al.</i> (2010), Pearson <i>et al.</i> (2013), Stewart <i>et al.</i> (2014)	
Access to energy or the production			Angeon and Bates (2015), Etsy <i>et al.</i> (2005), García-Sánchez <i>et al.</i> (2015), Hsu <i>et al.</i> (2013), Porter <i>et al.</i> (2013)	
Area of agricultural land			García-Sánchez <i>et al.</i> (2015), Zuideau and Lipovac (2009)	

Tableau II : Article 1 – Table II: Characterization of sub-dimensions and distribution of variables within dimensions (cont.)

Dimension	Sub-dimensions	Number of identified variables	Examples of variables	References
Environment	Biodiversity	10	Number of endangered species (amphibians, birds and mammals)	Etsy <i>et al.</i> (2005)
			Monetary value for ecosystem services	Estoque and Murayama (2014)
			Proportion or density, of the area that included protected zones	Angeon and bates (2015), Estoque and Murayama (2014, 2017), Etsy <i>et al.</i> (2005), García-Sánchez <i>et al.</i> (2015), Hsu <i>et al.</i> (2013), Zuindeau and Lipovac (2009)
			Natura 2000 classified land, total area of ecologically rich land	Braconnier <i>et al.</i> (2011), Zuindeau and Lipovac (2009)
			Total land artificialisation	Braconnier <i>et al.</i> (2011)
	Water quality	10	Concentration of dissolved oxygen	Etsy <i>et al.</i> (2005)
			Electric conductivity	Etsy <i>et al.</i> (2005)
			Phosphorus concentration	Etsy <i>et al.</i> (2005)
			concentration of suspended solids	Etsy <i>et al.</i> (2005)
			Average nitrate concentration	Zuindeau and Lipovac (2009)
			Proportion of good quality water	Braconnier <i>et al.</i> (2011), Kerényi (2011), Pearson <i>et al.</i> (2013)
			Percentage of the population that had access to a water network served by water systems with at least one health-based violation	Miller <i>et al.</i> (2017)
			Data pertaining to the organic water pollutants	Etsy <i>et al.</i> (2005), Porter <i>et al.</i> (2013)
			Areas that suffer from water scarcity	Hsu <i>et al.</i> (2013)
			Air quality	26
	SO ₂ : emissions and/or concentrations, population weighted, per unit of area, exportations	Etsy <i>et al.</i> (2005), Hsu <i>et al.</i> (2013), Zhang <i>et al.</i> (2015)		
	PM _{2.5} , PM ₁₀ , TSP : average annual concentration, population weighted or not	Etsy <i>et al.</i> (2005), Hsu <i>et al.</i> (2013), Kerényi (2011), Pearce <i>et al.</i> (2010), Pearson <i>et al.</i> (2013), Sadd <i>et al.</i> (2011), Su <i>et al.</i> (2009)		
	Weighted annual average of the ATMO index	Zuindeau and Lipovac (2009)		
	Average number of days per year in areas where ozone exceeds the safe threshold for human health	Braconnier <i>et al.</i> (2011), Sadd <i>et al.</i> (2011)		
	Indoor air pollution	Etsy <i>et al.</i> (2005), Hsu <i>et al.</i> (2013)		
	Carbon monoxide CO	Pearce <i>et al.</i> (2010)		
	Volatil Organic Compounds (VOC)	Etsy <i>et al.</i> (2005)		

Tableau II : Article 1 – Table II: Characterization of sub-dimensions and distribution of variables within dimensions (cont.)

Dimension	Sub-dimensions	Number of identified variables	Examples of variables	References
Environment	Environmental safety or environmental risks	21	Percentage of the population living less than 10 miles from a nuclear plant	Miller <i>et al.</i> (2017)
			Proximity measures for different pollution sources (industries, hazardous waste sites, transport infrastructures, etc.)	Braconnier <i>et al.</i> (2011), Huang and London (2012), Miller <i>et al.</i> (2017), Monteiro <i>et al.</i> (2012), Pearce <i>et al.</i> (2010), Sadd <i>et al.</i> (2011), Stewart <i>et al.</i> (2014), Zuideau and Lipovac (2009)
			Models of noise from different sources (road traffic, industry, ...)	Zuideau and Lipovac (2009)
			Natural risks (flooding and earthquakes, amongst others)	Cutter <i>et al.</i> (2010), Estoque and Murayama (2014), García-Sánchez <i>et al.</i> (2015), Miller <i>et al.</i> (2017)
Social	Social network	5	Number of associations or the proportion of the population involved in these associations, social defense organizations	Braconnier <i>et al.</i> (2011), Cutter <i>et al.</i> (2010), Miller <i>et al.</i> (2017), Zuideau and Lipovac (2009)
	Integration and anomie	21	Percentage of households that are considered linguistically isolated	Cutter <i>et al.</i> (2010), Huang and London (2012), Miller <i>et al.</i> (2017), Sadd <i>et al.</i> (2011), Stewart <i>et al.</i> (2014)
			Suicide rates, comparative suicide mortality index	Braconnier <i>et al.</i> (2011), Zuideau and Lipovac (2009)
			Violent crimes rates, homicide or criminality rate, number of police officers per 10,000 inhabitants	Braconnier <i>et al.</i> (2011), Estoque and Murayama (2014), Kerényi (2011), Miller <i>et al.</i> (2017), Porter <i>et al.</i> (2013), Zuideau and Lipovac (2009)
			Data related to ethnicity	Huang and London (2012), Pearson <i>et al.</i> (2013), Sadd <i>et al.</i> (2011), Su <i>et al.</i> (2009), Zuideau and Lipovac (2009)
		Data related to immigration or migratory history	Cutter <i>et al.</i> (2010)	
	Population sensitivity	6	Age-related data	Cutter <i>et al.</i> (2010), Huang and London (2012), Miller <i>et al.</i> (2017), Monteiro <i>et al.</i> (2012), Sadd <i>et al.</i> (2011), Stewart <i>et al.</i> (2014), Zhang <i>et al.</i> (2015)
		Part of the population with and without disabilities	Cutter <i>et al.</i> (2010), Miller <i>et al.</i> (2017)	
	Education level	18	Literacy or scolarity rate	Biggeri and Mauro (2018), Cutter <i>et al.</i> (2010), Lasso de la Vega and Urrutia (2001), Porter <i>et al.</i> (2013), Sadd <i>et al.</i> (2011), Togtokh and Gaffney (2010)
			Number of individuals that have continued to higher education and/or obtained further qualifications	Braconnier <i>et al.</i> (2011), Cutter <i>et al.</i> (2010), Huang and London (2012), Kerényi (2011), Zuideau and Lipovac (2009)
			Presence or distance to childcare facilities	Cutter <i>et al.</i> (2010), Estoque and Murayama (2014), Pearson <i>et al.</i> (2013), Zuideau and Lipovac (2009)
			Presence or distance to schools, state schools and /or higher education establishments	Pearson <i>et al.</i> (2013), Porter <i>et al.</i> (2013)
PISA score (Programme for International Student Assessment)			Kerényi (2011)	

Tableau II : Article 1 – Table II: Characterization of sub-dimensions and distribution of variables within dimensions (cont.)

Dimension	Sub-dimensions	Number of identified variables	Examples of variables	References
Social	Social resources	3	Access to contraception, childcare	Pearson <i>et al.</i> (2013), Porter <i>et al.</i> (2013)
			Proportion of the population that received public aid	Braconnier <i>et al.</i> (2011), Monteiro <i>et al.</i> (2012)
	Urbanization	4	Rural / urban status	Pearson <i>et al.</i> (2013), Zhang <i>et al.</i> (2015)
			Population density	Estoque and Murayama (2014, 2017), García-Sánchez <i>et al.</i> (2015), Monteiro <i>et al.</i> (2012), Pearson <i>et al.</i> (2013)
			Number of accidents or victims of road accidents	Braconnier <i>et al.</i> (2011), Zuideau and Lipovac (2009)
		Urban population growth	Etsy <i>et al.</i> (2005), García-Sánchez <i>et al.</i> (2015), Pearson <i>et al.</i> (2013)	
Economics	Employability	8	Employment and/or unemployment rate	Angeon and Bates (2015), Braconnier <i>et al.</i> (2011), Cutter <i>et al.</i> (2010), Estoque and Murayama (2014), Kerényi (2011), Miller <i>et al.</i> (2017), Zuideau and Lipovac (2009)
			Proportion of unstable contracts, insecurity rate of labour market	Estoque and Murayama (2014), Kerényi (2011), Zuideau and Lipovac (2009)
	Diversity and/or economic drive	15	Rate of job creation over three years and the company deregistration rate	Braconnier <i>et al.</i> (2011)
			Part of the population employed in different sectors	Etsy <i>et al.</i> (2005), Braconnier <i>et al.</i> (2011), Estoque and Murayama (2014), Miller <i>et al.</i> (2017), Zuideau and Lipovac (2009)
			Company creation rate, delisting of companies	Braconnier <i>et al.</i> (2011), Miller <i>et al.</i> (2017), Zuideau and Lipovac (2009)
			Training offer for the unemployed	Zuideau and Lipovac (2009)
	Equality of sexes	2	Diversity of job sectors	Miller <i>et al.</i> (2017)
			Rate of childcare for children under the age of three	Braconnier <i>et al.</i> (2011)
	Incomes	8	Proportion of women in the labour market	Cutter <i>et al.</i> (2010)
			Income per capita, gross available household revenue, median income per consumption unit	Angeon and Bates (2015), Biggeri and Mauro (2018), Braconnier <i>et al.</i> (2011), Estoque and Murayama (2014, 2017), García-Sánchez <i>et al.</i> (2015), Kerényi (2011), Lasso de la Vega and Urrutia (2001), Togtokh and Gaffney (2010), Zhang <i>et al.</i> (2015), Zuideau and Lipovac (2009)
			D9/D1 ratio of income distribution, Gini Index	Braconnier <i>et al.</i> (2011), Cutter <i>et al.</i> (2010), Miller <i>et al.</i> (2017)
			Household savings	Kerényi (2011), Miller <i>et al.</i> (2017), Zhang <i>et al.</i> (2015)
		Proportion of the population below the poverty threshold, or more than twice below the poverty threshold	Estoque and Murayama (2014), Huang and London (2012), Miller <i>et al.</i> (2017), Sadd <i>et al.</i> (2011), Stewart <i>et al.</i> (2014), Su <i>et al.</i> (2009)	

Tableau II : Article 1 – Table II: Characterization of sub-dimensions and distribution of variables within dimensions (cont.)

Dimension	Sub-dimensions	Number of identified variables	Examples of variables	References
Economics	Economic influence	7	Proportion of polluting goods and raw materials that had been imported and the ratio of gasoline price compared to the world average	Etsy <i>et al.</i> (2005)
		2	Sums invested in private investments	Angeon and Bates (2015)
			Current transfers and transport costs	Angeon and Bates (2015)
Services	Access to the homeownership market	3	Proportion of home owners who occupy their main dwelling address	Braconnier <i>et al.</i> (2011), Cutter <i>et al.</i> (2010), Miller <i>et al.</i> (2017), Sadd <i>et al.</i> (2011), Stewart <i>et al.</i> (2014), Zuideau and Lipovac (2009)
			Accessibility to affordable accommodation	Porter <i>et al.</i> (2013)
	Habitat characteristics	8	Age of the dwelling	Cutter <i>et al.</i> (2010), Miller <i>et al.</i> (2017), Monteiro <i>et al.</i> (2012)
			Proportion of the population that lived in an uncomfortable and/or overpopulated dwelling with the number of rooms per capita	Kerényi (2011), Zhang <i>et al.</i> (2015), Zuideau and Lipovac (2009)
			Median value of the dwelling	Sadd <i>et al.</i> (2011)
	Access to basic needs	7	Percentage of the population with or without access to fresh water, electricity or sanitation facilities	Angeon and Bates (2015), Etsy <i>et al.</i> (2005), García-Sánchez <i>et al.</i> (2015), Hsu <i>et al.</i> (2013), Kerényi (2011), Miller <i>et al.</i> (2017), Porter <i>et al.</i> (2013)
			Distance from ambulance and fire services	Pearson <i>et al.</i> (2013)
			Proportion of the population with or without covered by medical insurance	Cutter <i>et al.</i> (2010), Miller <i>et al.</i> (2017)
	Access to public services	1	Number of places for senior citizens in residential homes	Zuideau and Lipovac (2009)
	Access to information and communication	2	Number of mobile telephone or internet contracts, digital access index	Angeon and Bates (2015), Estoque and Murayama (2014), Etsy <i>et al.</i> (2005), Cutter <i>et al.</i> (2010), Porter <i>et al.</i> (2013)
			Proportion of population that had a telephone, number of people with landline telephones	Estoque and Murayama (2014)
	Mobility	8	Number of kilometres of the public transport network per citizen and the proportion of the road network that includes cycle paths	Zuideau and Lipovac (2009)
			Length of main and secondary roads per unit of area	Cutter <i>et al.</i> (2010), Estoque and Murayama (2014), Miller <i>et al.</i> (2017)
Proportion of the population that worked in the same employment area			Zuideau and Lipovac (2009)	
Proportion of the population that owned a vehicle			Cutter <i>et al.</i> (2010), Miller <i>et al.</i> (2017), Zuideau and Lipovac (2009)	
		Proportion of the population who still reside in the state where they were born	Cutter <i>et al.</i> (2010), Miller <i>et al.</i> (2017), Pearson <i>et al.</i> (2013), Zuideau and Lipovac (2009)	

Tableau II : Article 1 – Table II: Characterization of sub-dimensions and distribution of variables within dimensions (final section)

Dimension	Sub-dimensions	Number of identified variables	Examples of variables	References
Services	Land use planning	14	Presence or proximity of different infrastructures that had a positive or negative effect on the phenomena studied (cinemas, sports and leisure facilities, healthy and unhealthy living infrastructures)	Braconnier <i>et al.</i> (2011), Cutter <i>et al.</i> (2010), Miller <i>et al.</i> (2017), Pearson <i>et al.</i> (2013), Zuideau and Lipovac (2009)
Policy	Political commitment	1	Voting rate	Cutter <i>et al.</i> (2010), Estoque and Murayama (2014), Kerényi (2011), Miller <i>et al.</i> (2017), Sadd <i>et al.</i> (2011), Zuideau and Lipovac (2009)
	Civil and political rights	15	Individual liberties (speech, assembly rights, religion, ...) and tolerance	Angeon and Bates (2015), Biggeri and Mauro (2018), Etsy <i>et al.</i> (2005), Miller <i>et al.</i> (2017), Porter <i>et al.</i> (2013)
			Individual property rights	Porter <i>et al.</i> (2013)
	Integrity	2	Corruption, democracy or civil and political liberties	Angeon and Bates (2015), Estoque and Murayama (2017), Etsy <i>et al.</i> (2005)
			Government efficiency (defined as the quality of public services)	Angeon and Bates (2015), Etsy <i>et al.</i> (2005)
	Environmental governance	7	Effectiveness of environmental governance, knowledge creation in environmental, technological and political sciences	Etsy <i>et al.</i> (2005)
			Local Agenda 21 initiatives	Angeon and Bates (2015), Etsy <i>et al.</i> (2005)
			Number of environmental treaties	Angeon and Bates (2015)
	Political fragmentation	1	Jurisdictional overlap	Cutter <i>et al.</i> (2010), Miller <i>et al.</i> (2017)
	Anticipate and plan	5	Proportion of the population covered by protection programs	Cutter <i>et al.</i> (2010)
			Country's ability to cope with any kind of threat	Angeon and Bates (2015)
	Budget	5	Share of expenditures for local services	Angeon and Bates (2015), Cutter <i>et al.</i> (2010), Estoque and Murayama (2014), Miller <i>et al.</i> (2017)
			Subsidies, including those for agriculture	Etsy <i>et al.</i> (2005), Hsu <i>et al.</i> (2013)
Regional public spending			Braconnier <i>et al.</i> (2011)	
General government final consumption			Angeon and Bates (2015)	

The “health” dimension

The number of variables included in the “health” dimension of the composite indices studied varies from 1 for Angeon and Bates (2015), Biggeri and Mauro (2018), Hsu *et al.* (2013), Lasso de la Vega and Urrutia (2001), Pearce *et al.* (2010), Pearson *et al.* (2013), Su *et al.* (2009), Togtokh and Gaffney (2010) and Zhang *et al.* (2015) to 12 for Porter *et al.* (2013). The contribution of these variables to the final index varies from 3 % (Angeon and Bates, 2015; Cutter *et al.*, 2010) to 33 % (Huang and London, 2012).

For the measurement of **general population health**, the most frequently observed health variable is life expectancy at birth, at 60 or at 65 years of age (47.8 % of studies). Death rate is also found in works by Pearce *et al.* (2010) and Pearson *et al.* (2013). Hsu *et al.* (2013) used the child death rate to develop the EPI while Etsy *et al.* (2005) used the death rates for infectious intestinal disease, for deaths of children due to respiratory disease and the death rate for children under the age of five per 1,000 births in the “reducing human vulnerability” component. Mortality variables for children and/or children under the age of five are sensitive variables that reflect recent changes in environmental and economic conditions, and unlike adult deaths are less reliant upon the time between exposure and the declaration of the illness (Sherrieb *et al.*, 2010). Finally, some authors mention the fertility rate (Etsy *et al.*, 2005), to indicate social evolution when socio-economic, political or health conditions improve (Fekete, 2018).

More specific public health events are sometimes mentioned as either **environmental health variables** or **complementary general variables of general public health**. The Chronic Obstructive Pulmonary Disease (COPD) rate was for example used as a health variable (Monteiro *et al.*, 2012), as the appearance and aggravation of this disease is affected by individual lifestyle, gene pool, general state of health but also socio-economic or environmental factors.

Variables that give an idea of **health care facilities** can be used to complete the “health” component. This information about resources can be used to meet the daily needs of the population or solve public health events such as epidemics or terrorist attacks (Braconnier *et al.*, 2011; Cutter *et al.*, 2010; Estoque and Murayama, 2014; Huang and London, 2012; Miller *et al.*, 2017; Pearson *et al.*, 2013; Sadd *et al.*, 2011; Zuideau and Lipovac, 2009).

The “environment” dimension

The number of variables in the “environment” dimension ranges from 1 for Biggeri and Mauro (2018), Lasso de la Vega and Urrutia (2001) and Togtokh and Gaffney (2010) to 41 for Etsy *et al.* (2005). The contribution of these variables to the final index varies from 6 % (Cutter *et al.*, 2010) to 91 % (Pearce *et al.*, 2010). The low availability of certain variables in some countries appears to be an obstacle to obtaining relevant information (Braconnier *et al.*, 2011). This may explain the imbalance of representation and precision for this dimension in the indices studied.

In cases where just one environmental variable is considered, the variable that is used reflects greenhouse gas emissions (Biggeri and Mauro, 2018; Lasso de la Vega and Urrutia, 2001; Togtokh and Gaffney, 2010). This variable is one of the most representative variables of environmental damage, and this data is already available for most countries (Lasso de la Vega and Urrutia, 2001). Angeon and Bates (2015), Estoque and Murayama (2017), Etsy *et al.* (2005), García-Sánchez *et al.* (2015), Hsu *et al.* (2013) and Porter *et al.* (2013) also used this variable to complement other environmental variables. Estoque and Murayama (2017), Etsy *et al.* (2005) and Porter *et al.* (2013) also integrated the ecological footprint within the composite indices

they developed. The ecological footprint monitors the combined impact of **anthropogenic pressures** that are more typically evaluated independently (Galli *et al.*, 2012) and tracks humanity's demands on the biosphere (Estoque and Murayama, 2017).

Regarding to **natural resources**, works by Pearce *et al.* (2010) established a link between the presence of green spaces and beneficial effects on blood pressure, excess weight, obesity and self-perceived health. Closeness to green spaces, parks and open spaces is also mentioned in a number of studies (Monteiro *et al.*, 2012; Pearson *et al.*, 2013; Stewart *et al.*, 2014). Access to green spaces encourages physical activity (Humpel, 2002), provides psycho-social benefits (Pretty *et al.*, 2005) and communities with little access to green areas tended to be more socially disadvantaged (Comber *et al.*, 2008; Estabrooks *et al.*, 2003). Zuindeau and Lipovac (2009) and García-Sánchez *et al.* (2015) considered the area of agricultural land in the environment dimension in order to evaluate the negative effect of agriculture on natural resources.

The variables mentioned for the measurement of **biodiversity** are varied, although few indices consider this issue. In their study, Estoque and Murayama (2014) used methods produced by Costanza *et al.* (1998) and de Groot *et al.* (2012) to calculate the sub-index of biological integrity. A monetary value for ecosystem services was attributed to each spatial unit studied, and this value increased with the diversity and richness of the biome and ecosystem in question. The term “ecosystem services” describes all direct or indirect benefits and services that ecosystems generate or provide to individuals (Estoque and Murayama, 2013).

Other variables can be considered to provide information about the quality of the environments within the area where the studied population live. Data for **water quality** were included by Braconnier *et al.* (2011), Etsy *et al.* (2005), Kerényi (2011), Miller *et al.* (2017), Pearson *et al.* (2013), Porter *et al.* (2013) and Zuindeau and Lipovac (2009). To account for the possible evolution of pressure on water bearing resources, data pertaining to the organic water pollutants or linked to agriculture are sometimes mentioned (Etsy *et al.*, 2005; García-Sánchez *et al.*, 2015; Huang and London, 2012; Porter *et al.*, 2013; Stewart *et al.*, 2014; Zhang *et al.*, 2015). The data used to characterize **air quality** concern emissions and/or concentrations of different pollutants (SO₂, NO₂, NO_x, PM, CO, O₃, VOC) (Etsy *et al.*, 2005; Hsu *et al.*, 2013; Kerényi, 2011; Pearce *et al.*, 2010; Pearson *et al.*, 2013; Sadd *et al.*, 2011; Su *et al.*, 2009; Zhang *et al.*, 2015).

Information about **environmental safety** or **environmental risk** were also used for the development of a composite indice (Braconnier *et al.*, 2011; Huang and London, 2012; Miller *et al.*, 2017; Pearce *et al.*, 2010; Sadd *et al.*, 2011; Stewart *et al.*, 2014; Zuindeau and Lipovac, 2009). For example, density of sources of burdens in each spatial unit studied, distance to industrial facilities or percentage of population living in an “at-risk” area were included. Cutter *et al.* (2010), Estoque and Murayama (2014) and Miller *et al.* (2017) also included **natural risks** in the “environment” dimension (flooding and earthquakes, amongst others).

The “social” dimension

In order to interpret the interactions between variables and possible aggregations thereof, social factors should be considered when studying the link between health and environment as variables that are related to environmental exposure and play an important role in health inequalities (WHO, 2012a). The number of variables integrated within the “social” dimension of the composite indices studied varies from 1 for Lasso de la Vega and Urrutia (2001), Su *et al.* (2009), Togtokh and Gaffney (2010) and Biggeri and Mauro (2018) to 14 for Porter *et al.* (2013). The contribution of these variables to the final index varies

from 4 % (Etsy *et al.*, 2005) to 42 % (Monteiro *et al.*, 2012). Angeon and Bates (2015) state that many variables can be selected to characterize human capital, and for Braconnier *et al.* (2011), this dimension has very varied and complex aspects which can lead us to consider increasing numbers of variables. However, Angeon and Bates (2015) recommend the use of data that can be interpreted without ambiguity.

To characterise the **social network** of the area studied, Cutter *et al.* (2010) state that the number of associations or the proportion of the population involved in these associations should be included as they make it possible to identify the relationships between individuals, neighbourhoods and communities. Variables, defined in different ways (*e.g.* the percentage of households that are considered linguistically isolated), were integrated to determine the social fragmentation that can result from the population structure that can lead to a form of isolation (Huang and London, 2012; Miller *et al.*, 2017; Sadd *et al.*, 2011).

To account for **social integration difficulties and anomie**, suicide rates and violent crimes rates are included as variables of community deterioration (Braconnier *et al.*, 2011; Miller *et al.*, 2017). Age-related data are used to characterize **population sensitivity** in the studied area (Cutter *et al.*, 2010; Huang and London, 2012; Miller *et al.*, 2017; Monteiro *et al.*, 2012; Sadd *et al.*, 2011; Stewart *et al.*, 2014). Social capital is also characterized by the **education level** of the population. An other aspect of this “social” dimension concerns **social resources** related to population structure.

Several studies include data about the **level of urbanization** in their composite indices. Indeed, high population density makes it possible to reduce transaction costs, improves the profitability of public spending on infrastructure and services, and facilitates the production and dissemination of information, all of which lead to economic growth (WHO, 2010). Rural / urban status, population density, urbanisation ratio have therefore been used by Estoque and Murayama (2017), García-Sánchez *et al.* (2015), Monteiro *et al.* (2012), Pearson *et al.* (2013) and Zhang *et al.* (2015). However, for Estoque and Murayama (2017), this variable was used as a proxy measure for vulnerability to risk. García-Sánchez *et al.* (2015) supplemented the information provided by population density by adding urban population growth and Etsy *et al.* (2005) added the percentage difference in population change between 2004 and the projection for 2050.

The “economics” dimension

The number of variables included in the “economy” dimension on the composite indices studied here varies from 1 for Biggeri and Mauro (2018), García-Sánchez *et al.* (2015), Huang and London (2012), Lasso de la Vega and Urrutia (2001), Pearce *et al.* (2010), Sadd *et al.* (2011), Su *et al.* (2009) and Togtokh and Gaffney (2010), to 10 for Angeon and Bates (2015), Estoque and Murayama (2014), and Miller *et al.* (2017). The contribution of these variables to the final index varies from 4 % (Sadd *et al.*, 2011) to 34 % (Angeon and Bates, 2015). This dimension was not considered for 26.1 % of the studied indices (Hsu *et al.* (2013), Monteiro *et al.* (2012), Pearce *et al.* (2010), Pearson *et al.* (2013), Porter *et al.* (2013) and Stewart *et al.* (2014)).

Within this dimension, the sub-dimensions of **employability, diversity and/or economic drive** of the study area and the **equality of sexes** were characterized by different items. Amongst the variables used in this dimension, incomes (income per capita, gross available household revenue, median income per consumption unit) are mentioned. According to the spatial scale studied, the measure of the statistical

dispersion of income distribution (the Gini index) makes it possible to measure **income inequalities**. To highlight this sub-dimension, Braconnier *et al.* (2011) use the inter-decile ratio (D9/D1) of income distribution. Kerényi (2011), Miller *et al.* (2017) and Zhang *et al.* (2015) also consider household savings in their indices in order to reflect living standards.

To characterize the **economic influence** of the studied area, Etsy *et al.* (2005) included the proportion of polluting goods and raw materials that had been imported and also the ratio of gasoline price compared to the world average. Indices created by Angeon and Bates (2015) incorporated sums invested in private investments and current transfers as well as transport costs to account for inefficient logistics.

The “services” dimension

The number of variables included in the “services” dimension of composite variables varies from 1 for Monteiro *et al.* (2012), Stewart *et al.* (2014) and Zhang *et al.* (2015) to 14 for Zuindeau and Lipovac (2009). The contribution of these variables to the final index varies from 3 % (Etsy *et al.*, 2005) to 35 % (Pearson *et al.*, 2013). Biggeri and Mauro (2018), Estoque and Murayama (2017), Huang and London (2012), Lasso de la Vega and Urrutia (2001), Pearce *et al.* (2010), Su *et al.* (2009) and Togtokh and Gaffney (2010) did not use any data that can be considered to be part of the “services” dimension.

To provide information about **access to the homeownership market**, Braconnier *et al.* (2011), Cutter *et al.* (2010), Miller *et al.* (2017), Sadd *et al.* (2011) and Zuindeau and Lipovac (2009) included the proportion of home owners who occupy their main dwelling address, while Porter *et al.* (2013) used accessibility to affordable accommodation within the studied area. In addition to these variables, data is also provided about the **habitat characteristics**: the age of the dwelling (Cutter *et al.*, 2010; Miller *et al.*, 2017; Monteiro *et al.*, 2012), the median value of the dwelling (Sadd *et al.*, 2011) or the proportion of the population that lived in an uncomfortable dwelling (Kerényi, 2011; Zhang *et al.*, 2015; Zuindeau and Lipovac, 2009). **Access to basic needs** and public services is characterized by variables that indicates the percentage of the population that had access to fresh water, electricity or sanitation facilities (Angeon and Bates, 2015; Etsy *et al.*, 2005; García-Sánchez *et al.*, 2015; Hsu *et al.*, 2013; Kerényi, 2011; Miller *et al.*, 2017; Porter *et al.*, 2013), by the number of places for seniors citizens in residential homes (Zuindeau and Lipovac, 2009).

Access to information and communication was included in composite indices *via* the number of mobile telephone or internet contracts (Angeon and Bates, 2015; Porter *et al.*, 2013), while Etsy *et al.* (2005) included digital access index. Cutter *et al.* (2010) considered the proportion of the population that had a telephone, and Estoque and Murayama (2014) included data for the number of people with landline telephones.

The **mobility** and the **land use planning** were also considered as sub-dimensions that should be included in this dimension. Several studies include: the number of kilometres of the public transport network per citizen and/or the proportion of the road network that includes cycle paths (Zuindeau and Lipovac, 2009), the length of main and secondary roads per capita (Miller *et al.*, 2017), the proportion of the population that owned a vehicle and the proportion who still reside in the state where they were born (Cutter *et al.*, 2010; Miller *et al.*, 2017). The presence or proximity of different infrastructures that had a positive or negative effect on the phenomena studied (cinemas, sports and leisure facilities, healthy and unhealthy living infrastructures) have been also considered (Cutter *et al.*, 2010; Miller *et al.*, 2017; Pearson *et al.*, 2013; Zuindeau and Lipovac, 2009).

The “policy” dimension

The “policy” dimension was only found in 56 % of the studied indices. The number of variables included in the “policy” dimension varies from 1 for Biggeri and Mauro (2018), Braconnier *et al.* (2011), Estoque and Murayama (2017), Hsu *et al.* (2013), Sadd *et al.* (2011) and Zuindeau and Lipovac (2009) to 21 for Etsy *et al.* (2005). The contribution of these variables to the final index varies from 2 % (Zuindeau and Lipovac, 2009) to 28 % (Etsy *et al.*, 2005).

For several years now, the emphasis has been placed on the policy dimension (or governance), which is considered to be an essential factor for environmental improvement (Etsy *et al.*, 2005). When building the SESI, Estoque and Murayama (2014) considered the governance dimension to be as important as the environmental, social and economic dimensions. Angeon and Bates (2015) considered this dimension to represent both the institutional framework and the institutional capacity on a national level. However, the governance or policy is the least integrated aspect in the studies considered. For Modica and Reggiani (2015), this dimension could show the ability of the areas studied to implement operational plans or public investments.

Voter turnout can be seen as a proxy variable of **political commitment**, and may also indicate a tendency to lose trust in the political regime and thus lead to instability (Fekete, 2018). A voting rate was incorporated into many composite indices. However, the use of this type of variable as a variable of voter turnout and commitment to political life should be adjusted according to each case. Indeed, some countries such as Australia, Belgium, Brazil, Costa Rica, Greece or Turkey have a mandatory vote that imposes fines on those who abstain for voting. In these countries, low abstention in the vote does not necessarily show a greater commitment or participation in political life. Variables of **integrity** and **civil and political rights** were also used. Porter *et al.* (2013) included variables for individual liberties (speech and assembly rights) and tolerance, while Biggeri and Mauro (2018) incorporated the composite indicator for civil and political rights developed by Freedom House. Angeon and Bates (2015) and Miller *et al.* (2017) also used a variable of freedom.

In order to identify and quantify **environmental governance**, Etsy *et al.* (2005) incorporated variables for the effectiveness of environmental governance, knowledge creation in environmental, technological and political sciences, or for local Agenda 21 initiatives in which governmental competence strengthens the ability to monitor and respond to environmental challenges. Angeon and Bates (2015) used the number of environmental treaties in their index. Cutter *et al.* (2010) and Miller *et al.* (2017) incorporated variables to specify the **political fragmentation** of the geographical areas studied, which was considered a source of vulnerability. Indeed, although jurisdictional overlap may be a positive attribute in terms of the plurality of potential resources, the multitude of jurisdictions raises the question how these entities can collaborate to attain their targets (Miller *et al.*, 2017).

Variables to measure the ability to **anticipate** and plan for certain hazards had been integrated in some indices, and/or the proportion of the population covered by protection programs or the country's ability to cope with any kind of threat.

Angeon and Bates (2015), Cutter *et al.* (2010), Estoque and Murayama (2014) and Miller *et al.* (2017) incorporated a **budget** sub-dimension with the share of expenditures for emergency medical services, social services, schools, firefighters or the police, whose efficiency depends on several factors including communication, organization or their resources (Miller *et al.*, 2017).

For each dimension, Table 2 shows that several sub-dimensions exist and can be integrated in the index. To characterize these sub-dimensions, several variables are possible and can be chosen according to the availability and/or the accessibility of the data. Consequently, it may be considered to implement this structuring in dimensions and sub-dimensions as a transposable framework to constitute a dataset. The adaptation of the index to a specific territory or to a local issue would then be effective through the included variables.

Discussion

The subjective perception of certain concepts

Different appraisals can be observed concerning the definition of the variable or the impact (positive or negative) of the variable on environmental health issues (Pearce *et al.*, 2010; Pearson *et al.*, 2013). In the area of environmental health where there is an acute need of assessment method, certain aspects as the complexities of human activities, environmental processes and human well-being come together (Briggs, 2008). Since well-being values are not universal, it would be inappropriate to include solely subjective variables. Each variable should be objective. However, as mentioned by Nascimento and Carrage (2007) and Stiglitz *et al.* (2009), it is acceptable for a life quality index to incorporate both objective and subjective variables.

When the sensitive population group is used in environmental health studies, it is sometimes defined as a group of individuals at an “extreme” age. Subjectivity here is thus dependent on the definition of these age limits. Indeed, Huang and London (2012) and Sadd *et al.* (2011) considered the population aged below 5 or over 60, whilst Stewart *et al.* (2014) defined them to be under 14 and over 65, Shrestha *et al.* (2016) considered a population aged 6 to 11, and Miller *et al.* (2017) considered sensitive populations to be those under 18 years of age or over 65 years of age. The conclusions of these studies are difficult to compare because the populations concerned have different characteristics (physiology, behaviour, exposure profile).

The impact of population migration on quality of life can be negative or positive. Cutter *et al.* (2010) considered that net international migration had a negative impact on well-being, while Zuideau and Lipovac (2009) mention that the impact of the foreign population on well-being is perceived in different ways according to the population concerned. For Sherrieb *et al.* (2010), demographic changes or migration could indicate a stress on resources or a loss of human capital.

The effects of urbanization are equivocal in studies, and can be positive (access to infrastructure) or negative (chronic stress caused by noise and respiratory problems related to road traffic) for human health (Pearson *et al.*, 2013). Cutter *et al.* (2010), Estoque and Murayama (2014) and Miller *et al.* (2017) considered access to a road network to be beneficial for well-being through the connectivity between territories, which provides easier access to services and infrastructures. However, road traffic-related atmospheric pollution and exposition to noise may negatively affect the health of populations living close to roads, although there is some uncertainty regarding the estimation of population exposure (Host, 2013). Monteiro *et al.* (2012) included the density of the road network in their study, as they considered this variable to be a factor that was at the origin of, or aggravated, the public health event studied (COPD).

Population density is sometimes considered as an explanatory factor of the result, and sometimes as a factor that directly impacts environmental health. Thus, Cutter *et al.* (2010) did not use population density

in the BRIC and the analysis of their results showed a bias that revealed an urban/rural divide. Pearson *et al.* (2013) revealed unexpected results for regions that were characterized by a high population density close to downtown areas, linking a good state of health with a higher social deprivation index. The lower mortality rates could however be artificially low in these neighbourhoods, as their inhabitants live there for shorter periods of time.

The potential environmental benefits on health are widely accepted, especially in connection with the presence of green spaces. However, characterizing the association between green spaces and public health events seems to be complicated because it can vary considerably according to the definitions used for “green spaces” (Klompaker *et al.*, 2018). There is also no consensus in the literature regarding the measurement of access to green spaces when different distances are involved (Stewart *et al.*, 2014), with availability also playing a possible role in the notion of access (Flacke *et al.*, 2016; Kruize *et al.*, 2014; Yao *et al.*, 2014). Pearce *et al.* (2010) use the proportion of the smallest area classified as green space and Monteiro *et al.* (2012) use the percentage of green spaces in neighbourhoods. To elaborate their composite index, Pearson *et al.* (2013) mainly considered the means of accessibility and distances to different types of structures and amenities, which explains their inclusion of the distance to parks. However, none of these three studies specifies the characteristics of these green spaces in terms of (i) size, (ii) availability or (iii) accessibility, which could raise the question of the comparability of these variables. Stewart *et al.* (2014) chose to use buffer zones, with different weighting according to the type of park: urban parks, which generally had areas 0.05 miles², had a lower weighting and a smaller buffer zone, whilst the county parks and open spaces, which had areas of approximately 3 miles², had twice the weighting of urban parks and a buffer zone that was twice as large. These different interpretations of the “green spaces” variable emphasize the importance of defining the subject of the evaluation precisely.

Within the “environment” dimension, certain authors also seek to evaluate the quality of the environments in which the populations live, mainly using data related to air and/or water quality (Braconnier *et al.*, 2011; Etsy *et al.*, 2005; Hsu *et al.*, 2013; Kerényi, 2011; Miller *et al.*, 2017; Pearce *et al.*, 2010; Pearson *et al.*, 2013; Sadd *et al.*, 2011; Su *et al.*, 2009; Zhang *et al.*, 2015; Zuideau and Lipovac, 2009). The ground is a support for human activity and absorbs types of pollution that can expose human populations to health risks in more long-term manner than *via* occasional water and air pollution. However, soil quality is only mentioned in composite indices in data for polluting / polluted sites listed by Huang and London (2012), Stewart *et al.* (2014) and Zuideau and Lipovac (2009). A study by Panagos *et al.* (2012) notes the rarity of soil data, and states that any such data were difficult or impossible to access, especially at the scales needed by policy makers.

Transposability and influence of the spatial unit studied

In our context, transposability can be defined as the possibility to successfully implement an index outside the framework of study for which it was originally developed, for a different geographical location (other territory), spatial unit (smaller or larger) and research question (*i.e.* different from the initial question for which the index was developed). There are two possible principles of transposability. First of all, transposability can be considered *stricto sensu*, *i.e.* whatever the spatial unit considered, it should be possible to calculate the composite index by using all the defined variables and the chosen methodology (imputation of missing data, normalization, weighting, aggregation). The use of transposability in this strict sense is rare because of the compatibility of the index with the research question (Lasso de la Vega and Urrutia, 2001; Togtokh and Gaffney, 2010). An alternative method would be to transpose the initial

developed reasoning and the published overall methodology to develop a new composite index in environmental health. The dimensions and calculation methods are identical to the initial method used, and the variables are adapted to the context of the geographical study area. Although less conservative, this method has less constraints than the first and could thus be more widely accepted and used. The reflection conducted by each author is then perceived as a support for a possible optimization.

The first limitation of transposability is the contextualization of the variables integrated into the composite indices that were developed within a given territory. For example, some authors examined the degree of telecommunications development, with Estoque and Murayama (2014) using the number of land lines per 1,000 inhabitants in the Philippine administrative governance sub-index. While this type of variable would be suitable for countries with the same level of technological development, it would be more relevant to identify Internet access for other countries (Etsy *et al.*, 2005). In addition, certain representative national data (percentage of marae for Pearson *et al.* (2013) or data relating to ethnicity for studies conducted in the United States) would only be suitable in the context for which they were used.

A second limitation is the lack of accessible variables at the appropriate spatial scale (Maantay, 2007; Rothenberg *et al.*, 2015). Certain facets are characterized only at a global scale (*i.e.* national), and not defined on a local scale. The use of a macroscopic index that was designed for use on a nationwide scale for a smaller spatial unit is made difficult by the lack of data measured on a fine scale (Cutter *et al.*, 2010; Estoque and Murayama, 2014). For example, the PISA scores of 15-year-old students in reading, mathematics and science are not available at the local town area scale. Moreover, moving from a large unit to a more discriminating spatial unit may lead to the existing data being subject to statistical secrecy for some data in examples such as the proportion of agricultural area in organic farming in the total agricultural area on the municipal scale. Aggregating these data at a higher level may resolve this problem and make it possible to disseminate information at the smallest possible scale that respects the threshold of statistical confidentiality (Branchu *et al.*, 2018). Another phenomenon to consider is that what appears on one scale may disappear on another (Hamman, 2011). So, it is important to assess how far the choice of spatial scale can influence the research question for which it is being used (Rey *et al.*, 2009). The first step when defining the geographical space to be considered is to choose the most suitable investigation scale, giving it an operational socio-political dimension (Salem, 1995).

Recommendations and good practices

This review provides a complete overview of the variables included in environmental health indices and consolidate the approach leading to the selection of variables. This proposal complements the OECD's thinking on composite indices construction (European Commission *et al.*, 2008). The OECD handbook provides comprehensive advices about how to design and develop a composite index, without application domain specificity (Figure 4). This guide is mentioned as a reference by Angeon and Bates (2015) and García-Sánchez *et al.* (2015), Etsy *et al.* (2005) for their part refer to a version published before 2008. Hsu *et al.* (2013) detail the entire methodology used to develop their composite index, and their method is very close to that proposed in the OECD guide.

Theoretical framework

- Provide the basis for the selection and combination of single indicators

Selecting variables

- On the basis of their analytical soundness, measurability, coverage, relevance and relationship to each other

Imputation of missing data

- Provide the missing patterns and selection of appropriate imputation methods

Multivariate analysis

- Investigate the overall structure of indicators (correlation) and assess methodological choices to be made for weighting and aggregation

Normalisation of data

- Render the variables comparable

Weighting and aggregation

- According to the underlying theoretical framework and the data properties

Robustness and sensitivity

- According to, including or excluding single indicators, the normalisation scheme, the imputation of missing data, the choice of weights and the aggregation method

Figure 4 : Article 1 – Figure 4: Good practice for developing composite indices (inspired by European Commission et al. (2008))

Authors must commit to a strict and explicit methodological framework that defines the criteria for the inclusion of variables and spatial units. The challenge is to limit the impact of missing data on the quality of the database used for the development of the composite index. This framework can also be used as a basic structure to identify the most appropriate weighting and aggregation methods for the research question concerned.

The missing data profile must be specified in order to determine the most appropriate imputation methods. Missing data can be a challenge when analyzing and interpreting the results, thus weakening the final conclusions (Pedersen *et al.*, 2017). Our analysis shows that within the 23 studies, 22 % of composite indices reported a problem of missing data in the development of their database. The percentages attributed to the acceptability of missing data range from 0 to 41 %, and there is no consensus on an acceptable percentage. The imputation methods are mentioned and/or described in the publications. However, only one study mentions a “reliability/robustness” analysis that also considered the results with and without the imputed data (Etsy *et al.*, 2005).

70 % of the studied indices did not use any weighting scheme. For the remaining 30 %, the employed methods were the Analytic Hierarchy Process (AHP), the Budget Allocation Process (BAP), the Multi-Criteria-Decision-Analysis (MCDA), the Criteria Importance Through Inter-Criteria Correlation (CRITIC) or population weighted index. The recent review of Becker *et al.* (2017) showed that there are problems with understanding and appropriating this step. The main difficulty of weighting lies in the choice of the authors to give greater weight to variables considered important regarding the research question, which can be in contradiction with more statistical approaches (Burgass *et al.*, 2017).

A sensitivity and robustness analysis is necessary in order to (i) identify the type of variation of the composite index, and (ii) quantify the overall uncertainty that is inherent to the choices made during the development of the methodology. This sensitivity analysis was performed for just over 13 % of the indices (*i.e.* 3 indices), and only Cutter *et al.* (2010) considers the impact of missing data on the results. According to European Commission *et al.* (2008), the development of composite indices involves subjective evaluations of the selection of individual variables, the attitude towards missing data, and the choice of weighting or aggregation models.

The selection of the variables must be made in the light of (i) their relevance in the literature, (ii) their availability, and (iii) the robustness of the association of the data with health. The selection criteria for the selection of the variables to be included are mentioned in 20 out of the 23 indices: the selection of the variables reflecting the multiple facets of the environmental health was generally based on an analysis of the literature. Beccari (2016) state that the availability of variables on the desired geographic scale could even be the main selection factor as it is a major problem when developing sub-national methodologies to measure and compare territories.

Conclusion

This review aimed to identify and characterize the dimensions and variables that are included in environmental health indices. This analysis led to the identification of 23 spatialized composite indices that each answered a specific research question. A great diversity in the design of the indices was highlighted: these composite indices could be included from 2 to 6 dimensions and from 4 to 76 variables. In this review, a total of 329 variables was assessed in a systematic way: this multitude of variables demonstrates the need for classification and ordering of concepts. These indices had been developed by scientific or institutional bodies, and the diversity of information used by these two actors highlights the lack of consensus for an index that would be acceptable to both parties (Strezov *et al.*, 2017). The 23 composite indices were developed with the aim of characterizing the territorial inequalities in health environment. The results comparability, between territories and even sometimes within one same territory and their cross-discussion is limited given (i) the differences of research question, and (ii) the wide variability of the methodological design.

This present work could be a first step towards a proposal of guidelines designed to provide a consensual framework that could facilitate the exploitation of environmental health indices. We recommended harmonizing practices in the selection and characterization of the dimensions and sub-dimensions related to health, environment, social, economics, services and policy. The proposed framework aimed to take up the challenge of constituting a relevant dataset for the development of a composite index in environmental health. It could be used as a basis to guide authors in their choice to include or exclude certain variables and thus make interterritorial comparisons more robust.

The quality of the definition and measurement of the variables will influence the quality of the composite index developed. The definition is all the more important as the evaluation of certain phenomena involved in environmental health is inevitably subjective. The research for the most holistic view is important, for example through the integration of data on soils, that is still anecdotal. The will of transposability regularly mentioned in the conclusions of the studied articles is limited by the contextualization of certain variables (*e.g.* related to the degree of technological development of the studied geographical area). Spatial granularity can be the source of specific problems, either when selecting variables (accessibility and/or availability of data on a fine scale), or when developing the index and carrying out the subsequent analysis

(developed index without discriminating power at the chosen scale). When selecting variables, authors have to pay close attention to the accessibility of data related to their problematic and at the spatial scale considered. This practice sometimes encounters local obstacles, an improvement in the culture of information sharing would make it possible to optimize the characterization of territories.

Control of uncertainties may be sufficient in some composite indices, but others may require reworking in order to improve conclusions robustness. Only 5 studies mention the issue of missing data. The explanation of their number and their mathematical treatment is however essential to assess the reliability of the dataset. The weighting and aggregation steps, well known as crucial by the authors and more broadly, are generally better explained. With the exception of Cutter *et al.* (2010), none of them provided a sensitivity analysis, a step that can be used to assess the robustness of composite indices. Presenting this step systematically would make it possible to remove certain criticisms related to the unavoidable sources of uncertainties.

In view of these different aspects, it is therefore important that frameworks for variable integration and index development methodology are clearly described in publications. This bid for transparency could also increase the understanding and adoption of these tools by public authorities and general public.

L'analyse a permis (i) d'identifier les indices composites développés dans l'objectif d'évaluer les déterminants territoriaux dans une perspective de santé environnementale et (ii) de répondre à certains questionnements en lien avec la caractérisation du territoire à travers le développement d'indices composites. Nous avons mis en évidence que cette méthodologie était utilisée aussi bien par la communauté scientifique que par des institutionnels et que les indices développés répondaient à une question de recherche spécifique. La diversité des informations utilisées par ces deux profils d'acteurs souligne l'absence de consensus sur une méthodologie de développement d'indice composite qui serait acceptable par chacun de ces deux profils d'acteurs (Strezov *et al.*, 2017). De manière générale, les indices composites développés à grande échelle (*i.e.* le pays) sont à l'origine de comparaisons afin d'obtenir un classement des unités spatiales considérées (Rothenberg *et al.*, 2015). Les évaluations à échelle plus locale semblent plus difficiles à mettre en place, le verrou méthodologique portant principalement sur l'accessibilité à des données pertinentes pour l'analyse des inégalités territoriales de santé (Maantay, 2007; Rothenberg *et al.*, 2015). Ceccarelli *et al.* (2020) soulignent de plus le caractère urgent d'une plus grande précision dans les évaluations ayant trait au bien-être équitable et durable.

La revue s'est essentiellement concentrée sur les premières étapes de construction d'un indice composite (*i.e.* le choix des variables intégrées) afin de mettre en évidence les similitudes et dissemblances. Les déterminants territoriaux utilisés dans le développement d'indices composites dans une perspective de santé environnementale ont été identifiés et une importante variabilité sur le type d'informations incluses dans les indices composites a été soulignée. Une autre conclusion de la revue est qu'il a été mis en évidence une diversité de répartition des variables en lien avec la vulnérabilité ou la résilience selon les dimensions. Ainsi au sein de la dimension « Santé », presque deux variables sur trois sont en lien avec la vulnérabilité. En ce qui concerne la dimension « Environnement », une très grande variabilité du type de variables incluses a été constatée, mais avec une très large proportion de variables décrivant des impacts négatifs (pressions anthropiques en lien avec l'environnement industriel, indicateurs de contamination des compartiments aquatiques et atmosphériques). Les variables environnementales en lien avec la résilience sont plus faiblement représentées dans les indices composites identifiés et décrivent principalement la biodiversité ou les ressources naturelles, la proximité et/ou la présence d'espaces verts. Pour les dimensions « Social » et « Économie », la distribution des variables au sein de la résilience ou de la vulnérabilité est plus équilibrée. Il est à noter toutefois qu'au sein de la dimension « Social », des variables décrivant le territoire ou une caractéristique de population présentent un caractère équivoque et ne présentent pas de réel consensus dans la littérature sur l'attribution d'un impact clairement négatif ou positif (*e.g.* densité de population, statut urbain/rural et part de la population issue de l'immigration). Les dimensions « Services » et « Politique » sont les moins fréquemment retrouvées dans les indices composites (respectivement 70 % et 56 %). Toutefois parmi le nombre total de variables les 14 % et 11 % des variables incluses appartenant à ces deux dimensions présentent une distribution plus favorable à la résilience.

La variabilité des informations intégrées aux indices composites qui a été constatée est par ailleurs accentuée par une importante variabilité des méthodes de développement en elles-mêmes. Les principales limites et éléments de discussion sur les choix méthodologiques opérés pour le développement des indices composites ne sont toutefois pas spécifiques aux analyses en santé environnement. Le caractère équivoque de certains concepts et/ou phénomènes ainsi que la transparence méthodologique font également l'objet de points d'attention systématiquement considérés dans de nombreux autres domaines (*e.g.* socio-économie, psychométrie). La transparence sur les

méthodologies adoptées et sur les variables incluses permet également d'accroître la compréhension et l'adoption de ces outils par les acteurs publiques et à l'avenir par le grand public afin d'appréhender de manière plus efficace les inégalités territoriales en santé environnement. Les principales conclusions de cette revue sont que (i) cette grande diversité souligne le manque de cadre méthodologique commun et (ii) cette absence de consensus sur une méthodologie unique tend à limiter les comparaisons entre les territoires.

En s'appuyant sur ces conclusions, une nouvelle approche d'évaluation des inégalités territoriales en santé environnement par des indices composites a été proposée. Cette approche se veut transposable et repose sur l'évaluation simultanée de la résilience et de la vulnérabilité, ces deux aspects étant considérés avec la même importance. Pour le Haut Conseil de la Santé Publique (2021), cette approche santé environnement par les facteurs « délétères » et « protecteurs » de santé est complémentaire qui contribue à une démarche globale de la santé (Une seule santé – « One health »). Le cadre méthodologique est basé sur la caractérisation des variables catégorisées en dimensions et sous-dimensions (e.g. santé générale des populations au sein de la dimension « santé », niveau d'éducation au sein de la dimension « social »). Plusieurs variables peuvent alors être possibles pour caractériser une sous-dimension, le choix d'une variable plutôt qu'une autre pouvant être opéré selon différents critères (exhaustivité, qualité, disponibilité et/ou accessibilité, avis d'expert). L'adaptation d'un indice composite à un territoire spécifique serait alors effective à travers les variables incluses qui ont été sélectionnées pour caractériser une sous-dimension. Mettre l'accent sur la méthode plutôt que sur un contenu prédéfini pourrait permettre de lever le verrou de l'accessibilité à des données précises visant à mettre en évidence les liens entre des conditions environnementales et des résultats sanitaires (Rothenberg *et al.*, 2015). Knol *et al.* (2010) ont souligné dans leurs travaux qu'un cadre flexible était préférable à un cadre préconçu et rigide.

Les travaux présentés dans les articles 2 et 3 ont pour principal objectif de démontrer la faisabilité de cette méthodologie tout en soulignant l'intérêt d'une conception multidisciplinaire. Ce cadre méthodologique permettra d'améliorer la perception globale du territoire afin de mieux appréhender les processus à l'origine des inégalités territoriales en santé environnement.

L'Article 2 (page 58) présente le set de données utilisé pour caractériser les déterminants dans la région des Hauts-de-France. Afin de décrire les dimensions et sous-dimensions décrites dans la revue, des variables ont été collectées depuis des bases de données publiques, disponibles en *open data*. L'enjeu était également de proposer la description la plus holistique possible du territoire à travers l'intégration de données multidimensionnelles à l'échelle des 3817 communes que comportaient la région selon le découpage géographique de l'INSEE pour l'année 2017. Les travaux valorisés dans l'article 3 (page 77) présente (i) la méthodologie proposée pour le développement des indices composites de résilience et de vulnérabilité, (ii) le mode d'analyse conjointe des deux indices et (iii) les résultats issus du déploiement de la méthodologie sur les communes de la région des Hauts-de-France. L'analyse conjointe permet (i) d'interpréter simultanément la résilience et la vulnérabilité de chacune des communes et (ii) d'identifier des typologies de territoire en fonction des niveaux de résilience et de vulnérabilité. Cette analyse conjointe constitue un outil de diagnostic pour les parties prenantes qui cherchent à hiérarchiser les actions à mettre en œuvre tout en améliorant la contextualisation des politiques locales dans tous les domaines en lien avec la santé.

Article 2

Data for the assessment of vulnerability and resilience in the field of environmental health in the north of France

Delphine Brousmiche, Michaël Genin, Florent Occelli, Lukas Frank, Annabelle Deram, Damien Cuny, Caroline Lanier

Data in Brief 107220

<https://doi.org/10.1016/j.dib.2021.107220>



Data for the assessment of vulnerability and resilience in the field of environmental health in the north of France

Delphine Brousmiche¹, Michaël Genin², Florent Occelli¹, Lukas Frank¹, Annabelle Deram¹, Damien Cuny¹, Caroline Lanier¹

¹ Univ. Lille, IMT Lille Douai, Univ. Artois, Yncrea Hauts-de-France, ULR 4515 – LGCgE, Laboratoire de Génie Civil et géo-Environnement, F-59000 Lille, France

² Univ. Lille, CHU Lille, ULR 2694 - METRICS : Évaluation des technologies de santé et des pratiques médicales, F-59000 Lille, France

Corresponding author: caroline.lanier@univ-lille.fr

Abstract

The integration of multidimensional data is necessary to improve the understanding of environmental and social inequalities in health. The challenge is to define a dataset that provides the most holistic description possible of the territory. This article presents a relevant dataset to characterize the territorial accumulation of health determinants in the second most densely populated region of metropolitan France (Hauts-de-France Region, in the north of France). The multidimensional dataset combines data related to the economic, social, environment, services, health and policy dimensions at fine scale (*i.e.* each municipality). Data outlining a negative impact on health inequalities (*e.g.* anthropogenic pressures, socioeconomic factors related to vulnerability, *etc.*) are considered to be as important as data outlining a positive impact on health inequalities (*e.g.* natural resources, diversity and economic drive, *etc.*). The proposed theoretical framework relies on data reuse. Over one hundred variables covering a time frame from 2008 to 2017 were collected from a dozen public and national database providers. The use of official organizations ensured the quality of the collected data. The Geographic Information System, designed to map and catalogue ready-to-use data, was used to generate new data or to deal with missing data. Finally, 50 variables, including mostly quantitative but also qualitative data, were selected after application of inclusion and exclusion criteria. The resulting dataset provides a broad characterisation of the 3,817 municipalities in the Hauts-de-France Region. These data will help to discriminate the distribution pattern of vulnerability and resilience levels in this region. This novel approach is described in the paper “How can we analyse environmental health resilience and vulnerability? A joint analysis with composite indices applied to the north of France”, which provides a detailed description of the methodology used to develop composite indices. This research could therefore be of use to researchers, policy makers and stakeholders in the field of environmental health seeking to identify the weaknesses but also the strengths of municipalities.

Keywords

- Health determinants
- Inter-territorial comparison
- Environmental and social inequalities in health
- Resilience as a capacity to cope with health inequalities
- Vulnerability as a reflection of the harmful constraints and nuisances to which populations are subjected

Specifications Table

Tableau III : Article 2 – Table I: Specifications Table

Subject	Public Health and Health Policy
Specific subject area	Territorial characteristics and environmental and social inequalities in health
Type of data	Table
How data were acquired	Data <i>reuse</i> , collected from official national organizations: AGENDA21FRANCE.ORG, BPE (Permanent Facilities Database), CNAF (National Health Insurance Fund), DGCL (General Management of Local Authorities), DGFIP (General Directorate of Public Finance), DGPR (General Directorate for Risk Prevention), DREES (Directorate of Research, Studies, Evaluation and Statistics), Geofabrik GmbH, GEOIDD (National Observation and Statistics Service), IGN (National Institute for Geographic and forest information), INSEE (French National Institute of Statistics and Economic Studies), MEDDE (Ministry of Ecological and Solidarity Transition), ODT (French National Observatory of Territories) and REE (National Directory of Companies and Establishments).
Data format	Raw, calculated
Parameters for data collection	Three inclusion criteria: (i) a plausible association with environmental health in the literature, (ii) comprehensive, spatially contiguous and contemporary data for the entire study area, (iii) production and regular updating by a reliable organization (thereby ensuring acceptable data quality). One exclusion criterion: data associated with individual behavior (<i>i.e.</i> smoking, food habits, physical activities)
Description of data collection	36 % of raw data were integrated in the dataset 64 % of data have been calculated (evolution calculation, data converted to per capita, data per 1,000 inhabitants, raw data converted to rate, transdisciplinary adaptation of calculations, area calculations with GIS software: ArcGIS v10.7, Q-GIS v2.14) Data-management operations were made with Excel (2016) Statistical analysis were made with R software (R-3.3.2)
Data source location	“Hauts-de-France” Region, France
Data accessibility	Data are hosted on a public repository. Repository name: Zenodo Data identification number: 10.5281/zenodo.3701557 Direct URL to data: https://zenodo.org/record/4889374#.YLYg7agzY2w
Related research article	Delphine Brousmiche, Michaël Genin, Florent Occelli, Lukas Frank, Annabelle Deram, Damien Cuny, Caroline Lanier How can we analyse environmental health resilience and vulnerability? A joint analysis with composite indices applied to the north of France Science of The Total Environment https://doi.org/10.1016/j.scitotenv.2020.142983

Value of the Data

- The data are useful for environmental health issues as they allow users to map and discriminate the distribution pattern of vulnerability and resilience levels at a municipality scale.
- These data allow a cross-comparison between 3,817 municipalities based on 6 dimensions of health determinants. This can result in a more detailed identification of the strengths and weaknesses of municipalities faced with the challenges of social and environmental inequalities in health.
- Researchers, policy makers and other stakeholders may use these data to build new databases on health determinants at a fine scale to compare their findings in different regions of France or in different countries.
- These multidimensional data will help society to consider possible levers for action to improve public health and well-being for all in the short term, and to develop a “Health in All Policies” approach over the long term.

1. Data Description

Validated by the National Institute for Geographic and forest information (IGN) arrested at the date of 12/31/2016, the region had 3,817 municipalities, a municipality being a small administrative entity in France. 50 data were collected and are fully available in the Zenodo repository (accession number: 10.5281/zenodo.3701558, Table I “Data accessibility”). The 6 dimensions of health inequalities determinants were investigated: “Economics” dimension in Table II, “Environment” dimension in Table III, “Policy” dimension in Table IV, “Health” dimension in Table V, “Social” dimension in Table VI and “Services” dimension in Table VII. These tables include:

- the given acronym, the description of the data and specifications,
- the data unit,
- their(s) source(s),
- the reference year of the data,
- the rate of missing data,
- a mention if the variable is “raw” or “calculated”,
- a statistical description with (i) the median and the inter-quartile range [Q1-Q3] calculated after the imputation of missing data step for quantitative data and (ii) the modalities of qualitative data and their frequency distribution which is expressed as a percentage.

In the context of data *reuse*, the use of official organizations ensured the credibility of the collected data (Dekkers *et al.*, 2013). The good completeness of the data-set is pointed up regarding the low rate of missing data. 62 % of variables has no missing data (31/50 collected data). For the remaining 38 % (19/50), the percentage of missing data ranges from 0.03 % (1/3,817 values) to 2.07 % (79/3,817 values). 36 % of the raw data were integrated directly into the dataset, without computation. The other 64 % were obtained by calculation: some could be described as usual (*e.g.* percentage, data per inhabitant), others inspired by previous works specialised in socio-ecology systems (*e.g.* human to ecosystem service value ratio index) or resulting from a transdisciplinary adaptation (*e.g.* Shannon-Weaver index to calculate the representation of socio-professional categories within each municipality).

Tableau IV : Article 2 – Table II: Metadata related to the “Economics” dimension

Acronym	Data	Specifications	Unit	Source(s)	Time period	Missing data (%)	Raw / Calculated	Median – [Q1 ; Q3] or modalities (%)	Link to primary data
EC_5YSURVRATE	The 5-year survival rate of companies	Calculation of a survival rate between 2 non-consecutive years.	%	INSEE, ODT, REE	2009-2014	0.21	Calculated	11.1 – [0 ; 50]	https://www.insee.fr/fr/statistiques/2521005?sommaire=2544975
EC_EMPLOY	Employment rate	Proportion of economically active persons aged 15-64 (in employment) in the total population aged 15-64.	%	INSEE, ODT	2013	0	Raw	65.6 – [61.2 ; 69.1]	https://www.insee.fr/fr/statistiques/2861854
EC_EVEMPATTEND	Employment evolution for attendance activities	Attendance activities are defined as activities that are implemented locally for the production of goods and services designed to satisfy the needs of residents or tourists in the municipality.	%	INSEE	2009-2014	0.29	Calculated	0 – [-0.3 ; 0.4]	https://www.insee.fr/fr/statistiques/1893206
EC_EVLABFOR	Evolution of the labour force	Calculation of a change in the number of people aged 15-64 in the labour force between 2 non-consecutive years.	%	INSEE	2008-2013	0	Calculated	0.43 – [-2.1 ; 2.89]	https://www.insee.fr/fr/statistiques/2518836
EC_INCOME	Average income per consumption unit	The disposable income of a household comprises earned income (net of social contributions), income from assets, transfers from other households and social benefits (including pensions and unemployment benefits), and net of direct taxes. The consumption unit (CU) is an OECD unit of measurement that weights individuals in a household according to their age, allowing comparison between households. CU are generally calculated in the following manner: 1 CU is allocated to the first adult in the household, 0.5 CU to other persons aged 14 or over and 0.3 CU to children under 14.	€/CU	INSEE, ODT	2012	0	Raw	19,546.7 – [17,904.7 ; 21,412.7]	https://www.insee.fr/fr/statistiques/2388572 http://carto.observatoire-des-territoires.gouv.fr/#l=fr;i=filosofi.med_disp;v=map56
EC_MUNLIVEWORK	Share of the labour force living and working in the same municipality	The labour force considered is aged 15 and over.	%	INSEE	2014	0	Calculated	15.6 – [11.9 ; 20.5]	https://www.insee.fr/fr/statistiques/2861854
EC_PARTLABFOR	Evolution of the share of the labour force in the total population	Calculation of the evolution of the share of the labour force aged 15 and over within the total population between 2 non-consecutive years.	%	INSEE	2009-2014	0.29	Calculated	-1.2 – [-5.1 ; 3.3]	https://www.insee.fr/fr/statistiques/2861854 https://www.insee.fr/fr/statistiques/2862200
EC_PRECARIOUS	Precarious employment rate	Share of the labour force aged 15-64 in fixed-term contracts, temporary employment, assisted employment and apprenticeships among the labour force aged 15-64.	%	INSEE	2014	0	Calculated	7.6 – [6.1 ; 9.1]	https://www.insee.fr/fr/statistiques/2863622 https://www.insee.fr/fr/statistiques/2861854

Tableau IV : Article 2 – Table II: Metadata related to the “Economics” dimension (final section)

Acronym	Data	Specifications	Unit	Source(s)	Time period	Missing data (%)	Raw / Calculated	Median – [Q1 ; Q3] or modalities (%)	Link to primary data
EC_SHANSPC	Shannon-Weaver index on the representation of socio-professional categories	Use of the Shannon-Weaver index to describe a population in terms of the number of categories (category richness) and their relative abundance within those categories (species equitability). 15 socio-professional categories were considered for the calculation, identified at the place of residence.	Index	INSEE	2014	0.29	Calculated	0.7 – [0.5 ; 0.9]	https://www.insee.fr/fr/statistiques/1893206
EC_UNEMPLOY	Unemployment rate	Share of people aged 15-64 who are unemployed and looking for work.	%	INSEE	2014	0	Raw	8.5 – [6.6 ; 10.9]	https://www.insee.fr/fr/statistiques/2861854

Tableau V : Article 2 – Table III: Metadata related to the “Environment” dimension

Acronym	Data	Specifications	Unit	Source(s)	Time period	Missing data (%)	Raw / Calculated	Median – [Q1 ; Q3] or modalities (%)	Link to primary data
EN_ACCGSP ACE	Share of population with access to green space within 400 meters of their homes	Green spaces of less than 0.5ha have been excluded. The distance of 400 meters has been used as it is more widely accepted in the literature than the 300-meter distance recommended by the WHO. This criterion included parks, forests and playground. It excludes areas that are inaccessible to the general population (e.g. campsites, cemeteries, golf courses, grassy areas, stadiums and zoos). The population data are those of the INSEE's 200-meter grid.	%	INSEE, Geofabrik	2011-2017	0	Calculated	97.2 – [82.0 ; 100]	http://download.geofabrik.de/europe/france.html https://www.insee.fr/fr/statistiques/2520034
EN_GHGEM IS	Greenhouse gas (GHG) emissions	The Global Warming Potential (GWP) expresses, in thousands of tonnes of CO ₂ equivalent, the cumulative effect of the substances emitted into the air that contribute to the increase in the greenhouse effect. It is calculated on the basis of the warming power of each gas over the next 100 years compared to that of CO ₂ . The following direct greenhouse gases are taken into account: CO ₂ (non-biotic), CH ₄ , N ₂ O, HFCs, PFCs and SF ₆ . GWP expressed here excludes carbon sinks. These sinks mainly correspond to the absorption of CO ₂ for vegetation growth and methane absorption by forest soils. Similarly, emissions from biotic sources (non-managed forests, natural grasslands, wetlands and rivers, forest fires) are not included. Solely the emissions caused by human activities or managed by humans are taken into account. These exclude emissions from international shipping and aviation.	Thousands of TEQ CO ₂	ODT	2012	0.58	Raw	4 – [2 ; 8]	http://www.observatoire-des-territoires.gouv.fr/observatoire-des-territoires/fr/emissions-de-gaz-%C3%A0-effet-de-serre-hors-puits
EN_HESV	Monetisation of ecosystem services per capita	Use of the Corine Land Cover database to determine 4 land-use categories (grassland, urban, cultivated and/or arable land, forests) and The Economics of Ecosystems and Biodiversity valuation database to assign a monetary value to the ecosystem services provided by these 4 categories (de Groot <i>et al.</i> , 2012; van der Ploeg <i>et al.</i> , 2010).	USD/ha/year	INSEE, MEDDE	2012-2014	0	Calculated	5.8 – [2.6 ; 11.6]	https://www.statistiques.developpement-durable.gouv.fr/corine-land-cover-0 https://www.insee.fr/fr/statistiques/2862200
EN_M2NAT SPACE	Area (m ²) of natural spaces per inhabitant	Use of open street map data (land use, points of interest): areas that are inaccessible to the general public (parks, playgrounds, campsites, cemeteries, grassy areas, golf courses, stadiums, zoo) were also included. Water surface (lakes, ponds, rivers, etc.) data were added.	m ² /inhab	INSEE, Geofabrik	2014-2017	0	Calculated	1,872.4 – [648.5 ; 4,619.0]	http://download.geofabrik.de/europe/france.html https://www.insee.fr/fr/statistiques/2520034
EN_PSSNB	Number of contaminated sites (polluted sites and soils)	Polluted sites and soils are sites that present an actual or potential perennial risk to human health or the environment due to pollution caused by current or former human activity. The database registers all the sites that require or are the subject of public, preventive or curative action. These are therefore situations that are clearly identified, dealt with, being dealt with or about to be dealt with.	Count	GEOIDD, MEDDE, DGPR	2015	0	Raw	0 – [0 ; 0]	http://geoidd.developpement-durable.gouv.fr/geoclip_stats_o3/#l=fr;i=sol_pollution.sitpol_nb;v=map1

Tableau VI : Article 2 – Table IV: Metadata related to the “Policy” dimension

Acronym	Data	Specifications	Unit	Source(s)	Time period	Missing data (%)	Raw / Calculated	Median – [Q1 ; Q3] or modalities (%)	Link to primary data
PO_AGENDA21	Municipality covered by an Agenda 21	Agenda 21 are developed at the local authority or sub-departmental level as a measure of sustainable development. This action plan reflects the voluntary commitment of communities to sustainable development. Qualitative data recorded in a binary variable.	-	GEOIDD, MEDDE, AGENDA21FRANCE.ORG	2014	0	Raw	0: not covered by an Agenda 21 (81.3 %) 1: covered by an Agenda 21, local and/or sub-regional (18.7 %)	http://geoidd.developpement-durable.gouv.fr/geoclip_stats_o3/#s=2014;l=fr;i=agenda21.p_a21_p_p_infradep;v=map1
PO_EXPTOTINVEST	Share of expenditure allocated to municipal facilities and services compared to the total investment	The total investment corresponds to the municipality's investments (equipment, urban development, etc.) and loan repayments.	%	DGFIP	2014	0.68	Calculated	81.4 – [60.5 ; 93.4]	https://www.data.gouv.fr/fr/datasets/donnees-comptables-et-fiscales-des-collectivites-locales/
PO_FINANPOT	Financial potential per inhabitant	The financial potential measures the wealth of a municipality by calculating all the stable resources of a local authority in relation to the reference population. These resources include all tax resources and allocations paid by the State on an automatic basis or recurring basis, and which are essential elements for the balancing of local authority budgets.	€/inhab	ODT, DGCL, INSEE	2013	0	Raw	567 – [486.4 ; 723.2]	http://www.observatoire-des-territoires.gouv.fr/observatoire-des-territoires/fr/potentiel-financier-par-habitant
PO_HOUSINGTAX	Housing tax proceeds compared to the number of primary and secondary residences	The housing tax is a local tax which varies according to the characteristics of a home, its location and the personal situation of its occupants (income, composition of the household, etc.) on 1 st January every year. Owners, tenants and free occupants of their main dwelling must pay this tax, and a housing tax must also be paid for a secondary residence. The total sum is then divided by the number of primary and secondary residences counted within the municipality.	€/number of primary and secondary residences	DGFIP, INSEE	2015	0	Raw	472.4 – [392.8 ; 567.9]	http://carto.observatoire-des-territoires.gouv.fr/#s=2015;z=10411,6631335,650975,477012;l=fr;i=fisc_loc.th_prod_locaux;v=map1
PO_MUNDEBT	Proportion of medium- and long-term debt of the municipality, per inhabitant	It measures the medium and long-term debt burden of the municipality. The annuity of the debt is obtained by summing the interest on the loans of the municipality and the amount of the principal repayment. The addition of these two amounts permits the measurement of the exact weight of the long and medium-term debt for municipalities.	€/inhab	DGFIP, INSEE	2014	0	Calculated	44.7 – [18.8 ; 77.1]	https://www.data.gouv.fr/fr/datasets/donnees-comptables-et-fiscales-des-collectivites-locales/ https://www.insee.fr/fr/statistiques/2862200

Tableau VI : Article 2 – Table IV: Metadata related to the “Policy” dimension (cont.)

Acronym	Data	Specifications	Unit	Source(s)	Time period	Missing data (%)	Raw / Calculated	Median – [Q1 ; Q3] or modalities (%)	Link to primary data
PO_MUNEXPEQ UIP	Expenditure for municipal equipment, per inhabitant	This expenditure corresponds to the expenditure of the municipality on tangible and intangible fixed assets or fixed assets under construction (construction of a school, a tramway line or a sports facility), in relation to the reference population.	€/inhab	DGFIP, INSEE	2014	0	Calculated	139.8 – [59.2 ; 297.6]	https://www.data.gouv.fr/fr/datasets/donnees-comptables-et-fiscales-des-collectivites-locales/ https://www.insee.fr/fr/statistiques/2862200
PO_MUNSUBS	Amount of subsidies paid by municipalities, per inhabitant	The subsidies are paid for sports and cultural associations, communal centre for social welfare, local mixed-economy companies, structures such as tourist boards, etc, in relation to the reference population. The subsidies are part of the municipality's operating costs (costs of personal among others), which take priority over investment costs.	€/inhab	DGFIP, INSEE	2014	0	Calculated	13.1 – [6.6 ; 23.9]	https://www.data.gouv.fr/fr/datasets/donnees-comptables-et-fiscales-des-collectivites-locales/ https://www.insee.fr/fr/statistiques/2862200
PO_MUNRESULT	Overall financial result of the municipality, per inhabitant	This result corresponds to the difference between the accounting result (difference between total operating income and operating expenses) and the financing requirement or capacity of the investment section, in relation to the reference population.	€/inhab	DGFIP, INSEE	2014	0	Calculated	108.5 – [9.7 ; 249.5]	https://www.data.gouv.fr/fr/datasets/donnees-comptables-et-fiscales-des-collectivites-locales/ https://www.insee.fr/fr/statistiques/2862200
PO_RAZ	Eligible areas for regional aid zoning	Companies based in any municipalities located within an AFR ('Aide à Finalité Régionale') zone can be allocated aids for investment and job creation by the State and local authorities. Qualitative data (recoded in a binary variable).	-	ODT	2015	0	Raw	0: non-eligible (82.0 %) 1: partially or completely eligible (18.0 %)	http://www.observatoire-des-territoires.gouv.fr/observatoire-des-territoires/fr/zones-daide-%C3%A0-finalit%C3%A9-r%C3%A9gionale
PO_SUBSOPEXP	Share of subsidies paid within operating expenses	Operating expenses mainly include personnel costs, compulsory contributions, in particular to the departmental fire and rescue service, subsidies paid and interest on loans.	%	DGFIP	2014	0	Calculated	2.6 – [1.4 ; 4.4]	https://www.data.gouv.fr/fr/datasets/donnees-comptables-et-fiscales-des-collectivites-locales/

Tableau VI : Article 2 – Table IV: Metadata related to the “Policy” dimension (final section)

Acronym	Data	Specifications	Unit	Source(s)	Time period	Missing data (%)	Raw / Calculated	Median – [Q1 ; Q3] or modalities (%)	Link to primary data
PO_TAXHOUSE HOLDS	Share of taxable tax households	The share of tax households taxed is the percentage of households that are liable to pay income tax. The tax liability of a tax household is the sum of the taxes payable by each person in the household.	%	INSEE, ODT	2013	2.07	Raw	52.1 – [48.3 ; 56.4]	http://carto.observatoire-des-territoires.gouv.fr/#v=map8;i=filosofi.men_jmpos;l=fr;z=25439,6620976,570632,367308
PO_TAXPOT	Tax potential per inhabitant	Tax potential is an indicator that can be used to compare the potential tax wealth of communities. A theoretical tax product is calculated, corresponding to the amount of taxes that each community would collect if it applied average rates or tariffs to its net tax bases. This tax potential is related to the number of inhabitants in order to allow analyses and comparisons that are not possible using gross amounts alone.	€/inhab	ODT, DGCL	2017	0.21	Raw	476.7 – [387.4 ; 622.4]	http://carto.observatoire-des-territoires.gouv.fr/#l=fr;i=potentiel_fin.pot_fisc_hab;v=map43

Tableau VII : Article 2 – Table V: Metadata related to the “Health” dimension

Acronym	Data	Specifications	Unit	Source(s)	Time period	Missing data (%)	Raw / Calculated	Median – [Q1 ; Q3] or modalities (%)	Link to primary data
HE_ACCGP	Potential spatial accessibility to general practitioners	This indicator is based on three variables: demand, supply, distance. It is calculated at the municipal level, taking into account the supply and demand of neighbouring municipalities. This accessibility is considered ‘potential’ because it is based on the geographical location of demand (population) and supply (general practitioners). It measures the activity level of general practitioners, based on the number of consultations and visits recorded over one year. This number is converted into Full-Time Equivalents (FTE).	Full-time equivalent/inhab.	DREES	2013	0	Raw	47.7 – [31.9 ; 67.4]	http://www.data.drees.sante.gouv.fr/ReportFolders/reportFolders.aspx
HE_ACCPHARM	Potential spatial accessibility to pharmacist	This indicator is based on three variables: demand, supply, distance. It is calculated at the municipal level, taking into account the supply and demand of neighbouring municipalities. This accessibility is considered ‘potential’ because it is based on the geographical location of demand (population) and supply (pharmacist). It measures the activity level of pharmacists, based on the number of consultations and visits recorded during a year. These acts are converted into Full-Time Equivalents (FTE).	Full-time equivalent/inhab.	DREES	2013	0	Raw	23.2 – [15.6 ; 33.0]	http://www.data.drees.sante.gouv.fr/ReportFolders/reportFolders.aspx
HE_BIRTH	Proportion of births	Locally registered births correspond to births counted from civil status reports, and recorded at the mother's home address.	Number per 1,000 inhab.	INSEE	2014	0	Raw	1.1 – [0.7 ; 1.5]	https://www.insee.fr/fr/statistiques/2521169
HE_DEATH	Proportion of deaths	The place of death is the last home address of the deceased and not the place of death.	Number per 1,000 inhab.	INSEE	2014	0.47	Raw	0.7 – [0.4 ; 1.1]	https://www.insee.fr/fr/statistiques/2521169
HE_HEALTHPROF	Number of health professionals, per 1,000 inhabitants	Number of health professions, all specialities combined, counted at the zip-code of the place of activity, per 1,000 inhabitants.	Number per 1,000 inhab.	INSEE	2016	0.21	Calculated	0 – [0 ; 3.2]	https://www.insee.fr/fr/statistiques/3568614?sommaire=3568656
HE_SHANMED	Shannon-Weaver index on the representation of medical and paramedical professions	Use of the Shannon-Weaver index to describe a population in terms of the number of categories (category richness) and their relative abundance within those categories (category equitability). 25 medical and paramedical professions were considered for the calculation.	Index	INSEE	2016	0.03	Calculated	0 – [0 ; 0.3]	https://www.insee.fr/fr/statistiques/3568614?sommaire=3568656

Tableau VIII : Article 2 – Table VI: Metadata related to the “Social” dimension

Acronym	Data	Specifications	Unit	Source(s)	Time period	Missing data (%)	Raw / Calculated	Median – [Q1 ; Q3] or modalities (%)	Link to primary data
SO_DENSITY	Population density	Number of inhabitants per ha.	inhab/ha	INSEE	2014	0.21	Calculated	59.1 – [33.3 ; 131.0]	https://www.insee.fr/fr/statistiques/2862200
SO_EMPEDUC	Employment in education and training, per 1,000 inhabitants	Employment in education and training corresponds to jobs in school and university education (primary, secondary and higher education) and vocational training, including the organisation of such education.	Number per 1,000 inhab	INSEE	2014	0	Calculated	0 – [0 ; 9.0]	https://www.insee.fr/fr/statistiques/1893206
SO_GRIDDENS	Communal density grid	In order to consider the municipal population and its distribution in space, the communal density grid is based on the distribution of the population within the municipality by dividing the territory into one-kilometre squares. The size of these agglomerated areas within the municipalities will make it possible to characterize them (rather than the usual communal density). Qualitative data.	-	INSEE, ODT	2015	0.24	Raw	1: densely populated area (2.1 %) 2: intermediate densely populated area (13.6 %) 3: sparsely populated area (56.1 %) 4: very sparsely populated area (28.2 %)	http://www.observatoire-des-territoires.gouv.fr/observatoire-des-territoires/fr/grille-communale-de-densite
SO_IMMIG	Share of immigrants in the total population	An immigrant is a person born abroad and residing in France. A distinction is made between foreign and immigrant populations: an immigrant is not necessarily a foreigner and conversely, some foreigners were born in France. Immigrant status is permanent: an individual continues to belong to the immigrant population even if he/she becomes French by acquisition.	%	INSEE	2015	0	Raw	1.5 – [0.8 ; 2.8]	https://www.insee.fr/fr/statistiques/3564100?sommaire=3561107 https://www.insee.fr/fr/statistiques/3561125?sommaire=3561139
SO_LESSQUALIF	Share of working-age population with few or no qualifications	Ratio between the number of out-of-school persons aged 15 years or over with no diploma (or at most a BEPC which is a French certificate of general education, awarded by schools at the end of the first four years of general secondary education) and the total number of out-of-school persons aged 15 or over.	%	INSEE	2014	0.29	Calculated	34 – [28 ; 40]	https://www.insee.fr/fr/statistiques/2862015
SO_POP65	Share of the population aged over 65 years	Percentage of the population aged over 65 years in the total population.	%	INSEE	2015	0	Calculated	16.3 – [13.7 ; 19.3]	https://www.insee.fr/fr/statistiques/3564100?sommaire=3561107

Tableau VIII : Article 2 – Table VI: Metadata related to the “Social” dimension (final section)

Acronym	Data	Specifications	Unit	Source(s)	Time period	Missing data (%)	Raw / Calculated	Median – [Q1 ; Q3] or modalities (%)	Link to primary data
SO_PUBSCHOOL	Number of state schools, per 1,000 inhabitants	This correspond to the sum of the number of state nursery schools, state elementary schools and state elementary schools with a pre-elementary class.	Number per 1,000 inhab	INSEE, BPE, GEOIDD	2014	0	Calculated	1.4 – [0 ; 2.9]	http://geoidd.developpement-durable.gouv.fr/geoclip_stats_o3/#l=fr;v=map1
SO_SINGLEFAM	Share of single-parent families	A single-parent family includes a lone parent and one or more unmarried (childless) children.	%	ODT	2013	0.13	Raw	9.6 – [5.1 ; 14.3]	http://carto.obse-rvatoire-des-territoires.gouv.fr/#s=2015;l=fr;ji=insee_rp_hist_xxx_x.part_menfammono;v=map56
SO_SOCBENEF	Share of persons covered by a social benefit	Percentage of persons covered by a benefit from the family branch, such as family benefits, early childhood benefits, family supplement, education allowance for a disabled child, family support allowance, back-to-school allowance, family housing allowance, social housing allowance, etc.	%	CNAF	2014	2.07	Raw	44.6 – [39.6 ; 49.6]	http://data.caf.fr/dataset/population-des-foyers-allocataires-par-commune
SO_STUDY	Study continuation rate for 18-24 years old	Share of the population aged 18 to 24 pursuing higher education.	%	INSEE	2013	0.65	Calculated	39.1 – [30.0 ; 49.2]	https://www.insee.fr/fr/statistiques/2044692

Tableau IX : Article 2 – Table VII: Metadata related to the “Services” dimension

Acronym	Data	Specifications	Unit	Source(s)	Time period	Missing data (%)	Raw / Calculated	Median – [Q1 ; Q3] or modalities (%)	Link to primary data
SE_4GCOV	Part of area with 4G cover provided by the most efficient operator	This indicator represents the share of areas covered for 4G by the operator offering the best coverage rate in the municipality. It reflects the availability, outside buildings, of access to a service, as displayed by the operators on their coverage maps.	%	ODT	2016	0	Raw	92 – [57 ; 99]	http://www.observatoire-des-territoires.gouv.fr/observatoire-des-territoires/fr/node/2289
SE_EQUIP15MIN	Share of population with average access to 12- mid-range municipal facilities and services in 15 minutes or less	The equipment in the intermediate range are in common use, without being close by. These may include supermarkets, bookshops and stationery shops, clothing stores, college, opticians, MOT centres, etc.	%	ODT	2011	0	Raw	94 – [85.5 ; 97.7]	http://www.observatoire-des-territoires.gouv.fr/observatoire-des-territoires/de/voluntion-de-la-part-de-la-population-ayant-acc-s-en-moyenne-aux-12-quipements-de-la-gamme-interm-dia?rech=1
SE_LRH	Low-Rent Housing (LRH) per 1,000 inhabitants	Low-rent housing is housing units managed by a public or private low-rent housing organization, which benefits from partial, direct (subsidy) or indirect public funding (various advantages including credits, tax exemptions, etc.).	Number per 1,000 inhab	INSEE	2013	0	Raw	0 – [0 ; 18.9]	https://www.insee.fr/fr/statistiques/2044711
SE_OWNER	Proportion of people owning their main residence	Ownership status applies to owner, co-owner and first-time homeowner households.	%	INSEE	2014	0.21	Calculated	81.5 – [73.9 ; 86.5]	https://www.insee.fr/fr/statistiques/2862034
SE_PRIMROAD	Length of primary roads per capita	The primary road is defined as a portion of a roadway intended for pedestrian or cycles and includes only sections of major road networks.	m/inhab	IGN, INSEE	2016	0.08	Calculated	0 – [0 ; 4.1]	http://professionnels.ign.fr/bdtopo https://www.insee.fr/fr/statistiques/3677785?sommaire=3677855
SE_TENANT	Share of population who are tenants of their main residence	The rental stock comprises the public and the private sectors.	%	INSEE	2014	0.29	Calculated	16.7 – [11.7 ; 24]	https://www.insee.fr/fr/statistiques/2862034
SE_TENANTLRH	Share of tenants who rent their main residence in Low-Rent Housing	LRH dwellings offer rents well below those of the private rental sector.	%	INSEE	2014	0.29	Calculated	0 – [0 ; 4.8]	https://www.insee.fr/fr/statistiques/2862034

2. Experimental Design, Materials, and Methods

2.1. Evolution calculations

This section concerns four data related to “Economics” dimension (Table II: the 5-year survival rate of companies “EC_5YSURVRATE”, the Employment evolution for attendance activities “EC_EVEMPATTEND”, the evolution of the labour force “EC_EVLABFOR” and the evolution of the share of labour force in total population “EC_PARTLABFOR”). Only data corresponding to year n and $(n - 5)$ were collected for the evolution calculation. Considering a year n (year 2013: “EC_EVLABFOR”; year 2014: “EC_5YSURVRATE”, “EC_EVEMPATTEND”, “EC_PARTLABFOR”), the evolution was computed as follows:

$$\text{Evolution} = \frac{\text{Raw data (year (n))} - \text{Raw data (year (n - 5))}}{\text{Raw data (year (n - 5))}} * 100$$

2.2. Raw data converted to per capita

This section concerns one data related to “Environment” dimension (Table III: the area (m²) of natural spaces per inhabitant “EN_M2NATSPACE”), four data related to “Policy” dimension (Table IV: the expenditure for the municipality equipment per inhabitant “PO_MUNEXPEQUIP”, the amount of subsidies paid by municipalities per inhabitant “PO_MUNSUBS”, the overall financial result of the municipality per inhabitant “PO_MUNRESULT”, the proportion of medium- and long-term debt of the municipality per inhabitant “PO_MUNDEBT”) and one data related to “Services” dimension (Table VII: the length of primary roads per capita “SE_PRIMROAD”). Considering a year n , the per capita variable was computed as follows:

$$\text{Per capita variable} = \frac{\text{Raw data (year (n))}}{\text{Population reference (year (n))}}$$

2.3. Raw data per 1,000 inhabitants

This section concerns one data related to “Health” dimension (Table V: the number of health professionals, per 1,000 inhabitants “HE_HEALTHPROF”) and two data related to “Social” dimension (Table VI: the employment in education and training per 1,000 inhabitants “SO_EMPEDUC” and the number of state schools per 1,000 inhabitants “SO_PUBSCHOOL”). Considering a year n , the raw data per 1,000 inhabitants was computed as follows:

$$\text{Variable per 1,000 inhabitants} = \frac{\text{Raw data (year (n))}}{\text{Population reference (year (n))}} * 1,000$$

2.4. Raw data converted to rate

This section concerns two data related to “Economics” dimension (Table II: the share of the labour force living and working in the same municipality “EC_MUNLIVEWORK”, the precarious employment rate “EC_PRECARIOUS”), two data related to “Policy” dimension (Table IV: the share of expenditure allocated to municipal facilities and services to the total investment “PO_EXPTOTINVEST”, the share of subsidies paid within the operating expenses “PO_SUBSOPEXP”), three data related to “Social” dimension (Table VI: the share of working age population with few or no qualifications “SO_LESSQUALIF”, the share of the population aged over 65 years “SO_POP65”, the study continuation rate for 18-24 years old “SO_STUDY”) and three data related to “Services” dimension (Table VII: the proportion of people owning their main residence “SE_OWNER”, the share of the population who are tenants of their main residence “SE_TENANT”, the share of tenants who rent their main residence in

low-rent housing “SE_TENANTLRH”). Considering a year n , the raw data converted to rate was computed as follows:

$$\text{Percentage} = \frac{\text{Raw data (year (n))}}{\text{Global amount (year (n))}} * 100$$

2.5. Calculation of the share of population with access to green space within 400 meters of their homes (EN_ACCGSPACE)

This data was based on WHO recommendations (WHO, 2016), which proposes in 2016, an indicator related to the health benefits of green spaces: the share of the population living within 300 meters of an area of at least 0.5 ha. In our study, the distance of 400 meters was chosen for the calculation because this distance is currently consensual in the scientific literature (Table III). In the review of Labib *et al.* (2020), this distance of 400 meters is considered to reflect a walking distance of five to 10 min and is the median distance for some health studies.

We extracted data from Open Street Map (Geofabrik GmbH) to determine the green space surface area. Parks, forests and playgrounds greater than 0.5 ha have been included whereas areas not accessible to the general population (camping sites, cemeteries, golf courses, grassy areas, stadiums and zoos) have been excluded. We also used the 200*200m gridded population data from the INSEE. Geographic Information System (ESRI ArcGIS v10.7) were used to generate a 400-meters buffer zone around each green space. The point centroid of each population grid was then intersected with the buffers in order to calculate the proportion of population living in the buffer in a municipality.

2.6. Calculation of the monetisation of ecosystem services per capita (EN_HESV)

This section concerns the monetisation of ecosystem services per capita “EN_HESV”, which was inspired by the work of Estoque and Murayama, (2014) (Table III). We downloaded from MEDDE the Coordination of Information on the Environment (CORINE) Land Cover (CLC) database, a pan-European geographic database, which describes the biophysical land use. Among the 5 major classes (*i.e.* artificial territories, agricultural territories, forests and semi-natural environments, wetlands, water surfaces), 44 land cover types are identified by a three-digit notation. In our study, this database was used to create 4 specific land use categories:

- “urban”, including codes 111 (continuous urban fabric), 112 (continuous urban fabric), 121 (industrial and commercial areas), 122 (road and rail networks and associated areas), 123 (port areas) and 124 (airports),
- “cultivated and/or arable land”, including codes 211 (non-irrigated arable land), 212 (permanently irrigated perimeter), 213 (rice paddies), 221 (vineyards), 222 (orchards and berries), 231 (pastures), 241 (annual crops associated with permanent crops), 242 (complex cropping and parcel systems) and 243 (mainly agricultural areas, interrupted by large natural areas),
- “grasslands”, including codes 321 (natural lawns and pastures), 322 (heath and scrub) and 323 (sclerophyllous vegetation),
- “forests”, including codes 311 (deciduous forests), 312 (coniferous forests), 313 (mixed forests) and 324 (changing forest and shrub vegetation).

The area in hectares for each of these 4 categories was calculated for each municipality, using Geographic Information System tools. Thereafter, the Economics of Ecosystems and Biodiversity (TEEB) Valuation Database from works by de Groot *et al.* (2012) and van der Ploeg *et al.* (2010) was studied to establish a monetary value to the ecosystem services provided by these four categories. Within TEEB, priority has been given to publications and results from the European Union in order to

remain as close as possible to the environmental context of the Hauts-de-France Region. An exception has been made for “grasslands” since there has been only one study for Europe, located in Spain. Due to the difference in the definition of “grasslands” between Spain and France, the global value has been preferentially considered.

The monetary values that have been retained were as follows:

- Urban: 6,111 USD/ha/year,
- Cultivated and/or arable lands: 2,140 USD/ha/year,
- Grasslands: 2,789 USD/ha/year,
- Forests: 3,789 USD/ha/year.

The human-to-ecosystem service value ratio index (H-ESV) was computed as follows:

$$H-ESV = \frac{\sum_{i=1}^n \text{area}_i \times \text{Monetary Value}_i}{\text{Total population}}$$

Where $i, 1 \leq i \leq 4$, corresponding to the 4 land cover categories, area is expressed in hectare, the monetary value in USD/ha/year and the total population data comes from the 2014 population census.

2.7. Shannon-Weaver index

This section concerns the calculation of the Shannon-Weaver index (Shannon, 1948), applied to (i) the representation of socio-professional categories (Table II: the Shannon-Weaver index on the representation of Socio-Professional Categories “EC_SHANSPC”) and (ii) the representation of medical and paramedical professions (Table V: the Shannon-Weaver index on the representation of medical and paramedical professions “HE_SHANMED”). The Shannon-Weaver index was computed as follows:

$$H' = \sum_{i=1}^S p_i \times \log(p_i), \quad \text{with } p_i = \frac{n}{N}$$

Where S is the total number of categories, p_i is the proportional abundance or the percentage of importance of the category, n is the number of individuals in a category and N is the total number of individuals in all categories.

Shannon Weaver’s Index allows a stand to be described by taking into account the number of categories (categories richness) and their relative importance within those categories (categories equitability). The value of the index ranges from 0 (*i.e.* a single category, or one category dominating all others) to $\log S$ (*i.e.* all categories are equally abundant).

2.7.1. Calculation of Shannon-Weaver index on the representation of Socio-Professional Categories (EC_SHANSPC)

The data used for this calculation came from an official inventory of the Socio-Professional Categories (SPC) located in each municipality (INSEE). This functional job analysis was provided at workers’ home address and each class is presented below (in EC_SHANSPC, $S = 15$):

- Public administration: jobs related to the sovereign and administrative activities of the State and local authorities, excluding health and education services. In particular, this category includes all the professions of public security and justice,
- Agriculture and fisheries: all occupations directly involved in agricultural production, fishing or forestry,
- Building and public works: all occupations directly involved in the construction of buildings and public works,

- Business-to-business trade: occupations directly involved in wholesale trade and trade between businesses, whether for purchase or sale,
- Design/research: in industry, these professions are involved in the phases prior to manufacturing. This category differs from the function of intellectual services by the innovation dimension included in the work of the trades concerned,
- Culture/leisure: cultural and leisure occupations, including but not limited to sportsmen and women;
- Distribution: all professions involved in sales to private individuals, including commercial craftsmanship,
- Education/training: occupations in school and university education (primary, secondary and higher education) and vocational training, including the organisation of such education. This function does not include sports or leisure activity leaders, who are included in the culture/leisure category,
- Maintenance/repair: professions primarily oriented towards upkeep and maintenance (excluding building and public works), as well as waste treatment (and thus by extension, the environment),
- Manufacturing: all professions consisting of implementing technical equipment or processes, excluding agriculture and fishing and construction and public works. For the most part, these are occupations that contribute directly to the various stages of the production of material goods and energy,
- Management: professions in business management, banking and insurance,
- Transport/logistics: professions in the transport of people and the flow of goods,
- Intellectual services: professions providing specific knowledge for consultancy, analysis, expertise, *etc.*,
- Health/social work: health and social work professionals, including pharmacists,
- Community services: professions providing everyday services (excluding distribution, transport, education and health).

2.7.2. Calculation of Shannon-Weaver index on the representation of medical and paramedical professions (HE_SHANMED)

The data used for this calculation came from an official inventory of the 24 medical and paramedical professions located in each municipality (INSEE). The information for all medical and paramedical specialities was provided at the professional activity location. The following medical and paramedical professions were considered (in HE_SHANMED, $S = 24$):

- the general practitioners,
- the cardiology specialists,
- the specialists in dermatology venereology,
- the specialists in gynaecology,
- the specialists in gastroenterology hepatology,
- the specialists in psychiatry,
- the specialists in ophthalmology,
- the specialists in otorhinolaryngology,
- the specialists in paediatrics,
- the specialists in pulmonology,
- the specialists in x-ray diagnosis,
- the specialists in stomatology,
- the dental surgeons,

- the midwives,
- the nurses,
- the masseur-physiotherapists,
- the speech therapists,
- the orthoptists,
- the chiropodists,
- the audioprosthetists,
- the occupational therapists,
- the psychometricians,
- the dieticians,
- the psychologists.

Article 3

How can we analyze environmental health resilience and vulnerability? A joint analysis with composite indices applied to the north of France

Delphine Brousmiche, Michaël Genin, Florent Occelli, Lukas Frank, Annabelle Deram, Damien Cuny, Caroline Lanier

Science of The Total Environment 142983

<https://doi.org/10.1016/j.scitotenv.2020.142983>



Le « *supplementary material and results* » de cet article est disponible dans l'Annexe 1.

How can we analyze environmental health resilience and vulnerability? A joint analysis with composite indices applied to the north of France

Delphine Brousmiche¹, Michaël Genin², Florent Occelli^{1,3}, Lukas Frank¹, Annabelle Deram^{1,3}, Damien Cuny^{1,4}, Caroline Lanier*^{1,3}

¹ Univ. Lille, IMT Lille Douai, Univ. Artois, Yncrea Hauts-de-France, ULR 4515 - LGCgE, Laboratoire de Génie Civil et géo-Environnement, F-59000 Lille, France

² Univ. Lille, CHU Lille, ULR 2694 – METRICS, F-59000 Lille, France

³ Faculté d'Ingénierie et Management de la santé (ILIS)/Faculté de pharmacie de Lille – LSVF

⁴ Faculté de Pharmacie de Lille – LSVF

*Corresponding author: caroline.lanier@univ-lille.fr

Abstract

In environmental health, vulnerability reflecting the cumulative harmful constraints and nuisances to which populations are subjected and resilience defined as the capacity of a territory to cope with health inequalities have been little extensively investigated together with the same importance. Besides the diversity of factors involved, there is no consensual framework to develop composite indices, one recognized methodology to deal with a multifaceted issue. Therefore, this research aims to establish a new transferable approach to assess the spatial heterogeneity of territorial inequalities. This new strategy relies on the simultaneous evaluation of resilience and vulnerability and the joint analysis based on the cross-interpretation of the spatialized composite indices of resilience and vulnerability. A case study was conducted to demonstrate the feasibility of this methodology, using the municipality as a spatial unit of analysis within a region in the north of France. To provide the most holistic description possible of the 3,817 studied municipalities, 50 variables related to the economic, environment, policy, health, services and social dimensions were used to develop the composite indices. The vulnerability Index has a median value of 0.151 with an IQR of [0.126-0.180] and the Resilience Index has a median value of 0.341 with an IQR of [0.273-0.401]. The joint analysis was conducted to classify each municipality among four defined typologies: 1,687 municipalities (44.2%) belong to the “To monitor” category, 1,646 (43.1%) to the “Resilient” category, 329 (8.6%) to the “Have resources” category and 155 (4.1%) to the “Territorial blackspot” category. The methodology herein may be a diagnostic tool to identify and prioritize municipalities that could benefit from the implementation of specifically tailored public health policies.

Keywords

Environmental health determinants, territorial inequalities, multidimensional indices, territorial blackspot, diagnostic tool

1. Introduction

Territorial inequalities in environmental health are regarded as an accumulation of unfavourable health determinants in a given geographical area. The complex interactions between these determinants of health, which are interconnected and evolve over time and space, strongly influence the health of populations (Sarkar and Webster, 2017). The 2012 Lancet Commission report highlights the need for a complex systems approach to characterize the multi-level processes (*e.g.* environmental, social, urban planning) that influence the health of populations (Rydin *et al.*, 2012).

The diversity of factors involved in environmental health, the complexity of the interactions to be considered and the spatial and temporal variability of the phenomena observed require the validation and use of multidimensional approaches that are more than a simple aggregation of health determinants (Briggs, 2008). One recognized method used to analyse this type of complex problem is based on the creation of spatialized composite indices combining different indicators that measure and transcribe the multifaceted nature of a phenomenon (Beale *et al.*, 2008; European Commission *et al.*, 2008; Flacke, 2015; Nascimento and Carrage, 2007). We have identified 23 spatialized health-environment indices at the international level (Brousmiche *et al.*, 2020b). The indices published to date are consistent with the consensual definition of territorial inequalities. They focus mainly on the vulnerability of territories, and reflect the cumulative harmful constraints and nuisances to which populations are subjected. This solely permits the measurement of damage (*e.g.* greenhouse gas emissions) and of the elements characterising the societal factors that shape society's vulnerability to these hazards (*e.g.* the share of the population with a higher education qualification). These publications primarily seek to assess the accumulation of health determinants in a territory. A holistic description of a territory therefore requires an approach that is as global as possible and goes beyond the negative aspects to include the positive factors that are still poorly represented in composite indices.

Many studies deal with environmental health issues, focusing generally on vulnerability, *i.e.* on determinants that have a negative impact on territorial inequalities (Angeon and Bates, 2015). To a lesser extent due to a recent awareness, the positive impact of health determinants are also taken into account (Fong *et al.*, 2018; Hartig *et al.*, 2014). These studies are generally based on proximity to green spaces in municipalities and have revealed, for example, a consequent decrease in cardiovascular mortality and the occurrence of cardiovascular accidents (Rojas-Rueda *et al.*, 2019; van den Bosch and Ode Sang, 2017; Yeager *et al.*, 2019), or an improvement in mental health (Houlden *et al.*, 2019; Mukherjee *et al.*, 2017). To our knowledge, articles and research works combining these two groups of determinants with the same importance are scarce (Estoque and Murayama, 2014; Miller *et al.*, 2017). These are embedded in their territory and do not propose a conveniently transposable methodology. The concept of resilience, defined as the capacity of a territory to cope with health inequalities (Lanier *et al.*, 2020), is closely intertwined with vulnerability, but resilience is not the antonym of vulnerability (Fekete, 2018). The challenges are therefore to define a set of data to characterize resilience and to submit a methodological approach that would be easily transferable.

The information included in existing spatialized indices is disparate: it can describe environmental, health, social, economic, political or territorial management and organization dimensions. Only 9 articles present an index that is simultaneously based on these 6 dimensions (Angeon and Bates, 2015; Braconnier *et al.*, 2011; Cutter *et al.*, 2010; Estoque and Murayama, 2014; Etsy *et al.*, 2005; Kerényi, 2011; Sadd *et al.*, 2011; Zuideau and Lipovac, 2009). The need to develop a process of interoperability between data in different formats and the difficulty of accessing data in certain contexts means that it is practically impossible to obtain a holistic description of the territory. Authors often choose to focus their study on a national scale because standardized data are readily available and international

comparisons are facilitated (Biggeri and Mauro, 2018; Estoque and Murayama, 2014; García-Sánchez *et al.*, 2015). Overall trends are observable and can be tracked over time. However, this approach does not allow users to take account of territorial specificities or to integrate spatial variations for certain types of information. A finer spatial scale (*i.e.* city, neighbourhood) facilitates a more precise description of phenomena, with the possible use of data that may not exist in other territories (*e.g.* local initiatives). The local dimension of territorial inequalities particularly meets the expectations of political decision-makers within the framework of their concrete urban planning and public health actions (Hat and Stöglehner, 2019).

The aim of this research is to optimize the assessment of the spatial heterogeneity of territorial inequalities. We propose a new strategy based on a joint analysis of two optimized composite sub-indices. This balance between resilience and vulnerability allows the simultaneous evaluation of the negative aspects of the territory in the Vulnerability Index (VI) and the positive aspects of the territory in the Resilience Index (RI). On this basis, a cross-classification is validated with the objective to distribute spatial units within territorial typologies. These homogeneous groups can be analyzed in terms of health, environmental, political, services, social and economic dimensions. A case study was conducted to demonstrate the feasibility of the methodology, using the municipality as a spatial unit of analysis within a region in the north of France.

2. Material and Methods

2.1. Study setting

This study was carried out in the Hauts-de-France region, located in the north of France. This densely populated area of 31,813 km² with 6,009,976 inhabitants, *i.e.* an average of 189 inhabitants/km², is located at the crossroads of Europe and includes 3,817 municipalities that contrast in terms of their respective sizes, population diversity, socioeconomic backgrounds and levels of environmental pollution (Cuchère and Terra, 2017). These municipalities have a median [interquartile range (IQR)] population of 423 [211 – 1,014] inhabitants and a median surface area of 6.91 [4.73 - 10.30] km². The multiple sources of environmental pollution, including road traffic, residential heating, industries, contaminated lands and conventional agriculture, are heterogeneously distributed across the territory.

2.2. Methodology

The proposed methodological framework is as follows (Fig. 1):

- (i) data collection and analysis for each of the six dimensions (health, environment, social, economics, services, policy) that characterize the multifaceted phenomena involved in environmental health, all of which are recognized as playing a role in health inequalities,
- (ii) the calculation of the resilience and vulnerability indices,
- (iii) the evaluation of the spatial distribution of resilience and vulnerability indices,
- (iv) the discretization of the resilience and vulnerability indices,
- (v) the joint analysis of these two indices and the categorization of municipalities ($n = 3,817$) into different typologies.

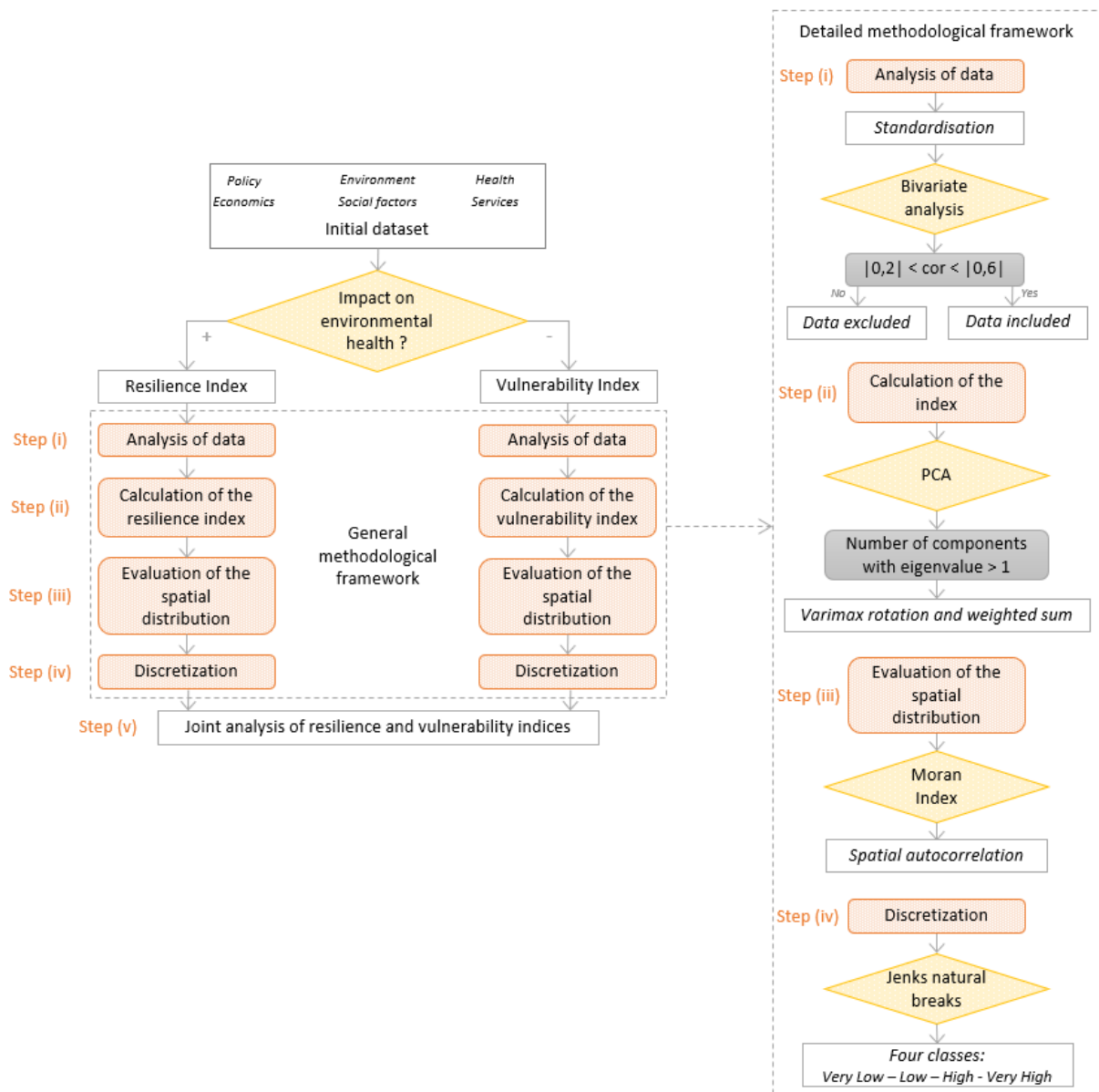


Figure 5 : Article 3 – Figure 1: A comprehensive methodology for the development of composite indices to assess environmental health inequalities. The general methodological framework presents the main steps to be followed for the development of composite indices, while the detailed methodological framework explains the different methods proposed for each step (step (i): analysis of data, step (ii): calculation of the index, step (iii): evaluation of the spatial distribution, step (iv): discretization and step (v): joint analysis of resilience and vulnerability indices)

2.3. Database development – inclusion/exclusion criteria

Data to be included were selected on the basis of three inclusion criteria:

- (i) the data was plausibly associated with environmental health in the literature and led to a characterization of a descriptor within a dimension,
- (ii) comprehensive, spatially contiguous and contemporary data were available for the entire study area,
- (iii) data monitoring was reliable and temporal monitoring was performed, thereby ensuring acceptable data quality.

The following exclusion criterion was applied: the data had been associated with individual behavior (*i.e.* smoking, food habits or physical activities).

As part of our study on the Hauts-de-France region, 50 variables were collected from public national databases (accessible in *Open data*) in order to describe multiple facets of each dimension (Table I). Of these 50 variables, 36 % (18 variables) were obtained directly from official and national databases and were integrated into the dataset for use with no modification. The remaining 64 % (32 variables) were (i) derived from a straightforward calculation (*e.g.* a proportion calculated in relation to the number of inhabitants), or (ii) derived from more complex calculations inspired by previous works such as the valuation of ecosystem services or the Shannon-Weaver index (Estoque and Murayama, 2014; Salvati and Carlucci, 2014). Further details concerning the dataset can be found in Brousmiche *et al.* (2020a) and in Brousmiche *et al.* (2021).

Tableau X : Article 3 – Table I: Presentation of sub-dimensions that aim to describe multiple facets for each of the six dimensions

Dimension	Sub-dimension
Health	General population health
	Health care facilities
Environment	Anthropogenic pressures
	Natural resources
	Biodiversity
	Environmental risks
Social	Integration and anomy
	Population sensitivity
	Level of education
	Social resources
	Urbanization
Economics	Employability
	Diversity and economic drive
	Incomes
Services	Access to the homeownership market
	Access to basic needs
	Access to information and communication
	Mobility
Policy	Environmental governance
	Budget
	Anticipation

The selected variables will be assigned to the resilience or vulnerability index according to the impact of an increase in each variable on environmental health. If an increase in the variable can be related to a broadly beneficial effect on environmental health, then the variable is considered to belong to the resilience index, while an increase that corresponds to a deleterious or negative effect of the variable leads to its classification in the vulnerability index (Fig. 2).

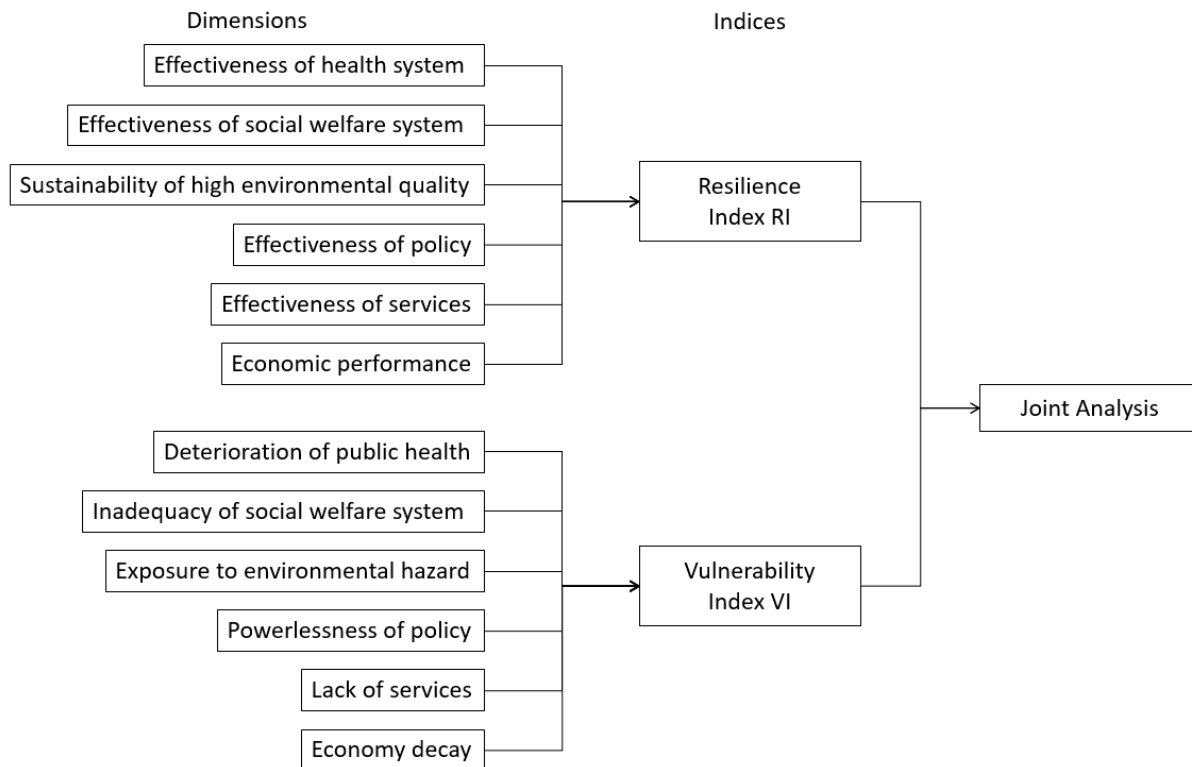


Figure 6 : Article 3 – Figure 2: Diagram illustrating the orientation of variables towards resilience or vulnerability index, according to the effect of an increase in the variable, for use in joint analysis

2.4. Construction of the resilience and vulnerability composite indices

2.4.1. Data pre-processing

Management of missing data

This step is necessary to provide a complete data set and avoid bias due to missing information (European Commission *et al.*, 2008). Missing data management was carried out using two imputation methods for two cases of missing data:

- Case 1: If (i) there was a history of a merger/creation of a new municipality and (ii) there was a value for the newly created municipality, this value was attributed to the those municipalities in the group for which data was lacking;
- Case 2: If there was no previous history of a merger / creation of municipality, the values of the neighboring municipalities (*i.e.* those sharing a border) were identified and the median of these values was attributed to the municipality for which the data was missing.

Standardization

Since data can be expressed in very different units and/or generated by different producers for different purposes and according to different reference systems, standardization is necessary to ensure that the indicators are comparable (Böhringer and Jochem, 2007; Saib *et al.*, 2015). Each of the variables was therefore standardized according to the z-score methodology, which is recommended before using Principal Component Analysis (PCA) (Marzi *et al.*, 2018). For a spatial unit i , the z-score z_i is defined by $z_i = (x_i - \bar{x})/sd$, where x_i corresponds to the raw value of the variable for spatial unit i , \bar{x} is the mean of the variable in all spatial units and sd corresponds to standard deviation. Thus, all variables are converted into a common scale with a mean equal to zero and a standard deviation equal to 1. This method is (i) less sensitive to extreme values than other methodologies and (ii) maintains

correlation structures between variables (European Commission *et al.*, 2008; Kolenikov and Angeles, 2004).

Bivariate correlation analysis

A selection of variables is necessary to obtain the most concise description possible of the initial dataset (European Commission *et al.*, 2008). This selection is based on a bivariate correlation analysis which identifies redundant variables and thus summarizes the major information contained in the initial dataset (King and Jackson, 1999). As mentioned by European Commission *et al.* (2008), the choice of variables is first guided by a theoretical criterion: at least one correlation coefficient in the range $[|0.2|; |0.6|]$. This made it possible to introduce into the PCA only variables having a minimum correlation structure with the other variables, thus avoiding that certain principal components are carried by a single variable (noise) or are strongly influenced by a few redundant variables. The balance and consistency in the diversity of dimensions and sub-dimensions considered in the dataset is then validated by experts representing a wide spectrum of knowledge and experience to ensure the implementation of an adequate characterization system. Particular care should be taken in selecting the variables that will minimize the mirror effect between resilience and vulnerability.

2.4.2. Weighting and aggregation

Resilience and vulnerability indices are constructed through a PCA-based weighting and aggregation process. Only the *p*-first principal components (PCs) with an eigenvalue above 1 were retained for the index construction. A rotation of the factorial axes is applied to components by means of a VARIMAX rotation to increase the readability and facilitate the interpretation of the outcomes (Kaiser, 1958). The index (either resilience index or vulnerability index) is thus expressed as follows:

$$Index = \sum_{j=1}^p w_j PC_j,$$

where *p* corresponds to the number of retained principal components, and w_j indicates the proportion of variance explained by the principal component PC_j .

Finally, to simplify the interpretation, the final index is standardized by the min-max method to obtain values in a common value range that is easily comprehensible (*i.e.* [0;1]). The min-max standardization is carried out as follows:

$$z_i = \frac{x_i - \min(x)}{\max(x) - \min(x)},$$

where z_i is the standardized value of the resilience or vulnerability index, x_i is the initial value of the resilience or vulnerability index, and $\min(x)$ and $\max(x)$ correspond respectively to the minimal and maximal values of the resilience or vulnerability indices.

2.5. Spatial assessment of the resilience and vulnerability indices

2.5.1. Mapping of the spatial distribution

To observe the spatial distribution of the composite indices within the study area, the indices are discretized using Jenks natural breaks classification method (Jenks, 1967). This method allows the creation of homogeneous classes with minimized intra-class variance and maximized inter-class variance. Three thresholds were calculated for each index, yielding four classes (Very Low, Low, High, Very High), making it possible to obtain a cartographic representation that can be rapidly interpreted without over-simplifying the information provided by the composite index. The maps were made using ArcGIS software (Environmental Systems Research Institute, 2011).

2.5.2. Evaluation of spatial auto-correlation

The presence or absence of spatial autocorrelations among the municipalities is quantified using Moran's index (Moran, 1948), in which a value greater than 0 indicates the presence of autocorrelation, and probed using Moran's test (Cliff and Ord, 1981). The statistical analyses are carried out using R software (version 3.4.3; package: DCluster). The threshold for statistical significance is set to $p < 0.05$.

2.5.3. Result appraisal of resilience and vulnerability indices

A comparison with other indicators is carried out to appraise the contribution of our approach to the strategies that are commonly used to assess the spatial heterogeneity of inequalities. According to their distribution, the composite indices will be compared through the calculation of the Spearman correlation coefficient. We appraise our results using the Human Development Index (HDI4) composite indices described by Marlier (2009), French Ecological Deprivation Index (EDI) as described by Pornet *et al.* (2012) and French social Deprivation index (FDep) as described by Rey *et al.* (2009). HDI4 reflects a concentration of populations with a certain degree of monetary, educational and health resources (Marlier *et al.*, 2015). French EDI aims to provide a methodology to develop an "intercultural" ecological indicator of socio-economic level, while FDep was created to provide a geographic index of social disadvantage specifically adapted to health studies on the French population (Ghosn, 2018a).

2.6. Joint analysis of resilience and vulnerability indices

The joint analysis aims to classify each municipality of the studied region among the four defined typologies. It is based on the cross-interpretation of resilience and vulnerability indices. The four classes of resilience and vulnerability are therefore grouped into two categories (Very High-High and Very Low-Low), leading to the four territorial profiles. Further details concerning the cross-interpretation of resilience and vulnerability indices can be found in supplementary material and results.

As regards vulnerable territories, two profiles can be distinguished according to their level of resilience: territorial blackspots, which do not have sufficient environmental, socio-economic or health resources, and territories that have resources they can implement to counteract sources of vulnerability (Maclachlan *et al.*, 2007). Given their lower levels of resources, territories with a lower RI should be nevertheless be monitored.

3. Results

Fig. 3 presents the results of the variable selection stages, from the initial dataset to the number of variables included in the resilience and vulnerability indices. Nine variables were excluded due to correlation values above $|0.6|$, seven others were discarded due to correlation values below $|0.2|$. VI was calculated from 13 variables, while 21 variables were included in RI. Supplementary material and results present the results of the correlation matrix, the correlogram of the 50 variables which illustrates the general structure of the correlations between each variable and PCAs prior to the Varimax rotation.

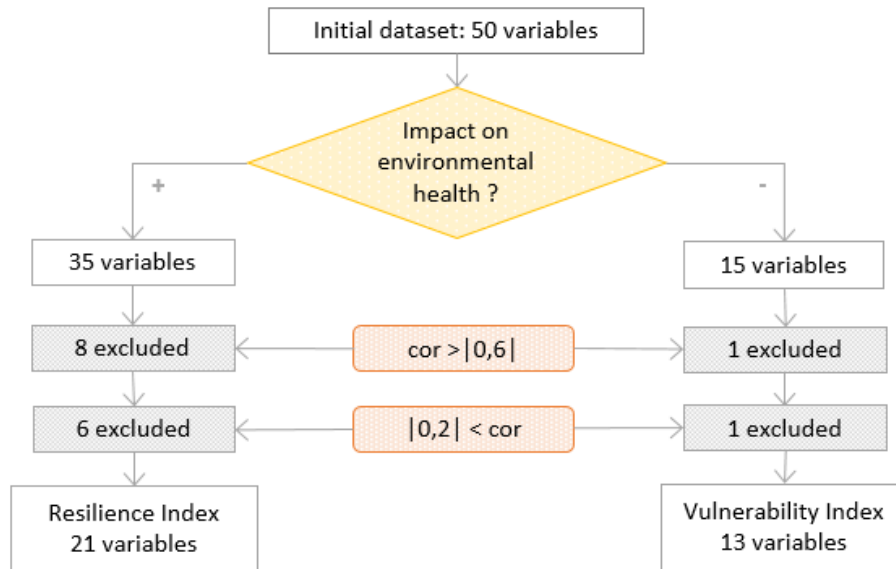


Figure 7 : Article 3 – Figure 3: Flowchart summarizing the variable selection steps, from the initial dataset (50 variables) to the Resilience Index (21 variables) and the Vulnerability Index (13 variables)

3.1. Vulnerability Indice (VI)

The VI values vary from 0 to 1, with 1 representing the highest possible vulnerability value in the area studied. The VI has a median value of 0.151 with an IQR of [0.126-0.180]. Fig. 4 (a) presents the spatial distribution of the VI calculated for the Hauts-de-France region and the threshold values used to create the vulnerability classes. A total of 1,623 (42.5%) municipalities belong to the Very Low vulnerability category, 1,710 (44.8%) to the Low vulnerability category, 479 (12.6%) to the High vulnerability category and 5 (0.1%) to the Very High vulnerability category. Moran Index results show a weak but significant spatial autocorrelation ($I = 0.063$, $p < 0.001$). A particular group of municipalities with a high level of vulnerability is located between Bethune, Lens, Douai and Valenciennes: this spatial trend is located on the former coal-mining basin, area for which multiple vulnerabilities have already been identified. Municipalities with higher vulnerability index values tend to gather around large urban areas (Dunkerque, Lille, Arras, Compiègne, Amiens), but to a small extent. Municipalities with lower vulnerability index values are more randomly located within the Hauts-de-France region.

3.2. Resilience Index (RI)

The RI values vary from 0 to 1, with 1 the highest possible resilience value in the area studied. The RI has a median value of 0.341 with an interquartile range of [0.273;0.401]. Fig. 4 (b) presents the spatial distribution of the RI calculated for the Hauts-de-France region and the threshold values used to create the resilience classes. A total of 519 municipalities (13.6 %) belong to the Very Low resilience category, 1,323 (34.7 %) to the Low resilience category, 1,401 (36.7 %) to the High resilience category and 574 (15.0 %) to the Very High resilience category. Moran index results show a significant spatial autocorrelation ($I = 0.597$, $p < 0.001$), indicating that similar resilience values tend to aggregate spatially. A geographical variation can be observed in the distribution of the municipalities within these four categories, with a specific group of municipalities belonging to the Very High resilience category around large urban areas. The large urban areas that include a number of municipalities with higher resilience index values are mainly located around Amiens, Arras, Compiègne, Douai, Dunkerque, Lille and Senlis. Municipalities with lower resilience index values tend to be located in less densely

populated areas, more particularly (i) along a north-west/south-east axis from Montreuil to Vervins and (ii) in the south-east of the region.

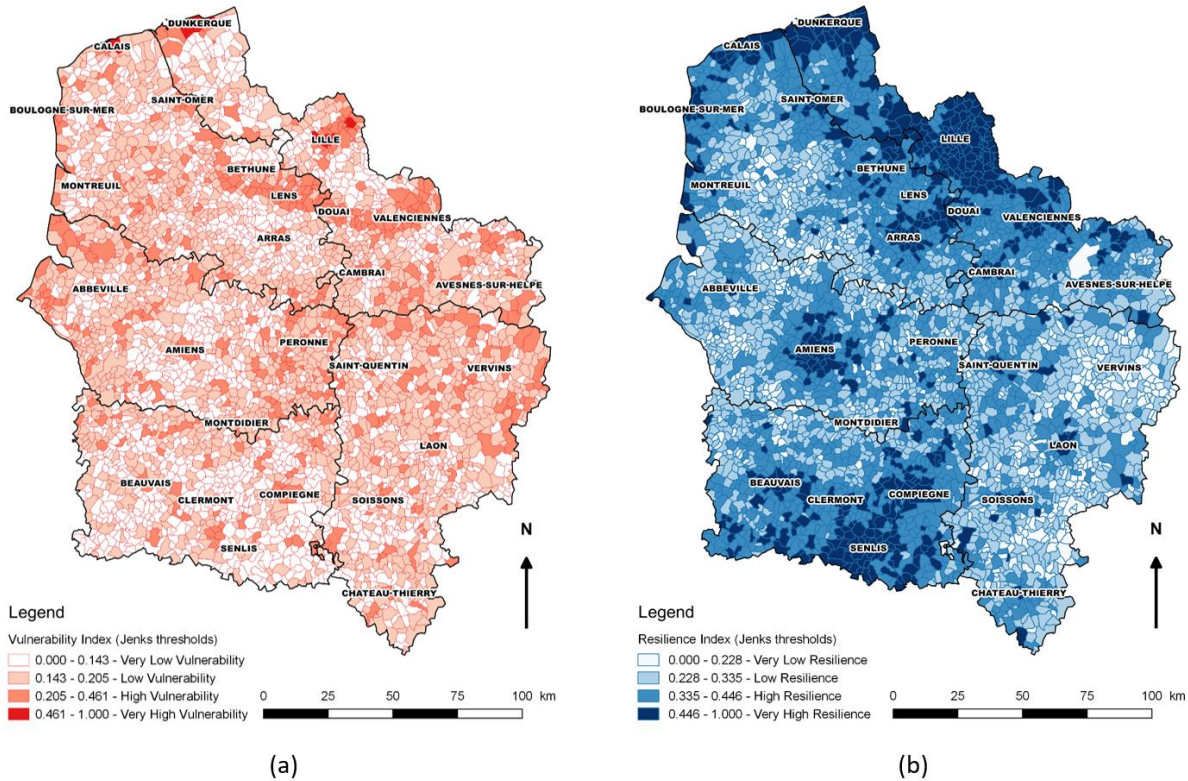


Figure 8 : Article 3 – Figure 4: (a) The Vulnerability Index for the municipalities in the Hauts-de-France region. Three Jenks’s natural breaks were calculated (0.143, 0.205, 0.461) to obtain four classes of vulnerability (very low, low, high, very high). (b) The Resilience Index for the municipalities in the Hauts-de-France region. Three Jenks’s natural breaks were calculated (0.228, 0.335, 0.446) to obtain four classes of resilience (very low, low, high, very high)

3.3. Comparison with others indicators

Table II presents the Spearman correlation coefficients which were calculated in light of the indices distributions. VI shows a relatively strong correlation with EDI and FDep (coefficient value of 0.70 and 0.65, respectively; for both, $p < 0.001$). These three indices roughly outline the same negative aspects, including unemployment rate, house tenancy rate, low level of education. However, the correlation is higher between VI and EDI than between VI and FDep, because a greater dichotomy occurs between urban and rural contexts in EDI than in FDep. RI has a moderate correlation with the other indices, suggesting that RI integrates distinct aspects. The architectural composition of RI actually seeks to be more holistic to take into consideration the multiple facets of phenomena involved in environmental health. The RI successfully integrates positive aspects, notably in the environmental and policy dimensions, which are not covered in the deprivation indices (FDep, EDI) and in the HDI4. Moreover, RI provides different information to that provided by VI (coefficient value of -0.01; $p = 0.57$), highlighting the weak mirror effect between resilience and vulnerability.

Tableau XI : Article 3 – Table II: Correlation matrix between the Vulnerability Index, the Resilience Index and other indicators (HDI4, French EDI, FDep). Results of the significant Spearman’s correlation tests (i.e. $p < 0.001$) are marked with an asterisk “*”

	VI	EDI	FDep	RI	HDI4
VI	1	0.70*	0.65*	-0.01	-0.58*
EDI		1	0.74*	-0.10*	-0.69*
FDep			1	-0.28*	-0.76*
RI				1	0.28*
HDI4					1

3.4. Joint analysis and interpretation

Each municipality of the studied region had been classified among the four defined typologies through the joint analysis of resilience and vulnerability. Thus, 1,687 (44. 2%) municipalities belong to the “To monitor” category, 1,646 (43.1 %) to the “Resilient” category, 329 (8.6 %) to the “Have resources” category and 155 (4.1 %) to the “Territorial hotspot” category (Fig. 5).

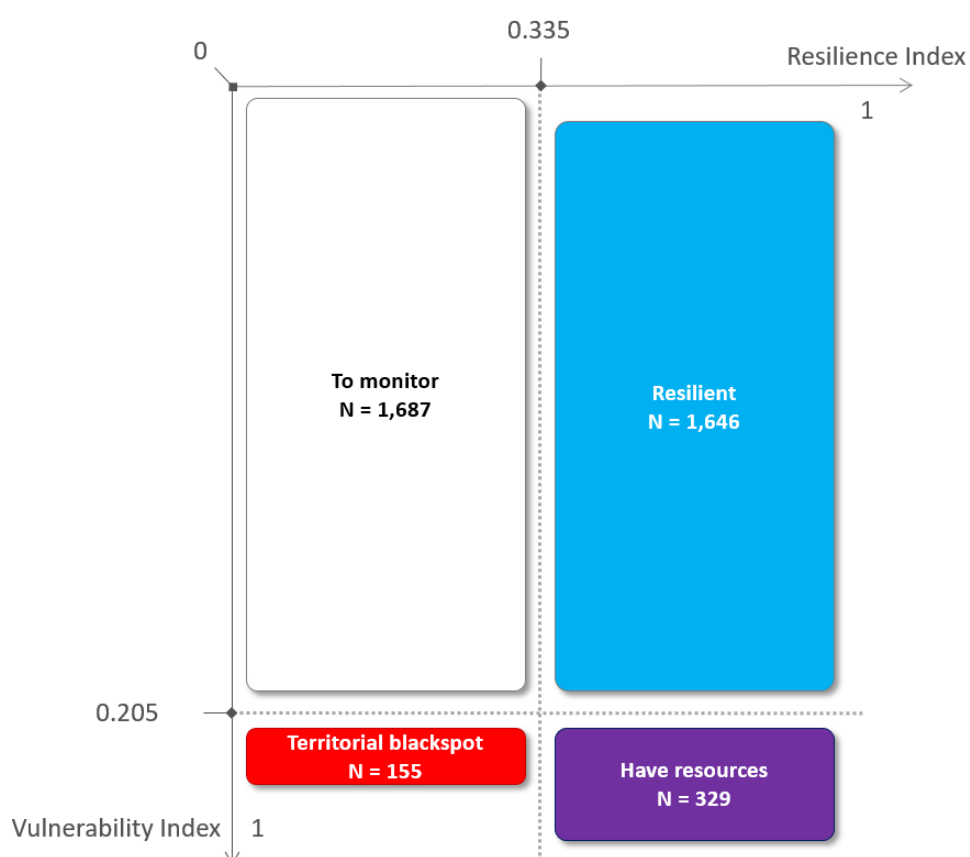


Figure 9 : Article 3 – Figure 5: Diagram showing the distribution of the number of municipalities among each typology (To monitor, Resilient, Have Resources, Territorial hotspot) after the cross-classification of resilience (RI) and vulnerability indices (VI) in the Hauts-de-France region

Fig. 6 presents the spatial distribution of the joint analysis of resilience and vulnerability indices in the Hauts-de-France region. Most of municipalities belonging to the “To monitor” category are located along a north-west/south-east axis, reflecting the results observed for resilience and vulnerability where lower values for both indices had been highlighted. Although other municipalities also had high vulnerability values, the integration of resilience revealed that they have resources. Municipalities

belonging to this “Have resources” category are mainly located in the north or south of the region, particularly in and around large urban areas (Compiègne, Douai, Dunkerque, Lens, Lille, Valenciennes) where resilience values were higher. Municipalities that belong to the “Resilient” category are also primarily located in the north and south of the region. The municipalities classified in the “Territorial blackspot” category tend to be more randomly distributed, confirming that management decisions must be taken at the municipality level.

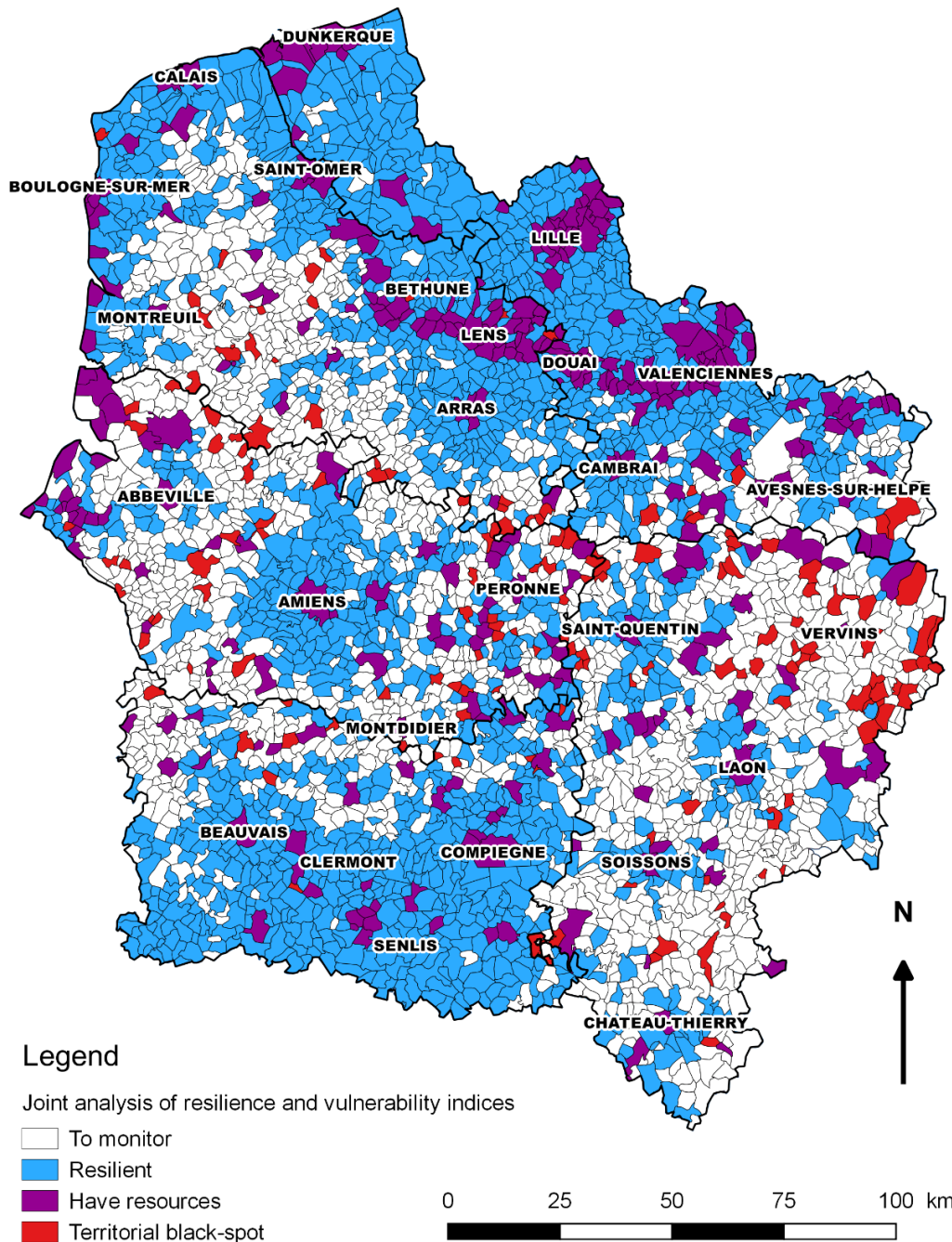


Figure 10 : Article 3 – Figure 6: Spatial distribution of the joint analysis of resilience and vulnerability indices for the 3,817 municipalities of the Hauts-de-France region. The “Territorial blackspots” are characterized by a high (or very high) vulnerability, paired with a low (or very low) resilience. The municipalities classified in the “Resilient” category are characterized by a low (or very low) vulnerability and a high (or very high) resilience. The municipalities classified in the “Have resources” category are characterized by a high (or very high) vulnerability and a high (or very high) resilience. The municipalities classified in “To monitor” are characterized by a low (or very low) vulnerability and a low (or very low) resilience

4. Discussion

This study establishes a new approach to assess the spatial heterogeneity of territorial inequalities among municipalities by integrating multiple health determinants. This novel approach based on an inventory of local data removes the difficulties of characterizing the accumulation of health determinants and provides an opportunity to build a diagnostic tool for stakeholders. To our knowledge, this is the first study to base the assessment the resilience / vulnerability balance on two spatialized sub-indices, combining the same dimensions for each sub-index. The transposability of the proposed framework is above all conceptual and methodological because the availability and / or accessibility of data may differ from one territory to another (Brousmiche *et al.*, 2020b). In a strategy of transposability, the points of attention relate in particular to the selection of specific data by including the data the most adapted to the spatial scale of study, demonstrating the strongest discriminating capacity between the spatial units, and the most in adequacy with territorial particularities (*e.g.* technological and / or economic development) in order to collect the greatest diversity and richness of information.

Although the level of vulnerability can be described at the local level, current methodology generally focuses on a small number of descriptors. Thus, previous French studies concerning our study area have already identified the former coal-mining basin, the Urban Community of Dunkerque and the Urban Community of Lille as areas of high and multiple vulnerability (Becuwe and Lecomte, 2015; Belhakem and Lermechin, 2018; Fleury *et al.*, 2010; Salem, 1995). These studies were based on the poverty rate, on the unemployment rate or on life expectancy. The different vulnerability assessment strategies used in these studies limit the comparison of the latter and hinder the ranking of municipalities. The present study has developed a multidimensional vulnerability index that successfully highlights these vulnerable territories. It allows comparison between municipalities and the monitoring of numerous parameters of interest through a construction from the same dataset, whatever the municipality.

Little attention has been paid to the impact of positive aspects of a territory on human health. Few studies have simultaneously considered resilience and vulnerability (Kumar *et al.*, 2019), and resilience has rarely been calculated in as much detail as vulnerability (Estoque and Murayama, 2014; Miller *et al.*, 2017). In the proposed framework, this issue is addressed by deriving and considering RI and VI in the same way. Initial attempts are being made to interpret the spatial distribution of the level of resilience, particularly by focusing on the link between resilience and demographics. For example, although the links between resilience and population density are equivocal in the literature, some authors confirm this relationship through different approaches of this demographic variable, *i.e.* either as a factor to be included in the study or as an explanatory factor for the level of resilience. Rural/urban status, population density and urbanisation ratio have been mentioned by Estoque and Murayama (2017), García-Sánchez *et al.* (2015), Monteiro *et al.* (2012), Pearson *et al.* (2013) and Zhang *et al.* (2015). In this study, we used a population density as a direct indicator, as described in Estoque and Murayama (2017) and Monteiro *et al.* (2012), and other indirect indicators related to the population of municipalities, such as the area of natural space per capita. When possible, considering data per inhabitant highlights the efforts made by a municipality for each of its inhabitants: although raw data for densely populated municipalities could have higher values than those of sparsely populated municipalities, data per inhabitant could actually be greater for municipalities with smaller populations (*e.g.* total expenditures for municipal equipment vs total expenditures for municipal equipment per inhabitant). However, our results reinforce all the evidence suggesting that resilience is higher in urban areas than in rural areas. The approach developed in this study meets the need for analytical tools by developing a composite resilience index that is adapted to European data and easy to appropriate.

To offset the classic approach based on the weaknesses of municipalities, our strategy of joint analysis aims to include the efforts that a municipality has made to improve the living environment of its inhabitants. The joint analysis provides an original and powerful approach that EDI, FDep and HDI4 failed to achieve, *i.e.* one that permits the assessment of the spatial heterogeneity of territorial inequalities. To a certain extent, VI identifies the same territories as indices based only on deprivation. VI thus highlights the accumulation of socio-economic, environmental and health determinants and the addition of RI sharpens the description of the territory in terms of resources; together, these indices summarize a wide range of data. The development of spatialized composite indices is also complicated by a lack of consensus on a methodology (Brousmiche *et al.*, 2020b). Certain weighting and aggregation methods that are commonly used to develop composite indices (arithmetic or geometric mean, weighted or not) do not successfully resolve (i) the issue of compensability between variables (*i.e.* is a given variable more influential than others?) and (ii) the objectification of a ranking which may be affected by the aggregation method. RI and VI obtained by PCA are based on the statistical structure of the data and are not automatically affected by the influence of individual data on the measured phenomenon. This methodology is notably suitable to develop composite indices insofar that it provides the best contrasts between spatial units, an advisable property for inter-territorial comparisons (European Commission *et al.*, 2008; Lalloué *et al.*, 2013). A flexible and adaptable framework that allows the integration of new or different data is required to take into account the specificities of the study area, namely (i) the contextualization of data related to technological development (*e.g.* number of landlines per 1,000 inhabitants vs internet access), (ii) the availability of representative data, which may be ample for some countries but less so or even non-existent for others (*e.g.* percentage of specific places of worship) and (iii) the absence of data for the geographical study area (*e.g.* data related to ethnicity in France) (Brousmiche *et al.*, 2020b). The power of the proposed framework is based on the respect of the dimensions to be explored, and allows users to freely integrate relevant and available data for the study area. The resulting composite indices demonstrate good discriminating power, making it possible to highlight disparities and to prioritize spatial units according to the vulnerability/resilience balance.

However, this approach does have some limitations that may actually be opportunities for future research work on composite indices. Some health determinants are only characterized at a global (*i.e.* national) scale and are not defined on a local scale, thus resulting in a lack of accessible variables at the appropriate scale. This lack of accessibility of fine-scale data meant that we did not include data that might have been relevant in RI or VI, such as the poverty rate, which measures the living standards in relation to the living standard of the population as a whole (relative approach) (Miller *et al.*, 2017; Stewart *et al.*, 2014). There is a particular need to include more data to characterize environmental quality. This issue has received less attention (Flacke, 2015; Nieuwenhuijsen, 2016; Sharifi, 2016) despite the fact that 22.7 % of the global disease burden was attributed to environmental factors in 2012 (Prüss-Üstün *et al.*, 2016). Despite the unavailability of this kind of data at a small scale for the region as a whole or the difficulty to obtain them reported by numerous authors (Huang and London, 2012; Sadd *et al.*, 2011; Zhang *et al.*, 2015), future composite indices should include more environmental data, primarily concerning air quality (indoor and outdoor), drinking water quality, pesticide measures or heavy metal contents in soil. Finally, an essential aspect of the evaluation of territorial inequalities is that the method implemented should take into account the relationships between health determinants to avoid over- or underestimating their respective importance. This requires the use of reliable statistical methods such as PCA, which take into account (i) the subjectivity in the classification of a given variable within a dimension and (ii) the interrelations of health determinants over time and space (Sarkar and Webster, 2017), and above all developing the

knowledge of the entire territorialized exposome. The concern for completeness supported by this notion of territorialized exposome is essential in all studies evaluating territorial inequalities.

Researchers and stakeholders lack screening methods, and the need for characterization tools is mentioned regularly (Cutter *et al.*, 2010; Rothenberg *et al.*, 2015; Saib *et al.*, 2015). The provision of a suitable process would facilitate the identification and prioritizing of locations, for example focusing on those with both higher vulnerability and lower resilience values compared to the region as a whole. The proposed cross-classification of RI and VI meets this challenge. In total, we collected a set of 50 robust (*i.e.* consistent, non-redundant, exhaustive) variables that are relevant to environmental health issues in the region and outline the environmental, health, social, economic, policy or service dimensions (Brousmiche *et al.*, 2021). We thus linked 3,817 municipalities to one of 4 typologies, taking into account the weight of the negative aspects of the environment (VI) and the weight of the positive aspects of the environment (RI) related to environmental health issues. This approach offers a more comprehensive view of how stakeholders are addressing social, health and environmental issues and could help the former to improve their public health policy. The interpretation of the results leads to different action according to the typology. The “Territorial blackspots” must benefit from a rapidly implemented action plan. The territories belonging to the categories “To monitor” and “Have resources” should be monitored to prevent their degradation into “Territorial blackspots”. Health policies should be encouraged in these territories so that they evolve towards the “Resilient” category. The territories belonging to the “Resilient” category could be considered as demonstrators of good practices that could be replicated. A further challenge of this approach is to encourage the exchange of good environmental health practices between territories (Nieuwenhuijsen *et al.*, 2017). This diagnostic tool can also become a prospective tool for the optimization of future preventive health policies or urban policies (Ramirez-Rubio *et al.*, 2019; Saib *et al.*, 2015) and a tool for monitoring the impact of the implemented actions.

Conclusion

This study aimed to establish a new transferable approach to assess the spatial heterogeneity of territorial inequalities and to provide a diagnostic tool for stakeholders. In the field of environmental health, no consensus has been made to date concerning the best methodology for developing composite indices, resulting in a subjectivity between the frameworks and limiting comparisons between different composite indices and thus between territories. To the best of our knowledge, this article presents the first joint analysis, at fine scale for a large study perimeter, of positive and negative impacts leading to four territorial profiles. The flexible and adaptable framework proposed for database development has overcome a main issue in developing composite indices, *i.e.* data availability. This development principle also offers the opportunity to improve the characterization of territories by integrating the most suitable information for the context of the geographical perimeter. To provide the broadest possible depiction of the territory, both resilience (the capacity of a territory to cope with health inequalities) and vulnerability (negative constraints and nuisances to which a population is subjected) were considered to be equally essential and important factors in the assessment of the resilience/vulnerability balance. Further studies should be conducted to confirm the benefit to draw on a joint analysis to assess the environmental and social inequalities of health.

To characterize health, environmental, social, service, policy and economic dimensions for the development of the two composite indices, a total of 50 variables were collected from public national databases. A methodology based on PCA was designed to develop the resilience and the vulnerability indices. PCA makes it possible to summarize a large amount of information while maximizing the proportion of total variance of the initial dataset and avoiding redundant information. The composite indices were discretized using the Jenks natural breaks classification method in order to observe their

spatial distribution. Thereafter, a cross-classification of the different groups of resilience and vulnerability made it possible to distribute each municipality among the four territory typologies: 1,687 municipalities belong to the “To monitor” category, 1,646 to the “Resilient” category, 329 to the “Have resources” category and 155 to the “Territorial blackspot”.

The joint analysis of resilience and vulnerability indices provides a powerful methodology to assess the spatial heterogeneity of territorial inequalities. It also facilitates the identification and the prioritization of municipalities that could benefit from the implementation of specifically tailored public health policies, *i.e.* an action plan for the territorial blackspots. This diagnostic tool contributes to the establishment of monitoring for the municipalities with lower levels of resilience. The comparison of our results with other major indicators shows a relatively strong correlation between vulnerability indices (FDep, VI, French EDI), reflecting the results of previous studies describing aspects such as unemployment rate or low educational levels. The results also showed a moderate correlation between HDI4 and RI and highlighted a weak correlation between RI and VI, confirming that (i) RI provides a more holistic characterization of the territory and (ii) RI and VI measure different phenomenon. Although the availability of spatialized information is increasing, certain data for health determinants (*e.g.* environmental data concerning factors such as drinking water quality, pesticide measures or the presence of heavy metals in soil) are still lacking or only accessible at a mid- or large-scale despite their recognized impact on health. The availability of this type of fine-scale indicator would substantially enrich current indices of resilience and vulnerability and meet the needs of decision-makers seeking to improve the contextualisation of local policies in all health-related domains.

Les travaux menés dans le cadre de l'analyse de la résilience et de la vulnérabilité ont permis d'éprouver la méthodologie de développement d'indices composites. La démarche par l'analyse de la balance résilience vulnérabilité basée sur le principe de multidimensionnalité des variables à intégrer aux indices composites permet d'affiner la caractérisation des inégalités territoriales en santé environnement. L'approche par la vulnérabilité en santé environnement est fréquemment mentionnée dans la littérature. Les déterminants de la vulnérabilité et leurs mécanismes dans la construction des inégalités territoriales en santé environnement sont mieux appréhendés que pour les déterminants en lien avec la résilience (Lanier *et al.*, 2020). L'indice de vulnérabilité qui a été calculé sur les Hauts-de-France a permis de mettre en évidence des zones de multiples vulnérabilités (sociales, économiques, environnementales) et a apporté une information plus complète que ce qui peut être plus classiquement mesuré par des indicateurs de défaveur socio-économique (*i.e.* FDep, french EDI). Sur le même précepte que la vulnérabilité, les déterminants en lien avec la résilience ont également été considérés comme multidimensionnels et ne relevant pas seulement d'un dynamisme économique, de la présence d'espaces verts ou de réservoirs de biodiversité. Cette approche par la résilience fait l'objet de développements moindres, se focalisant principalement sur le lien entre la présence d'aménités environnementales et leur impact sur les inégalités territoriales en santé environnement. La méthodologie d'analyse simultanée de la balance résilience/vulnérabilité proposée ici a permis de considérer avec la même importance les phénomènes ayant un impact positif et négatif sur les inégalités territoriales en santé environnement et d'ainsi apporter une vision élargie des problématiques de santé environnement. La hiérarchisation et la priorisation des communes de la région Hauts-de-France les unes par rapport aux autres pour des actions curatives en santé environnement ont été possibles par l'analyse conjointe qui a été menée. La stratégie globale qui a été valorisée permet donc de considérer l'impact parallèle des aspects négatifs et positifs du territoire dans la caractérisation et la priorisation des unités spatiales pour fournir aux parties prenantes un outil de diagnostic adaptable à différents contextes territoriaux.

Au regard (i) des cumuls de difficultés (taux élevé de non-diplômés, marché du travail dégradé, suroccupation des logements, monoparentalité importante) auxquels l'EPCI Dunkerque fait face (Bailleul et Lefèvre, 2019) et (ii) des résultats de l'analyse de la balance résilience/vulnérabilité menée à l'échelle des communes de la région Hauts-de-France, la Communauté Urbaine de Dunkerque (CUD) souhaitait poursuivre l'analyse des inégalités territoriales de santé à un niveau infra-communal. En effet, au cours de l'analyse menée à l'échelle des communes de la région Hauts-de-France, la CUD présentait des disparités en termes de vulnérabilité (certaines communes appartenant à la classe de très haute vulnérabilité étaient contiguës de communes identifiées très faiblement vulnérables) et une absence de variabilité en termes de résilience (très haute résilience).

La CUD est une aire urbaine densément peuplée avec un tiers de son territoire artificialisé. Il y existe également de fortes disparités de revenus, de conditions environnementales et de mortalité toutes causes confondues (AGUR, 2019; D2DPE, 2010). L'objectif était d'obtenir une analyse des inégalités territoriales de santé, qu'elles résultent des impacts des déterminants en lien avec la vulnérabilité et/ou la résilience, à l'échelle des îlots Regroupés pour l'Information Statistique (IRIS). Cette échelle très fine peut en effet permettre de mieux contextualiser les politiques à mettre en œuvre et proposer des actions spécifiques à certains quartiers. Toutefois, le choix d'une échelle spatiale lors d'une analyse peut être à l'origine de problèmes spécifiques, soit en préambule du développement des indices composites (accessibilité et/ou disponibilité des données à très fine échelle), soit lors du développement de l'indice puis de son analyse. En effet, un résultat obtenu avec une échelle spatiale peut ne pas permettre de discriminer les unités spatiales entre elles. Des indices composites de résilience et de vulnérabilité ont été développés simultanément pour différentes unités administratives et une discussion est menée sur les conséquences du choix de l'unité spatiale et de

l'aire géographique dans l'évaluation des inégalités territoriales de santé. L'Article 4 présente les résultats des travaux issus de la collaboration avec la CUD.

Article 4

Spatial scale and spatial effects: shedding light on how a change of scale can affect the characterization of environmental health inequalities

Brousmiche Delphine, Lanier Caroline, Ocelli Florent, Pochet Louis, Degezelle Samuel, Genin Michaël, Deram Annabelle, Cuny Damien

Environmental Science & Policy (*under review*)



Le « *supplementary materials and results* » de cet article est disponible dans l'Annexe 2.

Spatial scale and spatial effects: shedding light on how a change of scale can affect the characterization of environmental health inequalities

Brousmiche Delphine^{1*}, Lanier Caroline^{1,2}, Occelli Florent^{1,2}, Pochet Louis², Degezelle Samuel³, Genin Michaël⁴, Deram Annabelle^{1,2}, Cuny Damien^{1,5}

¹ Univ. Lille, Univ. Artois, IMT Lille Douai, JUNIA, ULR 4515 – LGCgE, Laboratoire de Génie Civil et Géo-Environnement, F-59000 Lille, France

² Univ. Lille, Faculté d'Ingénierie et Management de la Santé (ILIS), F-59000 Lille, France

³ Communauté Urbaine de Dunkerque, Direction Environnement, Territoire et Transition Energétique, Service Urbanisme et Environnement, F-59140 Dunkerque, France

⁴ Univ. Lille, CHU Lille, ULR 2694 – METRICS : Evaluation des technologies de santé et des pratiques médicales, F-59000 Lille, France

⁵ Univ. Lille, Faculté de Pharmacie de Lille – LSVF, F-59000 Lille, France

*corresponding author: delphine.brousmiche.etu@univ-lille.fr

Abstract

Resolving health inequalities continues to be a priority in improving quality of life around the world. Their impact is of increasing concern to public health governance, which considers the issue of environmental health inequalities to be paramount. To meet this challenge, we have developed a methodology based on the development of spatial composite index in order to assess these inequalities while also providing an alternative to classic approaches that are solely based on territory weaknesses. The aggregation of the widest possible range of factors involved in environmental health provide a clear picture of a territory. This study uses different cases to understand how the change of scale influences the analysis of environmental health inequalities. Our analysis is based on a densely populated area located in a region in the north of France where the standardized mortality ratio is 21.3% higher than the national average. Composite indices were developed simultaneously from datasets collected for different administrative units, and were used to compare levels of resilience and vulnerability within the geographical area. Joint analysis of composite indices led to the identification of territorial blackspots. Indices could therefore enable stakeholders to prioritize certain administrative units, implementing specific measures according to the specificities of each administrative unit. The study discusses the consequences of the choice of a spatial unit and a geographical area in the assessment of environmental social inequalities in health, particularly in terms of accessibility, availability and exhaustivity. The selection criteria also depend on the jurisdictional fields of action stakeholders.

Keywords

Environmental health determinants, 'Health in All Policies', specificity of data, urban community, spatial variability, territorial diagnosis

1. Introduction

Over a decade after the publication of the influential WHO report on Social Determinants of Health, health inequalities remain high on the research agenda (Gkiouleka *et al.*, 2018). The impact of health inequalities issues is attracting increasing levels of attention in both public and political spheres (Eckersley, 2015). Among these inequalities, those related to environmental health are currently a major issue for public health governance. Environmental health inequalities are defined as the avoidable and unfair distribution of outcomes that can be observed between social groups within a same population (McCartney *et al.*, 2019; WHO, 2012). The assessment of these territorial inequalities generally focuses on vulnerability, which reflects the cumulative harmful constraints and nuisances to which populations are subjected. A disproportionate exposure to hazards is indeed recognized as a key factor in the development of environmental health inequalities (Alvarez and Evans, 2021; Brousmiche *et al.*, 2020b; Petroni *et al.*, 2021). Recently, we have proposed the concept of territorial resilience as a means to include the actions carried out by public authorities for the benefit of public health and to highlight/consider the aspects of each territory that positively contribute to health (Lanier *et al.*, 2020). In this context, resilience can be defined as the capacity of a given territory to cope with health inequalities. It is important to consider all factors related to environmental health, whether they have a negative or a positive impact, in order to obtain a reliable picture that reflects the complexity of phenomena involved in territorial inequalities (Eckersley, 2015). A joint analysis of resilience and vulnerability therefore provides a clearer description of the territory by providing an alternative to classic approaches that are based solely on territorial weaknesses.

One of the main methodologies for the assessment of environmental health inequalities relies on the development of the spatial composite index (Cartone and Postiglione, 2021; European Commission *et al.*, 2008). A flexible and adaptable framework at a local (municipality) scale has been proposed to deal with (i) the specificities of a study area and (ii) the availability or absence of data that is representative of the geographical area of study at the spatial unit considered (Brousmiche *et al.*, 2020a). The previous study was carried out over a large perimeter and enabled regional stakeholders to identify the municipalities that required specific environmental health action. Although our results offer a partial understanding of the determinants at the origin of environmental health inequalities, they cannot simply be transposed to a lower spatial scale, due to a lower local variability. The latter is related to (i) the spatial autocorrelation of the data, the first law of geography being “Everything is related to everything else, but near things are more related than far things” (Tobler, 1970) and (ii) the choice of data, which is mainly dictated by its availability at the studied spatial scale and by specific territorial concerns. To address this issue and create an optimal decision support tool for local stakeholders, the methodology should be deployed specifically across a more restricted area or at a finer spatial scale, taking care to use only relevant data to characterize the territory.

In quantitative studies, geography of spatial phenomena prioritizes indicator selection and/or methodology over geographical variability and scale choice (Cebrecos *et al.*, 2018; Schuurman *et al.*, 2007). The use of different partitioning systems for the same analysis could indeed offer different results (Wong, 2004). This phenomenon, known as the Modifiable Area Unit Problem (MAUP), was first identified by Gehlke and Biehl (1934). It combines (i) a “zoning effect”, whereby new boundaries can be drawn to create new zoning configuration, and (ii) a “scale effect”, whereby smaller units can be aggregated into larger units, and vice versa, thus creating hierarchical systems (Wong, 2004). The MAUP is one of the most far-reaching and longstanding problems in geography, and it is a critical factor in quantitative studies on spatial phenomena (Dark and Bram, 2007).

This study therefore seeks to understand how a change of scale influences the analysis of environmental and social health inequalities. Our work uses a multi-scale analysis based on a resilience index and a vulnerability index. These indices were developed for two reasons, namely (i) to compare spatial units with each other in order to prioritize them within the analysis perimeter and (ii) to determine their affiliation within a class of resilience or vulnerability. The joint analysis of these indices then seeks to classify each spatial unit into distinct territorial typologies, each of which require specific measures in the field of environmental health (preventive or curative action according to their vulnerability and resilience profiles). To this end, three cases were investigated to develop these composite spatialized indices for different spatial scales based on administrative units in different geographical areas. A comparative analysis was then carried out to identify the strengths and weaknesses inherent in the choice of spatial unit and study area.

2. Material and methods

2.1. Study area

This study was carried out in the Dunkerque Urban Community (DUC), a large industrial zone surrounded by a densely populated area, and located on the coast in the Hauts-de-France (HDF) Region (Figure 1). The 17 municipalities cover an area of 299.9 km² and have 198,341 inhabitants. These 17 municipalities can be divided into 93 IRIS. The acronym IRIS was created by the French national institute for statistics and economic studies (INSEE), and indicates ‘aggregated units for statistical information’, which each represent approximately 2,000 inhabitants. This high spatial resolution is the most commonly fundamental administrative unit used for the dissemination of infra-communal data. More details about population and cover areas of municipalities and IRIS can be found in supplementary materials.

The main activities of local industry in the DUC include petrochemistry, chemistry, metallurgy and energy production, and produce numerous atmospheric pollutants. Thirty-three days of air pollutant (PM₁₀, PM_{2.5}, NO₂) threshold exceedances were recorded in 2018. Thirty-two sites are classified in the national Pollutant Release and Transfer Register (PRTR), a national register designed to meet the requirements of the European (E-PRTR) and international (PRTR) pollutant protocol. Despite the relatively small area of the DUC, 850 sites are listed in the historical inventory of industrial sites and service activities or in the database of polluted or potentially polluted sites and soils. There are strong disparities in terms of income, housing or unemployment within the DUC IRIS, and the standardized mortality ratio between 2009 and 2015 is 21.3 % higher than the national average for the period studied (Ocelli *et al.*, 2016; OR2S, 2018). The leading causes of deaths on this territory are cancers of the upper respiratory and gastro-intestinal apparatus (malignant tumor of lip, oral cavity, pharynx, larynx, trachea, bronchi and lung), accentuated by risky behaviors (tobacco and alcohol in particular). Furthermore, actions in the health system shows high and significant results on avoidable mortality within the DUC (AGUR, 2015).

The DUC was highlighted in a previous study analyzing environmental and social inequalities in health over a larger perimeter (Brousmitche *et al.*, 2020a). Several municipalities displayed high or very high classes of vulnerability, associated with a very high level of resilience compared to the results obtained for all the municipalities in the HDF region.

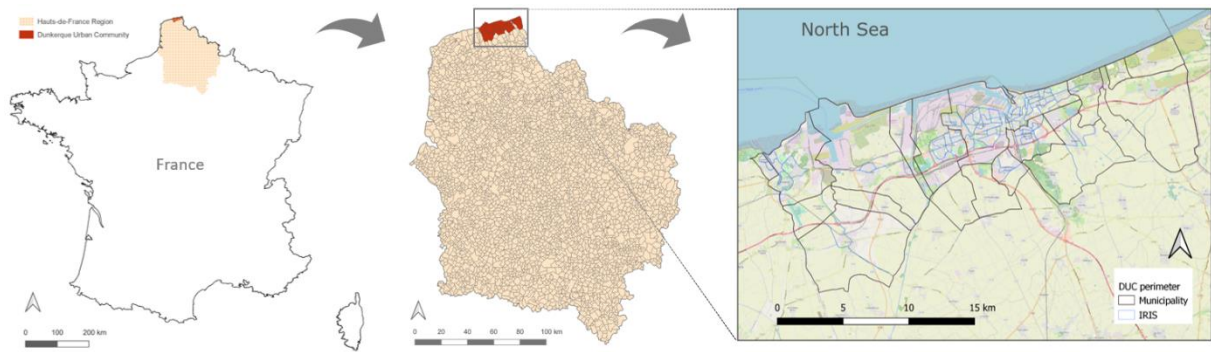


Figure 11 : Article 4 – Figure 1: Dunkerque Urban Community, located in the Hauts-de-France region (on the coast in the north of France)

2.2. [Methodological framework to develop the composite indices of resilience and vulnerability](#)

The proposed framework (Figure 2) was developed in Brousmiche *et al.* (2020a), where composite indices were used to assess environmental health inequalities in the municipalities of the “Hauts-de-France” region.

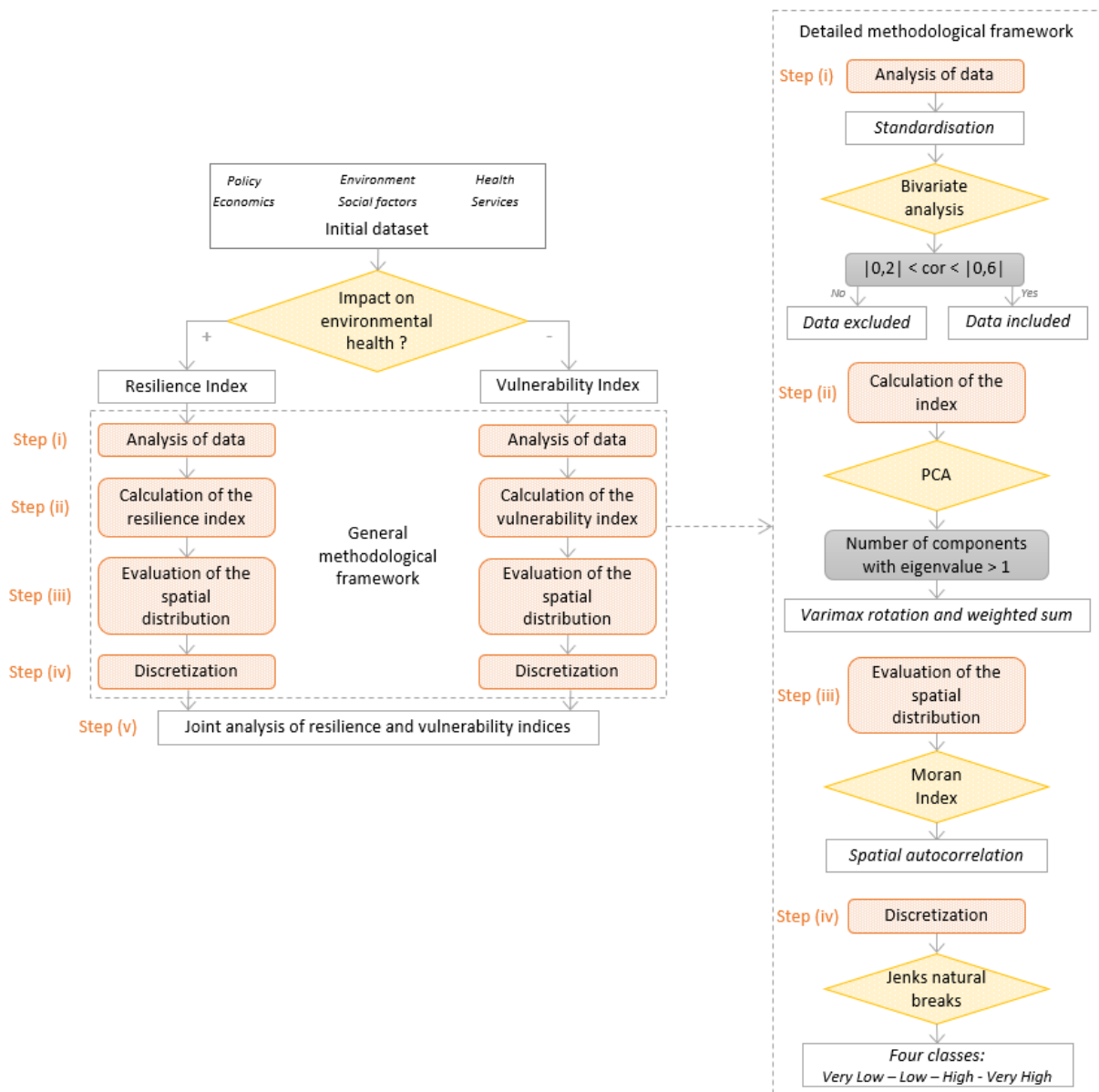


Figure 12 : Article 4 – Figure 2: Methodological framework used to develop composite indices (Brousmitche et al., 2020a)

Variables are collected in order to characterize the multifaceted phenomena involved in environmental health. A standardization step is then carried out to ensure that the variables are comparable (*i.e.* different units, different reference systems). A bivariate analysis is used to obtain a concise description of the initial dataset. This process enables us to (i) identify redundant variables and (ii) summarize the extensive information of the initial dataset. The composite indices are then constructed through a PCA-based weighting and aggregation process. A rotation of the factorial axes is applied to increase readability and discriminating power. The index is expressed as the weighted sum of the score for a principal component multiplied by the proportion of explained variance. The final indices are standardized by the min-max method to obtain values in a common, easily comprehensible range (*i.e.* [0-1]). The spatial autocorrelation is afterwards assessed by the Moran Index. For the subsequent joint analysis (*i.e.* the simultaneous interpretation of both discretized indices), a Spearman’s correlation test of the Resilience Index (RI) and the Vulnerability Index (VI) is used to identify the extent to which the indices provide different and complementary information.

Finally, Jenks natural breaks are used for discretization, leading to the classification of each spatial unit within one of the four following typologies:

- The “To monitor” typology corresponds to a Low or Very Low Resilience with a Low or Very Low Vulnerability,
- The “Resilient” typology, defined by a High or Very High Resilience and a Low or Very Low Vulnerability,
- The “Have resources” typology, combining a High or Very High Resilience and a High or Very High Vulnerability,
- The “Territorial blackspot” typology, characterized by a Low or Very Low Resilience and a High or Very High Vulnerability.

2.3. Methodology for joint analysis of indices at different scales

In order to consider the territorial specificities, three study cases are examined to specifically evaluate which of the determinants within the DUC perimeter have an impact on the characterization of health inequalities. A comparison in terms of accessibility, availability and exhaustivity is then carried out to study the influence of spatial scale on the development and results of composite indices (Figure 3).

	Case 1	Case 2	Case 3
Geographical area	Hauts-de-France Region	Dunkerque Urban Community	
Spatial unit	3,817 municipalities	17 municipalities	93 IRIS
Data set	50 variables collected from public national databases, accessible in open data (Brousliche et al, 2020b)		53 variables collected from public national databases, open data of the Dunkerque Urban Community or previous studies on this territory (Lanier et al., 2019, Occelli et al., 2020, 2016, 2014)
Methodology	Brousliche et al., 2020a Extraction of RI and VI values, classes of RI and VI and typology defined by the joint analysis of resilience and vulnerability indices	Development of new indices, adapted from Brousliche et al., 2020a	

Figure 13 : Article 4 – Figure 3: Synthesis of cases examined to assess the influence of the spatial unit on the development and results of composite indices that are used to study resilience and vulnerability in environmental health inequalities (RI: Resilience Index; VI: Vulnerability Index; IRIS: smallest aggregated unit for statistical information)

In Case 1, the classes for RI and VI and typologies are directly extracted from the previous study, where a dataset of 50 variables from public national databases had been used to develop the spatialized composite indices (Brousliche *et al.*, 2020a). For the second case, the entire methodology detailed in Brousliche *et al.* (2020a) is applied to the same dataset as in Case 1, but focuses on the 17 DUC municipalities in order to be more specific to the territorial context. Further details of the dataset at the municipality scale can be found in Brousliche *et al.* (2021). The third case is carried out at a higher spatial resolution (IRIS scale). Fifty-three infra-communal variables were collected from (i) public national databases, (ii) open DUC data and (iii) previous studies of this specific territory (Lanier *et al.*, 2019; Occelli *et al.*, 2020, 2016, 2014a). The supplementary materials provide all meta-data related to variables at the IRIS scale.

The variables were selected on the basis of their plausible association in the literature with environmental health. A wide variety of determinants of health such as social structure, neighborhood

characteristics and labor have a great impact on population health (Sorensen, 2000). For each case, the vulnerability indices therefore integrate data related to the multidimensional nature of the adverse environmental health determinants. Because poverty and social factors are main determinants of human health (Dunn and Hayes, 2000; Mackenbach *et al.*, 2008; Tzoulas *et al.*, 2007), vulnerability indices aggregate case-dependent data related to factors such as the unemployment rate, the precarious employment rate and the share of the population with few or no qualifications. Data related to housing (*e.g.* the share of the population who are tenants renting their main residence, or the dwelling construction date) have also been added, because housing is an expression of socio-economic status and plays a role in the production of social health inequalities (Dunn, 2002). Data related to environmental biomonitoring with plants were used in case 3 to broadly include environmental contamination (Cuny *et al.*, 2008). The share of population exposed to noise were added in case 3 given the immediate, cumulative and long-term impact of noise on health (Tao *et al.*, 2020).

The approach proposed in this methodology considers resilience and vulnerability to be equally important, with the challenge to provide a resilience index that would not be the simple antonym of the vulnerability index (Brousmiche *et al.*, 2020a). Identical efforts were therefore made to collect data (i) characterizing positive effect on environmental health and (ii) reflecting the same dimensions as the vulnerability indices (*i.e.* health, environment, economics, social factors, services). Green spaces could play a role in reducing air pollution and the urban heat island effect. They encourage people to be more physically active and are associated with a wide range of substantial health benefits (Aerts *et al.*, 2020; Bixby *et al.*, 2015; Twohig-Bennett and Jones, 2018; Tzoulas *et al.*, 2007). At a municipal level, health indicators appear to be more strongly associated with green environments than with urbanity (de Vries *et al.*, 2003). The area of natural spaces per inhabitant and the share of the population that has access to a green space within 400 m of their homes have therefore been aggregated in the RI. To describe the economics dimension, the evolution of the labor force over a five-year period and the average income per consumption unit were included in RI. A transdisciplinary adaptation was carried out to indicate the diversity of medical and paramedical professions within a Shannon-Weaver index. The latter is commonly used in ecology to describe a population in terms of the number of categories (category richness) and their relative abundance within those categories (species equity) (Shannon, 1948). The health dimension in our study was also characterized by the potential spatial accessibility to general practitioners. Others variables such as the number of state schools per 1,000 inhabitants, the share of the territory with 4G cover provided by the most efficient operator or the number of sports clubs, associations and cultural facilities per 1,000 inhabitants, are examples of determinants considered to represent social and services dimensions in RI.

Table 1 presents the distribution by index and by dimension of (i) the 50 variables at the municipality scale and (ii) the 53 variables collected at the IRIS scale. A dimension is a thematic classification of variables related to environmental health that partly reflect the health determinants presented by (Dahlgren and Whitehead, 1991). The selected dimensions are therefore health, environment, economics, social factors, services and policy.

Tableau XII : Article 4 – Table 1: Correlation matrix between the Vulnerability Index, the Resilience Index and other indicators (HDI4, French EDI, FDep). Results of the significant Spearman's correlation tests (i.e. $p < 0.001$) are marked with an asterisk

Dimensions	Dataset at the municipality scale (cases 1 and 2)		Dataset at the IRIS scale (case 3)	
	VI	RI	VI	RI
Economics	2	8	2	6
Environment	2	3	9	3
Policy	1	11	1	2
Health	1	5	3	3
Services (including housing)	4	3	8	5
Social factors (including education)	5	5	8	3
Total	15	35	31	22

The proposed methodology relies on a minimum of intercorrelations between variables to efficiently summarize the initial information. For each of the three cases, the Spearman correlation coefficients between VI and RI are calculated specifically for the DUC.

The statistical analyses are carried out using R software v1.3.1093. The indices are embedded into a Geographic Information System to observe the spatial distribution within the DUC, and the maps are produced in QGIS v3.16.15.

3. Results

3.1. Main statistical characteristics

Table 2 presents the results of the variable selection step (Step (i) in Figure 2) to build up resilience and vulnerability composite indices.

Tableau XIII : Article 4 – Table 2: For each of the three cases, results of the variable selection step to build up resilience and vulnerability composite indices

	Case 1		Case 2		Case 3	
	VI-1	RI-1	VI-2	RI-2	VI-3	RI-3
Variables included in the initial dataset	15	35	15	35	31	22
Non-discriminatory variables	-	-	-	2	-	-
Variables excluded due to $\rho > 0.6 $	1	8	8	14	15	7
Variables excluded due to $\rho < 0.2 $	1	6	-	-	-	-
Variables used in composite index	13	21	7	19	16	15

For Case 1, VI-1 was calculated from 13 variables among the 15 variables related to vulnerability, while RI-1 was composed of 21 of the 35 variables related to resilience that were initially present within the dataset of 50 variables. For Case 2, two variables (namely municipality covered by an Agenda 21 and the share of taxable tax households) were non-discriminatory at the DUC municipality scale (i.e. same value for all municipalities). Likewise, all the variables had a correlation coefficient greater than $|0.2|$ with at least one other variable. In addition, the correlation analysis revealed a higher number of coefficients above $|0.6|$, mainly among the vulnerability variables. This led to a lower number of variables being included in VI-2 (7 variables versus 13 in Case 1) and in RI-2 (19 variables versus 21 in Case 1). For example, the unemployment rate, the number of contaminated sites (polluted sites and soils) or the share of the population aged over 65 years were not retained for the development of VI-2. Variables with correlation coefficients above $|0.6|$ were also discarded. These included also the

potential spatial accessibility to general practitioners, the average income per consumption unit and the study continuation rate for 18-24-year-olds.

For Case 3, all the variables had at least a correlation coefficient greater than $|0.2|$ with at least one other variable. A greater number of correlations above $|0.6|$ is observed among the variables related to vulnerability (especially in the “environment” dimension). At the IRIS level, VI-3 includes 16 variables (15 variables are excluded, e.g. unemployment rate, share of population exposed to noise throughout the day, annual mean concentrations of SO₂ and share of single-parent families) while RI-3 has 15 variables (7 variables are excluded, e.g. monetization of ecosystem services per capita, evolution of the labor force and proportion of people owning their main residence).

Table 3 shows the distribution of spatial units (i.e. municipality in Cases 1 and 2, IRIS in Case 3) among the Very Low, Low, High and Very High classes (in number and percentage). Table 3 also presents the result of the Spearman’s correlation test between RI and VI and the distribution within the “To monitor”, “Resilient”, “Have resources” and “Territorial blackspot” typologies (in number and percentage).

Tableau XIV : Article 4 – Table 3: For each of the three cases, (i) distribution of spatial units among the Very Low, Low, High, Very High classes for the resilience and vulnerability indices, (ii) results of the Spearman’s correlation test between RI and VI, significant Spearman’s correlation tests (i.e. $p < 0.001$) are marked with an asterisk “”, and (iii) distribution municipalities and IRIS within the defined typologies*

		Case 1	Case 2	Case 3
Geographical area		HDF region	DUC	DUC
Administrative unit		Municipality	Municipality	IRIS
Number of spatial units		17 among the 3,817	17	93
Vulnerability Index (VI)	Very Low Vulnerability Class	4 (23.5%)	5 (29.4%)	16 (17.2%)
	Low Vulnerability Class	5 (29.4%)	4 (23.5%)	36 (38.7%)
	High Vulnerability Class	6 (35.3%)	5 (29.4%)	30 (32.3%)
	Very High Vulnerability Class	2 (11.8%)	3 (17.6%)	11 (11.8%)
Resilience Index (RI)	Very Low Resilience Class	-	3 (17.6%)	10 (10.8%)
	Low Resilience Class	-	1 (0.1%)	33 (35.5%)
	High Resilience Class	-	9 (52.9%)	38 (40.9%)
	Very High Resilience Class	17 (100%)	4 (23.5%)	12 (12.8%)
Joint analysis	Spearman’s correlation test between RI and VI	0.600*	0.748*	0.177
	To monitor	-	4 (23.5%)	29 (31.2%)
	Resilient	9 (52.9%)	5 (29.4%)	23 (24.7%)
	Have Resources	8 (47.1%)	8 (47.1%)	27 (29.0%)
	Territorial Blackspot	-	-	14 (15.1%)

In Case 1, the correlation test between VI-1 and RI-1 calculated solely for the DUC municipalities shows a relatively high correlation. This result underlines that RI-1 and VI-1 provide different information, albeit to a small extent. The extraction of the resilience classes from a previous study on a larger perimeter did not provide a discrimination of the territory in terms of resilience, as all municipalities belonged to the Very High Resilience class. Moreover, the joint analysis carried out in Case 1 did not reveal any “To monitor” or “Territorial blackspot” categories. In Case 2, there is a high correlation between RI-2 and VI-2. This finding highlights that the use of composite indices does not achieve to provide different information. Although all classes of vulnerability and resilience are represented, the joint analysis did not reveal any territorial blackspot when the municipality is used as spatial unit of study. In Case 3, the correlation between RI-3 and VI-3 is relatively moderate, the p-value being non-

significant. These two spatialized composite indices can therefore be considered to provide complementary information.

3.2. Results for municipalities of Dunkerque Urban Community

Composite indices in Case 1 were based on 34 variables that were selected after bivariate analysis of the 3,817 municipalities in the region. Figure 4 shows the spatial distribution of the vulnerability index (Figure 4a), the resilience index (Figure 4b) and the joint analysis of these two composite indices (Figure 4c) for the Hauts-de-France region and DUC. Composite indices in Case 2, developed specifically for the 17 municipalities of the DUC, were based on 26 variables. Figure 4 shows the municipalities cartography on DUC as a function of VI-2 ranking (Figure 4(d)), RI-2 ranking (Figure 4(e)) and the joint analysis of these two composite indices (Figure 4(f), Case 2). Supplementary results present the list of included variables in each composite index (Cases 1 and 2).

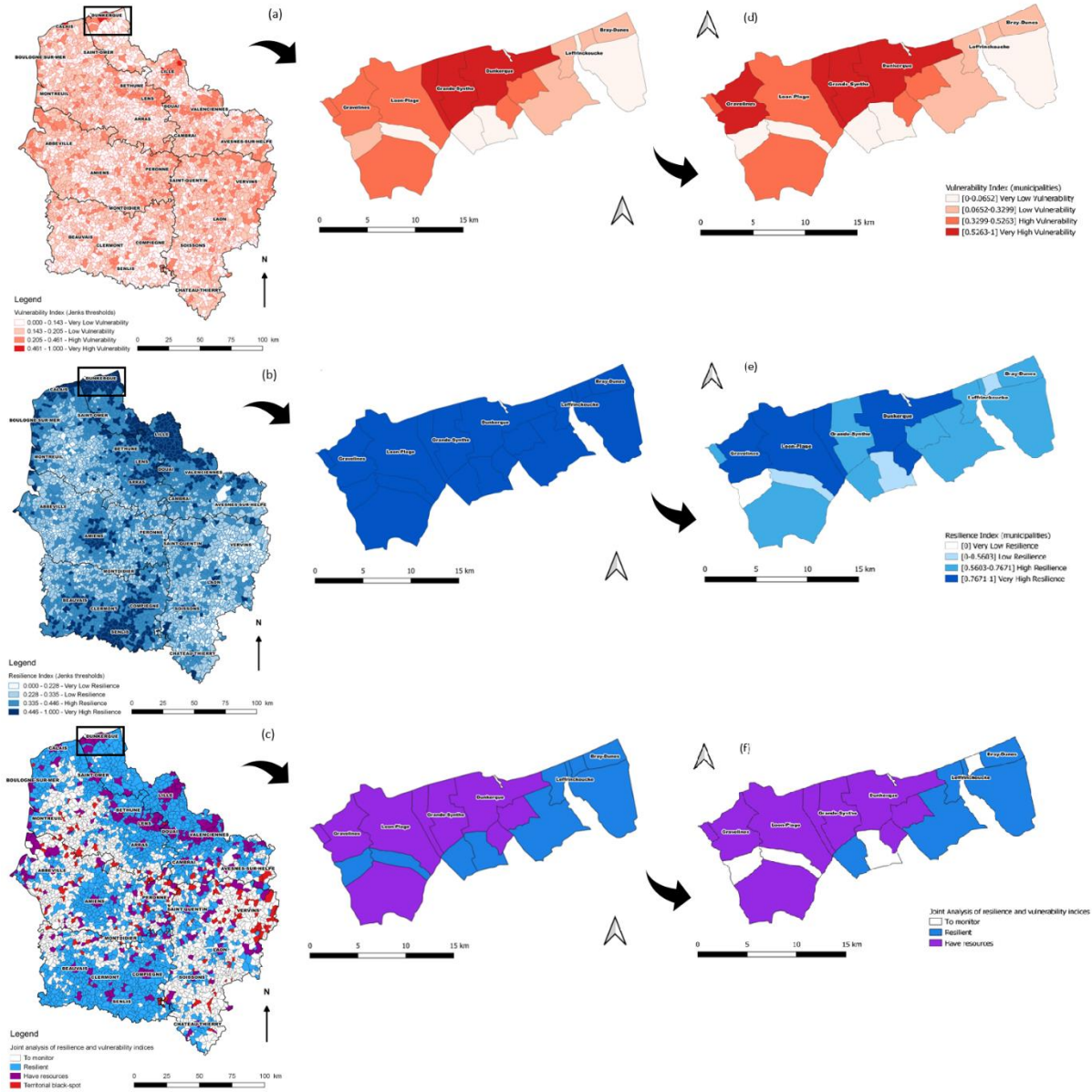


Figure 14 : Article 4 – Figure 4: Composite indices based on 34 variables – focus on the municipalities of the Dunkerque Urban Community (Brousmiche et al., 2020a); case 1: Vulnerability Index (a), Resilience Index (b), Joint analysis of Resilience and Vulnerability (c); Case2: Vulnerability Index (d), Resilience Index (e), Joint analysis of Resilience and Vulnerability (f)

In Case 1, the municipalities classed as Very High Vulnerability are mainly located in the center of the DUC and correspond to the municipalities of Dunkerque and Grande-Synthe. The municipalities classed as Low or Very Low Vulnerability are generally located in the more rural municipalities of the DUC, most often in the south-eastern area of the DUC. Extracting the RI does not make it possible to provide a spatial discrimination on the DUC: all municipalities are classified within the Very High Resilience group. The municipalities classified in the “Have resources” typology are therefore superimposed on the zones of High or Very High Vulnerability.

In Case 2, the municipalities with Very High Vulnerability levels are located in Dunkerque, Grande-Synthe and Gravelines, in central and western parts of the DUC. The municipalities classified as Low or Very Low Vulnerability are rather located in more rural areas, which are in the south-eastern part of the DUC. The municipalities classified as High or Very High Resilience are Dunkerque, Gravelines and Loon-Plage (center and west of the agglomeration). The four municipalities with Low or Very Low Resilience classes are located in the outskirts of the agglomeration. Municipalities classified as “Have resources” are located in the center-west of the agglomeration, while those classified as “Resilient” and “To monitor” are on the outskirts of the DUC.

3.3. [Case 3: Application of the methodology to the IRIS of Dunkerque Urban Community](#)

Composite indices in this case were based on 31 variables that were selected after bivariate analysis of the 93 IRIS in the DUC. Figure 5 shows the spatial distribution for the IRIS in the DUC for the Vulnerability Index (Figure 5(a)), the Resilience Index (Figure 5(b)) and their joint analysis (Figure 5(c)). Supplementary materials and results present meta-data and the list of included variables in resilience and vulnerability indices.

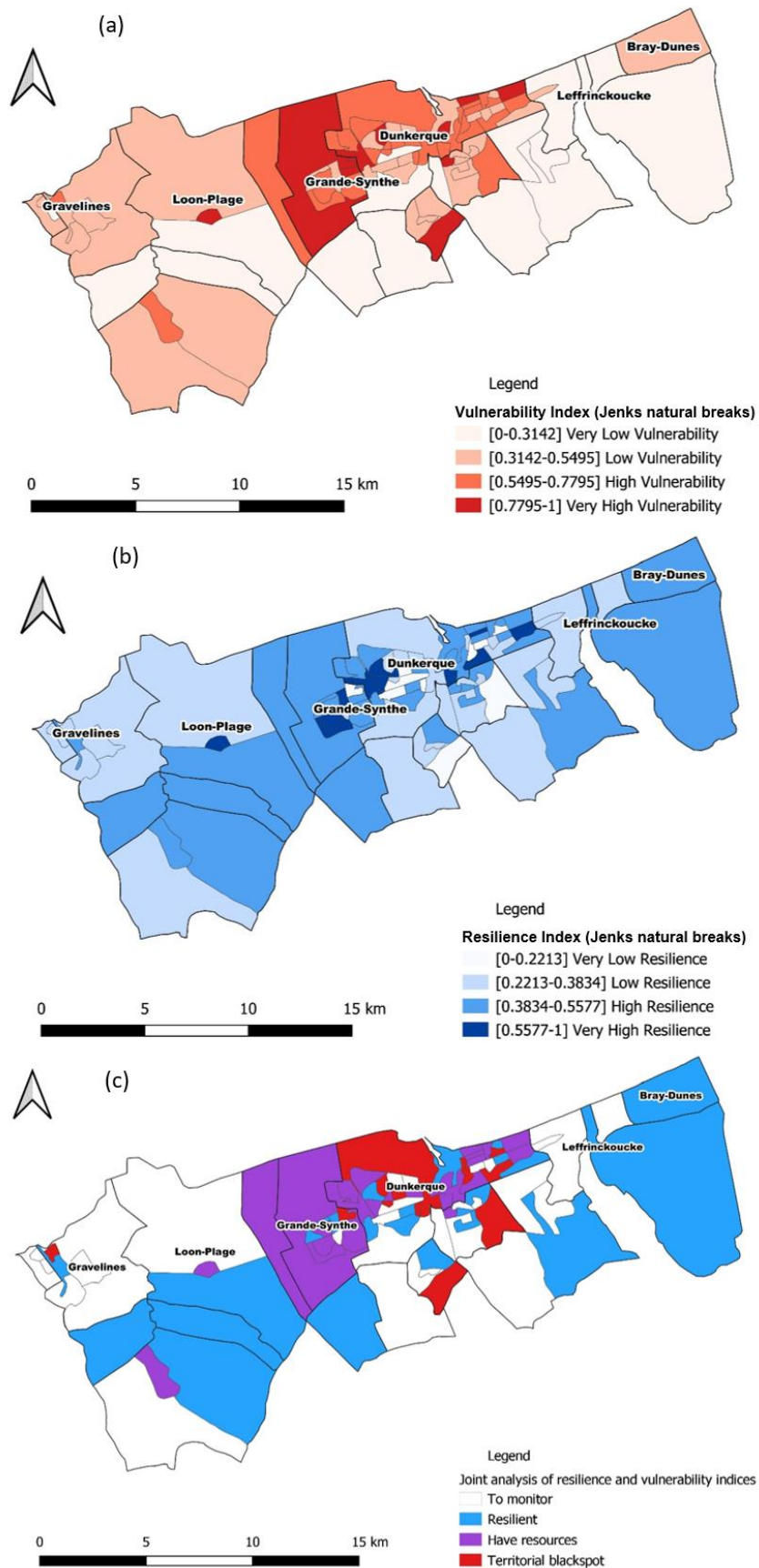


Figure 15 : Article 4 – Figure 5: Development of composite indices based on 31 variables carried out for the 93 IRISs of Dunkerque Urban Community. (a): Vulnerability Index; (b): Resilience Index; (c): Joint analysis of resilience and vulnerability indices

In Case 3, the IRIS classified as High or Very High Vulnerability are located in the center of the DUC, and are mainly located in the municipality of Dunkerque. The areas of High or Very High Resilience are also located in the center of the agglomeration, with most of the IRIS located in Dunkerque. The IRIS classified as “Have resources” and “Territorial blackspot” are therefore mainly located in the center of the DUC. The IRIS classed in the “To monitor” and “Resilient” categories seem to be a little more randomly distributed across the study territory, with a slightly higher tendency of location in the outskirts of the DUC.

4. Discussion

This research studied the implications of the choice of a spatial unit and a geographical area in the assessment of environmental and social inequalities in health. For this purpose, composite indices of vulnerability and resilience were developed for three specific cases. The joint analysis of these two indices sought to classify each spatial unit in a specific territory typology indicating the need for specific measures in the field of environmental health (*i.e.* preventive actions for “Have resources” typology, and curative actions for the “Territorial blackspot” typology).

The identification of the variables to be collected is the first challenge when developing composite indices that aim to provide an overall description of environmental health determinants. Biomonitoring with plants (poplar, lichen) is recognized as a reliable biological approach to characterize ambient contamination (Lanier *et al.*, 2019), and within the DUC geographical area results were highly correlated with the annual mean concentrations of atmospheric pollutants (PM₁₀, NO₂ and SO₂). Contributing to the overall atmospheric pollution, noise is also considered as an important factor impacting human health in urban areas (Cincinelli and Katsoyiannis, 2017; Khreis *et al.*, 2016). As these data (*i.e.* biomonitoring with plants, noise) are not available for a wider geographical area, they could not be integrated into the dataset for Cases 1 and 2 despite their relevance with environmental health issues.

The second challenge is to select the area and the spatial resolution of the study (Cebrecos *et al.*, 2018; Petroni *et al.*, 2021). Cases 1 and 2 were implemented at the same spatial scale, but initially deployed for different geographical areas. Although the variable selection step is among the main factors that can cause differences (Neves Almeida and García-Sánchez, 2016), the results for VI in Cases 1 and 2 are broadly analogous. The spatial distribution classes obtained in Case 2 coincides to a great extent with the spatial distribution of classes obtained in Case 1. Additionally, the classes of high/very high vulnerability match areas for which multiple vulnerabilities have already been identified. Environmental nuisances are mainly located in the center/extreme west, and the standard mortality ratio is more favorable at the east of the agglomeration and particularly negative in Grande-Synthe municipality (AGUR, 2019). The main difference between Cases 1 and 2 is explained by the spatial pattern of the Resilience Indices. Extracting resilience classes that were determined for a larger geographical area has no effect on discriminating between municipalities of the DUC. The joint analysis performed in Case 1 permitted the classification of each municipality in one of two typologies. Nevertheless, this has few interests for DUC stakeholders in analyzing and dealing with specific territorial concerns. Implementing the methodology to develop composite indices on a restricted geographical area (Case 2) allowed a more effective transcription of the resilience variability present within the territory, where variables were selected according to territorial specificities. The composite indices developed in Cases 1 and 2 provide different information, albeit to a small extent, given the relatively high correlation coefficients between resilience and vulnerability calculated for the DUC. The result of the correlation test between RI-1 and VI-1 is, however, much higher than what was obtained for a larger geographical area with 3,817 spatial units ($\rho = -0.01$ - $p = 0.57$, Brousmiche *et al.* (2020a)).

This high correlation is an indicator of the poor relevance of this approach (simple extraction of a result obtained from a more macroscopic approach): the information provided by RI-1 and VI-1 is very similar and does not successfully meet the discrimination objective. It is important to bear in mind that the reference in each case is invariably the implementation scope of the methodology. The index values are indeed specific to the context for which they were developed (number of spatial units, variables).

Cases 2 and 3 were developed for the same geographical area but for different spatial units. This makes it possible to compare the importance of the choice of the spatial unit on the depiction quality of the determinants at the origins of the environmental and social inequalities in health. The smaller number of spatial units in a limited geographical area in Case 2 led to the exclusion of some variables from the dataset due to their lower discriminatory power. The results of our study corroborate the findings of existing literature stating that larger spatial units tend to present less variability, because the original variability would have been averaged or smoothed (Buzzelli, 2020; Wong, 2004). Some environmental data over smaller areas can show a lower spatial variability, failing to highlight differences within spatial units (Cartone and Postiglione, 2021). This concerns for example the annual mean concentration of PM₁₀, which has a median of 25.09 µg/m³ with an IQR of [24.31;25.51] on the DUC perimeter. A lower variability is an issue when composite indices aim to prioritize and rank spatial units among themselves. Although the PCA-methodology is statistically robust to develop composite indices (European Commission *et al.*, 2008), less variability may result in an artificially lower weight. In Case 3, this question of the lower spatial variability of atmospheric pollution on the restricted area of the DUC has been overcome by using eutrophication and dust deposition variables, which are highly correlated with commonly monitored indicators (annual average concentration of NO₂, PM₁₀ and SO₂).

The indices at an infra-communal level highlight the great disparities within a municipality, with some IRIS classified as very low resilience or vulnerability levels, while others located nearby belong to the are classified at very high levels. The use of a higher resolution spatial unit in Case 3 makes it possible to analyze a greater spatial variability and to obtain a clear description of the health determinants, mainly in the most densely populated municipalities of the DUC (Dunkerque and Grande-Synthe are home to over half of the DUC inhabitants and include respectively 43 and 10 of the 93 IRIS). Nevertheless, the main limitation of a very small spatial scale is the availability of data for certain variables. This mainly concerns data related to the policy dimension, even though the latter is considered essential to improve environment (Etsy *et al.*, 2005). Disaggregating the information at the IRIS scale could have been an alternative means to complete the dataset in Case 3. However, zoning differences are one of the main causes of the MAUP phenomenon. The overlapping of artificial spatial units can lead to potential measurement errors that could affect result accuracy (Cebrecos *et al.*, 2018; Zhao *et al.*, 2021). In our study, the MAUP issue has been minimized by avoiding the use of several different scales within an analysis. All the variables in Case 3 have therefore been collected and calculated at the IRIS scale, and this small spatial scale allows a higher discriminatory character while preserving high resolution. Because of the scale-related effect of the MAUP, it is impossible to conclude which scale is “better” or “more correct” when searching for certainty (Buzzelli, 2020). Many studies have attempted to propose solutions, but there is no consensus to date on how best to deal with the MAUP phenomenon (Josselin and Louvet, 2019). Decreasing the effect of MAUP remains a challenge, as it could either have an influence on the association of environmental health indicators and health outcomes, or reduce evidence strength between them (Labib *et al.*, 2020; Lee *et al.*, 2020; Zhao *et al.*, 2021).

The proof of concept carried out on this agglomeration of 299.9 km² enables us to consider developing composite resilience and vulnerability indices at the IRIS level for a larger geographical area. There are a number of interesting points that should be considered, such as (i) the number of spatial units within

a given perimeter, (ii) the specificity of data to certain smaller perimeters (*e.g.* noise data only exists for certain urban areas) and (iii) the exhaustiveness of the data across the whole territory. The smaller number of spatial units in cases 1 and 2 does not allow to appraise the multicollinearity which might arise when too high intercorrelations exist in a dataset (European Commission *et al.*, 2008) and the spatial autocorrelation of composite indices. More entities than in cases 1 and 2 are advisable to obtain reliable results when appraising multicollinearity by the Variance Inflation Factor which is commonly employed (European Commission *et al.*, 2008; Frigerio and De Amicis, 2016; Kim, 2019; Saghapour *et al.*, 2021; Vorstius *et al.*, 2019), or to characterize spatial patterns by the global Moran's Index (Getis, 2007; Goodchild, 1986). The use of a more restricted area can ensure the accessibility of more numerous environmental data (environmental noise and biomonitoring data integrated in Case 3). Lastly, working at a very fine scale over a large geographical area can lead to a high rate of missing data. This completeness problem can complicate the analyses and the interpretation of the results, and could even weaken the final conclusions (Pedersen *et al.*, 2017).

The identification of the "Territorial blackspot" typology is not a goal in itself. It is fundamental to diagnose the variables that have led to this classification and the use of high-resolution spatial units is of real interest for urban and regional planning (Salvati and Carlucci, 2014). This deeper understanding of the territory could help practitioners and stakeholders to (i) prioritize the actions to be carried out, (ii) provide greater contextualization of the public health policies to be implemented, with the adaptation of action to issues that are potentially specific to a neighborhood and (iii) suggest more accurate modifications of the living environment (Cartone and Postiglione, 2021). This is in line with the World Health Organization 'Health in All Policies' approach, which aims to improve population health and health equity (Leppo, 2013). Finally, among the variables included in the composite indices, several fall within the jurisdiction of the agglomeration. This can concern issues related to the urban transportation plan, the identification of areas that should remain unbuilt or the creation and/or maintenance of sport facilities. These examples of mitigation actions that may lead to healthier cities could therefore be implemented faster by local jurisdictions than by national governments because of their unique features (Khomenko *et al.*, 2020; Nieuwenhuijsen and Khreis, 2019).

Discussion Chapitre I

1. Variables incluses aux indices composites

Les indices composites à fine échelle peuvent permettre de répondre aux besoins des décideurs pour caractériser finement un territoire et améliorer la contextualisation des politiques locales dans tous les domaines en lien avec la santé. En effet, les indices composites proposent une vision agrégée des variables intégrées et ont pour but de caractériser une problématique. L'analyse de la littérature qui a été menée a permis (i) d'identifier les variables couramment ou ponctuellement incluses dans des indices composites en santé environnement et qui reflètent la diversité des facteurs impliqués en santé environnement et (ii) d'orienter le choix des variables à inclure dans des indices composites de résilience et de vulnérabilité. De plus, l'assignation des variables à la résilience ou la vulnérabilité s'est effectuée selon l'information portée par la variable. Si une augmentation de la valeur de la variable est en lien avec un effet bénéfique sur la santé environnement, la variable est considérée comme faisant partie de la résilience tandis que si une augmentation peut être reliée à un effet délétère sur la santé environnement, la variable est considérée comme faisant partie de la vulnérabilité.

1.1. Indices de vulnérabilité

Les indices de vulnérabilité en santé environnement qui ont été développés dans les précédents travaux visaient à caractériser les contraintes et nuisances auxquelles sont soumises les populations. D'après Wagstaff (2002), les inégalités dans le secteur de la santé sont associées à la pauvreté. La relation entre ces deux aspects est considérée comme une injustice sociale, les inégalités de santé se produisant presque toujours au désavantage des populations pauvres (Wagstaff, 2002). Selon l'INSEE (2021), un ménage français est considéré comme pauvre lorsque le niveau de vie est inférieur à 60 % du niveau de vie médian de la population (*i.e.* le revenu disponible correspond à 1 102 € par mois pour une personne seule et 2 314 € pour un couple avec deux enfants âgés de moins de 14 ans). La pauvreté est toutefois un phénomène complexe et multidimensionnel qui ne peut être limitée à la mesure du niveau de revenus (Concialdi, 2014; Dollé, 2008; Mercier, 2005; Sen, 1999). Plusieurs aspects en lien avec la pauvreté ont été considérés dans les indices de vulnérabilité.

Parmi la dimension « Économie », des variables en lien avec la précarité et/ou le chômage sont fréquemment retrouvées dans la littérature dans l'analyse de la vulnérabilité d'un territoire (Angeon et Bates, 2015; Braconnier *et al.*, 2011; Estoque et Murayama, 2014). Le chômage, au-delà du manque de revenus qui peut, selon les législations nationales, être compensé en partie, contribue à une forme d'exclusion sociale et peut entraîner des problèmes de santé psychologique et physique (Sen, 1999). De plus, bien que l'activité ne soit pas un facteur protecteur vis-à-vis de la pauvreté, les chômeurs et les inactifs y sont particulièrement confrontés (Concialdi, 2014; Rollet et Thomas, 2021). La précarité est souvent considérée comme un synonyme de la pauvreté alors qu'elle devrait s'en distinguer. Le terme de précarité renvoie plutôt à une fragilité due à l'accumulation de conditions de vie instables possiblement temporaires et réversibles (Chauvin et Estecahandy, 2010; Wresinski, 1987). Pour Pierret (2013), une population précaire est en effet privée d'une ou plusieurs des sécurités (*e.g.* l'emploi) et pourrait constituer de potentiels pauvres en devenir, presque toutes les catégories sociales pouvant de plus y être confrontées. La précarité (part de contrat à durée déterminée, emplois aidés, intérim, apprentissage) a donc été distinguée de l'emploi et du chômage dans les indices composites. Ces déterminants de la santé ont un impact important sur l'amélioration de la santé, car au-delà des écarts

socio-économiques (*e.g.* revenus, situation familiale), les comportements favorables à la santé ou encore la sensibilité aux messages de prévention suivent un gradient social (Potvin *et al.*, 2010).

La situation sociale du territoire a également été caractérisée par l'intégration de variables en lien avec cette dimension. En effet, ces déterminants peuvent constituer des facteurs aggravants de la vulnérabilité. Bien que la part de familles monoparentales n'ait pas été retrouvée dans les indices composites identifiés dans la revue, elle peut constituer un déterminant important de la pauvreté au regard de la fragilité de cette structure familiale (Rollet et Thomas, 2021). Une autre dimension de la pauvreté concerne l'échec scolaire et la non-scolarisation, les niveaux de savoir, d'alphabétisation ou d'instruction tendant à être plus bas parmi des populations pauvres (Dollé, 2008; Henaff *et al.*, 2009; Wagstaff, 2002). La part des personnes pas ou peu diplômées a également été intégrée afin de mesurer dans l'unité spatiale le degré de moindre instruction ou de sortie précoce du système éducatif. De plus, le niveau d'éducation est un déterminant qui évolue peu dans le temps et qui s'avère facilement mesurable (Maresca and Helmi, 2014). La part de population immigrée au sein de l'indice de vulnérabilité permet de retranscrire l'isolement linguistique des ménages qui est parfois considéré dans le développement de certains indices composites. Cette variable reflète ici la possible fracture sociale qui peut résulter de la structure de la population conduisant potentiellement à une forme d'isolement (linguistique, social, économique). Cette variable est toutefois un exemple des variables à lecture équivoque dans la littérature (Zuindeau et Lipovac, 2009), la part de population immigrée pouvant parfois être considérée comme une source de richesse communautaire. Le degré d'urbanisation, la densité de population ou encore le statut urbain/rural sont d'autres exemples de variables à caractère équivoque dans la littérature. La densité de population a été intégrée à l'indice de vulnérabilité comme un proxy pour certains risques plus prépondérants en milieu urbain (*e.g.* bruit, trafic routier) bien qu'elle aurait pu être considérée comme un proxy pour une plus grande accessibilité à de nombreuses ressources en lien avec un degré d'urbanisation plus grand (WHO, 2010). Des données en lien avec l'âge de la population ont également été intégrées. En effet, une population plus âgée est déjà de santé potentiellement plus fragile, dans une situation instable entre bonne santé et altération du bien-être. Une population plus âgée présente en effet un accroissement du risque de maladie avec possiblement de moindres capacités d'y faire face (Cambois *et al.*, 2021; Dourlens, 2008). Ces populations peuvent également rencontrer des limitations fonctionnelles ou être considérées comme socialement vulnérables (Huang et London, 2012; Lang, 2011). Toutefois, l'analyse de la littérature a mis en évidence que les classes d'âge considérées pour caractériser cette fragilité diffèrent selon les études, la classe d'âge de 60 ans et plus est mentionnée pour Huang et London (2012), Monteiro *et al.* (2012) et Sadd *et al.* (2011), Stewart *et al.* (2014) ont considéré les plus de 65 ans dans leurs travaux, Cutter *et al.* (2010) ont intégré la part de population non âgée sans qu'il y ait une définition précise de la classe d'âge de la population non âgée. Cependant, les comparaisons nationales et internationales du vieillissement des populations semblent plus classiquement réalisées à partir des données sur les 65 ans et plus (de Lapasse *et al.*, 2018), l'âge de 65 ans a donc été retenu pour caractériser la fragilité du territoire en lien avec le vieillissement. La prise en compte de ces déterminants dans l'analyse permet de contextualiser et d'adapter les politiques de santé publique en fonction des spécificités territoriales qui auront été identifiées (*e.g.* part élevée de familles monoparentales, de population issue de l'immigration et/ou de population âgée).

D'autres déterminants de la santé ont également été intégrés aux indices composites. Il s'agit notamment de l'accès au logement et aux conditions d'habitat dans la dimension « Services ». En effet, ces déterminants peuvent générer ou contribuer aux inégalités de santé selon différents mécanismes, par exemple (i) le lien fort entre le niveau de vie et la catégorie de logement, (ii) les variations du prix du foncier ou (iii) des différences d'attractivité ou de mécanismes d'évitement (Ginot et Peyr, 2010; Laurent, 2013; Potvin *et al.*, 2010). Le mode d'occupation du logement est régulièrement mentionné

dans la littérature pour ses liens sur son rôle matériel, la richesse créée ou encore un sentiment de sécurité (Bentley *et al.*, 2016; S. Wang *et al.*, 2019). L'ancienneté de la construction peut également figurer parmi les causes possibles d'indécence des logements ou être un facteur aggravant pour certaines pathologies (Monteiro *et al.*, 2012; Rollet et Thomas, 2021). La qualité d'un logement (équipements, taille, situation géographique) peut également avoir des impacts sanitaires, par exemple sur la santé mentale (Rollet et Thomas, 2021; Wang *et al.*, 2019).

En ce qui concerne les caractéristiques physiques de la zone d'habitat, les populations les moins aisées ont tendance à vivre à plus faible distance de différentes sources de pollution environnementale qui peuvent entraîner des impacts directs et négatifs sur la santé en raison de la pollution atmosphérique, du trafic routier ou encore du bruit (Faburel, 2008; INERIS, 2014; Potvin *et al.*, 2010; WHO, 2019). Ces populations plus défavorisées seraient dans le même temps plus exposées mais également plus sensibles à ces impacts sanitaires résultant d'effets synergiques (Caudeville and Rican, 2016; WHO, 2019, 2012b). Des variables en lien la présence de routes primaires dans la zone de vie, de sites et sols pollués ou d'Installations Classées pour la Protection de l'Environnement (ICPE) ont été intégrées aux indices de vulnérabilité comme « proxy » de la concentration industrielle (Saib, 2015). Les pressions anthropiques ont également été intégrées aux indices de vulnérabilité, les variables intégrées étant différentes en raison de différences d'accessibilité (*e.g.* les émissions de Gaz à Effet de Serre à l'échelle des communes, des données de issues la biosurveillance de la qualité de l'air sur les IRIS de la CUD). L'approche par la biosurveillance permet d'apporter une information complémentaire par rapport aux données physico-chimiques mesurées ponctuellement et qui visent à caractériser la qualité de l'air (Cuny, 2012).

1.1. Indices de résilience

Les indices de résilience de santé environnement développés à l'échelle des communes des Hauts-de-France (Brousmiche *et al.*, 2020a) et des IRIS de la CUD (Brousmiche *et al.*, *under review*) visaient à retranscrire la capacité du territoire à faire face aux inégalités de santé. Parce que le patrimoine et les revenus constituent des ressources fondamentales des ménages (Wagstaff, 2002), ils ont été intégrés aux indices de résilience au sein de la dimension « Économie ». La variable « revenu disponible par unité de consommation » correspond au revenu disponible du ménage pour vivre en tenant compte de la structure du ménage (INSEE, 2021b). Les impôts directs (impôt sur le revenu, taxe d'habitation, prélèvements sociaux) sont déduits du total des ressources des ménages (revenus d'activité, indemnités, retraites, pensions, prestations sociales), ce revenu étant alors rapporté aux équivalents adultes que composent le ménage.

Les territoires, quels que soient les découpages administratifs, sont des objets de gestion et d'aménagement. Les différentes stratégies mises en place et les efforts consentis peuvent être reflétés par exemple la présence d'équipements collectifs ainsi que leur accessibilité (Lang, 2011; Moine, 2006). A l'échelle des communes, pour la dimension « Politique », les données financières du secteur public local en lien avec les budgets communaux permettent de retranscrire des orientations politiques pouvant être spécifiques pour certaines communes (subventions versées par la commune, dépenses pour les équipements). Pour les IRIS, en l'absence de ce type de données à très fine échelle une réflexion a été menée pour identifier des aspects quantifiables en lien avec la dimension « Politique ». En effet, la gouvernance et les politiques publiques sont pour Etsy *et al.* (2005) des aspects essentiels pour améliorer l'environnement. A l'échelle des IRIS, des variables ont été calculées à partir des Plans Locaux d'Urbanisme (PLU), ces derniers devant respecter des principes respectueux du Développement Durable pour les projets globaux d'aménagement et d'urbanisme pour réhabiliter, restructurer ou encore aménager le territoire pour la satisfaction des besoins (*e.g.* équipements publics, équipements sportifs et culturels). Les politiques publiques ont une influence sur les

équipements collectifs d'une zone de résidence ainsi que sur leur accessibilité, qui sont des facteurs déterminants pour favoriser des comportements sains (Lang, 2011).

Pour enrichir la dimension « Services », des données en lien avec les transports en commun ont été intégrées. Les différentes infrastructures de transport appartiennent pour Potvin *et al.* (2010) aux déterminants sociaux de la santé. De plus pour Nieuwenhuijsen (2020) et Rojas-Rueda *et al.* (2013), les politiques publiques qui viseraient à favoriser les transports publics et notamment les transports publics actifs permettraient d'améliorer la santé publique en réduisant par exemple la morbidité. De manière plus étendue, intégrer les équipements d'une zone géographique dans une analyse permet de représenter l'espace de vie collectif. Le cadre de vie d'un territoire est en lien avec sa capacité à assumer les besoins des individus et de leur entourage pour proposer des conditions de vie acceptables (accès à l'éducation, accès au sport, accès à la culture, accès aux soins, ...) (Rollet et Thomas, 2021). La dimension « Social » des indices de résilience comprend la poursuite d'étude des 18-24 ans afin de caractériser les inégalités d'accès à l'enseignement supérieur. Dans la région Hauts-de-France, les enfants issus des classes sociales défavorisées sont particulièrement moins bien représentés dans l'enseignement supérieur avec une sur-représentation des enfants des cadres et des professions libérales (51,4 % en 2019-2020) par rapport aux enfants des ouvriers (7,2 % en 2019-2020) (Rollet et Thomas, 2021).

En s'appuyant sur les conclusions de la revue avec notamment plusieurs indices composites ayant intégré des variables en lien avec la diversité et/ou le dynamisme économique, des variables traduisant ces aspects ont été intégrés aux sets de données (évolution de la population active, taux de survie à 5 ans des entreprises). La sphère présente selon la partition économique des activités de l'INSEE se distingue de la sphère productive dans le sens où ces activités (i) répondent à une demande locale, (ii) nécessitent une certaine proximité avec la population et (iii) ont pour but d'être mises en œuvre de manière locale, pour les besoins des populations présentes dans la zone (résidentes ou touristes) (INSEE, 2015; Reynard *et al.*, 2015). Cette variable traduisant le dynamisme local de l'activité économique a donc été ajoutée aux sets de données. De plus, le pourcentage de populations vivant et travaillant dans la même commune a également été intégré afin de pouvoir retranscrire plusieurs aspects. Tout d'abord celui de la mobilité quotidienne des populations qui est principalement liée au travail et aux études, 85 % des déplacements pouvant être associés au travail (Zelem, 2013). L'accès à l'automobile peut être source d'inégalités (i) en termes de mobilité urbaine quotidienne et ce d'autant plus si les alternatives modales sont peu présentes voir insuffisantes dans un territoire (Nicolas *et al.*, 2012), (ii) le poids financier de l'automobile peut être plus conséquent pour les ménages à revenus moins élevés ce qui peut contribuer à fragiliser davantage ces populations (Nicolas *et al.*, 2012; Orfeuill, 2006; Zelem, 2013). Des populations présentant des revenus et/ou des niveaux de diplôme plus faibles tendent en effet à vivre dans des zones où les coûts de logement sont moins élevés (banlieue, zone rurale) qui peuvent être moins bien desservies par les transports en commun ce qui contraint le recours à la voiture (Zelem, 2013). Cette variable peut également permettre de mesurer dans quelle mesure la commune offre une activité professionnelle aux personnes résidentes et rendre compte de la part de population exposée aux pressions anthropiques locales, l'ensemble des expositions étant déterminée au lieu de vie. Toutefois, l'accès à l'automobile en tant que source d'inégalités est à nuancer au regard d'un mode de vie plus citadin où les populations vivant dans les centres villes d'agglomérations urbaines auraient davantage recours à des modes déplacement vers le lieu de travail plus actifs (*i.e.* transports en commun, déplacements à pied ou à vélo) (Zelem, 2013).

Une réflexion transdisciplinaire a été menée afin de pouvoir proposer une variable traduisant de la diversité des emplois proposés au lieu de vie sur la base des catégories socio-professionnelles. L'indice de Shannon est très utilisé en écologie pour renseigner la diversité spécifique car il tient compte (i) du

nombre d'espèce et (ii) de leur abondance relative au sein d'un écosystème (Shannon, 1948). Une des limites de cet indice est qu'en écologie, le nombre total d'espèces présentes au sein d'un écosystème peut être complexe à déterminer. Toutefois, cette limite n'existe pas ici car le nombre d'espèces correspond au nombre de catégories socio-professionnelles. Pour Reynaud *et al.* (2012), la catégorie socio-professionnelle fait partie des variables « standard », avec le revenu et le niveau de diplôme, qui peuvent être utilisées pour évaluer la position sociale des individus.

Cette démarche transdisciplinaire a également été menée pour quantifier la diversité et la richesse spécifique des professions médicales et paramédicales qui peuvent exister au sein des différentes unités spatiales. Les inégalités d'utilisation des offres de soins peuvent être en effet accentuées par leurs disparités spatiales (Ducros *et al.*, 2015). Cet aspect est venu (i) enrichir la dimension « santé » du set de données développé à l'échelle des IRIS et (ii) apporter un éclairage complémentaire à l'Accessibilité Potentielle Localisée (APL) aux médecins généralistes à l'échelle des communes. L'APL permet de quantifier l'offre de soins en médecine générale en tenant compte de l'offre, de la distance potentielle entre les populations et le médecin généraliste et de la demande potentielle en s'affranchissant des limites administratives (Chevallard *et al.*, 2018; Ducros *et al.*, 2015). L'APL peut permettre d'identifier si des territoires sont concernés par la problématique des « déserts médicaux » qui peuvent être en lien avec un isolement du territoire, une faible densité de population ou encore le non-remplacement des médecins partant en retraite. L'isolement du territoire est de plus susceptible de se cumuler à l'isolement de populations (*e.g.* population âgée, issue de l'immigration et barrière de la langue) contribuant à accentuer un état de santé dégradé (Bourret, 2016). Bien que les résultats de l'APL tendent à souligner que certaines zones rurales isolées présentent de meilleurs résultats que des communes rurales en périphérie de pôles urbains, l'accessibilité potentielle ne représente pas nécessairement l'accessibilité réelle (Chevallard *et al.*, 2018; Ducros *et al.*, 2015). De plus, les « déserts médicaux » ne concernent pas uniquement des territoires ruraux isolés, les conditions d'exercice peuvent être difficiles dans des quartiers sensibles au sein des pôles urbains ce qui peut contribuer au départ des professionnels de santé (Bourret, 2016). Enfin, le recours aux soins et le non-recours aux soins ne sont pas uniquement associés au niveau socio-économique des territoires, des difficultés sociales ou culturelles peuvent exister et peut en outre relever de choix individuels (Chevallard *et al.*, 2018; Ducros *et al.*, 2015; Vaillant *et al.*, 2021). La diversité et la richesse spécifique des professions médicales et paramédicales recensées au lieu d'exercice retranscrite au travers de l'indice de Shannon permet donc de proposer une vision plus élargie des caractéristiques sanitaires d'un territoire.

Concernant les variables environnementales intégrées aux sets de données, de nombreuses études dans la littérature ont souligné les impacts positifs de certains indicateurs environnementaux sur la santé (Mukherjee *et al.*, 2017; van den Bosch et Ode Sang, 2017). Par exemple, Villanueva *et al.* (2015) relie les espaces verts publics avec la santé et le bien-être, Richardson *et al.* (2010) s'appuient sur le lien entre accessibilité aux espaces verts sur la santé perçue ou la mortalité tout risque, et pour Yeager *et al.* (2020) la proximité résidentielle à la végétation est associée à de nombreux impacts positifs sur la santé (réduction du risque cardiovasculaire, diminution de la mortalité respiratoire), la végétation du quartier encouragerait une plus grande cohésion sociale et la pratique sportive. De plus, pour Potvin *et al.* (2010), le niveau d'activité physique moyen du quartier est influencé avec la disponibilité et/ou l'accessibilité à des infrastructures sportives.

Enfin, certaines variables peuvent être classées dans plusieurs dimensions parce qu'elles peuvent avoir une lecture à caractère équivoque selon le point de vue considéré pour leur interprétation (*e.g.* la part de voies vertes et pistes cyclables est à la frontière de plusieurs dimensions : « environnement », « services », « politique », possiblement la « santé » ; le nombre d'écoles publiques à la frontière des dimensions « Social » et « Services » ; les municipalités couvertes par un Agenda 21 à la frontière entre

les dimensions « Environnement » et « Politique » ; accessibilité aux médecins généralistes à la frontière entre la « Santé » et les « Services »). Pour AUDIAR (2016), il peut être utile d'intégrer ce genre d'informations synergiques, car les réponses en lien avec ce type de variables impactent plusieurs dimensions.

2. Méthodologies d'analyse des inégalités territoriales de santé

Les travaux du PNUD ont largement contribué au développement des indices composites et suscitent toujours une littérature abondante qui met en lumière les bénéfices et les limites méthodologiques des indices composites (Dialga, 2016; Klugman *et al.*, 2011b). Les indices composites parce qu'ils proposent une valeur synthétique facilitent la lecture des résultats obtenus pour les décideurs et le grand public. Néanmoins, parce qu'ils agrègent différents indicateurs en une valeur unique, la causalité entre la valeur agrégée et le phénomène étudié au travers du développement de cette méthodologie pourrait être plus complexe à mettre en évidence (HCSP, 2021). La synthèse apportée par la valeur agrégée de l'indice composite ne doit toutefois pas donner lieu à des conclusions analytiques simplifiées voire simplistes (European Commission *et al.*, 2008). Les résultats doivent être appréciés au regard des unités spatiales considérées, des variables incluses et des choix méthodologiques qui ont été opérés. Pour le Haut Conseil de la Santé Publique, la question de la transparence pour toutes les étapes de développement d'un indice composite est fondamentale pour l'interprétation des résultats (HCSP, 2021).

Parce qu'il existe plusieurs méthodologies possibles pour chacune des étapes, une analyse critique présentant les avantages et les limites de chacune a été réalisée afin de pouvoir proposer une méthodologie adaptée et robuste pour caractériser la multidimensionnalité des phénomènes impliqués dans les inégalités territoriales en santé environnement. Lors de l'étape de sélection des variables, la disponibilité des données peut contraindre le choix d'une unité spatiale vers un niveau d'agrégation plus élevé afin d'intégrer des indicateurs plus pertinents (Exner *et al.*, 2016). De plus, pour des unités spatiales qui présentent des tailles inégales, des variables sous forme de taux, de proportion ou encore de moyennes seraient préférables (Guerrien, 2003). La qualité des variables incluses est à prendre en considération lors d'une analyse des forces et faiblesses d'un indice composite (European Commission *et al.*, 2008; Freudenberg, 2003; Mauro *et al.*, 2021). Cet aspect bien que pouvant avoir un impact non négligeable sur la fiabilité des indices composites développés est toutefois peu mentionné dans la littérature (Ceccarelli *et al.*, 2020; Mauro *et al.*, 2021). Pour l'European Commission *et al.* (2008) et Freudenberg (2003), l'identification, la disponibilité et la qualité des variables d'intérêt constituent un problème majeur lors du développement d'un indice composite quel que soit le champ disciplinaire. Pour des indices qui visent à être mis à jour régulièrement, il apparaît judicieux d'intégrer des données suivies qui permettront de faire évoluer les valeurs au fur et à mesure du temps. Lorsque les indices composites ont pour but d'établir des comparaisons internationales, il est fondamental de considérer dans la discussion que les indicateurs intégrés peuvent avoir été mesurés et calculés de manière différente, le choix d'un indicateur pouvant même être non pertinent selon le contexte (Daoui, 2018).

Parce que des indicateurs peuvent agréger des variables exprimées dans des unités très différentes (*e.g.* pourcentage, valeur monétaire, nombre pour 1000 habitants) ou dans des échelles différentes, une normalisation doit être réalisée afin de pouvoir obtenir des variables qui puissent être comparables (Böhringer and Jochem, 2007; Saib *et al.*, 2015). La normalisation sur la base des minimas et des maximas observés constitue une des méthodes reconnues. L'utilisation de seuils fixes basés sur des résultats historiques peut permettre de (i) pouvoir comparer les résultats d'une année sur l'autre et (ii) tendre vers une objectivation de cette étape. Toutefois l'utilisation de valeurs historiques pour la normalisation est controversée, les minimas et maximas historiques peuvent présenter « un

caractère aberrant » ce qui peut fortement influencer les résultats obtenus après la normalisation (Dialga, 2016).

Les étapes de pondération et d'agrégation doivent être choisies au regard du cadre théorique de développement de l'indice composite mais également en fonction des propriétés des données qui sont incluses à l'indice. L'étape de pondération est essentiellement un jugement de valeur, sur l'attribution d'une pondération ou non mais également sur le choix de la méthode qui est retenue (European Commission *et al.*, 2008; Gan *et al.*, 2017). La principale difficulté de cette étape repose sur l'attribution d'un poids plus important à certaines variables et/ou à certaines dimensions, parce qu'elles sont jugées plus importantes par rapport à la question de recherche (Burgass *et al.*, 2017). Néanmoins, le choix d'affecter une pondération arbitraire peut s'avérer complexe dans le cas de la santé environnementale parce que ses déterminants sont nombreux et que pour certains, les connaissances de leurs impacts sur les inégalités territoriales en santé environnement sont encore parcellaires. Sous-estimer méthodologiquement certains facteurs par l'attribution d'une pondération moindre peut alors sous-estimer des relations entre ces déterminants et les inégalités territoriales en santé environnement, concourant à moins prendre en considération l'impact d'un déterminant dans l'analyse des inégalités. L'agrégation, éventuellement pondérée, consiste en la synthèse des différents indicateurs en une valeur chiffrée unique (Joumard, 2016). L'agrégation linéaire, et notamment la moyenne algébrique, attribue mathématiquement une parfaite substituabilité et une compensabilité entre les différentes variables et/ou dimensions. Il est à noter également qu'un plus grand nombre de variables au sein d'une dimension peut contribuer dans le cas d'une moyenne arithmétique à mécaniquement minimiser l'influence de chacune de ces variables. De la même manière, considérer toutes les variables comme équivalentes (*i.e.* absence de pondération des variables) peut entraîner un déséquilibre de représentation des dimensions dans les indices composites. Dans le cas d'un nombre différent de variables par dimension, une pondération égale ne signifie donc pas une égale représentation des dimensions, les dimensions caractérisées par plus de variables ont *de facto* un poids plus important (European Commission *et al.*, 2008). En ce qui concerne l'agrégation géométrique, elle engendre une substituabilité moindre que l'agrégation linéaire mais est plus sensible à la présence de faibles valeurs et tend à favoriser une convergence des résultats. De plus, toutes les méthodes de pondération ne sont pas compatibles avec l'agrégation géométrique. Par rapport à la sélection des variables, s'il existe des corrélations relativement importantes entre des indicateurs, les modes de pondération et d'agrégation ont moins d'impact sur les résultats finaux (Dialga, 2016). Néanmoins, des corrélations élevées entre des variables visant à être intégrées par une agrégation linéaire induisent un phénomène de double comptage, un seuil de corrélation au-delà duquel écarter des variables peut être mis en place afin d'éviter ce phénomène (European Commission *et al.*, 2008). D'après la littérature, ces étapes de pondération et d'agrégation peuvent s'avérer être les plus décisives en termes de résultats (European Commission *et al.*, 2008; Gan *et al.*, 2017; Tešić *et al.*, 2018; Wiréhn *et al.*, 2015).

Après analyse critique, la méthodologie de développement d'indice composite qui a été retenue repose sur l'Analyse en Composantes Principales (ACP). Cette méthodologie factorielle permet de (i) s'affranchir des choix empiriques qui doivent être opérés au moment des étapes de pondération et d'agrégation, et (ii) de construire l'indice sur la base des relations statistiques qui existent entre les variables au sein d'une unité géographique (Pampalon et Raymond, 2000). L'ACP a de plus la capacité de synthétiser l'information de nombreuses variables quantitatives ce qui peut faciliter leur interprétation (Guerrien, 2003). D'après Marzi *et al.* (2018), la standardisation (*i.e.* normalisation par le z-score) est à privilégier avant une ACP, toutes les variables ayant alors une moyenne de 0 et un écart type de 1. Cette méthode de normalisation présente l'avantage d'être moins sensible aux valeurs extrêmes et conserve la structure des corrélations entre les variables (European Commission *et al.*, 2008; Kolenikov and Angeles, 2004). Afin de limiter la redondance des informations apportées dans les

indices composites, une gamme de coefficients de corrélation a été proposée lors de l'étape de sélection des variables afin (i) d'obtenir un nombre optimal de variables à intégrer aux indices composites ce qui minimise le phénomène de double comptage et (ii) d'introduire dans l'ACP des variables qui présentent un minimum de corrélation les unes avec les autres pour restreindre les composantes principales portées par une seule variable (bruit). L'objectif était de construire des indices composites riches en informations et pas uniquement riches en nombre d'indicateurs. Cette analyse des corrélations a été réalisée sur le set global de données (*i.e.* données en lien avec la vulnérabilité et données en lien avec la résilience) afin de pouvoir proposer une sélection de variables proposant une caractérisation du territoire la plus large possible et tenant compte de l'effet « miroir » qui peut exister entre certains déterminants. Cette méthode de sélection de variables a permis d'obtenir des indices de résilience et de vulnérabilité qui apportaient des informations complémentaires.

Parce que le développement d'indices composites implique une sélection de variables parmi un set initial, il est recommandé d'étudier la sensibilité qui est intrinsèquement liée aux incertitudes des indices développés (European Commission *et al.*, 2008). De plus, cette étape concourt à la transparence nécessaire lors de l'analyse des résultats. Deux méthodes d'agrégation linéaire par l'ACP ont été testées : une ACP par dimension et une ACP globale. L'analyse de sensibilité qui a été menée a considéré la part de données manquantes afin d'évaluer l'impact du pourcentage de données manquantes sur la qualité de la réponse au regard du pourcentage de classification correcte pour retenir la méthodologie qui serait la plus stable. La part de données manquantes peut en effet être une source de difficultés lorsque les données pour caractériser des déterminants de santé sont issues de sources dont le processus qualité pourrait ne pas être optimisé. De plus, si des données sont manquantes pour certaines unités spatiales cela implique (i) de mettre en œuvre des méthodes d'imputation entraînant une source d'incertitudes supplémentaires ou (ii) d'exclure ces unités spatiales pour le développement des indices composites, cette seconde alternative ne correspondant toutefois pas aux conditions initiales (*i.e.* proposer une caractérisation de l'ensemble du périmètre considéré). Les résultats de cette analyse sont présentés dans la Figure 16.

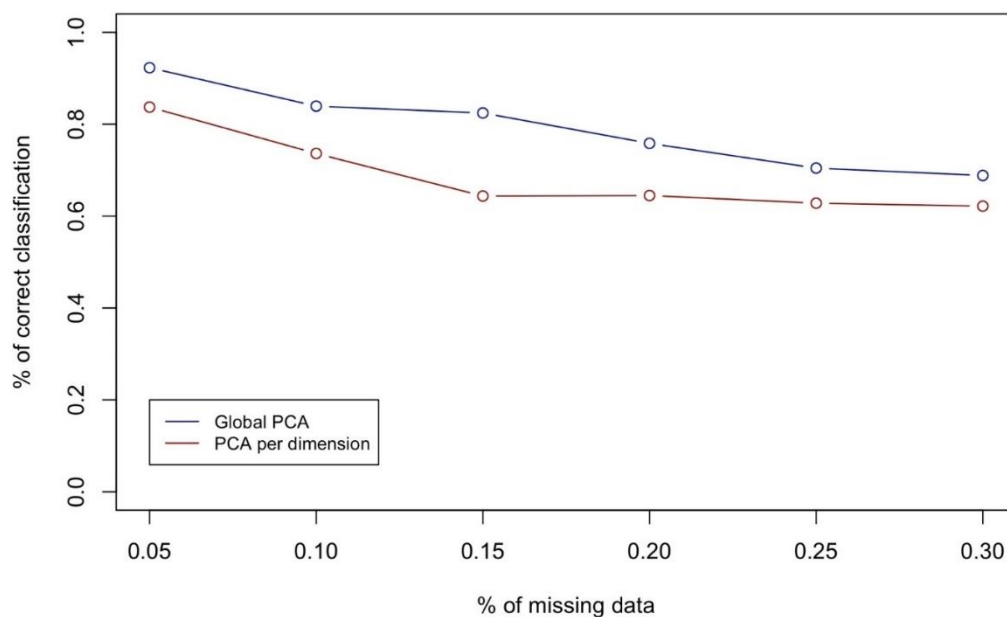


Figure 16 : Analyse de sensibilité vis-à-vis des données manquantes, comparaison de deux méthodes de développement d'indices composites

L'ACP globale (courbe en bleue sur la Figure 16) a présenté les résultats les plus stables en matière de classement lorsque le pourcentage de données manquantes augmentait. Cette méthodologie a donc été retenue pour le développement des indices composites. Cette méthodologie a considéré l'ensemble des déterminants des inégalités territoriales de santé de manière concomitante car ils ne sont en effet pas indépendants les uns des autres (Lang *et al.*, 2017) et l'ACP globale a permis de s'affranchir de la possible équivocité de classement d'un déterminant au sein d'une dimension. D'autres analyses de sensibilité doivent encore être menées pour compléter l'évaluation des incertitudes méthodologiques, notamment vis-à-vis des variables qui ont été intégrées.

La méthodologie qui a été proposée pour le développement d'indices composites spatialisés de résilience et de vulnérabilité s'appuie sur l'*open data*. L'*open data* en France est en plein développement depuis le milieu des années 2000 et les politiques d'ouverture ainsi que la progression du numérique ont renforcé l'accroissement de ces données (Jouve, 2013; Rebouillat, 2020). Selon le Journal Officiel, l'*open data* correspond à la mise à disposition des données principalement publiques sous forme numérique, le format de mise à disposition devant favoriser leur réutilisation (Chignard, 2012; Dulong de Rosnay, 2017; JORF, 2014). Le *data reuse* permet de proposer une méthodologie transposable pour d'autres territoires que ceux pour lesquels elle a été initialement proposée (*e.g.* la région des Hauts-de-France). Les bases de données de santé et environnementales mais également les progrès en lien avec les SIG ont contribué à développer l'accès à des informations détaillées pour la recherche sur les inégalités territoriales en santé environnement. L'amélioration de la caractérisation de l'exposition environnementale *via* des données de résolution plus élevée a fortement profité à l'épidémiologie environnementale (Chang, 2016; Maantay, 2007). En France, plusieurs lois se sont succédées sur l'accessibilité et la réutilisation des données (lois Valter, Macron, NOTRe), néanmoins l'*open data* s'est véritablement institutionnalisé depuis la loi pour une République numérique de 2016 (Mericksay, 2021). L'expansion de l'*open data* fait également suite aux actions de surveillance dans différents domaines (santé, environnement, recensement de la population, ...), ces données ouvertes pouvant réellement constituer une opportunité d'accroître les connaissances sur le lien entre la santé et l'environnement (Saib, 2015). La littérature scientifique souligne la disponibilité toujours plus accrue de données ouvertes au public ainsi que le développement croissant d'indicateurs à visée de recherche scientifique à partir de ces données secondaires (Harwell *et al.*, 2019). Les principales limites actuelles des données en libre accès portent sur des aspects fonctionnels et opérationnels pour la réutilisation, notamment d'un point de vue format de production et d'extraction qui pourraient encore être optimisés (Maday, 2016). Il existe en effet de multiples formats, de nombreuses sources de données (institutionnelles, privées, associatives, issues du crowdsourcing), de nombreuses plateformes, ce qui peut contribuer à complexifier l'accès à l'information, la mutualisation et la réutilisation (Jouve, 2013; Larroche et Dymytrava, 2018). De plus, d'après Larroche et Dymytrava (2018) et Mericksay (2018) (i) l'insuffisance voire l'absence de normalisation des données et des métadonnées associées et (ii) de culture de la donnée (notamment vis-à-vis de la structure) constituent un défi supplémentaire pour l'interopérabilité. Pour pallier ces différences, les données issues des organismes officiels (*e.g.* INSEE), préférentiellement nationaux, ont été privilégiées. Pour chacune donnée issue d'un organisme national, la mesure et/ou l'enregistrement sont effectués de manière identique ce qui permet de lever le verrou méthodologique de la possible variabilité inhérente en lien avec plusieurs sources pour une même donnée (variabilité en termes de mesure et variabilité qualitative). Toutefois, en l'absence de disponibilité nationale de certaines données (*e.g.* bruit), les institutions publiques locales ont été privilégiées sur la base de différents critères d'inclusion. La donnée décrivant un déterminant de la santé environnement devait (i) tendre à l'exhaustivité sur le périmètre considéré et/ou présenter un pourcentage de données manquantes peu élevé avec une distribution spatiale homogène, (ii) être contemporaine par rapport aux autres données incluses dans le set de données ou (iii) être suivies

dans le temps. L'application de ces critères ainsi que des analyses menées pour chacune des données décrivant un déterminant de la santé environnementale ont permis de produire des sets de données fiables pour le développement des indices composites de vulnérabilité et de résilience visant à caractériser les inégalités territoriales en santé environnement sur l'ensemble du périmètre considéré.

Chapitre II : Caractéristiques du territoire et maladies chroniques

Introduction

Près d'un quart des décès prématurés dans le monde était imputable à des déterminants environnementaux évitables (WHO, 2018a). Les infections, la pollution des milieux (eau, air, sol), l'exposition aux substances chimiques ou encore le changement climatique constituent les principaux facteurs de risques environnementaux influençant l'état de santé des populations (Ministère des Solidarités et de la Santé, 2022). En fonction du niveau d'industrialisation et du développement des pays, la nature et les sources de pollution évoluent (Landrigan *et al.*, 2018). Les études de santé environnement sont complexes car (i) il n'existe pas de spécificité des impacts sanitaires de l'environnement, (ii) les populations sont exposées à de multiples substances dont les effets combinés sont encore mal connus et (iii) il peut y avoir un délai entre l'exposition et l'apparition d'une maladie, les effets sanitaires pouvant apparaître à plus long terme (Ministère des Solidarités et de la Santé, 2022). L'évaluation des liens entre environnement dégradé et effet sanitaire peut être difficile au regard de la sensibilité individuelle en lien avec le patrimoine génétique, la durée entre l'exposition et la survenue de pathologies ou encore la présence de facteurs de confusion (Ministère des Solidarités et de la Santé, 2022; Vineis *et al.*, 2020). De plus, il existe une très grande variabilité des environnements naturels mais également de diversité des structures sociales ou en encore des choix individuels (Bhatnagar, 2016). Toutefois, l'environnement n'est pas seulement à considérer comme un facteur de risque, certains déterminants peuvent avoir en effet des impacts bénéfiques sur la santé (*e.g.* respiratoire, cardiovasculaire, mentale) (Porcherie *et al.*, 2017).

Parmi les déterminants de santé environnement, la pollution atmosphérique est désignée comme une problématique majeure de santé publique, plus de la moitié des décès en lien avec des conditions environnementales évitables étant due à celle-ci (WHO, 2022). En 2019, la quasi-totalité de la population mondiale respirait un air ne respectant pas les seuils de pollution au-delà desquels des effets sur la santé sont observés (WHO, 2021a, 2022). La pollution atmosphérique est constituée d'un mélange de différents polluants dont les origines peuvent être naturelles ou anthropiques et les particules (notamment les PM₁₀ et les PM_{2.5}), l'ozone troposphérique (O₃), les oxydes d'azote (NO_x), le monoxyde de carbone (CO) et les oxydes de soufre (SO_x) font partie des formes de polluants principalement étudiés dans la littérature (Dominski *et al.*, 2021; Landrigan *et al.*, 2018; Loomis *et al.*, 2013). Le trafic routier est le principal contributeur de la pollution atmosphérique et à ce titre les populations urbaines sont particulièrement affectées (Manisalidis *et al.*, 2020). Depuis plusieurs décennies, de nombreuses revues systématiques et méta-analyses sont conduites pour identifier les liens entre pollution atmosphérique et ses nombreux impacts sanitaires. Cette abondante littérature souligne l'intérêt toujours prépondérant suscité par cette problématique majeure de santé publique. Une exposition aiguë à la pollution atmosphérique peut être reliée à la survenue ou l'aggravation d'effets notamment respiratoires (toux, essoufflement, asthme), de la bronchopneumopathie chronique obstructive (BPCO) et à une augmentation des taux d'hospitalisation (Manisalidis *et al.*, 2020). Les effets à long terme de la pollution atmosphérique sont notamment associés à des problématiques respiratoires (asthme chronique, insuffisance pulmonaire) et aux maladies cardiovasculaires (Manisalidis *et al.*, 2020). La pollution atmosphérique a été classée comme cancérigène avéré pour l'Homme en 2013 (cancer du poumon) et une association positive avec un risque accru de cancer de la vessie a également été démontrée (Dominski *et al.*, 2021; Hassan Bhat *et al.*, 2021; IARC, 2013). De plus, une exposition à long terme à la pollution atmosphérique peut être

associée à la prévalence du diabète, aux naissances prématurées ou encore à des pathologies en lien avec le système nerveux central (autisme, démence) (Eze *et al.*, 2014; Landrigan *et al.*, 2018). Quels que soient les pays et leur niveau d'industrialisation, les populations pauvres et marginalisées présentent des effets plus fréquents et/ou plus sévères (Landrigan 2018).

En ce qui concerne le changement climatique, ses effets négatifs sur la santé sont déjà observables (Haines and Ebi, 2019). Au cours des 30 dernières années, plus de 150 000 décès par an ont pu en effet être associés à l'augmentation des températures et/ou des précipitations (Song *et al.*, 2017). La survenue de certaines maladies serait de plus particulièrement affectée par des variations climatiques sur le long terme ou lors d'évènements climatiques extrêmes à court terme (Watts *et al.*, 2018). De plus, les populations âgées et les personnes souffrant de maladies préexistantes seraient particulièrement vulnérables à ces effets (Linares *et al.*, 2020). L'augmentation des températures peut (i) exacerber des problèmes de santé existants contribuant à une augmentation de la morbidité (*e.g.* exacerbation des effets sanitaires de la pollution atmosphérique), (ii) nuire au bien-être des populations vieillissantes et (iii) induire de nouvelles menaces de santé (Argacha *et al.*, 2018; Chen *et al.*, 2019; Rajagopalan and Landrigan, 2021; Song *et al.*, 2017; Watts *et al.*, 2018). Toutefois, les résultats de nombreuses études épidémiologiques ont également souligné l'impact sanitaire de températures plus froides, avec des courbes en « J » en « V » ou en « U » (Song *et al.*, 2017). La question de la relation entre santé humaine et climat constitue une problématique environnementale et médicale complexe (McGeehin and Mirabelli, 2001).

La caractérisation des liens entre les inégalités de santé et ses facteurs contributifs est essentielle pour réduire les disparités qui sont observées (Juhász *et al.*, 2010). La santé figure en effet parmi les dimensions du développement et du bien-être de manière prépondérante (Dahlgren and Whitehead, 2007). Afin d'avancer dans la connaissance des liens entre géographie, santé humaine et pathologies les Systèmes d'Information Géographique (SIG) apparaissent comme un outil important (Bowe *et al.*, 2017). La spatialisation des informations en amont d'études plus précises permet de formuler des hypothèses étiologiques. (Fromageot *et al.*, 2005). L'analyse cartographique des pathologies offre de plus aux gestionnaires territoriaux la possibilité (i) de cibler spécifiquement les programmes de prévention et (ii) de suivre et d'évaluer des programmes de santé publique (Fromageot *et al.*, 2005; Maclachlan *et al.*, 2007).

Partie II-1 : l'Insuffisance Rénale Chronique Terminale

Introduction

Les Maladies Rénales Chroniques (MRC) sont définies comme des anomalies de la fonction rénale ou de la structure du rein présentes depuis plus de trois mois et qui ont des implications pour la santé des patients (KDIGO, 2013). Elles constituent depuis 2002 un problème mondial de santé publique avec une incidence et une prévalence en augmentation (Tanner *et al.*, 2013; Zhang and Rothenbacher, 2008). Les MRC sont des pathologies à évolution lente aboutissant à l'insuffisance rénale lorsque les reins sont irréversiblement détruits, et le nombre de patients concernés par les MRC est difficile à évaluer au regard de cette lente évolution (INSERM, 2017; KDIGO, 2013). Elles sont largement dues au vieillissement de la population et à la réduction physiologique normalement constatée (Jungers *et al.*, 2011). En outre, les premiers stades de ces maladies sont souvent asymptomatiques, la pathologie se manifestant généralement quand un stade relativement avancé a été atteint (HAS, 2021). L'insuffisance rénale comporte cinq stades, établis en fonction de la perte progressive de la fonction rénale qui est quantifiée par l'évaluation du Débit de Filtration Glomérulaire (DFG) (Jungers *et al.*, 2011). Il correspond au volume de plasma filtré par les reins par unité de temps normalisé à 1,73 m² en fonction du sexe, de l'ethnie, du poids et de la taille des patients (IRDES and ABM, 2018; Jungers *et al.*, 2011; Rouse, 2017). La valeur normale du DFG pour un adulte de type caucasien sans hypertension artérielle est d'au moins 90 mL/min/1,73 m². L'insuffisance rénale est considérée comme chronique lorsque le DFG est inférieur à 60 mL/min/1,73 m² et terminale lorsque celui-ci est inférieur à 15 (INSERM, 2017; IRDES and ABM, 2018). L'Insuffisance Rénale Chronique Terminale (IRCT) constitue le stade ultime de l'altération de la fonction rénale et l'incidence moyenne en Europe en 2015 s'élevait à 119 par million d'habitants (pmh), avec de fortes disparités selon les pays s'échelonnant de 24 en Ukraine à 232 pmh en République Tchèque (Kramer *et al.*, 2018). La France se situe dans la moyenne haute européenne avec une incidence moyenne à 166 pmh, le nombre de patients suivant un traitement de suppléance augmentant de 4 % chaque année (HAS, 2021). Une importante variabilité régionale existe également et les taux d'incidence les plus élevés se situent dans l'ancienne région du Nord-Pas de Calais, ceux-ci variant de 35 à 396 pmh (Collister *et al.*, 2016; IRDES and ABM, 2018).

Le diabète de type 2 ou l'hypertension artérielle, dont les prévalences augmentent également naturellement avec l'âge, sont les principales causes de l'évolution des MRC vers l'insuffisance rénale sévère (Bowe *et al.*, 2017; INSERM, 2017; Jungers *et al.*, 2011; Nortier *et al.*, 2017). En 2015, 22 % des cas incidents d'IRCT ont été causés par l'hyperglycémie diabétique qui a entraîné une détérioration des capillaires glomérulaires (néphropathie diabétique). La néphropathie hypertensive est diagnostiquée dans 25 % des cas incidents d'IRCT. Elle associe deux mécanismes différents : d'une part des rétrécissements des artéioles du rein et d'autre part une réduction de la vascularisation conduisant possiblement à la défaillance rénale (INSERM, 2017). D'autres mécanismes (inflammatoire, dégénératif, génétique) peuvent être impliqués dans les autres néphropathies qui sont souvent détectées lors d'analyses sanguine ou urinaire. Les MRC sont des maladies qui requièrent des soins tout au long de la vie des patients et la prise en charge est le plus souvent réalisée par transplantation rénale ou par des méthodes d'épuration du sang (Bilgel, 2019). Il s'agit par exemple de l'hémodialyse dans le cas d'une épuration *via* une dérivation extracorporelle ou de la dialyse péritonéale lorsque le péritoine est utilisé comme membrane d'échange (INSERM, 2017). En France en 2014, l'hémodialyse constitue la principale méthode de dialyse (93,5 %) et 55,2 % des patients suivent leur dialyse en centre lourd contre 18,9 % en Unité de Dialyse Médicalisée (UDM) (Rouse, 2017). De plus, d'importantes disparités régionales existent que ce soit sur les stratégies de traitements qui sont mises en place pour

pallier l'insuffisance rénale (greffe, dialyse) ou sur l'accessibilité géographique aux centres de dialyse (EGR, 2013; Rousse, 2017).

D'autres facteurs sont également associés à la progression des MRC (*e.g.* sexe, ethnie, dyslipidémie, tabagisme) (KDIGO, 2013) et certaines causes de l'IRCT sont déjà bien identifiées (*e.g.* facteurs génétiques et/ou dégénératifs, facteurs socio-démographiques, exposition à des agents néphrotoxiques). Toutefois pour certains patients, l'étiologie demeure indéterminée (INSERM, 2017; KDIGO, 2013). Les causes d'IRCT pour ces patients pourraient être attribuées à l'environnement bien que les connaissances de la contribution environnementale sur l'IRCT soient encore parcellaires (Bowe *et al.*, 2019; Kamel and El-Minshawy, 2010). Les populations rurales et/ou agricoles semblent en effet plus affectées quand les facteurs de risques habituels ne sont pas en cause, avec une prévalence plus élevée (Almaguer *et al.*, 2014; Glaser *et al.*, 2016) ou une distribution des patients en zones rurales significative pour Kamel and El-Minshawy (2010). Les MRC sont également des pathologies impactées par le changement climatique, notamment en raison des températures plus élevées. Selon Glaser *et al.* (2016), la néphropathie due au stress thermique pourrait même représenter l'une des premières épidémies due au réchauffement climatique. L'identification des causes encore inconnues des MRC permettrait (i) d'identifier l'influence relative de ces facteurs de risque associés et (ii) d'ajuster le traitement (KDIGO, 2013).

Cet axe de la thèse fait l'objet d'une collaboration avec le registre Néphronor, qui est la déclinaison régionale du Réseau, Epidémiologie, Information, Néphrologie (REIN). Le registre existe depuis 2005 et recense l'ensemble des patients atteints d'IRCT. Il vient poursuivre les travaux valorisés dans la publication de Ocelli *et al.* (2014b) qui ont été menés à l'échelle des cantons du Nord-Pas de Calais. Après ajustement sur l'âge et le sexe, des disparités significatives de l'incidence de l'IRCT ont été soulignées au sein des cantons du NPC. Une association significative avait également été mise en évidence entre l'incidence de l'IRCT et la défaveur socio-économique approchée par l'indice de Townsend. Nos travaux visent à explorer à fine échelle l'influence des caractéristiques territoriales, qu'elles soient en lien avec la vulnérabilité et/ou la résilience, sur la part restante de la variabilité spatiale de l'incidence de l'IRCT qui ne serait pas expliquée par des paramètres connus (*e.g.* paramètres individuels, défaveur socio-économique). Ces travaux visent à (i) mettre en évidence si les caractéristiques territoriales sont identiques d'une zone de sur-incidence (ou de sous-incidence) à une autre et (ii) identifier des facteurs de risque de sur-incidence (ou de sous-incidence) par rapport à l'incidence moyenne constatée au sein du Nord-Pas de Calais.

Matériel et méthodes

Population d'étude

Le périmètre de recherche porte sur les départements du Nord et du Pas de Calais. La région comporte environ 4,2 millions d'habitants et couvre une superficie de 12 481 km² qui inclue des régions industrielles, urbaines mais également rurales. La région est affectée par de fortes disparités socio-économiques mais également environnementales. L'unité spatiale de la commune sur ce périmètre a une superficie moyenne de 8,08 km² avec une étendue de 0,18 à 97,61 km².

Entre le 1^{er} Janvier 2005 et le 31 Décembre 2018, 10 757 patients âgés de plus de 18 ans ont été inclus dans le registre à la suite d'un diagnostic d'IRCT. Ils ont été enregistrés à leur commune de résidence au moment de leur enregistrement. Différents critères d'exclusion ont été appliqués : ont été exclus les patients pédiatriques, les patients non-résidents dans les départements du Nord-Pas de Calais, les patients ayant un DFG supérieur à 20, les patients pris en charge pour insuffisance cardiaque

réfractaire et sans DFG, les patients pour lesquels les données insuffisance cardiaque réfractaire et DFG étaient manquantes. Les données pour les 9 339 patients répondant aux critères d'inclusion ont été agrégées à l'échelle communale selon la géographie de l'INSEE en vigueur en 2018 (*i.e.* 1 539 communes). Les patients ont été catégorisés par sexe et âge (intervalle de 5 ans, la dernière étant les plus de 95 ans). La population moyenne sur la période 2005-2018 a également été stratifiée par sexe et âge selon les mêmes définitions de classes comme que la population à risque.

Facteurs de confusion

Pour cette étude, des facteurs de confusion en lien avec l'accessibilité au système de soin, les pratiques de soin et les caractéristiques socio-économiques ont été considérés.

Accessibilité au système de soin

En France, il existe plusieurs types de centre de dialyse :

- Les centres de dialyse dits lourds : ils sont destinés aux patients hémodialysés dont la présence d'un néphrologue est requise pendant toute la durée de la séance ;
- Les Unités de Dialyse Médicalisées (UDM) : ces unités sont réservées aux patients qui ont été formés à l'hémodialyse. Dans ces UDM, le patient participe plus ou moins activement à la réalisation de son traitement en collaboration avec le personnel soignant sans que le néphrologue ne soit présent durant toute la séance. Le néphrologue effectue toutefois une visite au moins hebdomadaire ;
- Les Unités d'Auto-Dialyse : ces unités s'adressent à des patients formés à l'hémodialyse qui participent à la réalisation du traitement. De la même manière qu'en UDM, le néphrologue n'est pas présent durant toute la séance et effectue une visite au moins mensuelle. Ces UAD s'adressent à des patients dont l'état de santé est stable.

Pour le calcul du temps de trajet, seuls les centres de dialyse lourds ont été considérés. L'ensemble des centres de la région des Hauts-de-France (*i.e.* Aisne, Nord, Oise, Pas de Calais, Somme) ont été considérés pour les calculs des temps de trajet afin d'éviter les effets de bord (notamment ceux des départements de l'Aisne et de la Somme qui sont limitrophes du Nord et du Pas de Calais) (liste disponible en Annexe 3). Pour chaque commune des départements du Nord et du Pas de Calais, le temps de trajet par la route a été déterminé entre le centroïde des communes et le centre lourd de dialyse le plus proche *via* l'interface *OSRM Batch Routing* qui utilise les données d'*Open Street Map* pour déterminer les itinéraires.

Pratiques de soin

Selon les centres de dialyse, des différences de pratique peuvent être observées sur l'initiation à la dialyse. Les équipes de néphrologues initient plus ou moins tôt la dialyse, le DFG peut être dès lors plus ou moins élevé. Une moyenne du DFG par centre lourd a été réalisée à partir des données du registre. En effet, le code équipe renseigné dans le registre a permis de rattacher les patients au centre dans lequel ils ont initié le traitement de suppléance. Pour les 1 039 communes qui recensent des patients, la moyenne du DFG au centre lourd a été pondérée par la part de patients effectivement suivis dans ce centre lourd, puis été interpolée afin d'obtenir une valeur (eDFG) pour les 1 539 communes du Nord-Pas de Calais.

Défaveur socio-économique

Pour cette première approche, l'âge, le sexe et la défaveur socio-économique ont été considérés. Pour ajuster sur la défaveur socio-économique, le FDep a été utilisé (Rey *et al.*, 2009). Cet indice est habituellement utilisé en France pour refléter l'hétérogénéité spatiale socio-économique. Il est

construit à partir de quatre variables : le revenu médian des ménages, le pourcentage de bacheliers dans la population âgée de 15 ans et plus, le pourcentage d'ouvriers dans la population active et le taux de chômage. Plus la valeur du FDep est grande, plus le niveau de défaveur est élevé. Le FDep moyen sur l'ensemble de la période 2005-2018 a été utilisé pour ajuster sur l'incidence sur ce facteur de confusion.

Déterminants territoriaux et indices composites de résilience et de vulnérabilité

Un set de données a été développé afin de pouvoir caractériser le territoire du Nord-Pas de Calais. Pour cela, des données en lien avec des déterminants pour lesquels les liens sont avérés et suspectés avec l'IRCT et/ou les MRC ont été collectées et/ou calculées à partir de données disponibles en *open data*. Les indices composites de résilience et de vulnérabilité ont été développés selon la méthodologie qui a été présentée dans le Chapitre I. Une variable est considérée comme appartenant à la résilience lorsque son augmentation peut être reliée à un effet bénéfique tandis que si l'augmentation de la valeur de la variable est susceptible de porter préjudice à la santé, elle est considérée comme appartenant à la vulnérabilité. Deux ACP avec des rotations VARIMAX ont donc été réalisées pour pouvoir déterminer les indices composites de résilience et de vulnérabilité. Le tableau XV (page suivante) présente par indice et dimension une description des variables qui ont été intégrées aux indices de résilience et de vulnérabilité.

Tableau XV : Liste des données qui ont été retenues pour le développement des indices composites de résilience et de vulnérabilité

Indice	Dimension	Descriptif	Médiane [Q1-Q3]
Vulnérabilité	Environnement	Emissions des Hydrocarbures Aromatiques Polycycliques (HAP), le total des 8, moyenne sur 2004-2008-2012 (kg)	1,44 [0,82-2,64]
		Moyenne annuelle des PM ₁₀ sur l'ensemble de la période 2005-2018 (µg/m ³)	23,71 [23,23-24,34]
		Moyenne de l'indicateur SOMO ₃₅ sur l'ensemble de la période 2005-2018	2348,32 [2194,1-2509,4]
		Nombre de jours où la température maximale est supérieure à 25°C, moyenne annuelle sur la période 1981-2010	31,7 [26,7-34,5]
		Achats totaux de pesticides (toute substance) : moyenne de 2014 à 2018, rapportés à la Surface Agricole Utile (SAU) de 2010 (kg/km ² de SAU)	381,65 [244,08-550,81]
	Economie	Part des emplois précaires (CDD, emplois aidés, apprentissage, intérim), moyenne sur la période 2006-2018 (%)	11,68 [9,70-11,42]
		Part des emplois en agriculture, enregistrés au lieu de travail, moyenne sur la période 2006-2018 (%)	10,04 [2,13-23,98]
	Social	Part de la population âgée de 65 ans et plus vivant seule, moyenne sur la période 2006-2018 (%)	5,55 [4,53-6,83]
		Part de la population immigrée parmi la population immigrée parmi la population totale, moyenne sur la période 2006-2018 (%)	1,32 [0,79-2,40]
	Services	Nombre de débits de tabac pour 1 000 habitants âgés de 15 ans et plus dans la population totale (2018)	0,51 [0,36-0,70]
Résilience	Social	Part de la population âgée de 15 ans et plus vivant en couple, moyenne sur la période 2006-2018 (%)	67,38 [63,74-70,45]
		Part des licenciés sportifs dans la commune, moyenne sur les périodes 2011-2012 et 2014-2018 (%)	17,44 [14,09-21,41]
	Economie	Part des emplois dans la sphère présentielle, moyenne sur les années 2008, 2009, 2010, 2013, 2014, 2015, 2018 (%)	62,22 [45,37-75,02]
	Services	Part des résidences principales occupées par des propriétaires, moyenne sur la période 2006-2018 (%)	80,11 [71,81-85,23]
	Santé	Accessibilité Potentielle Localisée (APL) aux médecins généralistes, moyenne sur la période 2015-2018 (nombre de consultations ou visites de médecine générale par an et par habitant)	3,87 [3,17-4,68]
	Environnement	Part de la végétation dans la trame urbaine (2018) (%)	15,14 [11,57-18,65]
		<i>Normalized Difference Vegetation Index</i> (NDVI), moyenne sur la période 2014-2018 (sans unité)	0,56 [0,52-0,60]
		Part de la population ayant accès à un espace vert d'au moins 0,5 ha dans un rayon de 400 mètres (2014) (%)	93,54 [77,00-100]

Analyses statistiques

Ratios d'Incidence Standardisés (SIRs)

La distribution spatiale de l'incidence de l'IRCT a été évaluée grâce au calcul des SIRs pour chaque commune. Le SIR est défini par le ratio entre le nombre de cas observés d'IRCT et le nombre de cas

attendus après une standardisation indirecte sur l'âge et le sexe selon la méthodologie présentée dans les travaux de Waller and Gotway (2004).

Hétérogénéité spatiale et autocorrélation spatiale

L'hétérogénéité spatiale de l'incidence de l'IRCT a été évaluée par le test de Potthoff-Whittinghill (Potthoff and Whittinghill, 1966) tandis que l'autocorrélation spatiale a été quantifiée par l'indice de Moran (Moran, 1948) et évaluée par le test de Moran (Cliff and Ord, 1981). Une valeur positive de l'indice de Moran indique la présence d'une éventuelle et positive autocorrélation signifiant que des unités spatiales adjacentes tendent à présenter une incidence similaire d'IRCT.

Régressions écologiques

Pour analyser les relations entre l'incidence de l'IRCT et l'exposition environnementale, des modèles de Poisson log-linéaires bayésiens hiérarchiques ont été mis en utilisés. Le SIR étant un estimateur reconnu pour son instabilité, notamment dans le cas d'un faible nombre de cas attendus, les modèles de poisson hiérarchique log-linéaire ont permis d'inclure dans le modèle les effets aléatoires spatialement structurés et non structurés. Ces deux effets permettent de prendre en compte dans le même temps la possible autocorrélation spatiale et l'hétérogénéité spatiale des SIRs d'IRCT sur la période d'étude. Ces effets aléatoires sont décrits en utilisant le re-paramétrage de la distribution autorégressive conditionnelle de Besag-York-Mollié (Besag *et al.*, 1991; Simpson *et al.*, 2017). Cinq modèles ont été considérés :

- Modèle 1 : uniquement l'effet spatial ;
- Modèle 2 : les effets aléatoires spatiaux et les facteurs de confusion liés à la santé (eDFG, temps de trajet vers le centre lourd de dialyse) ;
- Modèle 3 : les effets aléatoires spatiaux, l'ensemble des facteurs de confusion (liés à la santé, FDep)
- Modèle 4 : les effets aléatoires spatiaux, l'ensemble des facteurs de confusion et l'indice de résilience ;
- Modèle 5 : les effets aléatoires spatiaux, l'ensemble des facteurs de confusion et l'indice de vulnérabilité

Il est à noter que tous les modèles sont ajustés sur l'âge et le sexe *via* le calcul des SIRs d'IRCT. Pour chaque co-variable, le Risque Relatif (RR) d'incidence d'IRCT et son intervalle de confiance à 95 % ont été calculés. Le pourcentage de variation spatiale du RR spatial d'IRCT expliqué par les co-variables a également été déterminé. Les différents modèles ont été ajustés à l'aide d'une approximation de Laplace emboîtée et intégrée (Blangiardo and Cameletti, 2015). Les détails sur les modèles de poisson hiérarchique log-linéaire sont disponibles dans l'Annexe 4.

Logiciels

Les analyses statistiques ont été menées avec le logiciel R-Studio (version 4.0.3, packages : R-INLA). Les seuils de significativité statistique ont été fixés à $p < 0,05$. Les cartes ont été réalisées à l'aide de Q-GIS (version 3.16.11-1).

Résultats

Incidence brute d'IRCT

Sur les départements du Nord-Pas de Calais entre 2005 et 2018, 9 339 patients ont eu un diagnostic d'IRCT (5 442 chez les hommes et 3 897 chez les femmes). Sur toute la période d'étude, la moyenne annuelle de l'incidence d'IRCT pour 100 000 habitants est de 16,45 sur la population totale [IC-95% : 15,85-17,04], 19,89 chez les hommes [18,00-21,77] et 13,20 [12,00-14,38] pour les femmes. La Figure 17 présente l'incidence brute de l'IRCT selon l'âge et le sexe et souligne l'importance de prendre en compte ces facteurs dans le calcul de l'incidence.

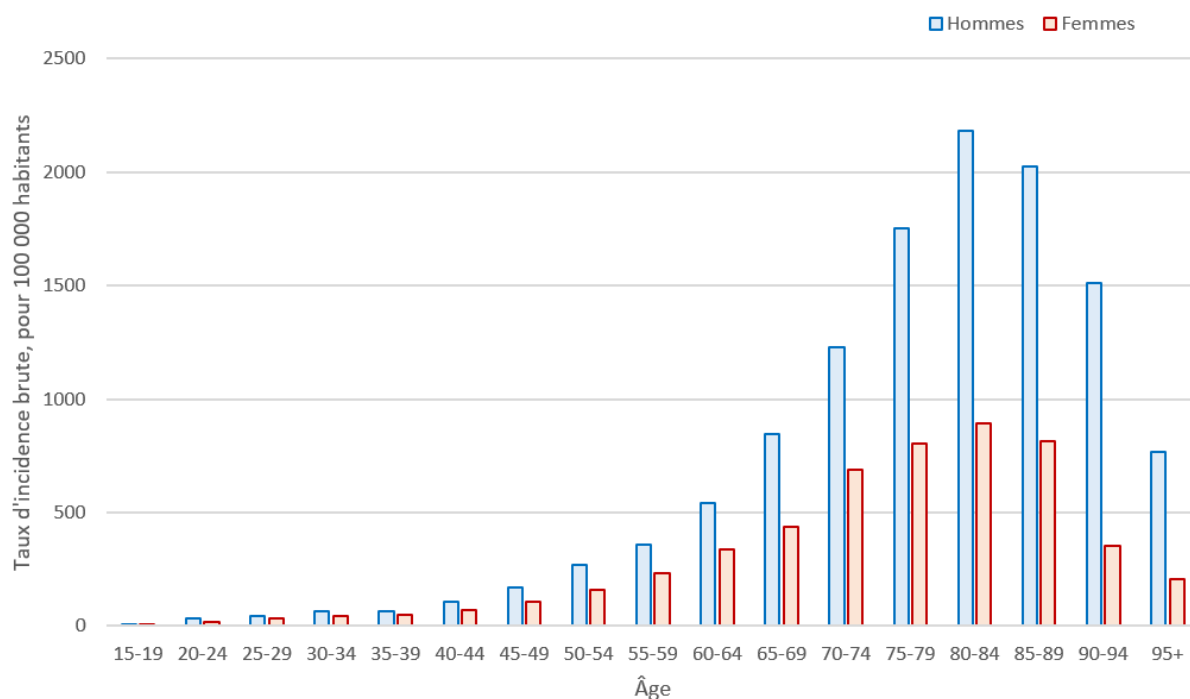


Figure 17 : Taux d'incidence brute d'IRCT sur 2005-2018, selon l'âge et le sexe

Facteurs de confusion

Temps de trajet vers le centre lourd de dialyse le plus proche

Sur les départements du Nord-Pas de Calais, la distance moyenne s'échelonne de 0,33 min à 51,47 min, la valeur médiane étant de 21,18 minutes et un intervalle interquartile de [13,37-26,55]. La distribution spatiale du temps de trajet moyen vers le centre lourd le plus proche est présentée dans la Figure 18 (page suivante).

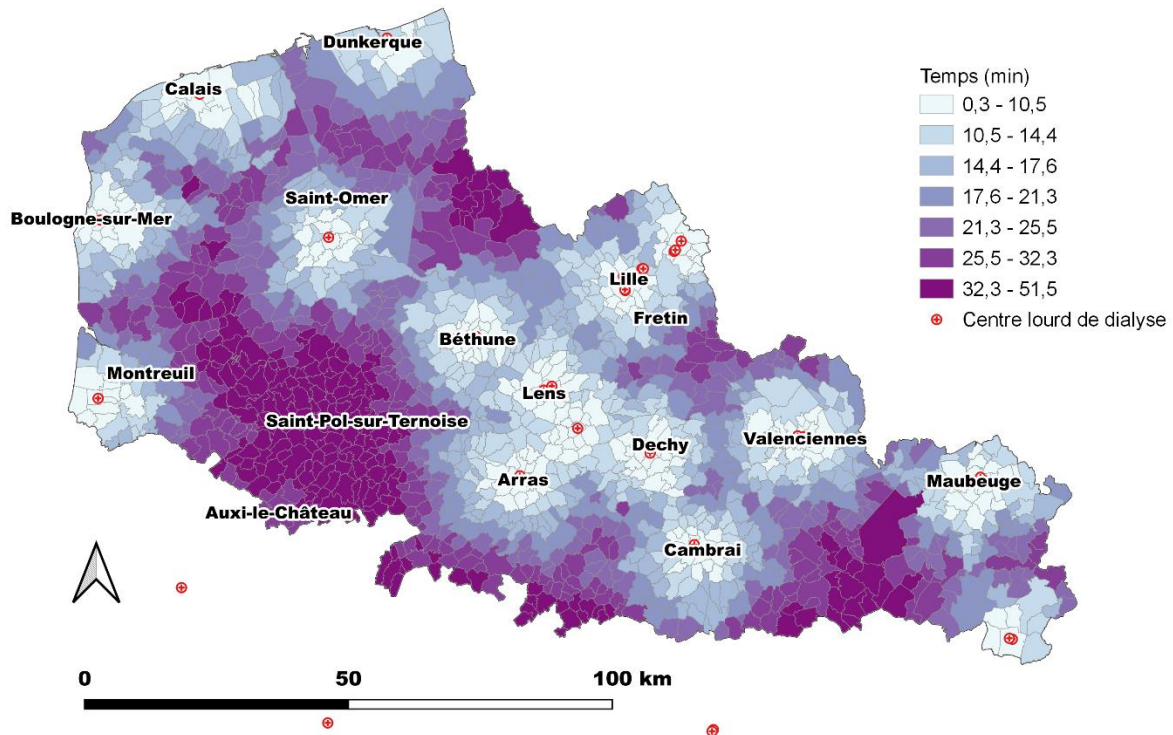


Figure 18 : Clusters d'IRCT obtenus après ajustement sur l'âge, le sexe et la défaveur socio-économique et le temps de trajet par la route en minutes vers le centre lourd le plus proche

La Figure 18 illustre la forte hétérogénéité spatiale du temps de trajet vers le centre lourd de dialyse. Les grandes agglomérations des départements présentent dans l'ensemble des temps de trajet plus courts. Toutefois, au sein du Pas de Calais et notamment autour de Saint-Pol-sur-Ternoise, de nombreuses communes présentent des temps de trajet supérieurs à 32 minutes.

Pratiques d'initiation de la dialyse

La valeur médiane du eDFG pour les 1 539 communes du Nord-Pas de Calais est de 9,62 avec un écart interquartile de [8,95-10,14] et s'étend de 8,23 à 11,35 mL/min/1,72m² (Figure 19).

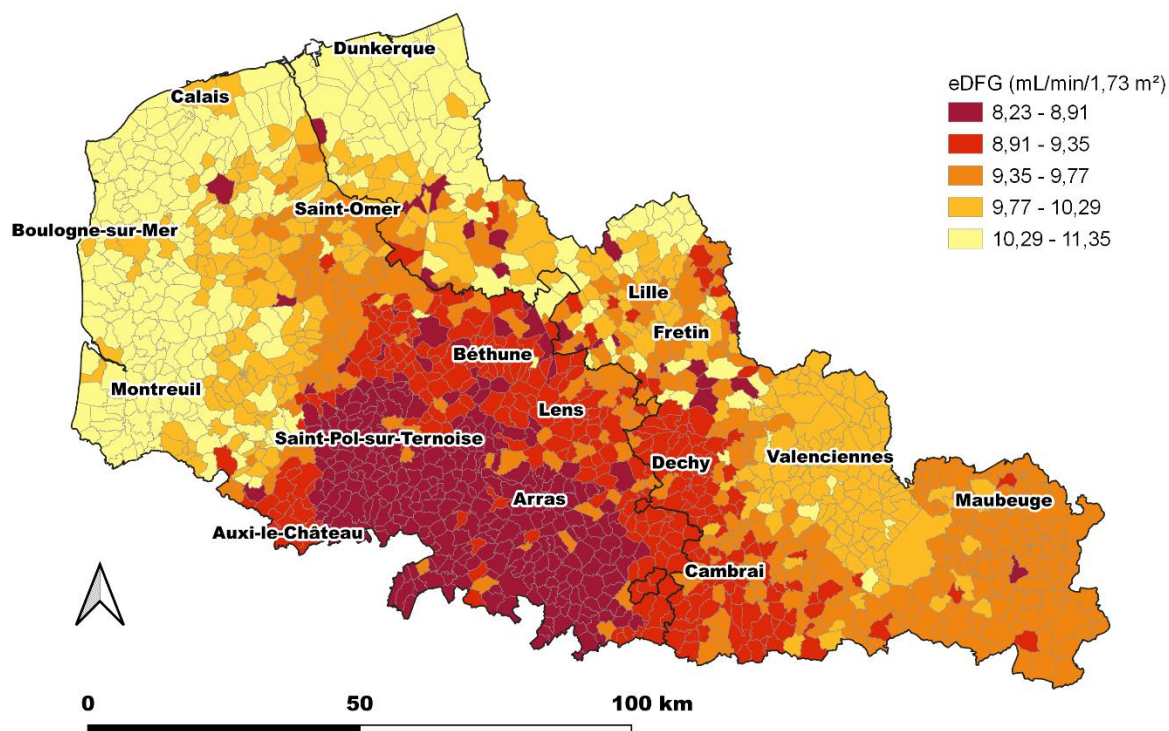


Figure 19 : Distribution spatiale de la moyenne pondérée du DFG à l'initiation des patients atteints d'IRCT (2005-2018)

De fortes disparités sur le niveau de DFG à l'initiation ont été constatées, soulignant les différences de pratique d'une équipe à une autre pour l'initiation de la suppléance de l'IRCT au sein des différentes communes du Nord-Pas de Calais. Le traitement de suppléance par la dialyse pour les patients résidant dans les communes au sud du Pas de Calais (entre Saint-Pol-sur-Ternoise et Arras) est initié le plus tardivement (eDFG bas à très bas) tandis qu'il est initié à des niveaux de eDFG plus élevés pour les patients résidents dans les communes situées sur la côte (de Dunkerque à Montreuil) et au Nord de la Métropole Lilloise. La variabilité résiduelle de l'IRCT ajustée sur le eDFG permet d'écarter une sous-incidence ou une sur-incidence « artificielle », et qui correspondrait plutôt à un « sous- ou à un sur-enregistrement » qu'à une réelle sur- ou sous-incidence.

Défaveur socio-économique (FDep)

Sur les départements du Nord-Pas de Calais, la moyenne du FDep pour la période 2005-2018 s'échelonne de -5,02 à 4,86 et présente une valeur médiane de 0,39 avec un intervalle interquartile de [-0,61 ;1,29]. La Figure 20 présente la distribution spatiale de cet indicateur de défaveur socio-économique sur les départements du Nord-Pas de Calais.

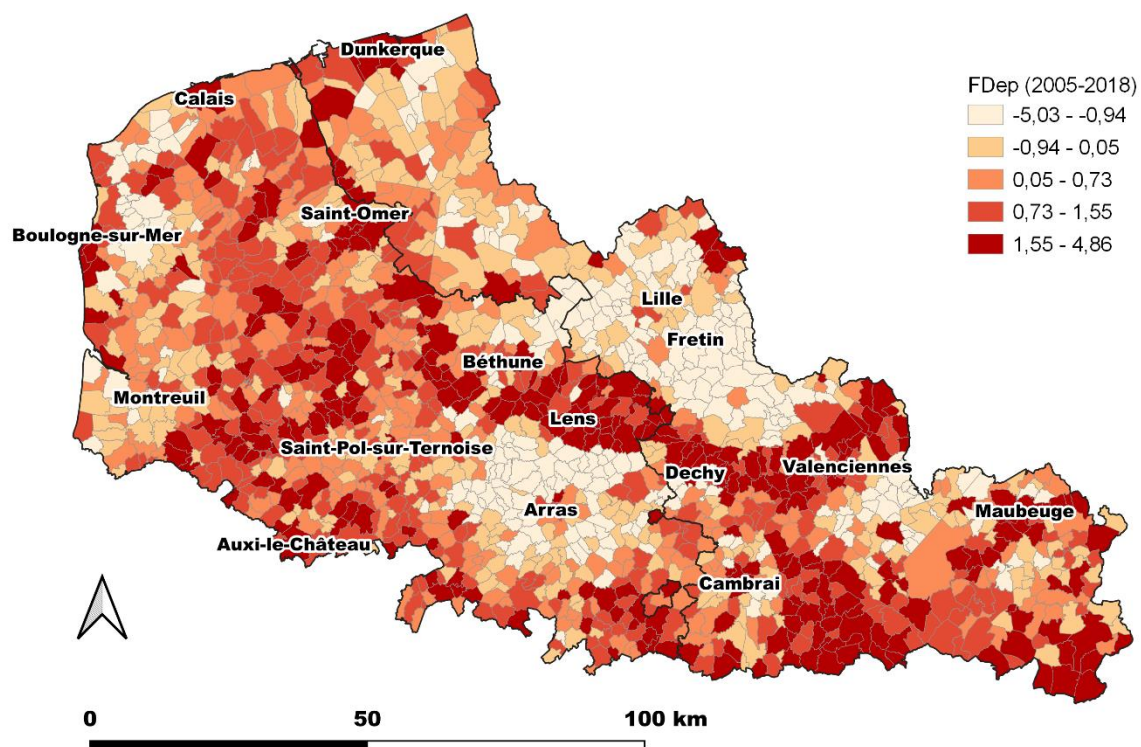


Figure 20 : Distribution spatiale du FDep sur le Nord-Pas de Calais, moyenne sur la période couverte par le registre (2005-2018)

Une grande hétérogénéité spatiale peut être observée au sein de ces deux départements. La principale zone de défaveur socio-économique se situe au niveau de l'ancien bassin minier (de Béthune à Valenciennes). Les communes qui présentent des populations non défavorisées sont plutôt situées au Nord et au Sud de l'ancien bassin minier mais également autour de Boulogne-sur-Mer et Montreuil. Une association significative a été mise en évidence entre le niveau de défaveur socio-économique et l'incidence de l'IRCT ajustée sur l'âge et le sexe (Tableau XVI).

Tableau XVI : Influence du FDep sur le RR de l'IRCT sur 2005-2018

FDep	RR	IC à 95 %	
		Borne inf.	Borne sup.
Quartile 1 (référence)	1		
Quartile 2	1,239	1,136	1,350
Quartile 3	1,334	1,216	1,463
Quartile 4	1,510	1,391	1,640

Comme le montre le Tableau XVI, plus les populations du Nord-Pas de Calais sont défavorisées, plus le RR d'IRCT augmente. Cette relation significative entre le FDep et l'incidence d'IRCT ajustée sur l'âge et le sexe a notamment été observée sur les zones de Dunkerque, Roubaix et Maubeuge.

Indices composites de résilience et de vulnérabilité

Indice de résilience

L'indice de résilience a agrégé huit variables et trois composantes principales ont été retenues pour son calcul. La première composante était portée par la part de la population vivant en couple ainsi que la part de propriétaires, ces variables ayant été déterminées sur la période de 2006 à 2018. La deuxième composante de l'ACP était principalement portée par l'accessibilité potentielle localisée

(déterminée à partir des données de 2015 à 2018), par la part de licenciés sportifs et dans une moindre mesure par la part des activités présentielle déterminées entre 2006 et 2018. Les déterminants en lien avec la résilience environnementale (*i.e.* part de la végétation dans la trame urbaine, NDVI, accessibilité aux espaces verts) ont quant à elles contribué majoritairement à la troisième composante de l'ACP. La valeur médiane de l'indice de résilience est de 0,39 avec un intervalle interquartile de [0,49-0,61]. La Figure 21 en présente la distribution spatiale sur le Nord-Pas de Calais.

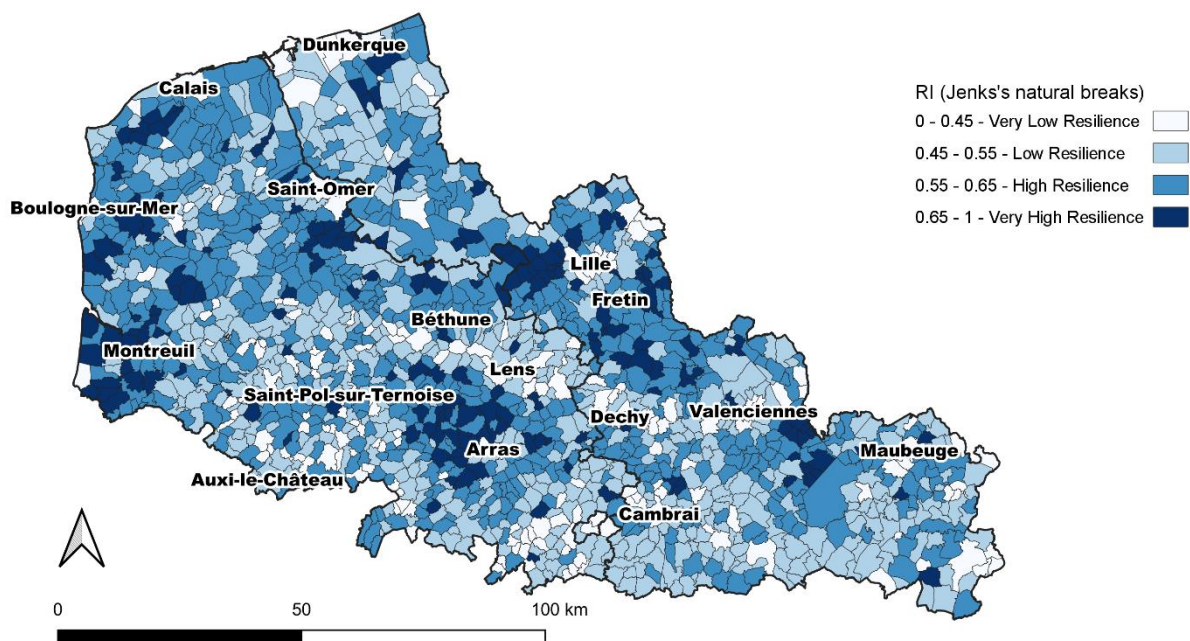


Figure 21 : Distribution spatiale de l'indice de résilience sur les 1 539 communes du Nord-Pas de Calais

Sur les départements du Nord-Pas de Calais, les communes présentant une plus faible résilience vis-à-vis de l'IRCT sont principalement localisées au niveau de l'ancien bassin minier (de Béthune à Valenciennes) et également dans le Sud du Nord-Pas de Calais. Les communes qui présentent une résilience élevée sont principalement localisées à l'Est de la MEL, entre Arras et Saint-Pol-sur-Ternoise mais également sur une large zone qui s'étend du Nord du bassin minier jusque Montreuil et Boulogne-sur-Mer et Calais.

Indice de vulnérabilité

L'indice de vulnérabilité a agrégé 10 variables et quatre composantes principales ont été retenues pour son calcul. La première composante était principalement portée par la concentration moyenne annuelle en PM₁₀ sur 2005-2018 et par la part des emplois dans l'agriculture entre 2006 et 2018. Le SOMO₃₅ et le nombre de jours où la température dépassait les 25°C sur 1981-2010 contribuaient principalement à la deuxième composante, avec une moindre mesure les émissions des 8 HAPs. La troisième composante était portée par la part de population précaire entre 2006 et 2018 et la part de ménage seul composé d'une personne de plus de 65 ans. La quatrième composante était portée quant à elle par les achats de pesticides, la part de population immigrée et le nombre de débit de tabac pour habitants. La valeur médiane de l'indice de vulnérabilité est de 0,34 avec un intervalle interquartile de [0,27-0,40]. La Figure 22 en présente la distribution spatiale.

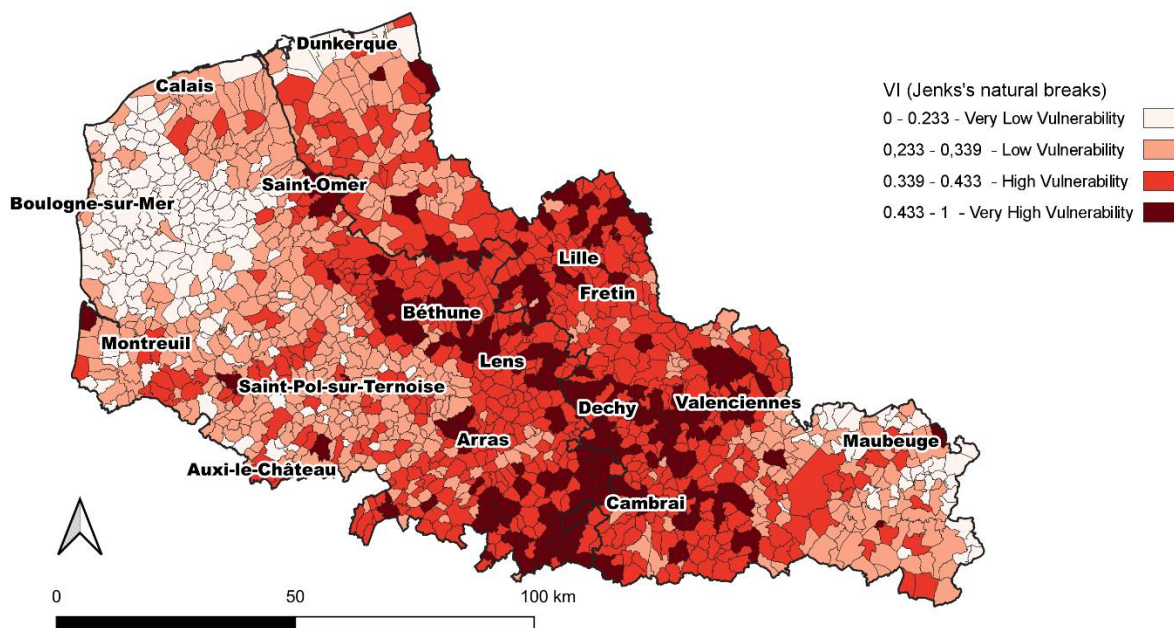


Figure 22 : Distribution spatiale de l'indice de vulnérabilité sur les 1 539 communes du Nord-Pas de Calais

Sur les départements du Nord-Pas de Calais, les communes présentant une plus faible vulnérabilité sont principalement localisées dans le Sud-Ouest de la région, au niveau de la Communauté Urbaine de Dunkerque et dans l'Avesnois (à l'Est de la région). Les communes qui présentent une vulnérabilité élevée sont principalement localisées au niveau de la MEL, de l'ancien bassin minier mais également selon un axe allant du Sud d'Arras vers Valenciennes. Le coefficient de corrélation de Spearman s'est avéré modéré entre les indices de résilience et de vulnérabilité qui ont été développés pour analyser leur influence sur l'IRCT ($\rho = -0,199$, $p < 0,001$) ce qui souligne qu'ils apportent des informations complémentaires.

Variations géographiques de l'incidence d'IRCT

Modèle 1

L'analyse des taux d'incidence d'IRCT ajustés uniquement sur l'âge et le sexe a montré la présence d'une hétérogénéité spatiale significative. Le test de Potthoff-Whittinghill présentait un $p < 0,05$ ainsi que des preuves d'autocorrélation spatiale selon les localisations (I de Moran = 0,04 avec un $p < 0,05$). La distribution spatiale du RR d'IRCT ajusté sur les seuls âge et sexe (modèle 1) sur le Nord-Pas de Calais est présenté dans la Figure 23 (page suivante). Il varie de 0,59 [IC : 0,38-0,88] à 1,95 [IC : 1,35-2,86].

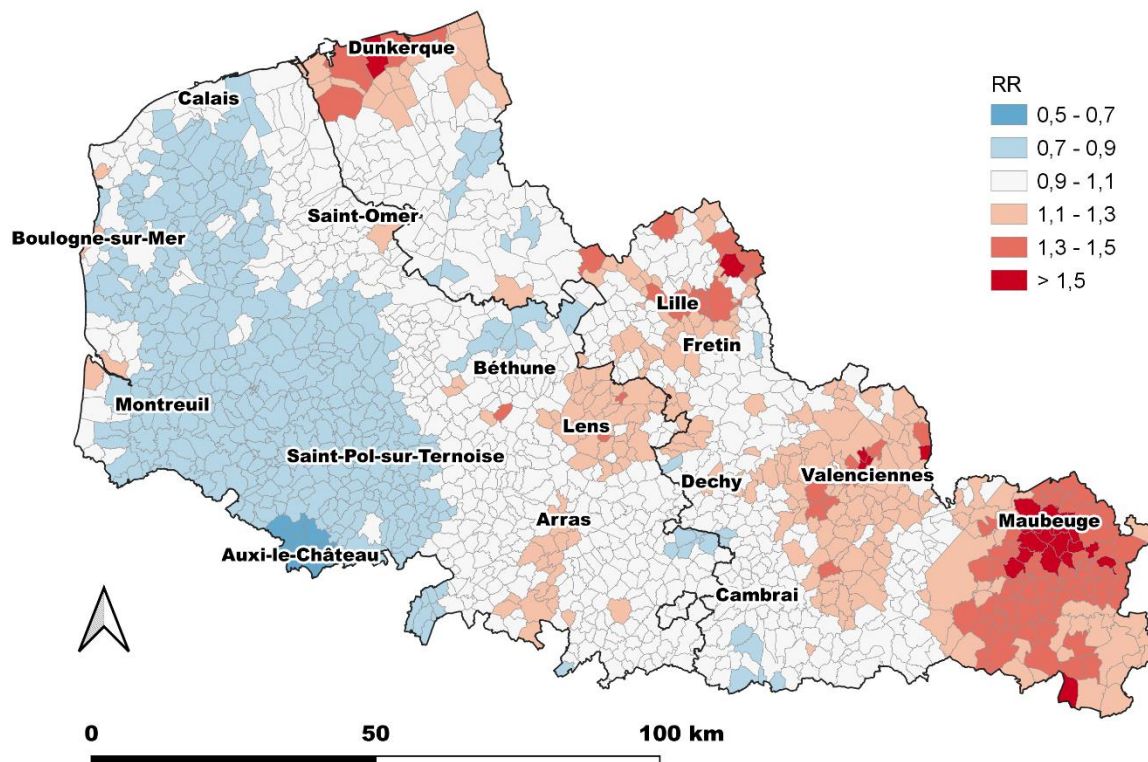


Figure 23 : RR d'IRCT après ajustement sur l'âge et le sexe sur 2005-2018 (modèle 1)

Trois zones de sur-incidence se distinguent et sont principalement localisés (i) au Sud-Est autour de Maubeuge (RR de Maubeuge = 1,89 [IC : 1,60-2,23]), (ii) sur l'agglomération lilloise (RR de Roubaix = 1,81 [IC : 1,59-2,04] et (iii) sur la CUD (RR de Dunkerque = 1,38 [IC : 1,22-1,55]). Des RR supérieurs à 1 et exprimant une sur-incidence sont également localisés de manière étendue autour de Lens et de Valenciennes. La Figure 23 présente également une zone étendue de sous incidence dans le Sud-Ouest (RR de Auxi-le-Château = 0,59 [IC : 0,38-0,88]).

Modèle 2

Après ajustement sur les facteurs de confusion liés à la santé (modèle 2), le RR spatial d'IRCT s'échelonne de 0,72 [IC : 0,48-1,03] à 1,71 [IC : 1,21-2,46] (Figure 24). Les facteurs de confusion liés à la santé permettent d'expliquer 26 % des variations spatiales observées pour le RR d'IRCT. Cette variation spatiale est devenue moins prononcée notamment sur certaines zones géographiques, *i.e.* la CUD et l'ancien bassin minier.

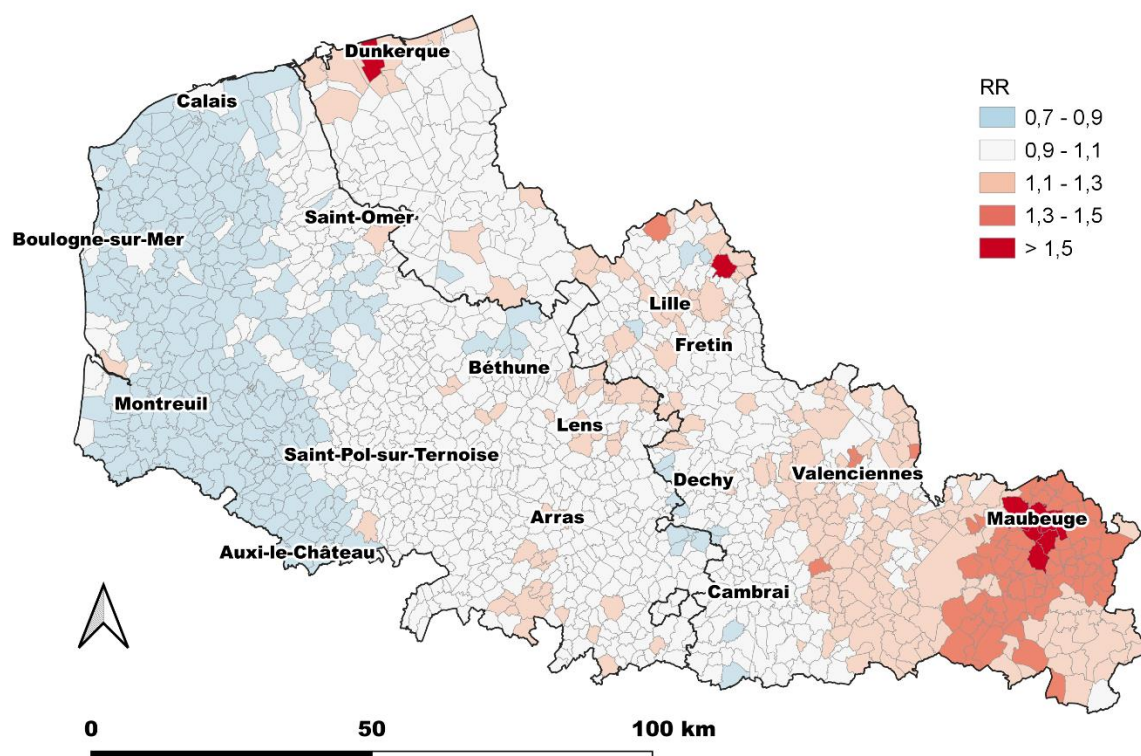


Figure 24 : RR d'IRCT après ajustement sur l'âge et le sexe et les facteurs de confusion liés à la santé sur 2005-2018 (modèle 2)

Par rapport à la Figure 24 qui présentait la distribution spatiale du FDep sur 2005-2018, des zones de défaveur socio-économique se superposent à des zones où les RR d'IRCT obtenus après ajustement sur les facteurs de confusion liés à la santé sont plus élevés.

Modèle 3

Le modèle 3 souligne des effets significatifs entre l'incidence de l'IRCT et l'ajustement sur les facteurs de confusion liés à la santé et la défaveur socio-économique. Pour chaque augmentation d'une unité du FDep, une augmentation de 15,1 % [IC : 12,2-18,2] de l'incidence de l'IRCT a été observée. Après prise en compte du FDep, le RR spatial d'IRCT varie de 0,74 [IC : 0,53-0,98] à 1,44 [IC : 1,12-1,89] (Figure 25). La prise en compte de ces facteurs de confusion (modèle 3) permet d'expliquer 55 % de la variation spatiale observée au sein des RR spatiaux d'IRCT. Cela signifie également que le FDep permet d'expliquer 40 %⁵ de la variation spatiale de l'incidence d'IRCT après qu'elle ait été ajustée sur les facteurs de confusion liés à la santé. La variation spatiale du RR d'IRCT ainsi obtenue est venue renforcer la structure de deux zones en sur-incidence (Figure 25).

⁵ Le gain du FDep a été déterminé de la manière suivante :

$$\frac{\text{Part expliquée par le modèle 3} - \text{part expliquée par le modèle 2}}{\text{Total de la variabilité} - \text{part expliquée par le modèle 2}} = \frac{55\% - 26\%}{100\% - 26\%}$$

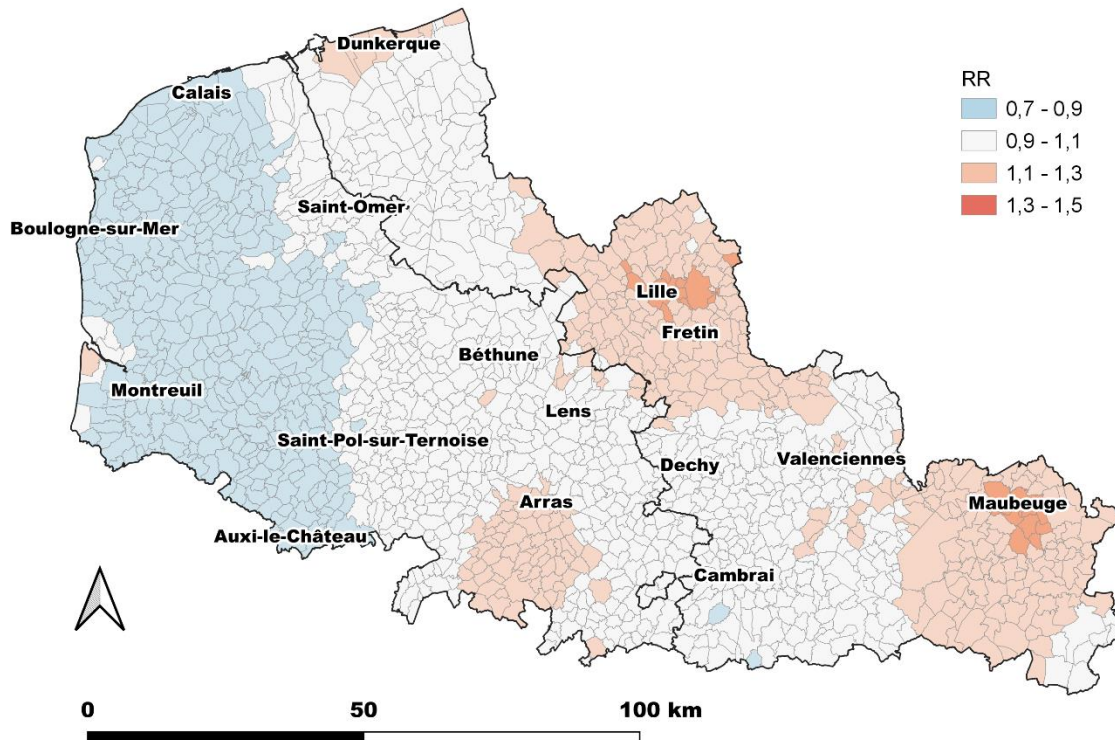


Figure 25 : RR d'IRCT après ajustement sur l'âge, le sexe, les facteurs de confusion liés à la santé et le FDep sur 2005-2018 (modèle 3)

Avec le modèle 3 (ajustement sur l'ensemble des facteurs de confusion considérés), les zones de sur-incidence sont principalement localisés autour de Maubeuge et de l'Avesnois, mais également au niveau de la MEL et s'étend jusqu'à la limite de l'ancien bassin minier. Les communes situées au Sud d'Arras présentent également des RR supérieurs à 1, soulignant une possible sur-incidence dans ces communes. Toute la partie Ouest de la région présente un RR compris entre 0,7 et 0,9, ce qui met en évidence que ces communes présentent toujours une sous-incidence en dépit d'une distance plus grande au centre lourd le plus proche et d'un eDFG plus faible.

Association de l'incidence de l'IRCT avec les indices de résilience (modèle 4) et de vulnérabilité (modèle 5)

Pour chaque augmentation d'une unité de l'indice de résilience (standardisé, $sd = 0,095$), une diminution de l'incidence de l'IRCT de 6,7 % [CI : 10,3,7] est observée, et ceci après ajustement sur tous les facteurs de confusion (modèle 4). La prise en compte de l'indice de résilience en plus de ces facteurs de confusion a permis d'expliquer 67 % de la variation totale du RR de l'IRCT, ce qui confère à l'indice de résilience un gain de presque 27 % par rapport à la part de la variabilité spatiale expliquée avec le modèle 3. En ce qui concerne l'indice de vulnérabilité (modèle 5), aucune association significative n'a pu être mise en évidence. La variabilité spatiale résiduelle obtenue après ajustement sur le modèle 4 est présentée dans la Figure 26.

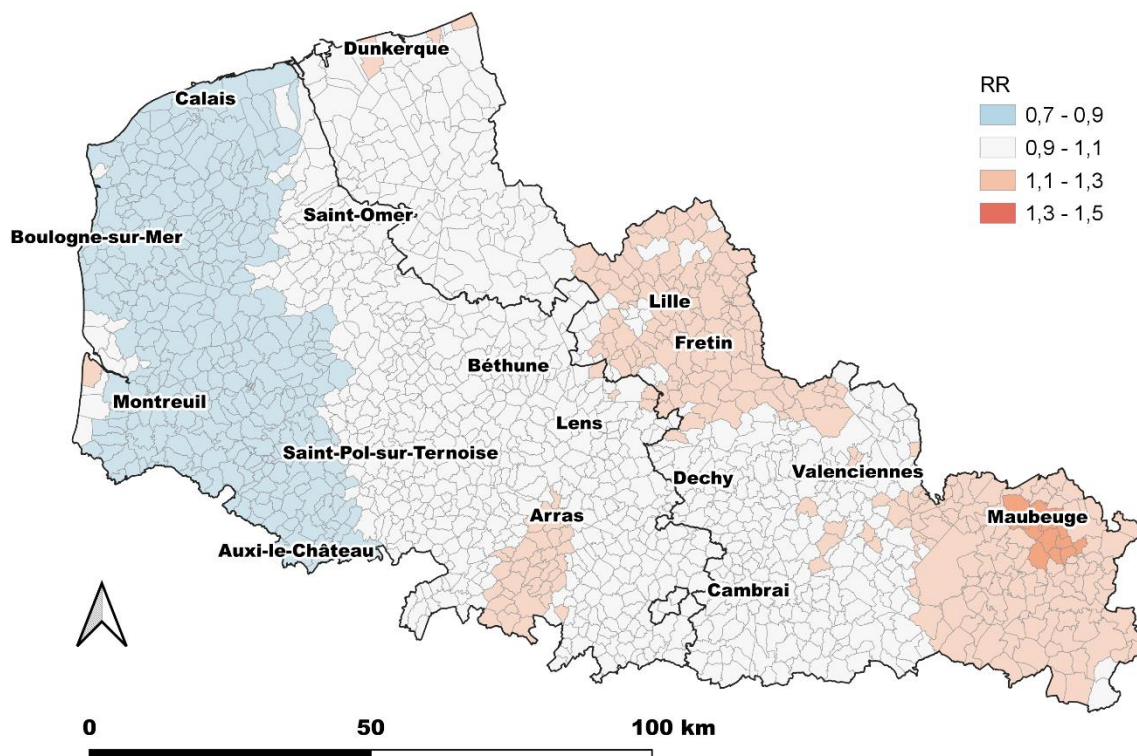


Figure 26 : RR d'IRCT après ajustement sur l'âge, le sexe, les facteurs de confusion liés à la santé, le FDep et l'indice de résilience sur 2005-2018 (modèle 4)

Comme le montre la Figure 26, les communes présentant un RR compris entre 0,7 et 0,9 et qui correspondent à une sous-incidence prenant en compte l'âge, le sexe, les facteurs de confusion liés à la santé (*i.e.* éloignement au centre lourd de dialyse, eDFG), le FDep et l'indice de résilience sont situés dans le département du Pas de Calais, de Calais à Auxi-le-Château. Les RR les plus élevés sont quant à eux retrouvés autour de Maubeuge. Les RR compris entre 1,1 et 1,3 sont principalement situés dans le département du Nord, sur une zone étendue autour de Maubeuge et la quasi-totalité de l'Avesnois, au niveau de la MEL jusqu'au Valenciennois et qui s'étend jusqu'au Nord de l'ancien bassin minier. Sur le département du Pas de Calais, le Sud d'Arras présente des communes en sur-incidence ainsi que de manière ponctuelle une commune à l'Ouest de Montreuil.

L'ensemble des analyses qui ont été réalisées sur les différents modèles considérés sont rassemblés dans le Tableau XVII.

Tableau XVII : RR, intervalle de confiance à 95 % et part de la variation spatiale observée expliquée par les covariables pour les différents modèles

	Modèle 1		Modèle 2		Modèle 3		Modèle 4		Modèle 5	
	RR	95% IC	RR	95% IC	RR	95% IC	RR	95% IC	RR	95% IC
Intercept	0,859	[0,825-0,894]	0,840	[0,807-0,874]	0,798	[0,766-0,830]	0,806	[0,774- 0,837]	0,795	[0,763-0,828]
eDFG	-	-	1,027	[0,979-1,078]	1,050	[1,003-1,098]	1,058	[1,013- 1,104]	1,050	[1,003-1,099]
Temps de trajet	-	-	0,906	[0,868-0,945]	0,927	[0,892-0,964]	0,943	[0,908- 0,979]	0,928	[0,893-0,965]
FDep	-	-	-	-	1,151	[1,122-1,182]	1,091	[1,054- 1,131]	1,149	[1,119-1,181]
Indice de résilience	-	-	-	-	-	-	0,931	[0,900- 0,963]	-	-
Indice de vulnérabilité	-	-	-	-	-	-	-	-	1,009	[0,976-1,044]
Part de la variabilité spatiale observée qui est expliquée par les covariables			26 %		55 %		67 %		55 %	

Association de l'incidence de l'IRCT avec les variables de l'indice de résilience

Afin de comprendre plus avant les mécanismes sous-jacents de relation entre l'indice de résilience et l'incidence de l'IRCT ajustée sur l'ensemble des facteurs de confusion, des analyses complémentaires ont été menées. Ces résultats sont présentés dans le Tableau XVIII.

Tableau XVIII : RR et IC de l'incidence d'IRCT ajustée sur l'ensemble des facteurs de confusion de l'ensemble des variables intégrées à l'indice de résilience

Variable	RR	IC 95%
Part de la population vivant en couple	0,951	[0,924 - 0,980]
Part des emplois dans la sphère présentielle	0,999	[0,965 - 1,034]
Part des résidences principales occupées par des propriétaires	0,948	[0,920 - 0,976]
Accessibilité potentielle localisée aux Médecins généralistes	0,986	[0,957 - 1,016]
Part de licenciés sportifs	1,013	[0,969 - 1,060]
Part de la végétation dans la trame urbaine	0,953	[0,914 - 0,994]
NDVI	0,947	[0,919 - 0,975]
Accessibilité à des espaces verts d'au moins 0,5 ha dans un rayon de 400 m	0,986	[0,953 - 1,020]

La part des emplois dans la sphère présentielle, la part de licenciés sportifs ainsi que l'accessibilité à des espaces verts d'au moins 0,5 hectare dans un rayon de 400 mètres n'ont pas montré de résultats significatifs. Par contre, les RR pour les cinq autres variables mettent en évidence une influence significative sur l'incidence résiduelle de l'IRCT. Les associations inverses les plus importantes concernent le NDVI, un indicateur de végétation, la part des résidences principales occupées par des propriétaires et la part de population vivant en couple.

Discussion

Les travaux visaient à mettre en évidence si les caractéristiques territoriales étaient identiques d'une zone de sur-incidence (ou de sous-incidence) à une autre par rapport à la variabilité résiduelle de l'incidence obtenue après ajustement sur des facteurs de confusion.

Les facteurs de confusion en lien avec la santé

Au regard des différences de pratiques qui peuvent être observées sur le territoire, l'éloignement au centre lourd de dialyse a fait l'objet d'une analyse exploratoire afin d'identifier s'il existait une relation entre ce paramètre et l'incidence de l'IRCT. En France, la dialyse est classiquement initiée en centre lourd, choisi par le médecin généraliste et/ou le patient, avant d'être potentiellement poursuivie dans un centre plus proche du domicile des patients. Les médecins généralistes de certains patients selon (i) l'avancement et le pronostic de l'IRCT et/ou (ii) l'éloignement du centre lourd de dialyse et le profil des patients (âge, autonomie) peuvent cependant choisir de ne pas les orienter vers la dialyse pour leur bien-être de fin de vie. La dialyse entraîne en effet des contraintes qui impactent sévèrement la qualité de vie, les patients traités en hémodialyse suivant généralement trois séances de quatre heures par semaine (HAS, 2010b). Une plus grande distance entre le centre de dialyse et le domicile du patient est alors susceptible d'impacter les parcours de soin des patients et l'incidence standardisée sur l'âge, le sexe et la défaveur socio-économique. Un point d'attention est à souligner pour la comparaison des temps de trajet moyens avec des résultats nationaux et internationaux. Le Nord-Pas de Calais présente

en effet des temps de trajet moyens vers les centres de dialyse (lourds et ambulatoires) inférieurs à ceux qui peuvent être généralement observés en France (*i.e.* 17,5 minutes vs 23,5 minutes) bien que des disparités au sein de la région existent (ABM, 2015). L'éloignement au centre lourd de dialyse ne reflète toutefois pas l'offre réelle de dialyse du territoire car elle n'inclut pas les centres satellites ambulatoires.

La région présente également des disparités géographiques sur le DFG à l'initiation, le Nord-Pas de Calais étant de plus la région qui initie le plus tôt la dialyse. Il est important de souligner qu'il n'existe pas une relation linéaire entre le DFG et le déclin de la fonction rénale : de faibles variations du DFG peuvent engendrer des déclins importants. Ces différences de pratiques en lien avec les équipes rattachées aux centres lourds engendrent de fait des DFG moyens à l'initiation plus élevés et donc une incidence artificiellement plus élevée pour les centres lourds qui initient plus tôt la dialyse. La prise en compte du eDFG dans les modèles a permis de considérer une sous-incidence ou une sur-incidence en lien avec des différences de pratiques dans les prises en charge de l'IRCT. Les ajustements sur les facteurs de confusion en lien avec la santé ont permis de confirmer que ces communes présentent une réelle sous-incidence et que les RR plus faibles ne peuvent pas être considérés comme du « sous-enregistrement ».

Les facteurs de confusion en lien avec la défaveur socio-économique

Sur les départements du Nord-Pas de Calais il existe une importante variabilité de la défaveur socio-économique, celle-ci pouvant avoir un impact fort sur l'incidence des MRC (Bilgel, 2019; Bowe *et al.*, 2017; Couchoud *et al.*, 2012; Nicholas *et al.*, 2015; Ward, 2008). La défaveur socio-économique constitue une importante voie de développement de facteurs de risque de la survenue ou de la progression de l'IRCT (Nicholas *et al.*, 2015). Gutiérrez (2015) a de plus mis en évidence une importante relation entre les facteurs socio-économiques et la qualité de l'alimentation. Le diabète, l'obésité ou encore l'hypertension apparaissant au sein de populations défavorisées seraient induits par exemple par de moins bonnes conditions de vie, une accessibilité à des soins de qualité moindre et des connaissances plus limitées en matière de santé (Nicholas *et al.*, 2015). Lors des travaux menés sur la période 2005-2011 à l'échelle des cantons du Nord-Pas de Calais, une relation positive et significative avait été mise en évidence entre la défaveur socio-économique retranscrite à l'aide de l'indice de Townsend et l'incidence standardisée sur l'âge et le sexe (Occelli *et al.*, 2014b). Cette approche écologique avait été mise en place afin de pallier l'absence de données sur le niveau socio-économique des patients du registre. Pour la période 2005-2018, le FDep a été privilégié pour approcher le niveau socio-économique en raison de sa meilleure capacité à retranscrire la défaveur socio-économique sur différents types de territoires (urbain vs rural). De la même manière qu'avec l'indice de Townsend à l'échelle des cantons sur 2005-2011, une relation positive et significative entre le FDep et l'incidence de l'IRCT avait été mise en évidence à l'échelle des communes pour la période 2005-2018.

Toutefois, un des principaux résultats attendus en termes d'incidence de l'IRCT n'a pas été constaté : l'ancien bassin minier qui s'étend de Bruay-la-Buissière jusqu'à Valenciennes n'est pas identifié comme étant en sur-incidence alors que de nombreuses fragilités socio-économiques y sont constatées (*e.g.* taux de pauvreté de 22 % en 2015, scolarité courte, accès difficile à un logement décent) (Marot, 2019). Une des hypothèses émises permettant d'expliquer cette absence de sur-incidence au sein du bassin minier est que les patients qui auraient été susceptibles de déclarer une MRC évoluant vers l'IRCT sont probablement décédés avant l'apparition et/ou l'aggravation des premiers. En effet, la mortalité y est supérieure de 27 % à la mortalité nationale, la mortalité pour les hommes dépassant même de 30 % le niveau national (ARS, 2017).

Indices composites et variabilité résiduelle de l'IRCT

Les zones géographiques qui ont été identifiées tiennent de ce fait compte des facteurs confondants suivants : l'âge, le sexe, le niveau de défaveur socio-économique, l'éloignement au centre lourd de dialyse le plus proche et le eDFG. L'incidence résiduelle ainsi obtenue a été mise en relation avec les indices composites de résilience et de vulnérabilité qui ont été développés en s'appuyant sur (i) la méthodologie développée dans le Chapitre I et (ii) sur la base de leurs liens avérés ou suspectés avec les MRC. Au regard de la plage temporelle relativement longue couverte par le registre et du type de données de santé utilisé (*i.e.* donnée en lien avec l'incidence moyenne sur 2005-2018), les variables qui ont été intégrées aux indices, en fonction de leur disponibilité, ont été calculées sur une période se rapprochant le plus de 2005-2018 pour obtenir une adéquation entre les données de santé et les données territoriales.

L'indice de vulnérabilité

Le RR de l'indice de vulnérabilité n'a montré aucune association significative avec l'incidence résiduelle de l'IRCT (RR = 1,009, [IC : 0,976-1,044]), l'ajout de l'indice dans le modèle 5 n'a pas contribué à augmenter la part expliquée de la variabilité spatiale observée de l'IRCT. Bien que la méthodologie d'analyse de l'association entre les déterminants en lien avec la vulnérabilité et l'IRCT soit différente, la littérature avait mis en évidence des associations entre des variables intégrées à l'indice de vulnérabilité et les MRC et/ou l'IRCT.

Au sein des départements du Nord-Pas de Calais, les concentrations moyennes annuelles pour les PM₁₀ mesurées sur 2005-2018 qui ont été intégrées dans l'indice de vulnérabilité présentaient des coefficients de corrélation de Spearman élevés avec d'autres polluants atmosphériques classiquement mentionnés dans la littérature, *i.e.* 0,74 avec les concentrations de NO₂ et 0,79 avec les concentrations de PM_{2,5}. Les concentrations moyennes annuelles de PM₁₀ sur la période 2005-2018 étaient de 23,8 µg/m³ tandis que les PM_{2,5} présentaient une moyenne de 15,0 µg/m³ sur la période 2009-2018. En France, les PM_{2,5} ne sont mesurées de manière exhaustive que depuis 2009, les moyennes annuelles de PM₁₀ ont donc été privilégiées pour obtenir une donnée en lien avec la pollution atmosphérique couvrant toute la période du registre (2005-2018). Les moyennes annuelles observées sur les communes du Nord-Pas de Calais présentent des résultats supérieurs (i) aux recommandations de l'OMS⁶ et seraient donc susceptibles d'avoir des impacts délétères sur la santé des populations et (ii) aux concentrations mentionnées dans la littérature pour lesquelles des associations entre pollution atmosphérique et incidence de l'IRCT avaient été mises en évidence. Dans leur étude sur le lien entre proximité résidentielle aux grands axes routiers et déclin de la fonction rénale, Lue *et al.* (2013) ont mis en évidence après ajustement sur différents facteurs (défaveur socio-économique du quartier de résidence) un déclin exponentiel de la fonction rénale avec la proximité à certains axes routiers. L'exposition au NO₂ et aux PM_{2,5} a été positivement associée à des groupes pour lesquels la fonction rénale était inférieure comparé à ce qui pouvait être observé au sein d'autres groupes, les résultats pour les PM_{2,5} n'étant toutefois pas significatifs (Cai *et al.*, 2019). Des excès de risque de MRC et d'IRCT ont également été mis en évidence pour des populations exposées à des concentrations élevées de polluants atmosphériques (*i.e.* NO, NO_x, SO₂ et PM_{2,5}), l'exposition pour les PM_{2,5} ayant été considérée comme élevée lorsqu'elle était supérieure à la valeur médiane de 36,27 µg/m³ (Lin *et al.*, 2020). De la même manière, l'augmentation des concentrations moyennes élevées de PM_{2,5} (*i.e.* la médiane était de 11,8 µg/m³ [10,1-13,7]) a été associée au déclin de la fonction rénale *via* la mesure du DFG (Bowe *et al.*, 2018). Des niveaux élevés de pollution atmosphérique par les particules (et notamment les PM_{2,5}) ont été associés à la progression des MRC, les femmes ou encore les populations pour lesquelles

⁶ 15 µg/m³ pour les PM₁₀ et 5 µg/m³ pour les PM_{2,5} (WHO, 2021b)

une néphropathie diabétique avait été diagnostiquée semblant les plus vulnérables (Feng *et al.*, 2021). Une exposition à long terme à de faibles doses de PM_{2,5} a été reliée au développement des MRC et de l'IRCT, plus particulièrement pour les moins de 65 ans dans les travaux de Wu *et al.* (2022) et pour les plus de 75 ans pour Bowe *et al.* (2018).

Les résultats pour la variable SOMO₃₅ (« *Sum of Ozone Means Over 35 ppb* ») rejoignent les résultats des travaux menés Hwang *et al.* (2021) où aucun lien significatif n'avait été mis en évidence entre une exposition à long terme à la pollution atmosphérique qui a considéré les concentrations moyennes d'ozone et d'autres polluants atmosphériques (*i.e.* PM₁₀, SO₂, NO₂, CO) et l'incidence de MRC entre 2006 et 2013 lorsque les données avaient été ajustées sur différents facteurs (*i.e.* âge, sexe, couverture médicale, aire de résidence, *Charlson Comorbidity Index*⁷). Toutefois, l'indicateur pour l'ozone intégré à l'indice de vulnérabilité diffère de celui mentionné dans les travaux de Hwang *et al.* (2021). Le SOMO₃₅ correspond à la somme des différences calculées entre le maximum journalier de la moyenne glissante sur huit heures et 35 ppb (*i.e.* 70 µg/m³), sur une année pour tous les jours (Amann *et al.*, 2008). De ce fait, seules les valeurs supérieures à 70 µg/m³ sont utilisées pour son calcul. Bien que non réglementaire, l'OMS recommande d'utiliser cet indicateur pour évaluer les effets sanitaires de l'ozone. C'est en effet un polluant secondaire qui se forme par exemple entre (i) des oxydes d'azote (NOx), principalement émis par le trafic routier et l'industrie et (ii) des Composés Organiques Volatils (COV) issus de différentes sources. Les périodes de temps ensoleillé constituent les périodes au cours desquelles les pics de concentration d'ozone peuvent être observés, la réaction étant activée par le rayonnement solaire (WHO, 2021a). Bien que les liens entre l'ozone et les MRC et/ou l'IRCT soient moins fréquemment considérés dans la littérature, des atteintes rénales aiguës avaient pu être mises en évidence pour des expositions à court terme à l'ozone (les concentrations moyennes en ozone de 70,5 µg/m³ [IQR : 59,0-84,0] ayant été considérées comme élevées) (Lee *et al.*, 2022). Il est également important de souligner que la littérature a mis en évidence des liens significatifs entre l'exposition à l'ozone et la survenue du diabète, une importante comorbidité pour les MRC (ABM, 2015; Afsar *et al.*, 2019; Li *et al.*, 2021; Yu *et al.*, 2021). Ce lien apparaîtrait plus important pour les femmes et les personnes de moins de 50 ans (Afsar *et al.*, 2019).

Certaines études dans la littérature ont mis en évidence des sur-incidences de MRC lorsque les températures étaient plus élevées (Johnson *et al.*, 2019). Des températures plus extrêmes, possiblement en lien avec le changement climatique, pourraient en effet entraîner des déshydratations répétées exacerbant des maladies rénales sous-jacentes. Cette exposition récurrente à des températures plus élevées entraînerait donc la survenue de MRC qui ne seraient alors pas dues au diabète ou à l'hypertension (Glaser *et al.*, 2016). D'autres études ont souligné que l'exposition à des températures plus élevées couplée à l'usage de substances potentiellement néphrotoxiques dans le cadre d'activités agricoles augmenterait le risque de néphropathies (Wilke *et al.*, 2019). Dans la littérature, les populations agricoles sont régulièrement mises en évidence comme potentiellement à risque de développer des MRC quand les facteurs de risque habituels ne sont pas en cause (Almaguer *et al.*, 2014; Couchoud *et al.*, 2012; Glaser *et al.*, 2016; Wilke *et al.*, 2019). Des MRC à l'étiologie inconnue (*i.e.* « *Chronic Kidney Disease of unknown etiology* ») émergeraient en effet au sein de populations agricoles et concerneraient plutôt des populations âgées de moins de 50 ans et les hommes (Wilke *et al.*, 2019). Les « *MesoAmerican Nephropathies* » constituent également des MRC à étiologie inconnue reconnues comme des formes endémiques affectant les populations agricoles

⁷ Le *Charlson Comorbidity Index* est un outil d'évaluation du niveau de comorbidité et de mortalité à long terme qui offre une évaluation de la situation clinique du patient. Le CCI est une variable catégorielle qui correspond à un indice pondéré en fonction de l'association avec la mortalité après un an de 19 comorbidités (Charlson *et al.*, 1987)

principalement localisées dans des pays d'Amérique Centrale (e.g. Nicaragua, Salvador) (Afsar *et al.*, 2019). Les populations agricoles ou vivant en zones rurales tendent en effet à être plus exposées à de nombreux Produits Phyto-Pharmaceutiques (PPP) (e.g. herbicides, fongicides, pesticides) (Kamel and El-Minshawy, 2010; Wilke *et al.*, 2019). Les connaissances sur les impacts à long terme des PPP sur la fonction rénale doivent encore être étudiées et/ou approfondies, certains PPP spécifiques (notamment l'alachlore parmi les herbicides ou encore la perméthrine qui est un insecticide) tendant à montrer des liens sur la survenue de MRC ou de l'IRCT (Lebov *et al.*, 2016).

Les ETMs (e.g. le cadmium, le mercure et le plomb) sont des contaminants avec des effets néphrotoxiques bien connus, bien que les effets à faibles doses en population générale soient plus controversés. Hellström *et al.* (2001) ont conclu que l'exposition à des niveaux professionnels mais également relativement faibles dans l'environnement semblait être un facteur déterminant pour le développement de l'IRCT. Une association de type dose réponse significativement accrue avait en effet été mise en évidence pour la population exposée au cadmium par rapport au groupe non exposé. Les présences de plomb et de mercure à faible doses dans l'environnement n'ont pas été associées aux MRC quand les données avaient été ajustées sur des facteurs de confusion (e.g. âge, IMC, statut tabagique, hypertension) tandis que le cadmium serait un facteur contribuant aux MRC dans les populations diabétiques ou hypertendues (Kim *et al.*, 2015). Plusieurs de ces métaux (e.g. mercure, cadmium, plomb, arsenic) présentent en effet un caractère néphrotoxique avéré : une augmentation des concentrations pourrait être associée à un déclin rapide du DFG (Afsar *et al.*, 2019; Bowe *et al.*, 2017; Tagne Fotso, 2016).

Dans le cadre de cette étude, nous nous sommes intéressés au REM des ETMs, calculé d'après la méthodologie présentée dans les travaux de Occelli *et al.* (2016) à partir des données d'émissions dans l'air, toute source confondue⁸, de neuf ETMs (*i.e.* arsenic, cadmium, chrome, cuivre, mercure, nickel, plomb, sélénium et zinc) pour les années 2004, 2007 et 2012 qui permet de retranscrire en une valeur agrégée les émissions multi-métalliques dans l'atmosphère.

Pour la construction de l'indice composite de vulnérabilité, l'étude des corrélations nous a amené à sélectionner les émissions dans l'air des HAPs (le total des huit⁹, tout secteur confondu⁸), au détriment du REM. Le choix de l'une ou l'autre de ces variables fortement corrélées s'est principalement porté sur le fait que la variable en lien avec les émissions de HAPs permettait de garder un plus grand nombre de variables dans les indices composites de résilience et de vulnérabilité concourant à une caractérisation plus étendue des territoires. La pertinence de la donnée sur les émissions de HAPs au regard des questions de santé a été validée : bien que des études supplémentaires demeurent nécessaires pour approfondir les connaissances sur les effets entre une exposition aux Hydrocarbures Aromatiques Polycycliques (HAPs) de manière générale et les atteintes néphrotoxiques (e.g. baisse de la créatininémie et altération du DFG), certaines études ont déjà montré que l'exposition aux HAPs pouvaient avoir des conséquences sur la fonction rénale (Kataria *et al.*, 2015).

⁸ Tous les secteurs d'activités recensés sur l'Inventaire National Spatialisé ont été considérés pour les émissions de polluants en lien avec des activités anthropiques ou les émissions naturelles dans l'air : combustion dans les industries de l'énergie et de la transformation de l'énergie, combustion hors industrie, combustion dans l'industrie manufacturière, procédés de production, extraction et distribution de combustibles fossiles, utilisation de solvants et autres produits, transport routier, autres sources mobiles et machines, traitement et élimination des déchets, agriculture et sylviculture et autres sources (forêts, volcans, ...)

⁹ Ce total comprend les émissions dans l'air des huit principaux HAP : le fluoranthène, le benzo[a]anthracène, le benzo[a]pyrène, le benzo[b]fluoranthène, le benzo[g,h,i]perylène, le benzo[k]fluoranthène, le indéno[1,2,3-cd]pyrène et le dibenz[a,h]anthracène

Des zones géographiques comportant une part élevée de personnes issues de l'immigration ou de minorités avaient été associées à des zones de déclin rapide du DFG (Bowe *et al.*, 2017). Au-delà des aspects d'isolement linguistique ou de fracture sociale pour ces populations, certaines ethnies apparaissent plus impactées par l'IRCT (notamment dans la littérature prenant place aux Etats-Unis). Les différences de prévalence constatées entre ethnies pour les MRC restent encore à clarifier mais pourraient être associées à des différences de prévalence pour le diabète et l'hypertension (Bilgel, 2019; Kovesdy, 2022). Les mécanismes en lien avec ces différences sont complexes et pourraient inclure les facteurs socio-économiques ou l'accès aux systèmes de santé, certaines ethnies pouvant être associées à des populations plutôt défavorisées (*e.g.* revenus moins élevés) (Bilgel, 2019; Nicholas *et al.*, 2015). Pour le développement de l'indice de vulnérabilité, la part des ménages composés d'une personne de plus de 65 ans vivant seule a été intégrée pour retranscrire la vulnérabilité sociale d'une population qui cumulerait deux déterminants pouvant impacter la qualité de vie. A notre connaissance, il s'agit du premier développement de ce type de déterminant composite pour une étude en lien avec les MRC et/ou l'IRCT. Ce type d'approche demande à être promu afin d'évaluer dans quelle mesure ce déterminant composite pourrait avoir une influence sur la survenue de l'IRCT.

La littérature a souligné la relation inverse entre la défaveur socio-économique et des habitudes de vie saines. Le tabagisme affecterait la fonction rénale au travers de la BPCO, mais également *via* l'hypertension (augmentation de la tension artérielle et du rythme cardiaque, vieillissement prématuré des artères (Bilgel, 2019; Madika and Mounier-Vehier, 2017). De plus, la proximité et/ou la densité des débits de tabac encourageraient le tabagisme bien que des disparités existent selon les conditions socio-économiques des populations ou encore selon l'appartenance à certaines minorités (Glasser *et al.*, 2022; Golden *et al.*, 2020). Toutefois dans les travaux de Golden *et al.* (2020), l'association significative entre densité de points de vente et prévalence du tabagisme n'a été mise en évidence que dans des zones urbaines ce qui souligne des mécanismes sous-jacents pour la consommation tabagique différents en zones rurales. Dans l'étude de Schold *et al.* (2018), il a été souligné que la part de patients dont l'hypertension était une cause primaire de l'apparition de l'IRCT était la plus importante parmi les comtés où l'espérance de vie était la plus faible, ces comtés étant également ceux qui présentaient la défaveur socio-économique la plus importante mais également la consommation de tabac la plus élevée. Bien que la littérature tende à montrer une association entre défaveur socio-économique et/ou un faible niveau d'éducation, la corrélation entre la densité des points de vente et le FDep est modérée au sein des départements du Nord-Pas de Calais ($\rho = 0,198$; $p\text{-value} < 0,001$).

L'indice de résilience

Au sein de la littérature sur les déterminants qui ont un impact positif sur la survenue des MRC et/ou de l'IRCT, les études se focalisent principalement sur la végétation ou sur la disponibilité et/ou l'accessibilité au soin. De manière générale, les espaces verts sont associés à de nombreux effets bénéfiques sur la santé (*e.g.* santé mentale, santé cardiovasculaire). Au sein des départements du Nord-Pas de Calais, le RR le plus important et significatif a été souligné pour le NDVI ce qui concorde avec les résultats de Liang *et al.* (2022) où une association inverse avait été mise en évidence entre le NDVI et la prévalence des MRC. Cette association s'était avérée plus importante pour (i) les populations avec des statuts socio-économiques plus élevés et (ii) au sein de régions qui présentent des concentrations plus importantes de $PM_{2,5}$. Dans d'autres travaux, aucune association significative n'a cependant pu être mise en évidence entre le NDVI et les MRC (Chien *et al.*, 2021). Toutefois, ils ont souligné que les populations vivant à proximité de petits espaces verts urbains (*i.e.* superficie de moins d'un hectare avec au moins un des trois éléments suivants : végétation, installation sportive, banc) présentaient de plus faibles risques de MRC après ajustement sur de nombreux facteurs (*e.g.*

paramètres individuels, pollution atmosphérique, statut tabagique). Ces deux études sont les seules qui ont analysé l'influence de cet indicateur de végétation sur la survenue des MRC alors que pour d'autres pathologies la littérature est plus riche (e.g. maladies cardiovasculaires, santé mentale). L'association entre l'accessibilité à un espace vert dans un rayon de 400 mètres et l'incidence de l'IRCT n'a pas montré de RR significatif au sein du périmètre d'analyse. Ce résultat ne concorde pas avec les travaux de Wu *et al.* (2020) qui avaient mis en évidence une association significative et inversement proportionnelle entre les espaces verts (i.e. les parcs et écoles dans un rayon de 2 000 mètres, permettant de s'entraîner à proximité de zones résidentielles) et l'incidence de l'IRCT ou du diabète.

La part des propriétaires dans les résidences principales et la part de population vivant en couple ayant permis de retranscrire un aspect de « bien-être de vie au quotidien » ont montré une association inverse significative avec l'incidence de l'IRCT. L'analyse de la littérature a montré que pour ce type de déterminant, les résultats pouvaient être équivoques selon les études et que des différences selon le sexe pouvaient être observées. Pour certains travaux, le soutien social, notamment familial et conjugal, apparaît comme primordial au regard du vieillissement des populations (Williams *et al.*, 2017). Les populations non mariées, divorcées ou veuves pourraient être associées à un bien-être moindre en raison de la solitude ou d'un manque de soutien dans la vie quotidienne (Lindström and Rosvall, 2012). Le statut marital constitue un important déterminant de la santé publique, et notamment pour la santé mentale. Néanmoins, ces différences peuvent être un peu modulées lorsque les facteurs socio-économiques sont considérés comme facteur de confusion. Les populations mariées (ou vivant en cohabitation) tendent à avoir une hygiène de vie meilleure (e.g. arrêt du tabac) que des populations non mariées bien que certaines études aient montré que le statut marital pouvait être associé à l'obésité (Lee *et al.*, 2020). Dans les travaux de Barzegar *et al.* (2021), le statut marital avait été considéré comme facteur d'ajustement dans l'analyse de l'association entre le niveau d'éducation et les MRC, et une incidence plus élevée avait mise en évidence pour les niveaux d'éducation les plus élevés soulignant que le statut marital pouvait dans ce cas refléter des choix alimentaires (e.g. alimentation riche en protéines) induisant des hyperfiltrations glomérulaires et de l'hypertension. Des études en lien avec le statut d'occupation du logement avait également souligné que celui-ci pouvait favoriser le bien-être des populations et créer un sentiment de sécurité (Bentley *et al.*, 2016; S. Wang *et al.*, 2019).

Le résultat non significatif pour l'APL pour les médecins généralistes sur la période 2015-2018 ne concorde pas avec les résultats d'autres travaux dans la littérature. Kihal-Talantikite *et al.* (2013) avaient conclu qu'au-delà de l'âge et du sexe, les caractéristiques du quartier en termes d'offres de soins (i.e. densité de médecins pour 100 000 habitants et nombre de postes de dialyse par habitant) pouvaient également expliquer une partie de la distribution spatiale de l'incidence de l'IRCT. Les travaux de Bilgel (2019) avaient considéré la disponibilité de l'offre de soins, en termes de densités de médecins et de nombre de centres de dialyses au sein des comtés des Etats-Unis. Pour certains comtés une meilleure offre de soins avait été associée à une diminution de la prévalence de l'IRCT tandis que pour d'autres la relation était positive. Ce résultat non attendu dans leurs travaux avait souligné que le nombre réel de cas d'IRCT pourrait avoir été sous-estimé et que des patients vivaient sans savoir qu'ils sont atteints d'une MRC. Dans d'autres études, un déclin plus rapide du DFG a été associé à la distribution spatiale de différentes spécialités médicales (e.g. médecins généralistes, dentistes) tandis que la densité de médecins généralistes considérée comme un *proxy* de l'accessibilité aux systèmes de soins a été significativement associée à une diminution de l'incidence de l'IRCT (Bowe *et al.*, 2017; Joyce Fan *et al.*, 2007). Dans l'étude de Couchoud *et al.* (2012), plusieurs déterminants en lien avec le système de soins avaient été analysés et seule la proportion de néphrologues dans le secteur privé a pu être associée à une augmentation significative de l'excès de risque d'IRCT.

Bien que le faible niveau ou l'absence d'activité physique fasse partie des facteurs de risque modifiables associés à une plus faible qualité de vie jouant un rôle crucial dans la plupart des maladies rénales (Mallamaci *et al.*, 2020; Wilund *et al.*, 2021), les résultats sur le Nord-Pas de Calais n'ont pas permis de souligner l'effet bénéfique de la part des licenciés sportifs dans la commune de résidence. Différents aspects en lien avec l'activité physique peuvent être considérés dans la littérature. Les liens entre la pratique d'une activité physique et une meilleure santé sont généralement reconnus. Un mode de vie plus sédentaire et une activité physique plus réduite pourraient être en lien avec une augmentation de la mortalité ou de la morbidité chez des patients atteints de MRC ou d'IRCT. Une activité physique plus importante impliquerait en effet des mécanismes en lien avec la réduction de la pression artérielle et une diminution de l'IMC et *de facto* du diabète (Cheung *et al.*, 2021; Mallamaci *et al.*, 2020). Pour Bilgel (2019), l'absence de pratique sportive ainsi que l'obésité entraînerait une hypertension qui serait être associée à une néphrosclérose des artéioles induisant alors (i) un déclin de la fonction rénale ou (ii) des complications des MRC préexistantes. Une autre problématique porte sur les liens entre l'accessibilité à des équipements et la pratique sportive réellement effectuée dans un premier temps et l'amélioration de la santé dans un second temps. En effet, l'accessibilité à des équipements de sports et loisirs dans un périmètre de vie ne renseigne pas nécessairement sur la pratique sportive de la population riveraine. L'accessibilité à la possibilité de pratiquer une activité physique n'avait pas montré de résultats significatifs quant au déclin de la fonction rénale dans les travaux de Bowe *et al.* (2017).

Limites et perspectives

Ces travaux de thèse se sont principalement focalisés sur l'exposition aux polluants atmosphériques qui constitue le principal facteur de risque sanitaire reconnu par de nombreux travaux et commissions d'experts (Daiber *et al.*, 2019). Néanmoins, les pollutions environnementales sont ubiquitaires et peuvent contaminer aussi bien des compartiments aquatiques que terrestres.

Jayasumana *et al.* (2013) et Kamel and El-Minshawy (2010) avaient étudié le lien entre qualité de l'eau potable et les MRC, avec une association significative pour Kamel and El-Minshawy (2010). Une association entre une exposition à l'arsenic dans l'eau potable et l'IRCT a également été mise en évidence dans les travaux de Cheng *et al.* (2018), l'effet ayant été modifié par des comorbidités. Les travaux de Chang *et al.* (2018) ont souligné l'association entre la contamination des eaux souterraines par l'arsenic et la prévalence des MRC et de l'IRCT (celle-ci serait plus importante pour des concentrations plus élevées) tandis que le zinc et l'ammoniac seraient des indicateurs de la prévalence des MRC uniquement. Une relation dose réponse a également été mise en évidence entre l'arsenic provenant de l'eau potable et des maladies rénales (Hsu *et al.*, 2017). Les mécanismes de néphrotoxicité impliqueraient une réduction de l'arsenic en arsénite puis une méthylation dans le foie avant son excrétion dans l'urine. Il pourrait alors se concentrer dans le rein et affecter les fonctions des tubules proximaux et du glomérule. En France, les données en *open data* sur les Eaux Destinées à la Consommation Humaine (EDCH) à l'échelle des communes recensent des paramètres physico-chimiques, bactériologiques et certains contaminants (*e.g.* sous-produits de chloration, ETMs). Toutefois pour la période 2016-2020 sur le périmètre couvert par le registre Néphronor, seuls six éléments-traces (*i.e.* antimoine, chrome, cuivre, fer, nickel, plomb) présentent des concentrations qui pourraient être utilisées. Pour les éléments-traces suspectés ou avérés d'avoir un lien sur les MRC et/ou l'IRCT dans la littérature, aucune des concentrations relevées ne s'est avérée supérieure à la limite de quantification.

Les travaux étudiant les liens entre les contaminations du sol et les MRC et/ou l'IRCT sont rares. Il a été mis en évidence que les expositions à long terme aux contaminants métalliques des sols étaient associées à un déclin rapide de la fonction rénale entraînant une progression des MRC vers l'IRCT (Tsai

et al., 2018). De la même manière, les populations qui vivaient dans des zones où les concentrations de chrome, cuivre, nickel et zinc ont été associées à des diagnostics de diabète et d'hypertension. Le Réseau de Mesure de la Qualité des Sols (RMQS) propose des prédictions sur une grille d'un kilomètre des fractions totales et biodisponibles pour huit éléments-traces (i.e. arsenic, cadmium, chrome, cuivre, mercure, nickel, plomb, thallium). Elles sont issues de traitements géostatistiques appliqués aux résultats obtenus lors de la campagne 2000-2009 à partir des 2 200 points de prélèvements situés au centre de mailles de 16 kilomètres de côté. Parce qu'elles ne permettent de déterminer que la teneur maximale qui est habituellement constatée dans les sols français¹⁰ et qu'elles ne renseignent pas sur la possible contamination des sols par les activités anthropiques¹¹, ces données ont été écartées de la caractérisation des territoires. De nouvelles approches permettant de traduire les pollutions du sol devront être développées.

Les études épidémiologiques apportent une autre démarche, complémentaire, en mettant en place des modèles qui ont permis de considérer l'imprégnation de l'individu sans la relier spécifiquement à un compartiment environnemental. Dans l'étude des causes des « *MesoAmerican Nephropathies* », les travaux de Fischer et al. (2020) ont souligné l'influence du nickel dans l'insuffisance rénale aigüe avec une relation dose réponse inverse entre les concentrations de nickel mesurées dans des échantillons d'ongle et la filtration rénale. Liu et al. (2020) avaient mesuré les concentrations plasmatiques en ETMs afin de pouvoir analyser la relation entre la fonction rénale et l'exposition multiple et simultanée aux éléments-traces. Ils ont mis en évidence l'influence de l'aluminium, de l'arsenic et du plomb sur le déclin de la fonction rénale. L'étude de Su et al. (2022) a mis en évidence que la co-exposition aux substances perfluoroalkylées (PFAS) et aux éléments-traces présentait une association négative et significative avec les niveaux de DFG et une augmentation des MRC. Les concentrations urinaires en éléments-traces (i.e. arsenic et vanadium) ont été positivement et significativement associées aux MRC, aussi bien séparément qu'en mélange (Zhou et al., 2021).

Selon Wu et al. (2020), des données en lien avec certaines comorbidités (e.g. hypertension, diabète, hyperlipidémie) peuvent également être associées à l'IRCT selon les modèles méthodologiques utilisés. Au sein du registre Néphronor, des données individuelles sont recueillies au moment de l'enregistrement du patient (i.e. IMC, statut tabagique, antécédents de bronchopneumopathie chronique obstructive (BPCO), insuffisance cardiaque, insuffisance coronarienne, infarctus du myocarde et cancer). Toutefois, les modèles de régression écologique qui prennent en compte la structure spatiale des phénomènes ne permettent pas de considérer des paramètres individuels. Des modèles de Poisson hiérarchiques mixtes auraient pu permettre d'ajuster sur ces paramètres, avec l'inconvénient majeur de la perte de l'information spatiale. De plus, l'intégration de données individuelles dans des modèles spatiaux auraient nécessité des informations qui n'étaient pas disponibles. En effet parmi certaines publications faisant référence sur l'IRCT (Bowe et al., 2017), les méthodologies utilisées s'apparentent à des études cas-témoin, où les « cas » sont par exemple des patients pour lesquels un déclin rapide du DFG est constaté, tandis que les « témoins » sont des patients pour lesquels ce déclin pourrait être qualifié de « normal ». Pour l'intégration dans les modèles spatiaux, il faudrait pouvoir avoir accès aux moyennes de ces paramètres soit dans la

¹⁰ Les points où les relevés sont réalisés sont situés sur des portions agricoles ou en forêt, aucun prélèvement de sol n'est réalisé sur du tissu urbain

¹¹ Pour le cadmium, la valeur maximale prédite pour la fraction totale qui a été observée sur le Nord-Pas de Calais est de 0,979 mg/kg (la valeur maximale de la fraction biodisponible est 0,477 mg/kg). Dans l'arrêté du 08/01/98 fixant les prescriptions techniques applicables aux épandages de boues sur les sols agricoles pris en application du décret n°97-1133, la valeur limite de concentration de cadmium dans les sols est fixé à 2 mg/kg ps (tableau 2).

population générale soit parmi les populations de patients dans les clusters de sous-incidence. Toutefois, ces données en population générale ne sont pas disponibles et par essence, les zones de sous-incidence ne regroupent que peu voire pas de patients ce qui ne permet pas d'avoir une représentativité des paramètres mesurés au sein de cette population. De plus, pour certains paramètres, les pourcentages de données manquantes constatés peuvent être relativement élevés et des disparités géographiques sont également observés en fonction des centres lourds (e.g. jusque 40 % de données manquantes pour l'IMC dans certaines zones de sur-incidence). La seule donnée individuelle qui a pu être intégrée au modèle est le DFG comme variable d'ajustement permettant de souligner les différences de pratique entre les centres de dialyse dits lourds.

Les travaux qui ont été menés dans le cadre de la collaboration avec le registre Néphronor ont permis l'identification des spécificités régionales en termes d'accessibilité aux soins et de pratiques de soins. Ces spécificités géographiques pouvant avoir un impact dans la détermination des variations spatiales de l'incidence, elles ont été intégrées en tant que facteurs de confusion dans les modèles statistiques. Les résultats du modèle 2 sont venus souligner ces hypothèses : l'intégration des facteurs de confusion en lien avec la santé a contribué à expliquer plus d'un quart de la variabilité spatiale de l'IRCT obtenue après ajustement sur l'âge et le sexe. L'ajout du FDep comme *proxy* de la défaveur socio-économique a de la même manière augmenté la part expliquée de la variabilité observée de l'incidence. L'ensemble des facteurs de confusion a contribué à expliquer plus de la moitié de la variabilité de l'incidence de l'IRCT. Une analyse comparative de la part de la variabilité expliquée par les indices composites de vulnérabilité et de résilience a été menée au travers des modèles 4 et 5. Singulièrement, l'indice de vulnérabilité qui a agrégé des déterminants pour lesquels une influence sur les MRC et/ou l'IRCT était avérée ou suspectée dans la littérature n'a pas contribué à augmenter la part de variabilité observée de l'IRCT. L'apport principal de ces travaux repose sur la mise en évidence de l'importance des déterminants en lien avec la résilience au regard du gain de 25 % de l'indice composite de résilience sur la variabilité expliquée de l'incidence de l'IRCT, et notamment du NDVI un indice de végétation classiquement mentionné dans la littérature comme ayant une influence favorable sur la survenue et/ou la prévalence d'autres événements de santé (e.g. les maladies cardiovasculaires, la santé mentale).

L'ensemble de ces résultats fait l'objet d'un article qui sera soumis prochainement « *Spatial variability of End-Stage Renal Disease incidence: what impacts of confounding factors, vulnerability and resilience?* ». Le journal envisagé pour la soumission est le « *Journal of the American Society of Nephrology* ». Pour cet article, l'ordre des auteurs est le suivant :

Delphine Brousmiche, Mohamed Ahmed-Salem, Michaël Genin, Sébastien Gomis, Caroline Lanier, Damien Cuny, Florent Occelli, François Glowacki

Mohamed Ahmed-Salem étant premier co-auteur

Partie II.2 : les maladies coronariennes

Introduction

Chaque année, les Maladies Cardiovasculaires (MCV), le cancer, les maladies respiratoires et le diabète figurent parmi les principales Maladies Non Transmissibles (MNT) responsables de 71 % des décès à travers la planète (WHO, 2021c). De plus, ces MNT ont des impacts socio-économiques, particulièrement considérables pour les pays en voie de développement. Avec 31 % des décès enregistrés en 2016, les MCV sont récemment devenues la principale cause de mortalité et d'invalidité dans le monde (Davoodabadi *et al.*, 2019). Parmi ces décès, 7,4 millions ont été estimés en lien avec une cardiopathie coronarienne. En France, les MCV constituent la première cause de décès chez les femmes et la deuxième pour les hommes. En 2016, plus d'un million de patients ont été hospitalisés pour raison cardiovasculaire. Des disparités régionales existent pour les MCV : la région des HDF présente une mortalité avant 65 ans supérieure de 38,4 % à la moyenne nationale. Au sein de cette région, il s'agit de la troisième cause de surmortalité prématurée. La ville de Lille est la capitale de région la moins bien classée de France en terme de mortalité cardiovasculaire prématurée (Devilleine, 2016; WHO, 2021d).

Les cardiopathies coronariennes appartiennent aux maladies cardiovasculaires et touchent les vaisseaux sanguins qui alimentent le muscle cardiaque. Parmi celles-ci, les infarctus du myocarde font partie des événements qui sont généralement aigus. Ils sont dus à l'athérosclérose qui se caractérise par l'obstruction d'une artère le plus souvent par une plaque d'athérome sur les parois internes des vaisseaux sanguins. Des facteurs individuels comme l'hypertension, l'hyperglycémie, l'hyperlipidémie, le surpoids ou le diabète peuvent en constituer des facteurs aggravants. L'athérosclérose peut également résulter de facteurs comportementaux (mauvaise alimentation, manque d'activité physique, tabagisme, consommation d'alcool). Toutefois, l'étiologie des maladies coronariennes est complexe et multifactorielle et de nombreuses études ont souligné l'importance des facteurs socio-économiques, des habitudes de vie, de l'environnement ou encore du patrimoine génétique (Steptoe, 2015; Zhang *et al.*, 2022). Selon Dahlgren and Whitehead (2007), les inégalités face aux MCV expliqueraient presque la moitié de la surmortalité dans les populations socio-économiquement défavorisées.

Parmi les déterminants environnementaux, la pollution atmosphérique dans son ensemble est reconnue comme le facteur de risque environnemental le plus couramment associé aux MCV (Davoodabadi *et al.*, 2019 ; Eze *et al.*, 2014 ; IARC, 2013 ; Luo *et al.*, 2017). Des populations plus vulnérables (*e.g.* les personnes âgées, les diabétiques, les femmes enceintes) présentent un risque cardiovasculaire plus élevé suite à une exposition à la pollution atmosphérique (Tibuakuu *et al.*, 2018). De plus, plusieurs facteurs de risque pour les maladies cardiovasculaires se développent avec l'âge et/ou certaines habitudes de vie (*e.g.* hypertension, hyperlipidémie, athérosclérose) et sont également associés avec la pollution atmosphérique particulaire (Landrigan *et al.*, 2018). Des études récentes suggèrent également que la pollution atmosphérique (et notamment les PM_{2,5}) aurait également un impact sur des MNT très répandues (*e.g.* le diabète) qui sont des facteurs de risque pour les MCV en impliquant des mécanismes en lien avec le stress oxydant. Le diabète, les MCV et la pollution atmosphérique évolueraient concomitamment (Eze *et al.*, 2014 ; Landrigan *et al.*, 2018). Toutefois, des déterminants environnementaux en lien avec la résilience (*e.g.* la proximité et/ou la présence d'espaces verts) sont également reconnus comme étant associés aux MCV. Ils entraîneraient entre autre une diminution de la mortalité cardiovasculaire et de la survenue d'accidents cardiovasculaires

(van den Bosch and Ode Sang, 2017; Yeager *et al.*, 2019). En effet, ces déterminants seraient susceptibles de réduire l'exposition à la pollution atmosphérique, les îlots de chaleur urbains ou encore l'exposition au bruit mais également d'encourager la pratique sportive (Aerts *et al.*, 2020; Mukherjee *et al.*, 2017; Yeager *et al.*, 2019). Toutefois, les impacts cumulatifs de l'environnement et le risque cardiovasculaire sont complexes à évaluer et certains mécanismes impliqués demeurent encore obscurs (Bhatnagar, 2016).

Cet axe de la thèse fait l'objet d'une collaboration avec l'Unité d'Epidémiologie et de Santé Publique INSERM UMR744 de l'Institut Pasteur de Lille – Registre MONICA (Multinational MONItoring of trends and determinants in CARDiovascular disease). Il vient poursuivre les travaux valorisés dans la publication de Occelli *et al.* (2020) qui avait pour principal objectif d'explorer l'existence de liens entre l'incidence des maladies coronariennes et différents éléments de la pollution atmosphérique entre 2008 et 2011. Une association entre une exposition à de multiples polluants atmosphériques et l'incidence de maladies coronariennes a été mise en évidence, et ce, quel que soit le niveau de défaveur socio-économique. L'exposition aux polluants atmosphériques a considéré le NO₂, les PM₁₀ et les PM_{2.5} (couramment retrouvés dans la littérature) mais également des métaux (et plus particulièrement le cuivre, le cadmium, le nickel et le palladium) au travers de données issues de la biosurveillance lichénique de la qualité de l'air sur la Métropole Européenne de Lille.

Dans le cadre de ces travaux de thèse, nous émettons l'hypothèse que les caractéristiques territoriales en lien avec la vulnérabilité et la résilience ont une influence sur l'incidence des maladies coronariennes après qu'elle ait été ajustée sur deux facteurs de confusion bien connus des MCV (*i.e.* l'âge et le sexe). Les inégalités de santé cardiovasculaire nécessitent en effet des approches multidimensionnelles intégrant une caractérisation la plus large possible de l'environnement et des déterminants en lien avec la résilience ou la vulnérabilité qui peuvent être associés aux MCV. L'incidence standardisée sera mise en relation avec les caractéristiques territoriales évaluées selon la méthodologie détaillée dans le chapitre 1 (développement d'indices composites de vulnérabilité et de résilience).

Matériel et méthodes

Périmètre

L'analyse est menée à l'échelle des IRIS des 85 communes de la MEL présentes à l'origine de la création du registre. Les IRIS de type A (regroupent environ 1 000 salariés et comptent au moins deux fois plus d'emplois salariés que de population résidente) et D (occupés par de grandes zones spécifiques peu habitées et ayant une superficie importante type parcs de loisirs, zones portuaires, forêts, ...) ont été exclus, aboutissant à un total de 472 IRIS.

Population d'étude et ratios d'incidence standardisés (SIRs)

Le registre MONICA recense les cas de syndromes coronariens aigus pour une population de 35 à 74 ans au sein des 85 communes de la MEL présentes lors de la création du registre. Les événements sont recensés aussi dans les urgences et hôpitaux des secteurs publics et privés. Ont été intégrés dans l'étude tous les cas correspondant aux dénominations suivantes : infarctus du myocarde, syndrome coronarien aigu, angor instable ou autres épisodes d'insuffisance coronarienne aigüe (symptômes de l'ischémie associée à des anomalies de l'électrocardiogramme et/ou une élévation des troponines et/ou des résultats anormaux à l'angiographie). L'étude a utilisé les données de 2012 à 2016 et seuls les cas incidents ont été considérés (n = 4 464). Les ratios d'incidence standardisés (SIRs) exprimant le

rapport des cas incidents observés sur les cas incidents attendus ont été déterminés par standardisation indirecte.

Indices composites de vulnérabilité et de résilience

La méthodologie de développement des indices composites s'est appuyée sur la méthodologie présentée dans le Chapitre I. Les indices composites ont été mis en relation avec les SIRs lissés qui ont été déterminés selon la méthodologie présentée dans le paragraphe « régressions écologiques » dans la partie 1 du Chapitre II sur l'IRCT.

Résultats – Discussion

Incidence brute des maladies coronariennes pour la période 2012-2016

L'incidence brute a été déterminée sur l'ensemble de la population comprise entre 2012 et 2016 et les résultats sont présentés dans la Figure 27.

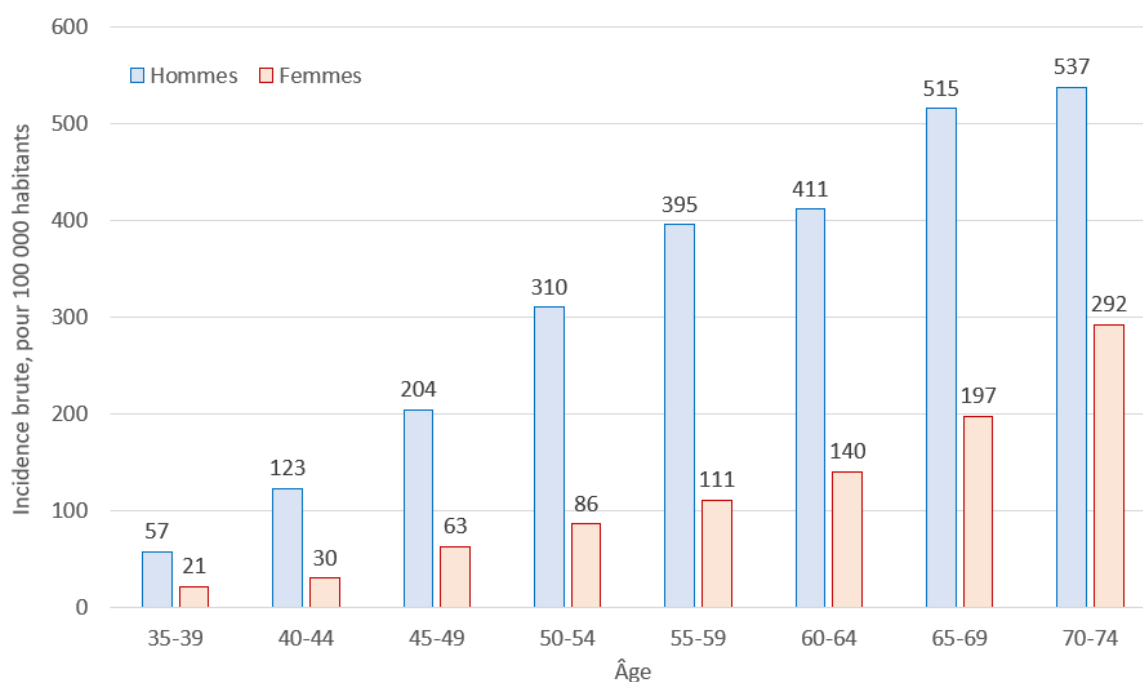


Figure 27 : Incidence brute des maladies coronariennes, pour 100 000 habitants (2012-2016)

L'incidence qui a été déterminée sur les 4 464 cas incidents du registre pour la période 2012-2016 souligne l'importance de considérer l'âge et le sexe comme facteurs de confusion pour les maladies coronariennes. En effet pour presque toutes les tranches d'âge, il existe un facteur trois entre les cas incidents enregistrés pour les hommes et ceux enregistrés chez les femmes. De plus, il apparaît au regard de la Figure 27 que l'incidence de ces pathologies augmente avec l'âge (d'un facteur 10 entre les premières et dernières classes d'âge). Après ajustement sur ces facteurs de confusion par standardisation indirecte, le SIR lissé présentait une médiane de 1,017 avec un intervalle interquartile de [0,859-1,188]. Un SIR de 1,188 exprime une morbidité de 18,8 % supérieure par rapport à la population de référence qui a été considérée pour son calcul, ici la population de référence correspond à la population totale de la MEL entre 2012 et 2016. La distribution spatiale des SIRs lissés sont présentés dans la Figure 28.

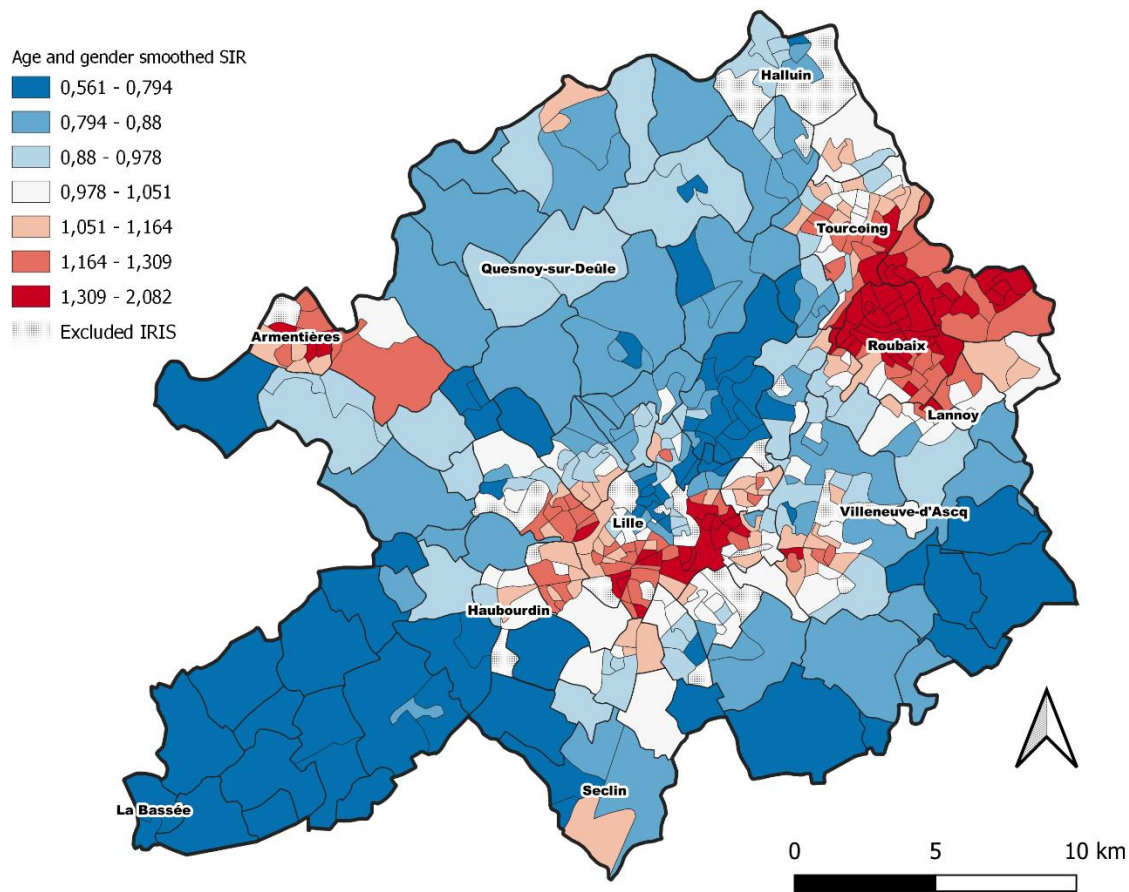


Figure 28 : Ratios d'Incidence Standardisé sur l'âge et le sexe pour les maladies coronariennes sur 2012-2016, après lissage spatial (cette figure sera publiée dans l'article 5)

Les zones de sur-incidence sont principalement localisées dans le sud de la ville de Lille, au Nord-Est de la MEL sur Roubaix, Tourcoing et Wattrelos ainsi qu'au Nord-Ouest autour d'Armentières.

Indice de Vulnérabilité

L'indice de vulnérabilité a agrégé 13 variables, et présente une médiane de 0,612 et un intervalle interquartile de [0,523-0,706]. Le Tableau XIX présente les contributions de chacune des variables aux quatre axes de l'ACP qui ont été retenus pour la rotation VARIMAX.

Tableau XIX : Contribution des 13 variables en lien avec la vulnérabilité aux quatre composantes retenues après la rotation VARIMAX

Variable	RC1	RC2	RC3	RC4
FDep	0,09	0,10	0,83	-0,20
Part des peu diplômés dans la population âgée de 15 à 64 ans	0,56	0,03	0,46	-0,10
Part des ménages composés d'une personne seule	-0,80	0,01	-0,14	-0,12
Part des ménages pauvres	-0,44	0,14	0,08	-0,52
Part de la population exposée à un bruit supérieur à 60 dB(A) sur 24h (L _{DEN})	0,00	0,85	0,08	-0,04
Part de la population exposée à un bruit supérieur à 45 dB(A) la nuit (L _N)	0,22	0,83	-0,02	0,04
Concentration moyenne annuelle en NO ₂	-0,39	0,60	-0,10	0,06
Empoussièrement des feuilles de peupliers	-0,31	0,02	-0,08	0,70
Imprégnation multi-métallique des lichens	-0,05	-0,19	-0,07	-0,58
Part des résidences principales construites avant 1970	-0,13	-0,13	0,76	0,15
Distance à l'équipement sportif le plus proche	0,57	-0,07	-0,45	0,18
Temps d'accessibilité par la route au service d'urgence le plus proche	0,32	-0,27	-0,12	0,52
Temps d'accessibilité par la route au médecin généraliste le plus proche	0,56	0,06	-0,12	-0,04
Proportion de variance expliquée :	0,30	0,26	0,24	0,20

La première composante de l'ACP retranscrit les parts des ménages composés d'une personne vivante seule, la distance à l'équipement sportif le plus proche et également les temps de trajet par la route au médecin généraliste le plus proche. La deuxième composante est principalement portée par des déterminants environnementaux (bruit et pollution atmosphérique). La troisième composante reflète la défaveur socio-économique associée à un habitat plus ancien. La quatrième composante quant à elle retranscrit principalement d'autres sources de nuisances environnementales au travers des données en lien avec la biosurveillance de la qualité de l'air, le temps d'accessibilité aux services d'urgence mais également la part de ménages pauvres. La Figure 29 présente la distribution de l'indice de vulnérabilité vis-à-vis de la santé cardiovasculaire sur les 472 IRIS de la MEL.

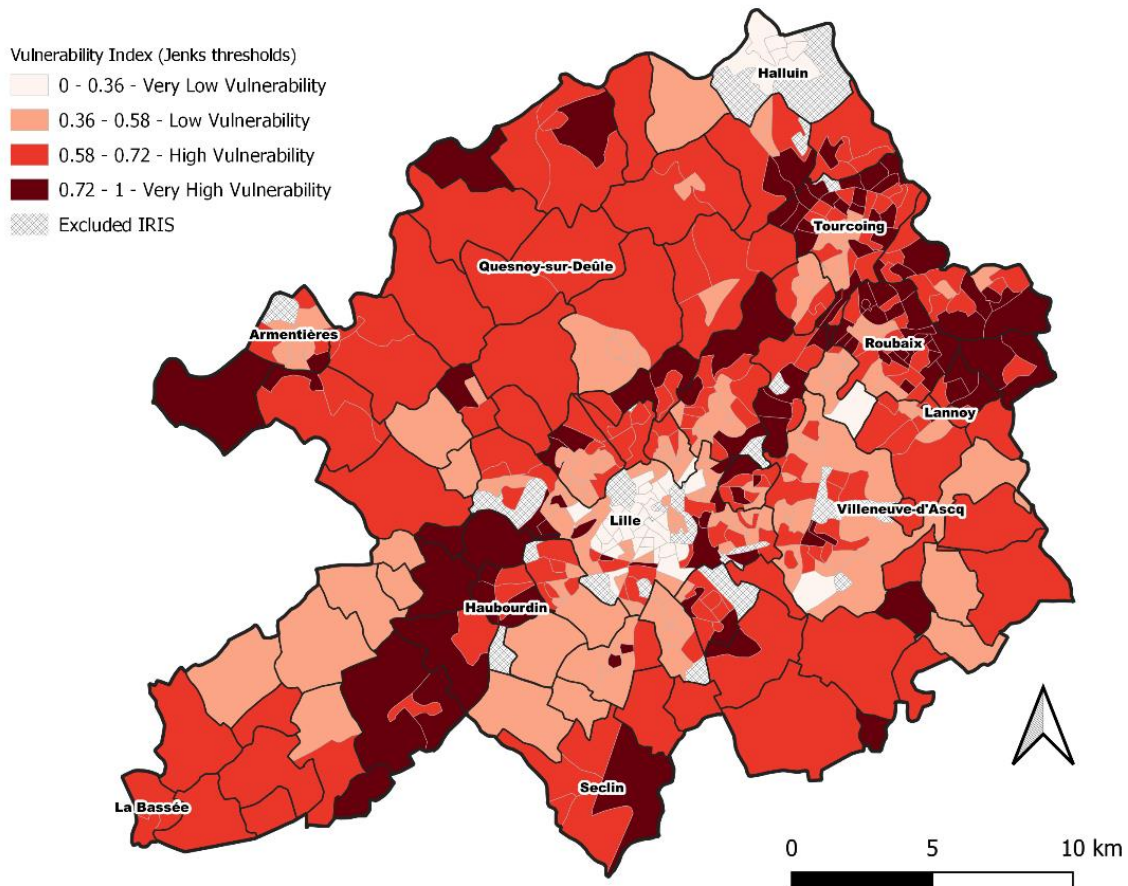


Figure 29 : Distribution spatiale de l'indice de vulnérabilité vis-à-vis de la santé cardiovasculaire sur les 472 IRIS de la MEL

Les IRIS qui présentent des indices de vulnérabilité faibles et très faibles sont principalement localisés dans la ville de Lille et dans les communes aux alentours. Les IRIS qui présentent une vulnérabilité vis-à-vis des maladies coronariennes sont plutôt localisés en périphérie nord de la ville de Lille, au sud d'Armentières mais également sur Roubaix, Tourcoing.

Indice de résilience

L'indice de résilience a agrégé 11 variables, et présente une médiane de 0,496 et un intervalle interquartile de [0,395-0,583]. Le Tableau XX présente les contributions de chacune des variables aux trois axes de l'ACP qui ont été retenus pour la rotation VARIMAX.

Tableau XX : Contribution des 11 variables en lien avec la résilience aux quatre composantes retenues pour la rotation VARIMAX

Variable	RC1	RC2	RC3
Part de la population active dans la population totale	0,01	0,81	0,10
Evolution de la part de population active dans la population totale	0,09	-0,68	0,20
Part des actifs occupés vivant et travaillant dans leur commune de résidence	-0,84	-0,01	-0,05
Accessibilité à des espaces verts, boisés ou des forêts d'au moins 0,5 ha dans un rayon de 400 m	0,25	0,18	-0,61
<i>Normalized Difference Vegetation Index</i>	0,81	-0,02	-0,25
Monétarisation des services écosystémiques	0,47	0,13	-0,06
Part des résidences principales construites après 2006	0,20	0,48	-0,14
Indice de Shannon pour la diversité spécifique des CSP au lieu de vie	0,51	0,45	0,37
Part des pistes cyclables dans la voirie totale	0,00	-0,11	0,70
Indice de Shannon pour la diversité spécifique des professions médicales et paramédicales au lieu d'exercice	-0,03	0,41	0,47
Part des déplacements actifs dans le cadre de la mobilité quotidienne	-0,89	0,06	0,10
Proportion de variance expliquée :	0,47	0,30	0,23

La part des actifs occupés vivant et travaillant dans leur commune de résidence, le NDVI, la part de personnes utilisant un mode de déplacement actif dans le cadre de la mobilité quotidienne, la diversité spécifique des CSP au lieu de vie ainsi que dans une moindre mesure la monétisation des services écosystémiques contribuent à la première composante de l'ACP. La deuxième composante principale de l'ACP est principalement portée par des variables économiques, avec la part de population active dans la population totale ainsi que son évolution entre deux années non consécutives, et dans une moindre mesure par une variable en lien avec l'habitat (part des résidences principales construites après 2006). La dernière composante de l'ACP retranscrit l'accessibilité à des espaces verts, la part des pistes cyclables dans la voirie totale et dans une moindre mesure la diversité spécifique des professions médicales et paramédicales recensées au lieu d'exercice. Au regard de ses variables explicatives, cette dernière composante pourrait retranscrire des « équipements » de santé environnementale. La Figure 30 présente la distribution spatiale de l'indice de résilience développé à l'échelle des IRIS de la MEL.

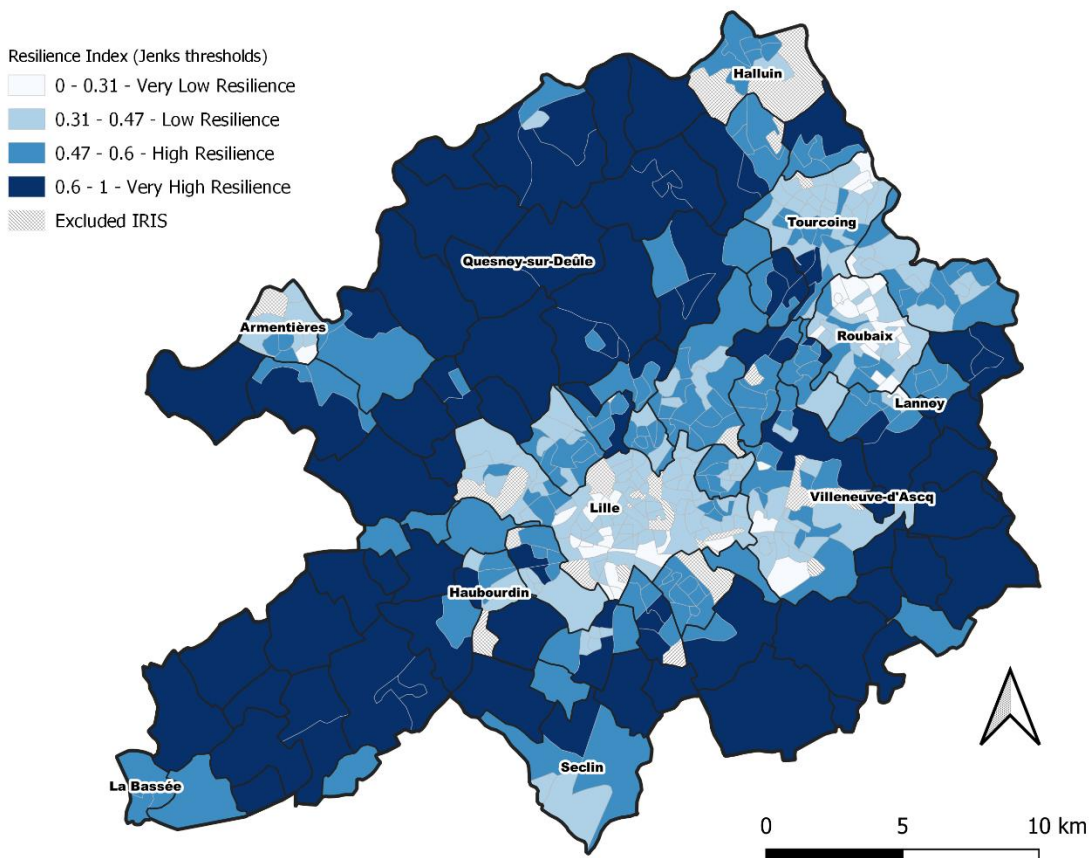


Figure 30 : Distribution spatiale de l'indice de résilience sur les 472 IRIS de la MEL

Les IRIS appartenant aux classes de faible résilience sont situées dans les principales villes de la MEL (Armentières, Lille, Roubaix, Tourcoing, Halluin). Les IRIS appartenant aux classes de résilience plus élevées sont principalement situées dans des communes non IRISées en périphérie des principales villes de la MEL. Bien que le résultat ne soit pas significatif, le coefficient de corrélation de Spearman entre l'indice de vulnérabilité et l'indice de résilience est de -0,04 ce qui souligne que ces deux indices mesurent des phénomènes différents.

Influence de la résilience et de la vulnérabilité sur les maladies coronariennes

Les SIRs ont été mis en relation avec les classes de vulnérabilité et de résilience qui ont été déterminées à partir des seuils de Jenks qui permet de minimiser la variance intra-classe et de maximiser la variance intra-classe. Les résultats sont présentés dans la Figure 31.

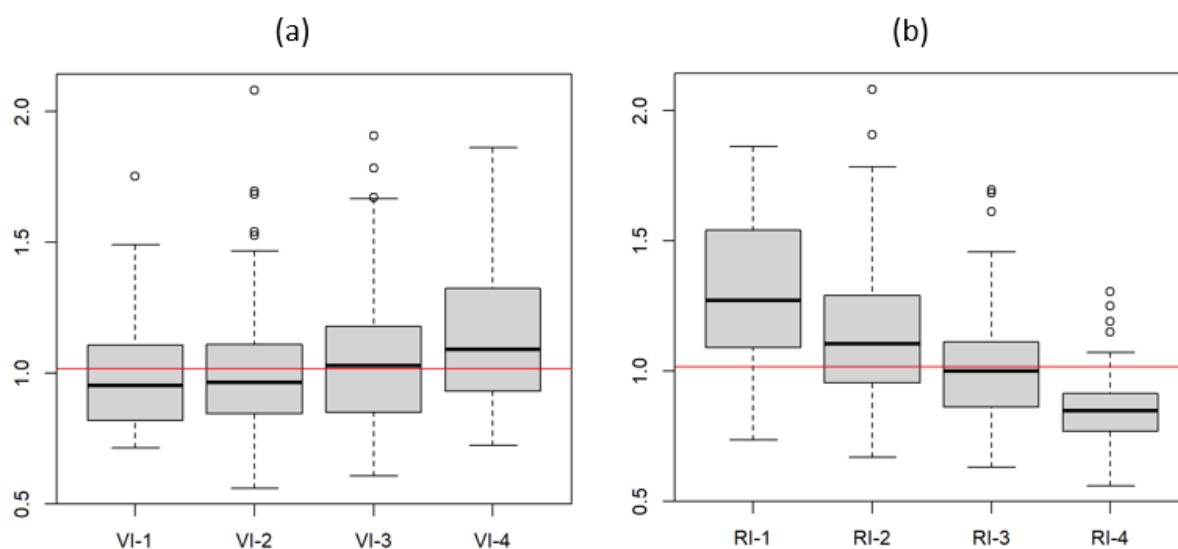


Figure 31 : Distribution des SIRs lissés en fonction des classes (a) de vulnérabilité et (b) de résilience, la ligne rouge représentant la médiane du SIR lissé

Il existe une relation positive entre les SIRs lissés et l'indice de vulnérabilité, tandis que l'association est inverse et plus forte avec l'indice de résilience. Les classes de vulnérabilité et de résilience ont été définies à partir des seuils de Jenks afin d'obtenir des groupes homogènes. Ces observations sont soulignées au regard des coefficients de corrélation : entre les SIRs et l'indice de vulnérabilité, le coefficient de Spearman est de 0,192 tandis qu'avec l'indice de résilience il est de -0,511, ces deux coefficients étant significatifs ($p < 0,001$).

Ces résultats sont cohérents avec ceux obtenus précédemment lors de l'analyse des déterminants en lien avec la variabilité résiduelle de l'IRCT. En effet, les indices de résilience apparaissent comme peu ou pas associés aux SIRs de ces maladies chroniques alors que les indices de résilience ont présenté une influence relativement importante. Pour ces deux indices de résilience, il a été mis en évidence l'importance du NDVI dans les résultats. De manière générale, ces résultats soulignent l'importance des déterminants territoriaux en lien avec la résilience dans l'étude du lien santé-environnement et l'analyse conjointe indispensable de la vulnérabilité et de la résilience.

Pour aller plus loin dans l'analyse des interactions entre les différents déterminants territoriaux, des profils complexes de territoires vis-à-vis des maladies coronariennes ont été développées à partir des variables sélectionnées pour les indices composites de vulnérabilité et de résilience. En effet, cela a permis d'analyser plus finement l'influence des déterminants en considérant leurs interactions possibles en fonction du sens de variation des variables intégrées dans la caractérisation du territoire. Ces profils complexes ont également été mis en relation avec les SIRs lissés des maladies coronariennes déterminés sur la période 2012-2016 afin d'identifier si des profils complexes pouvaient être associés à des zones de sur- ou de sous-incidence. Les résultats de ces travaux sont présentés dans l'Article 5.

Article 5

How do territorial characteristics affect spatial inequalities in the risk of coronary heart disease?

D. Brousmiche, C. Lanier, D. Cuny, C. Frevent, M. Genin, C. Blanc-Garin, P. Amouyel, A. Deram, F. Occelli, A. Meirhaeghe



Le « *supplementary materials and results* » de cet article est disponible dans l'Annexe 5.

(Article en relecture interne pour une soumission prochaine)

How do territorial characteristics affect spatial inequalities in the risk of coronary heart disease?

Authors: D. Brousmiche^{1*}, C. Lanier^{1,2}, D. Cuny^{1,3}, C. Frevent⁴, M. Genin⁴, C. Blanc-Garin⁵, P. Amouyel⁵, A. Deram^{1,2}, F. Occeili^{1,2}, A. Meirhaeghe⁵

Institutions:

¹ Univ. Lille, Univ. Artois, IMT Lille Douai, JUNIA, ULR 4515 – LGCgE, Laboratoire de Génie Civil et Géo-Environnement, F-59000 Lille, France

² Univ. Lille, UFR3S-Faculté d'Ingénierie et Management de la Santé (ILIS), F-59000 Lille, France

³ Univ. Lille, UFR3S-Faculté de Pharmacie de Lille – LSVF, F-59000 Lille, France

⁴ Univ. Lille, CHU Lille, ULR 2694 – METRICS : Evaluation des technologies de santé et des pratiques médicales, F-59000 Lille, France

⁵ Univ. Lille, CHU Lille, Institut Pasteur de Lille, Inserm UMR1167 - RID-AGE (Risk Factors and Molecular Determinants of Aging-Related Diseases), F-59000 Lille, France

Corresponding author: delphine.brousmiche@univ-lille.fr

Abstract:

Cardiovascular diseases remain the leading cause of death and disabilities worldwide and among them coronary heart diseases are the most diagnosed. Due to their multifactorial etiology that involves individual, behavioral and territorial determinants, multidimensional approaches need to be implemented to assess links between territorial characteristics and the incidence of coronary heart diseases. The aim of this research was therefore to establish complex territorial profiles that have been characterized by the standardized incidence. This led to identify the influences of determinants that can be related to a beneficial or a deleterious effect on cardiovascular health. 44 variables related to economic, social, health, environment and services dimensions for which their impact on cardiovascular health were established or suspected were used to describe the multidimensional characteristics involved in cardiovascular health. 3 complex territorial profiles had been highlighted and characterized by the standardized incidence rate of coronary heart diseases after adjustment on age and gender. Profile 1 was characterized by and SIR of 0.895 [CI-95: 0.876-0.914] and a higher number of determinants that revealed favorable territorial conditions. Profiles 2 and 3 were characterized by SIRs of respectively 1.225 [CI: 1.187-1.263] and 1.119 [CI: 1.028-1.075]. Territorial characteristics among these profiles of over-incidence were nevertheless dissimilar. Profile 2 revealed higher deprivation, lower vegetation and lower atmospheric pollution while profile 3 displayed population rather privileged with contrasted territorial conditions. This methodology empowered to characterize the multidimensional determinants involved in cardiovascular health, whether they have a negative or a positive impact and would provide to stakeholders a diagnostic tool to implement contextualized public health policies to prevent coronary heart diseases.

Keywords:

Complex territorial profile; multidimensional assessment; resilience; vulnerability

Introduction

Despite medical advances in the previous century, cardiovascular diseases (CVD) remain the leading cause of death and disabilities worldwide and among them Coronary Heart Diseases are the most frequently diagnosed diseases (Davoodabadi *et al.*, 2019; Schultz *et al.*, 2018; WHO, 2021d; Wolf *et al.*, 2021). Some risk factors are considered as individuals (genetic predisposition, diabetes, age, sex, dyslipidemia, hypertension) while others are considered as behavioral (unhealthy diet, tobacco use, physical inactivity, alcohol harmful consumption) (WHO, 2021d). Nevertheless, these individual and behavioral risk factors alone cannot explain all the spatial variability of CVD incidence. Simultaneously, other health determinants measured at a population level (low education level, socioeconomic status) have an increasing acknowledgment as affecting the cardiovascular health of population (Carter *et al.*, 2019; Dhindsa *et al.*, 2020; Nash *et al.*, 2021). In addition, it has been highlighted that deprived neighborhoods can be related to well-established behavioral risks (Kleinschmidt *et al.*, 1995; Thatipelli *et al.*, 2022; Xiao *et al.*, 2017). The income inequality, the poverty, the proximity to industrial activities are also part of social determinants of health that are associated with a higher pollutant exposure and pollution-related to CVD (Rajagopalan and Landrigan, 2021).

Independently of the socioeconomic status, different environmental conditions are associated with death and disabilities and among them the atmospheric pollution has a major impact on cardiovascular health (Argacha *et al.*, 2018; Davoodabadi *et al.*, 2019; Occelli *et al.*, 2020). The environmental conditions could however be beneficial for health (Bixby *et al.*, 2015). Different studies have analyzed the positive impact on health and CVD of green spaces proximity, the importance of the tree cover and/or the greenness (Aerts *et al.*, 2020; Bambra *et al.*, 2015; Yeager *et al.*, 2020; Yitshak-Sade *et al.*, 2017). To explain this relation for which the processes remain confused, different mechanisms are suggested, the natural environment can contribute to (i) decrease air pollution, (ii) encourage the physical activity, (iii) reduce stress or (iv) promote social contacts and cohesion (Aerts *et al.*, 2020; Donovan *et al.*, 2015; Hystad *et al.*, 2014; Yeager *et al.*, 2020). Although there are numerous publications about the impact of certain environmental conditions, the burden of the broad environment in the etiology of chronic diseases such as CHD still need to be deepened.

The cardiovascular health determinants are actually multidimensional. This complex character is difficult to consider because their key contributors and predictors are still insufficiently understood, particularly given the reciprocal interactions between the individual conditions, the behavioral risk factors, the socioeconomic conditions and the living environment (Hu *et al.*, 2020; Patino *et al.*, 2021; Vasan *et al.*, 2020). In public health research, the understanding on how social and environmental characteristics influence population's health has become a decisive objective (Xiao *et al.*, 2017). Nevertheless, a joint analysis of deleterious and beneficial factors involved in cardiovascular health has been so far developed scarcely. Addressing cardiovascular risk inequalities at a population level requires indeed multidimensional and holistic approaches which can promote the implementation of public health policies the more specific and contextualized as possible (Brousriche *et al.*, 2020b; Cartone and Postiglione, 2021; Katz *et al.*, 2021; Vasan *et al.*, 2020).

The primary purpose of this research was therefore to develop complex territorial profiles at a very small scale that aim to describe the multidimensional characteristics involved in cardiovascular health, whether they are established or suspected. Furthermore, these complex territorial profiles also had to consider the determinants that can be related to beneficial or deleterious effects on cardiovascular health with an equal importance. We compared afterwards the spatial distribution of the territorial profiles and the CHD rate after adjusting for the age and gender.

Material and methods

Study area and spatial unit

This study was carried out in the European Lille Metropolis (ELM), a densely populated area located in the north of France (1,765 inhab./km²). This is an urban area which presents multiple sources of pollutants (road traffic, residential heating, industrial activities, intensive farming) and the socioeconomic deprivation is higher there than in other regions of France. Within the region to which the ELM belongs, the mortality before 65 years is 38.4 % higher than the national average. Cardiovascular diseases are the third leading cause of premature death there and the municipality of Lille has the weakest decreases in case-fatality by acute myocardial infarction (Devillaine, 2016).

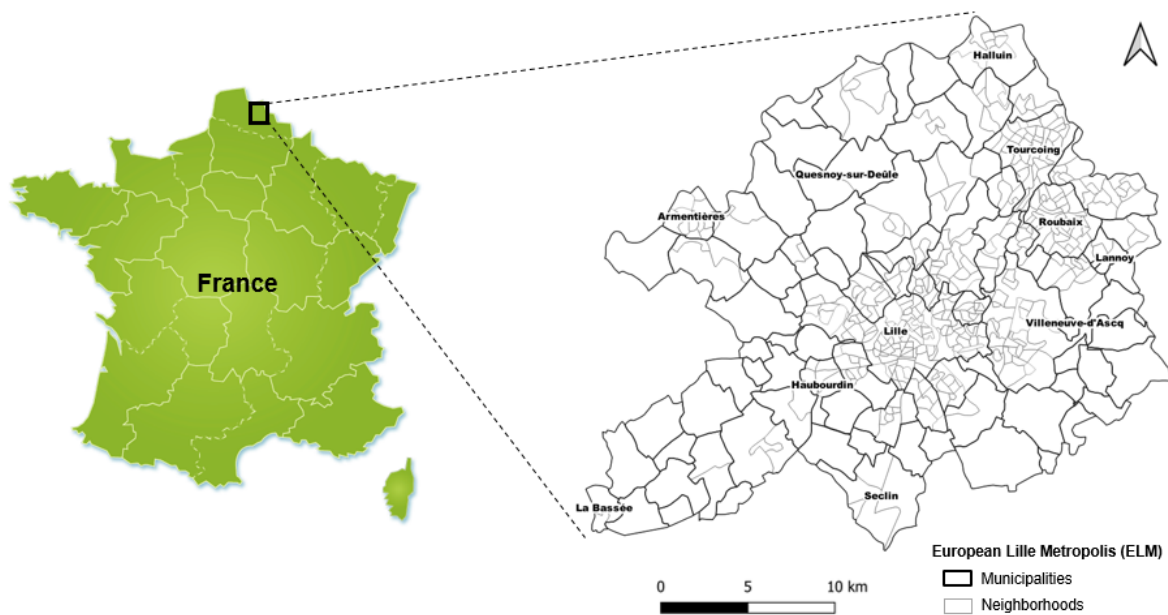


Figure 32 : Article 5 – Figure 1: European Lille Metropolis (ELM), located in the north of France

The spatial unit used for statistical analysis is the IRIS, an administrative unit created by the French national institute for statistics and economics studies (INSEE), where the population generally falls between 1,800 and 5,000 inhabitants. This 'aggregated unit for statistical information' is the most commonly fundamental administrative unit used for the dissemination of infra-communal data. 472 residential IRISs were included in our study, the following exclusion criteria were applied: (i) the IRIS is located in a municipality which was not located in the ELM when the health registry was created, (ii) the business IRIS with more than 1,000 employees, with at least twice as many employees as other residents and (iii) the miscellaneous IRIS corresponding to large surface areas which are sparsely inhabited (*e.g.* leisure parks, forests, ...). Descriptive measures for the population and cover areas are available in Supplementary Materials (Table SM-I, *Annexe 5*).

Incident cases of CHD

The study was performed within the framework of the French MONICA registry (Ducimetière *et al.* (2011). The population sample comprised men and women aged between 35 and 74 and residing in the ELM. Cases of acute coronary syndromes were identified in all the public- and private-sectors hospitals and emergency departments. Multiple sources (discharge letters, computerized lists covering the hospital discharge diagnosis, emergency department computer lists, death certificates, *etc.*) were reviewed and cross-checked to ensure that all relevant information had been included. The acute

coronary syndromes were documented on the basis of the clinical history and hospital records. Clinical diagnoses were obtained from medical records. Only patients with clinical diagnosis of acute coronary syndrome with one of the following denomination were included in the study: acute myocardial infraction, acute coronary syndrome, unstable angina or other episodes of acute coronary insufficiency (symptoms of ischemia associated with ECG abnormalities and/or raised troponins and/or abnormal angiography findings). The present study used data from 2012 to 2016 (*i.e.* 4,512 events). Only incident (*i.e.* first-ever) acute coronary syndrome in patients with no history of acute or chronic CHD were analyzed in the present study (n=4,464 incident cases). All outcomes were geocoded at the home address.

Ethics

Implementation of the French MONICA registry has been approved by the French National Registry Committee (Paris, France). The present study was registered with the French National Data Protection Commission ("*Commission Nationale de l'Informatique et des Libertés*" (Paris, France)). All patients were informed that their data could be used for research purposes, as approved by the research ethics committee. All data were anonymized prior to analysis.

Ecological data

A dataset leading to provide the widest description of the determinants was developed on the basis of their recognized or suspected impacts on CHD. The initial dataset combined 44 multidimensional variables collected from (i) public national databases accessible in Open Data, (ii) open ELM data and (iii) previous studies within this specific territory (Occelli *et al.*, 2020, 2014a). These variables were collected or calculated at the IRIS scale and aimed to describe the multidimensional characteristics of the territory that were related to cardiovascular health. Because a data is basic element allowing to understand a dimension (Brousmiche *et al.*, 2020b), the distance to emergency services is an example of data included in health dimension, the proportion of poor households in economic dimension, the Normalized Difference Vegetation Index in environmental dimension, the marital status in social dimension and the distance to sport facilities in services dimension. Each variable had been classified into resilience and vulnerability according to the impact of an increase in each variable on CHD. A variable belongs to resilience when the increase in the variable can be related to a beneficial impact on cardiovascular health whereas a variable belongs to vulnerability when the increase in the variable corresponds to a deleterious impact on cardiovascular health. More details about the initial dataset can be found in Supplementary Materials (Table SM-II to SM-VI, *Annexe 5*). The inclusion/exclusion criteria are detailed in Brousmiche *et al.* (2020a). Briefly, data to be included was related to a cardiovascular health determinant established or suspected, had to be available for the entire perimeter and reliable, data associated with individual behavior had been excluded.

Statistical analyses

Standardized Incidence Ratios (SIRs)

The Standardized Incidence Ratios (SIR) were computed for each IRIS to assess the spatial distribution of CHD, considering two well-known confounding factors (*i.e.* age and sex). The SIRs were calculated according to the indirect standardization method (Waller and Gotway, 2004), as follows:

$$SIR_j = \frac{O_j}{E_j} \quad \text{where} \quad E_j = \sum_{i=1}^n N_{ij} * T_i$$

With:

- j: the spatial unit,

- O_j : the number of observed cases of CHD in the spatial unit j ,
- E_j : the number of expected cases of CHD in the spatial unit j ,
- i : the stratum index (16 strata):
 - Men aged 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74,
 - Women aged 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74,
- N_{ij} : the population size corresponding to stratum i in the spatial unit j ,
- T_i : the specific incidence rate of the reference population for stratum i .

Due to the fine scale of analysis and the heterogeneous sizes of the IRIS, the SIRs could be influenced by (i) the low frequency of cases and an unequal distribution of population between the IRIS and (ii) the non-inclusion of spatial autocorrelation between adjacent spatial units (Clayton and Kaldor, 1987; Colonna and Sauleau, 2013). The presence or absence of spatial autocorrelations among the IRIS is quantified using Moran's index (Moran, 1948). A value greater than 0 probed using Moran's test indicates the presence of autocorrelation (Cliff and Ord, 1981), the threshold for statistical significance being set to $p < 0.05$. To consider the instability of SIRs related to spatial autocorrelation and a possible smaller number of incident cases in spatial units, Bayesian Poisson regression model were used to smooth the SIRs (Besag *et al.*, 1991).

Development of the complex territorial profiles

This research aimed in a first step to establish complex territorial profiles from a data-set including 44 multidimensional data by a factorial analysis. The methodological framework to assess the territorial characteristics relied firstly on the methodology proposed by Brousmiche *et al.* (2020a) to select variables. The redundant information of the initial dataset was identified by a bivariate analysis. A variable was included if there was at least one correlation coefficient in the range $[|0.2| - |0.6|]$ with another variable. This can lead to (i) obtain a concise description of the initial dataset by summarizing the major information and (ii) introduce in factorial analyses only variables with a minimum correlation structure limiting Principal Components carried by a single variable or strongly influenced by a few redundant variables. Because variables can be expressed in different units or reference systems, a step of standardization according to the Z-score methodology had been therefore performed to ensure that variables are converted into a common scale (European Commission *et al.*, 2008). A Principal Component Analysis (PCA) and a Hierarchical Clustering on Principal Components (HCPC) with a consolidation of classes had been carried out on the 24 selected variables to search similarities within the IRISs and assign IRISs among homogeneous territorial profiles. The Ward's minimum variance method was implemented to reach minimal differences within a class and maximal differences between classes (Ward, 1963). The quantitative variables that best describes each complex profile were identified by the results of the analysis of variance p-value (F-test in one way). Table I presents the distribution by dimension and sub-dimension and according to their affiliation to resilience or vulnerability of the 24 selected variables.

Tableau XXI : Article 5 – Table XXII: Distribution by dimension and sub-dimension of the selected variables in the dataset, each had been assigned to resilience or vulnerability

Dimension	Sub-dimension	Number of data related to vulnerability	Number of data related to resilience
Economics		2	4
	Economic diversity		1
	Economic drive		1
	Economic influence		1
	Employability		1
	Deprivation	2	
Environment		5	3
	Air quality	3	
	Environmental risks	2	
	Natural resources		3
Health		2	1
	Health care facilities	2	1
Services		2	3
	Habitat characteristics	1	1
	Land use planning	1	
	Mobility		2
Social		2	
	Education level	1	
	Social resources	1	
	Total	13	11

Associations between complex territorial profiles and smoothed SIRs

The age- and gender- smoothed SIRs were compared to classes obtained after PCA and HCPC. The significant variables in each class help to define complex territorial profiles that are either related to higher or lower incidence CHD ratios. In a first step, an ANOVA was conducted to identify if the smoothed SIRs in classes are equal. Tukey Honestly Significant Difference (HSD) test was then computed for pairwise comparisons. Within each class, an analysis of the means of significant variables compared to what was observed within the ELM allowed to identify the determinants which can influence a higher or lower incidence.

Software

R software was used to carry out all statistical analyses (version 4.1.2). Spatial statistics were performed using Arc GIS software (version 10.7), maps were produced using Q-GIS software (version 3.16.11-1). The threshold for statistical significance was set to 0.05.

Results

Geographical variation in the incidence of CHD

The median age of incident CHD cases was 59 years [61-66] and 71.8 % (3,206 incident cases) have concerned men. The crude CHD incidence rate was 186 per 100,000 inhabitants within the geographical area studied, 279 per 100,000 men and 100 per 100,000 women. The age and gender SIR allowing to consider these confounding factors has a Global Moran's Index of 0.182 ($p < 0.001$) which

corresponds to a moderate spatial autocorrelation (more details can be found in supplementary results, Figures SM-1 and SM-2, *Annexe 5*). After the use of Bayesian Poisson regression model, the age- and gender- smoothed SIRs ranged from 0.561 to 2.082, the median being 1.017 [0.859;1.188]. Figure 2 presents the spatial distribution of smoothed SIRs within the studied perimeter.

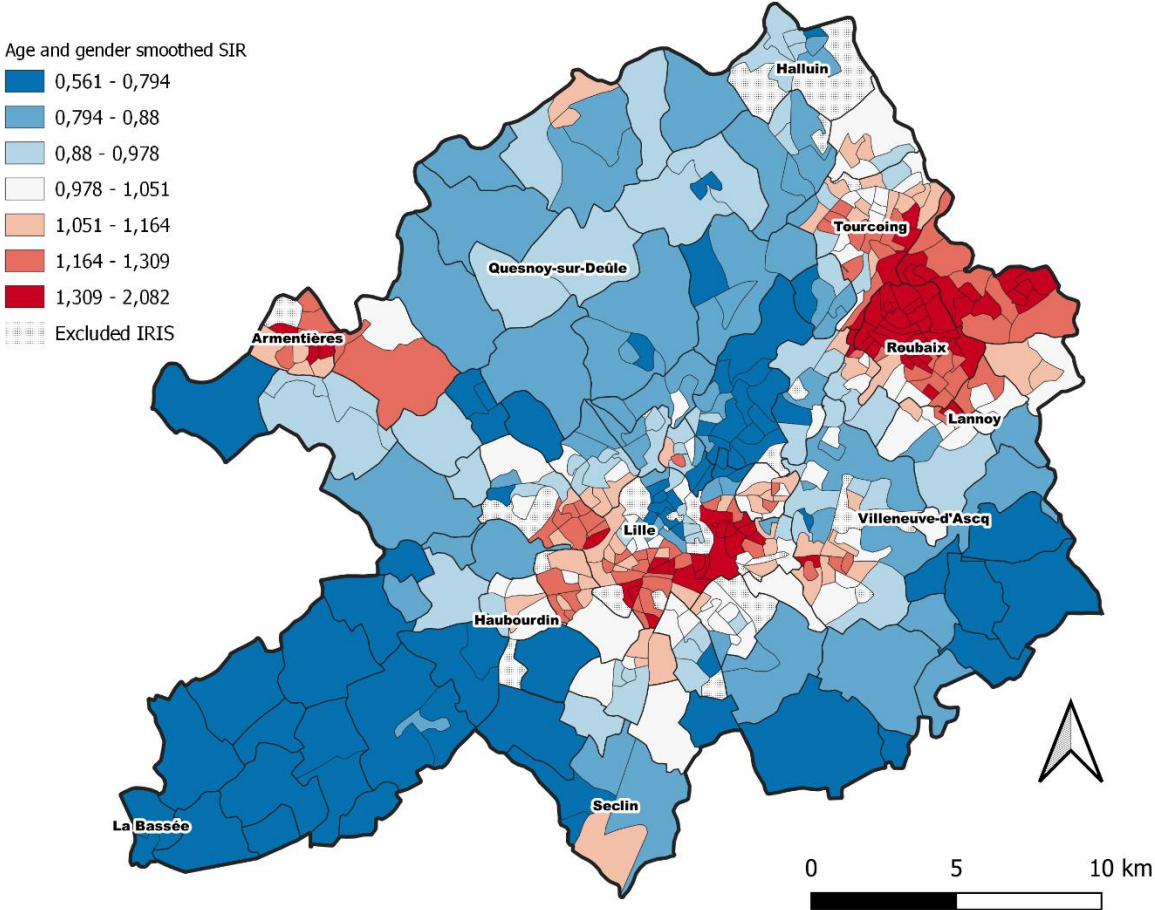


Figure 33 : Article 5 – Figure 2: Age and gender smoothed Standardized Incidence Ratios for CHD (2012-2016)

The IRISs that present higher SIRs (*i.e.* > 1.000) were for the most part located in the North-East, the Center and the North-West of the ELM within the main municipalities of the ELM (*i.e.* Lille, Armentières, Roubaix, Tourcoing, Wattlelos). 69 IRISs belong to the class of highest incidence of CHD (*i.e.* SIRs between 1.309 and 2.082), these ratios expressing that the morbidity is higher by 30.9 to 108.2 % compared to the reference population that was considered for SIRs calculation. The IRISs displaying lower SIRs (*i.e.* < 1.000) were rather located on the outskirts of the large municipalities of the ELM, in IRISs and municipalities that are less densely populated, mainly from South-West to South-East. Although the spatial distribution of the SIRs seemed to be shaped by urbanization, the correlation between SIRs and population density was moderate (Pearson rho = 0.395, p value < 0.001).

Complex territorial profiles

After PCA and HCPC, 3 classes have been obtained and each IRIS of the ELM has been distributed among these complex territorial profiles (Figure 3). Thus, 217 IRISs (45.98 %) belong to complex profile 1, 158 to complex profile 2 (33.47 %) and 97 (20.55 %) to complex profile 3.

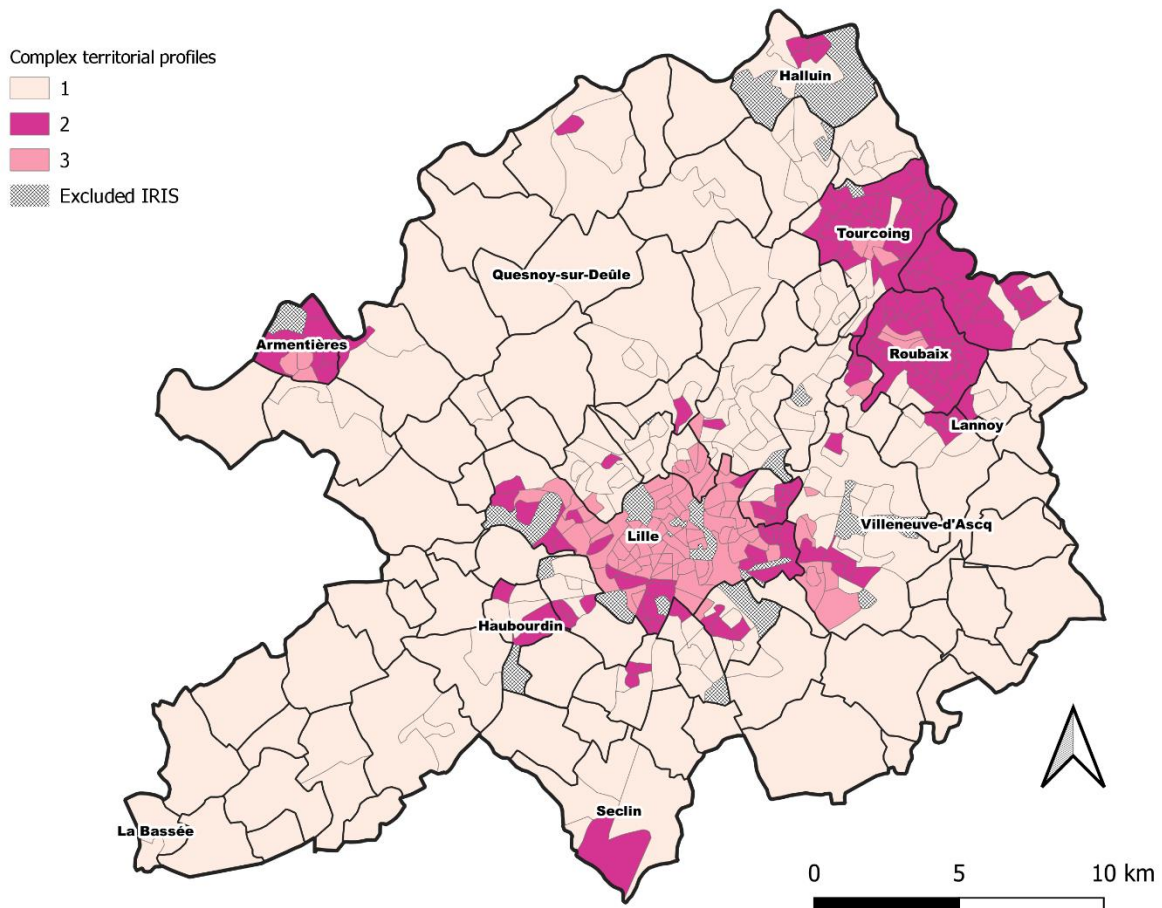


Figure 34 : Article 5 – Figure 3: Spatial distribution of classes obtained after PCA-HCPC based on the 24 selected variables

Class 1 (in light orange on Fig. 3) was located in the outskirts of three main areas (Lille, Roubaix-Tourcoing, Armentières). Class 2 (in purple on Fig. 3) and class 3 (in salmon-pink on Fig. 3) were essentially located in the main municipalities of the ELM.

Links between territorial characteristics and SIRs

The boxplots of smoothed SIRs according to the class highlighted the differences between complex territorial profiles (Fig. 4) ($p < 0.05$). More details can be found in Supplementary Results about the Tukey HSD test which have highlighted significant differences between territorial profiles (Table SM-IX, Figure SM-4, *Annexe 5*).

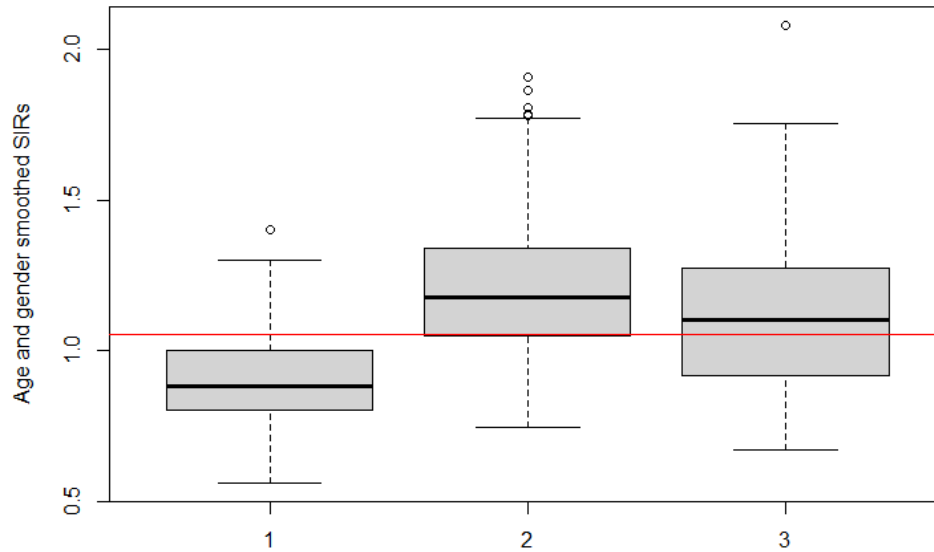


Figure 35 : Article 5 – Figure 4: Averaged age and gender smoothed SIRs for CHD according to class obtained after PCA and HCPC; line in red represents the overall mean of SIRs on ELM

The common and specific characteristics of each complex territorial profile are highlighted by the significant variables within complex profiles (Table II). For each selected variable, means within each profile were compared to what was observed on the ELM. This allowed to identify the levels of predominance of certain determinants which differ according to the complex profiles which were characterized by an under- or an over-incidence.

Tableau XXIII : Article 5 – Table II: Mean (standard deviation) of each of the 24 variables within the complex territorial profiles obtained after PCA and HCPC. A mean (sd) or a variable in grey was non-significant while significant means (sd) in class are marked with an asterisk “*” (i.e. $p < 0.05$)

	Profile 1	Profile 2	Profile 3	ELM
Number of IRIS	217	158	97	472
Mean of age and gender smoothed SIRs [CI 95%]	0.895 [0.876-0.914] *	1.225 [1.187-1.263] *	1.119 [1.064-1.174] *	1.105 [1.028-1.075] *
“Economics” dimension				
Deprivation Index	-0.9720 (1.0893) *	1.8629 (1.3115) *	-0.5167 (1.5231) *	0.0705 (1.8011)
Proportion of poor households	3.73 (4.86) *	<i>7.77 (4.87)</i>	14.36 (13.45) *	7.27 (11.38)
Share of the labor force in the total population	47.23 (3.75) *	42.36 (5.69) *	49.72 (7.20) *	46.11 (3.26)
Evolution of the share of the labor force in the total population	<i>0.13 (2.93)</i>	1.04 (3.05) *	-0.64 (3.93) *	0.28 (5.99)
Share of the labor force living/ working in the same municipality	18.56 (6.97) *	32.96 (13.40) *	47.35 (14.73) *	29.30 (15.86)
Shannon-Weaver index on the representation of socio-professional categories	0.6325 (0.0303) *	0.5835 (0.0459) *	<i>0.6039 (0.0343)</i>	0.6102 (0.0429)
“Social factors” dimension				
Share of working-age population with few or no qualifications	<i>4.69 (1.11)</i>	5.21 (1.49) *	3.21 (1.33) *	4.56 (1.48)
Proportion of single households	29.31 (8.09) *	32.42 (7.50) *	52.19 (11.63) *	35.06 (12.43)
“Environment” dimension				
Annual mean concentrations NO ₂	32.96 (8.50) *	32.72 (6.03) *	38.62 (7.56) *	34.05 (7.91)
Dust deposition	<i>1954.40 (221.17)</i>	1916.41 (176.41) *	2037.13 (137.71) *	1958.68 (196.40)
<i>Mean Impregnation Ratio of trace elements</i>	<i>1.76 (1.29)</i>	<i>1.92 (1.75)</i>	<i>1.99 (0.28)</i>	1.86 (1.34)
<i>Share of population exposed to noise throughout the day</i>	<i>27.14 (19.69)</i>	<i>27.10 (23.44)</i>	<i>28.46 (26.57)</i>	27.40 (22.48)
Share of population exposed to noise at night	66.08 (22.45) *	<i>58.99</i>	53.30 (28.93) *	61.08 (25.86)
Share of population with access to green space within 400m	75.15 (25.81) *	<i>64.38 (37.03)</i>	54.95 (39.75) *	67.39 (33.91)
Normalized Difference Vegetation Index	0.49 (0.06) *	0.39 (0.06) *	0.34 (0.07) *	0.43 (0.09)
Monetization of ecosystem services per capita	210.86 (340.11) *	78.34 (50.55) *	70.91 (51.17) *	137.74 (243.17)
“Services” dimension				
Share of main residences built before 1970	44.77 (16.76) *	68.38 (17.40) *	49.75 (18.14) *	53.70 (20.25)
Distance to sport and leisure facilities	545.61 (224.22) *	342.91 (140.26) *	313.52 (116.66) *	430.06 (209.68)
Share of main residences built after 2006	10.35 (8.16) *	5.42 (6.32) *	<i>7.32 (6.53)</i>	8.08 (7.58)
Share of bicycle paths in the total road	76.24 (10.82) *	<i>78.91 (13.52)</i>	<i>78.97 (12.06)</i>	77.70 (12.10)
Share of physically active commute modes	17.41 (7.99) *	<i>29.37 (11.72)</i>	49.50 (11.12) *	28.01 (15.74)
“Health” dimension				
Shannon-Weaver index on the representation of medical and paramedical professions	0.6107 (0.2750) *	0.3845 (0.2635) *	0.7062 (0.2400) *	0.5546 (0.2927)
Travel time to emergency services	10.25 (3.43) *	<i>8.98 (3.31)</i>	7.77 (3.44) *	9.31 (3.52)
Travel time to general practitioners	1.79 (0.63) *	<i>1.78 (0.56)</i>	1.43 (0.44) *	1.72 (0.59)

In complex profile 1, the mean of age- and gender- smoothed SIRs for CHD between 2012 and 2016 was 0.895 [CI 95%: 0.876-0.914]. Population in this profile 1 was mainly advantaged (lowest FDep, proportion of poor households and single households), with the highest mean of economic diversity (in terms of socio-professional categories) and share of labor force in the total population. The annual mean concentration of NO₂ was below the ELM mean and data related to environmental resilience (accessibility to green spaces, NDVI, monetization of ecosystem services per capita) presented the highest results compared to what is observed on ELM. Concerning the characteristics of habitat, complex profile 1 showed the highest share of main residences built after 2006 and the lowest share of main residences built before 1970. The specific diversity of medical and paramedical professions provided at the professional activity location was above the ELM mean. Nevertheless, certain variables presented results whose mean could influence cardiovascular health. This could be partly due to the localization of IRISs in the outskirts of the main municipalities of ELM. The share of population living and working in the same municipality as well as the share of physically active commute modes were indeed the lowest in this profile, while the distances to the nearest (i) sport and leisure installations, (ii) emergency services and (iii) general practitioners were the highest. IRISs belonging to this profile were also characterized by a highly share of population exposed to noise at night and the lowest proportion of bicycle paths in the total public road within the spatial units.

The mean of the smoothed SIRs in complex profile 2 is 1.225 [1.187-1.263] which means that the 158 IRISs presented a higher coronary morbidity of 22,5 % between 2012 and 2016 compared to the reference population. The determinants that could have a deleterious influence on cardiovascular health were related to the socioeconomic vulnerability, with the highest (i) deprivation index and (ii) share of working-age population with few or no qualification, but also to the socioeconomic resilience with the lowest (i) share of labor force in the total population and (ii) economic diversity. The specific diversity of medical and paramedical professions presented also the lowest mean within the ELM. The characteristics of the habitat in this class were the opposite of what was observed in profile 1, namely the share of main residences built before 1970 was the highest in the ELM while the share of recent main residences was the lowest. Due to their localization in densely populated municipalities, the variables related to environmental resilience (*i.e.* NDVI, monetization of ecosystem services) showed relatively low means. However, some determinants have presented in the same time results that could have a positive influence on cardiovascular health. These determinants were related to the economic resilience (*i.e.* the highest (i) evolution of the share of the labor force and (ii) share of population living and working in the same municipality), to the services resilience with the lowest distance to distance to sport and leisure facilities and to the social resilience with the lowest proportion of single households. Concerning the environmental vulnerability, the annual mean concentrations of NO₂ and dust deposition displayed the lowest mean observed within the ELM.

The mean of the smoothed SIRs in complex profile 3 for its part is 1.119 [1.064-1.174]. All the variables belonging to “Services” dimension (*i.e.* share of main residences built before 1970, accessibility to sport equipment, share of physically active commute modes) and “health” dimension (*i.e.* specific diversity of medical and paramedical professions, accessibility to emergency and general practitioners) presented means that can positively influence incidence of CHD. Concerning the socioeconomic variables, results were more contradictory. The population was relatively advantaged, with the lowest rate of “uneducated” population but in the same time there were the highest proportion of poor households and single households in this class. Although the economic drive there was negative, the share of labor force and the labor force living and working in the same municipality were the highest. Regarding the “environment” dimension, most of the means can be related to a negative influence on CHD. The annual mean concentrations of NO₂ and dust deposition

presented the highest means on the ELM. Data related to environmental resilience displayed additionally the lowest means in this profile. However, the share of population exposed to noise at night presented the lowest result compared to the what was observed on the ELM.

Discussion

The integrative strategy developed in this study aimed to propose an aggregated view of the multidimensional determinants which may be associated with the risk of CHD. We assumed that this more global and diversified characterization of a territory leads to better understand the causes of cardiovascular health inequalities which are not the consequence of only one determinant or one dimension. The cardiovascular health within a territory results from the simultaneous interactions of determinants which can either positively or negatively influence the incidence of CHD. In our study among the 24 selected variables to develop the complex territorial profiles, ten were identified as discriminating because of significative differences among the three complex territorial profiles. Five were related to vulnerability and might have a negative influence on incidence of CHD within our study perimeter (*i.e.* socioeconomic deprivation, proportion of single households, annual mean of NO₂, proportion of main residences built before 1970, distance to sports and leisure equipment). The five others were related to resilience and might influence positively the incidence of CHD. These resilience determinants were the share of the labor in the total population, the share of the labor that live and work in the same municipality, the NDVI, the monetization of ecosystem services and the specific diversity of medical and paramedical professions at the professional activity location (described through the Shannon-Weaver index from a transdisciplinary adaptation). Our results emphasized the importance to consider in a territory not only recognized or suspected determinants that may influence negatively the incidence of CHD but also determinants that might have a positive impact.

Over incidence and complex territorial profiles

Our results showed that highest SIRs observed in complex profile 2 seemed to be rather related to socioeconomic conditions. Significant associations between socioeconomic status and CHD in high-income countries have been previously highlighted in literature, mainly through the use of basic determinants such as educational attainment, income level or neighborhoods socioeconomic factors (Bevan *et al.*, 2022; Schultz *et al.*, 2018). The share of the single households in this profile was relatively low which is in accordance with previous studies. The marital status has indeed been mentioned as a source of beneficial impact for cardiovascular mortality, with possibly stronger effects for men (Dhindsa *et al.*, 2020; Kim *et al.*, 2018; Wong *et al.*, 2018). Nevertheless, the cardiovascular inequities would solely explain about half of the excess mortality in low socioeconomic groups, with an association which seems stronger for men (Dahlgren and Whitehead, 2007; Rey *et al.*, 2011). In previous studies, cardiovascular health disparities across neighborhoods have also been shown to be shaped by physical environment (Hu *et al.*, 2020). Results in this complex profile have highlighted also a lack or an insufficiency of resources to tackle the coronary heart events. The environmental conditions in this complex profile were contrasted: on one hand the annual mean concentration of NO₂ and dust deposition are the lowest within the ELM. On the other hand, data related to environmental resilience was mainly unfavorable which reflected the rather urban location of this profile. In our study, the annual mean concentration of PM₁₀ (34.21 µg/m³) has fallen into our exclusion criteria. However, the annual mean concentrations of NO₂ which was included in PCA-HCPC and PM₁₀ over the ELM were almost perfectly correlated (*i.e.* cor = 0.99). The biomonitoring approach to characterize air quality provides an additional information to physical and chemical data (Cuny, 2012). This original approach although less mentioned in the literature could make it possible to better tract pollution over several years than a single measure (Occelli *et al.*, 2020).

The over-incidence in complex profile 3 could be related to socioeconomic and also environmental conditions. The socioeconomic results in this profile (*i.e.* higher rates of poor households and also of single households) have to be qualified having regard to the presence of the 100,000 students registered at the University of Lille. These students live for the most part alone and in private accommodation near the various university campuses (Lille, Villeneuve d'Ascq, Roubaix and Tourcoing). The over-incidence in this complex profile would therefore seem to be more related to environmental conditions, with (i) the highest results of atmospheric pollution and (ii) the lowest greenness compared to what was observed on ELM. Currently, air pollution is the most recognized environmental risk factor for cardiovascular health and some populations (*e.g.* elderly, diabetics) are more susceptible to these effects (Davoodabadi *et al.*, 2019; Luo *et al.*, 2017; Tibuakuu *et al.*, 2018). Among the multiple biological mechanisms, atmospheric pollution leads to oxidative stress and systemic inflammation affecting the cardiovascular system and the development of atherosclerosis which contribute to CHD (Wolf *et al.*, 2021). The annual mean concentrations of NO₂ within the ELM (34.05 µg/m³) are far above the World Health Organization limit (*i.e.* 10 µg/m³), even in profile 2 in which the annual mean is the lowest within the ELM (*i.e.* 32.72 µg/m³). The environmental data related to resilience displays the lowest results in profile 3 which also contributes to the unfavorable environmental conditions. A recent review has highlighted that (i) the links between green spaces and cardiovascular diseases are still inconsistent and (ii) the indicators measuring greenness are various (Liu *et al.*, 2022). The use of the NDVI is however the most commonly index exploited to analyze the association of the local greenness with cardiovascular diseases, and an inverse relationship exists between this greenness indicator and the coronary outcomes (Aerts *et al.*, 2020; Guo *et al.*, 2022; Hu *et al.*, 2022; Liu *et al.*, 2022). It is also important to point out that some studies have suggested different susceptibility according the sex and that the beneficial impacts of green spaces could be more advantageous for men (Bambra *et al.*, 2015). Although knowledge between greenness and beneficial impacts on health still needs to be deepen, this constitutes a lever for action for stakeholders that aim to improve, *inter alia*, cardiovascular health within their territory.

Urban rural gap

Variables related to “health” and “services” dimensions present contrasted results that are partly inconsistent with other studies on the same subject in the literature. We had highlighted significant lower distances to the nearest sport and leisure facilities in profiles with over-incidence, even in complex profile 2 in which population was highly deprived. In the works of Petrovic *et al.* (2018) it had been underlined that there were fewer opportunities for a healthy life in deprived neighborhoods. Others studies have associated the local built environment (*e.g.* transportation infrastructure, urban design, ...) and the neighborhood social environment with a better cardiovascular health (Briggs *et al.*, 2019; Tamura *et al.*, 2019). The available literature suggests also that (i) the physical activity in a neighborhood may be influenced by the availability and/or the accessibility of sports and leisure facilities and (ii) the level of physical activity and the urbanization seems to follow inverse gradients (Miao and Wu, 2016; Potvin *et al.*, 2010). Within our studied perimeter, the sport and leisure facilities were mainly located in the outskirts of the main municipalities which are densely populated. This rather urban location of sport facilities was also reflected by the significantly higher results within profile 1 where population had a greater distance to the nearest installation. In our study, higher distances to health services were also observed in profile where deprivation was lowest whereas Briggs *et al.* (2019) had shown neighborhoods with lower income could present fewer health services. Greater share of population which use active commute mode were found in complex profiles 2 and 3, these profiles being located in more urban neighborhoods with over-incidence of CHD. Laverty *et al.* (2013) and Patterson *et al.* (2020) have conversely emphasized that population with active commute modes was more likely to present

better results in terms of overweight, diabetes and hypertension than population who used private transport. Moreover, van den Berg *et al.* (2016) have pointed out that the use of active commute modes, particularly public transport, would decrease feelings of loneliness by greater interactions with others. These results emphasize 'atypical' characteristics of the ELM and an urban/rural gap. According to Briggs *et al.* (2019), unexpected results might be observed between urban and rural areas in which the environmental and territorial influences can present strong differences. Although the SIRs seem to follow an urban/rural gradient, the correlation with the population density is only significant within the overall studied perimeter and in profile 1. For higher population densities, the incidence of incident coronary events appeared to be independent of this factor.

Non-discriminatory data

The Mean Impregnation Ratio (MIR) allowed to transcribe multi-metallic atmospheric contamination (Occelli *et al.*, 2016). Although a growing body of evidence related to the metal exposure and their impacts on cardiovascular diseases is available (Rajagopalan and Landrigan, 2021), there is to date no consensus on metals to be considered. Most of the studies related to the impact of metal exposure on cardiovascular health mainly exploits metal concentrations in particles. Significant associations were found between lead, selenium and arsenic and Ischemic Heart Disease in Thurston *et al.* (2016). Guo *et al.* (2022) have found higher excess risk for CVD related to lead, cadmium, selenium, thallium and aluminum while Badaloni *et al.* (2017) have associated copper, zinc, iron and nickel with Ischemic Heart Disease. The integration of the multi-metallic atmospheric burden from biomonitoring data had overcome this lack of consensus on metals. In the previous study within the same perimeter, rises of incidence risks had been highlighted for cadmium, palladium, nickel and copper (Occelli *et al.*, 2020). These metals, considered in tertiles, were evaluated through ecological regression and displayed higher risk relative with the CHD-incidence. Although our study was carried out both on the same perimeter and spatial unit, the multi-metallic burden did not reveal any significant results among the three complex territorial profiles. This could be partly due to the use of the MIR which aggregated data related to 16 metals with heterogeneous results (*i.e.* only four displayed significant results with CHD-incidence while the 12 others not), a higher number of cardiovascular determinants considered in the methodology and/or to the different methodologies developed.

Concerning data related to environmental noise, there is no real consensus in the literature neither on a noise level nor on a specific indicator (Andersson *et al.*, 2020). In their works on road traffic noise and its impact on cardiovascular health, Cai *et al.* (2018) have considered the L_{DEN} above 55 dB(A), Begou *et al.* (2020) have divided the L_{DEN} into 6 dB categories (<55, 55-59, 60-64, 65-69, 70-74, >75), Andersson *et al.* (2020) the L_{DEN} above 60 dB while for Tobías *et al.* (2015) used the value of 60 dB(A) as a No Observed Adverse Effect Level (NOAEL). WHO (2011) recommends however to distinguish the noise exposure during the day and at night. In our study, we retained the value of 60 dB(A) for L_{DEN} and 45 dB(A) for L_N (from 22:00 to 6:00) to assess the part of population exposed to noise (road traffic, railway, aircraft, industrial facilities). The value of 45 dB(A) has been considered to follow the WHO Environmental noise guidelines for the European Region (WHO, 2018b). Noise throughout the day was not identified as significant in any of the three complex territorial profiles unlike noise at night which have displayed significant results in (i) profile 1, where the highest share of population exposed to night noise was observed and contrasting its rather favorable environmental conditions and (ii) profile 3 which have revealed the lowest results.

Strengths and limitations

In ecological studies, it is fundamental to have an exhaustive list of cases of a pathology in a given territory to formulate etiologic hypothesis. In our study the collaboration with the French MONICA

registry had made it possible to obtain exhaustive and robust data for CHD on the ELM perimeter. This allowed us to improve knowledge about the positive or negative influence of determinants on CHD, whether they were related to resilience or vulnerability. According to Jomier and Lassarade (2021), the use of health registries should be developed to better assess links between the incidence of a pathology and environmental and more broadly territorial factors that are suspected to impact this. Because ecological studies rely on the use of aggregated data, individuals may be misrepresented making impossible to directly transpose results of ecological studies to individuals. The use of a very fine spatial unit (*i.e.* the IRIS) in our study allowed us to reduce this well-known phenomenon of ecological fallacy without however being able to avoid it. The use of this very fine spatial unit has empowered also to accurately characterize the multidimensional determinants involved in cardiovascular health. The development of this methodology at very fine scale allowed to develop complex territorial profiles that combined in the same time determinants that were related to vulnerability and resilience. This offered a contrasted depiction of the phenomena that may be intertwined. Furthermore, the relatively long latency period for CHD provides for health care professionals and stakeholders an opportunity to (i) promote health policies to detect, treat and prevent the individual risk factors of CHD (Rajagopalan and Landrigan, 2021; Vasan *et al.*, 2020) and (ii) implement territorial policies to enhance the overall environmental conditions.

This study has however limitations which form areas for improvement in further researches on influence of territorial characteristics on CHD. Some lifestyle factors (smoking, diabetes, size and weight allowing to calculate the Body Mass Index (BMI)) and individual risk factors for CHD (high blood pressure, dyslipidemia) could not be considered in our analysis. Data are recorded in the registry since 2013 and display a relatively high rate of missing values (*e.g.* 23.7 % for diabetes to 31.0 % for BMI). Although there is a V-like relationship between the cardiovascular mortality and temperature, data related to this environmental risk factor could not be explored in our analysis. Both cold and hot temperature would exacerbate air pollution effects on health through different mechanisms (Argacha *et al.*, 2018; Chen *et al.*, 2019; Rajagopalan and Landrigan, 2021). Moreover, a greater vulnerability to temperatures have been observed among single households or with low educational level (Tibuakuu *et al.*, 2018). Within the ELM, data related to (i) temperature (annual mean over 30 years) was only available at the municipality level and displays a relatively low variability due to the restricted analysis perimeter (the mean is 10.6 °C and the standard deviation is 0.08) and (ii) urban heat island was available only for less than half of the study perimeter. A last key limitation of neighborhood research relies on the use of the home address as the geographical unit in which population is exposed to environmental conditions (Tamura *et al.*, 2019). Most of studies assess the exposition at the place of residence although population only spend less than one-third of their daily life at home (Lan *et al.*, 2022). Population may indeed move in different places which could differ according to the daytime or the day. This could lead to over- or under-estimate association between an environmental exposition determined by the place of residence and a health outcome (Kwan, 2018). It is mainly due to the spatial and temporal uncertainties linked to the lack of information about where population experienced the environmental influences (Kwan, 2012a; Tamura *et al.*, 2019). In epidemiological studies, dealing with this effect of Neighborhood Effect Averaging Problem (NEAP) is crucial when certain environmental contexts highly influenced by place are assessed (*e.g.* green spaces, noise or air pollution (Kwan, 2018)). In our study, the variable related to the population living and working in the same municipality aims to identify in which extent the population tends to move across others municipalities to work. Almost half of the population in complex profile 3 work and live in the same municipalities while in complex profile 1 this represents less than 20 %. Because workplace data is missing, environmental exposures of population within complex profile 1 may have been therefore underestimated.

Conclusion

The purpose of this research was to develop complex territorial profiles at a very small scale that aim to describe the multidimensional characteristics involved in cardiovascular health. Since determinants can have simultaneous beneficial or deleterious effects on cardiovascular health, the developed methodology has considered them together because of their possible interactions. 3 profiles have been highlighted: the first was characterized by (i) a significantly lower incidence rate of coronary heart diseases after being adjusted on age and gender and (ii) a higher number of determinants with a positive influence on incidence. Although the second and third profiles were characterized by significantly higher rates of coronary heart diseases, the territorial characteristics were noticeably different. The second profile presented higher number of determinants that had a negative influence on CHD, with socioeconomic deprivation and contrasted environmental conditions. The atmospheric pollution and the impregnation of the overall environment there were the lowest while determinants related to greenness displayed low results. The over-incidence in the third profile tended to be more associated with higher degraded atmospheric conditions whereas determinants related to others dimensions underlined potential beneficial effects (lower distance to emergency services and to sport and leisure facilities, population mainly socioeconomically privileged, higher diversity of medical and paramedical professions). Our results brought a broader view of the environmental and territorial issues that need to be tackled to improve cardiovascular health. Future studies should consider lifestyle or individual factors and the effects of temperature which might provide an even deeper examination of the overall determinants involved in CHD and which are closely intertwined.

Discussion Partie II.2 : profil territorial complexe et maladies coronariennes

Les zones de sur-incidence issues du calcul des SIRs lissés pour la période 2012-2016 présentent une distribution spatiale globalement centralisée sur certaines agglomérations de la MEL (Armentières, Lille, Roubaix, Wattrelos). Ces localisations coïncident partiellement avec ce qui avait été souligné sur la période 2008-2011 (Occelli *et al.*, 2020). Sur les périodes 2008-2011 et 2012-2016, la couronne Sud de la ville de Lille, la quasi-totalité des villes de Roubaix et Wattrelos présentent des morbidités supérieures à celles observées sur la MEL sur les deux périodes temporelles considérées. En revanche, certains IRIS au Sud-Ouest (*i.e.* Fournes-en-Weppes, Hantay, Sainghin-en-Weppes), au Nord de la MEL (*i.e.* Halluin) ainsi qu'au Nord-Est de la MEL (*i.e.* Tourcoing) pour lesquels une morbidité supérieure de 11 à 50 % avait été mise en évidence sur la période 2008-2011 présentent sur 2012-2016 une morbidité inférieure de 6 à 44 % par rapport à la population de référence (*i.e.* déterminée au sein des 85 communes de la MEL présentes à la date de création du registre MONICA). A l'inverse, certains IRIS des communes d'Armentières (*i.e.* Nord-Ouest de la MEL) et Wattrelos (*i.e.* Nord-Est de la MEL) présentaient sur 2008-2011 une morbidité inférieure de 10 à 44 % par rapport à la population de référence alors que sur 2012-2016 la morbidité coronarienne y est supérieure de 5 à 31 %.

Les travaux menés sur la période 2012-2016 ont dans un premier temps mis en évidence des IRIS pour lesquels le contexte socio-économique est connu comme défavorable au sein de la MEL (couronne au Sud de la Ville de Lille, Roubaix, Tourcoing, Armentières). Les facteurs socio-économiques jouent en effet un rôle important dans la survenue des maladies coronariennes (Havard, 2008; Kastorini *et al.*, 2015). De plus, ils sont généralement considérés comme variable d'ajustement dans les études de santé environnement, au même titre que des déterminants plus classiquement mentionnés (Saib, 2015). La mesure du niveau socio-économique du lieu de résidence des populations constitue un enjeu majeur en santé au regard de l'absence possible de données individuelles (Ghosn, 2018b). Pour intégrer la défaveur socio-économique dans la détermination des profils territoriaux complexes, deux indices composites ont été envisagés : le *French Deprivation index* (FDep) développé par Rey *et al.* (2009) et le *french European Deprivation Index* (EDI) proposé par Pornet *et al.* (2012). Le FDep est positivement associé à la mortalité standardisée et permet de retranscrire les inégalités sociales de santé de manière homogène entre les zones urbaines et rurales (Ghosn, 2018b; Windenberger *et al.*, 2012). Le *french* EDI propose quant à lui un indice composite socio-économique interculturel avec des variables sélectionnées en fonction des perceptions spécifiques du désavantage social de chaque pays (Ghosn, 2018b). Au sein de la région Hauts-de-France, ces deux indices de défaveur socio-économique s'avèrent fortement associés eu égard leur coefficient de corrélation significatif ($\rho = 0,74$, Tableau II dans Brousmiche *et al.* (2020a)). Néanmoins, certaines variables de l'EDI peuvent selon les zones géographiques ne pas être en lien avec une défaveur (*e.g.* le fait de ne pas posséder une voiture dans une unité spatiale très urbanisée avec une forte accessibilité/disponibilité aux transports en commun n'est pas nécessairement synonyme de défaveur). Le FDep a donc été privilégié pour sa meilleure adaptabilité vis-à-vis de la dichotomie urbain/rural (Ghosn, 2018b). Lors de la détermination des profils complexes des IRIS de la MEL, la classe avec une morbidité significativement supérieure de 18,7 % à 26,3 % présentait un FDep moyen le plus défavorable sur la MEL (classe 2) tandis que la classe où la morbidité coronarienne est significativement inférieure affiche une défaveur socio-économique faible (classe 1). Ces résultats correspondent (i) aux travaux précédemment menés sur 2008-2011 où les SIRs lissés avaient significativement été associés à la défaveur socio-économique et (ii) aux conclusions de la littérature sur son influence sur la survenue des maladies coronariennes.

Les profils complexes ont également permis de souligner que d'autres aspects socio-économiques pouvaient être également défavorables au sein des classes présentant une sur-incidence coronarienne. Selon Kastorini *et al.* (2015) et Lahelma *et al.* (2004), il apparaît en effet important de considérer plusieurs déterminants en lien avec la défaveur socio-économique afin d'en proposer une caractérisation plus étendue. Les différents facteurs socio-économiques (*e.g.* revenus, éducation) peuvent de plus présenter des interrelations, les uns pouvant interagir avec les autres, les mécanismes par lesquels ils sont susceptibles d'impacter les inégalités de santé pouvant être différents selon les déterminants considérés (Kastorini *et al.*, 2015). La part de population âgée de plus de 15 ans pas ou peu diplômée¹² pouvant refléter une sortie précoce du système éducatif est de la même manière la plus élevée au sein de la classe 2. Ce résultat concorde avec les données de la littérature où l'éducation avait été mise en évidence comme une composante importante du gradient social : plus les populations sont diplômées, plus ces populations déclarent un meilleur état de santé avec de moindres limitations d'activité et de maladies chroniques (Bauwelinck *et al.*, 2021; Danet, 2010). De plus, le type de profession (ou la catégorie socio-professionnelle), la satisfaction de la situation financière et également certaines habitudes de vie (*e.g.* alimentation, tabac) sont des paramètres sur lesquels le niveau d'éducation a une influence significative (Kastorini *et al.*, 2015). Bien que la part de population pas ou peu diplômée tende à se rapprocher du FDep qui agrège notamment la part de bacheliers dans son calcul, le coefficient de corrélation entre ces deux variables s'est avéré modéré (*i.e.* $\rho = 0,305$) ce qui souligne la complémentarité des informations apportées dans l'établissement des profils complexes. Néanmoins, la classe 3 où une sur-incidence significative a également été mise en évidence présente des résultats nuancés par rapport aux aspects socio-économiques. Les populations de cette classe présentent (i) une défaveur socio-économique relativement faible, (ii) la part de population peu diplômée la plus faible de la MEL et (iii) les parts de ménages pauvres et vivant seuls soient les plus élevés. Ces résultats spécifiques pourraient être expliqués en partie par la présence de campus universitaires (et donc une part plus importante d'étudiants). Les principaux facteurs de risque de sur-incidence au sein de ce troisième profil complexe semblent de ce fait moins en lien avec des aspects socio-économiques. Toutefois, les facteurs psychosociaux (et notamment le fait de vivre seul) peuvent avoir un impact sur les maladies coronariennes (Dhindsa *et al.*, 2020; Wong *et al.*, 2018; Zhang *et al.*, 2011). Ils impliqueraient en effet des mécanismes en lien avec l'appauvrissement des relations sociales ce qui induirait un sentiment de solitude et de stress chronique (Steptoe, 2015). Les travaux de Silventoinen *et al.* (2022) ont souligné que le type d'union pouvait également avoir un impact, les populations mariées sans historique de divorce présenteraient de plus faibles risques de maladies coronariennes par rapport aux populations en union libre. Le statut civil ou encore la part de population vivant en couple avaient été intégrés dans les travaux de Bauwelinck *et al.* (2021) et de Héritier *et al.* (2018) en tant que facteurs de confusion dans l'analyse du lien entre le bruit et différences pathologies cardiovasculaires. Zhang *et al.* (2011) avaient quant à eux mis en évidence que les populations non mariées présentaient plus d'infarctus du myocarde fatals. Les résultats du premier profil complexe concernant la part de ménage vivant seul concordent avec les données de la littérature au regard de la plus faible part de population vivant seule dans cette classe de sous-incidence significative.

En s'appuyant sur les travaux présentés dans le chapitre 1, des déterminants territoriaux dont les aspects ne reposent pas uniquement sur la vulnérabilité ont été intégrés également dans les profils complexes qui ont ensuite été mis en relation avec l'incidence standardisée des maladies coronariennes. En effet, la prise en considération de la diversité économique et culturelle constitue un des piliers sur lequel s'appuyer pour développer des stratégies de prévention plus efficaces vis-à-

¹² Un individu est défini comme pas ou peu diplômé lorsqu'il est titulaire au mieux d'un Brevet d'Etudes du Premier Cycle (BEPC), d'un Diplôme National du Brevet (DNB) ou du brevet des collèges

vis des facteurs de risque modifiables, notamment dans les zones géographiques où les revenus sont les plus faibles (Parini *et al.*, 2021; Walker and Crotty, 2015). De plus, les équipements (*i.e.* sportifs, médicaux et paramédicaux) ont été inclus afin de pouvoir représenter l'espace de vie collectif dans lequel s'inscrivent les populations. En l'absence de données individuelles sur l'activité physique effective ou non des populations au sein du registre, l'accessibilité aux installations de sports et loisirs a été utilisée comme *proxy* de l'activité physique. En effet, elle constitue un élément essentiel aussi bien pour le maintien que pour l'amélioration de la santé cardiovasculaire, l'absence de pratique sportive étant un contributeur majoritaire des MCV (Arnett *et al.*, 2019; Halonen *et al.*, 2015). Différentes études ont suggéré qu'une augmentation de la distance aux équipements sportifs diminuait la pratique sportive ce qui impacterait significativement la survenue des MCV (Halonen *et al.*, 2015; Howard *et al.*, 2010; Patino *et al.*, 2021). Patterson *et al.* (2020) ont quant à eux mis en évidence une diminution de la mortalité cardiovasculaire chez les populations utilisant le vélo pour se rendre au travail et Laverty *et al.* (2013) ont associé significativement l'utilisation des transports publics et le risque de surpoids, de diabète et d'hypertension. Les résultats de nos travaux ne concordent pas totalement avec ce qui est mentionné dans la littérature. En effet, les plus grandes distances aux installations de sports et loisirs sont retrouvées dans le profil où la morbidité coronarienne est inférieure à ce qui est observée dans la population de référence. De plus, dans ce profil la part de populations utilisant les transports actifs pour se rendre au travail est la plus faible. Les résultats pour ces deux variables (*i.e.* distance aux équipements sportifs et part de population utilisant les transports actifs dans le cadre de la mobilité quotidienne) soulignent les localisations majoritairement péri-urbaines des IRIS appartenant à ce profil complexe, les installations de sports et loisirs étant plutôt localisées à proximité ou dans les centres urbains de la MEL. De plus, la comparaison des résultats entre IRIS du profil n°1 et ceux appartenant aux profils 2 et 3 majoritairement localisés dans les centres urbains renforcent cette dichotomie urbain/rural. En effet, le profil n°1 présente la plus faible part de population qui utilise des modes de transport actifs (marche à pied, vélo, transports publics) et des indicateurs en lien avec la végétation qui présentent de meilleurs résultats. En outre, la distance à parcourir pour se rendre aux services d'urgences est la plus grande au sein du profil n°1, alors que les SIRs de mortalité coronarienne sont les plus élevés dans les communes où sont implantés les principaux services d'urgences de la MEL (*i.e.* Lille, Roubaix, Tourcoing et Armentières) (la carte du SIR brut non lissé de la mortalité coronarienne est disponible en Annexe 6). Ces résultats viennent contraster les résultats de certaines études qui avaient mis en avant un lien entre le pronostic « fatal » d'un évènement coronaire et la distance aux services d'urgences, la population ayant fait l'objet d'une caractérisation en termes de comorbidités (*e.g.* hypertension, BPCO) et sociodémographiques (*i.e.* âge, sexe, éducation, statut marital, revenus) (Avdic, 2016). Au regard des relations positives mais modérées constatées dans nos travaux entre le SIR brut de la mortalité coronarienne et (i) le FDep, (ii) la part de personnes vivant seule et (iii) la part de population peu diplômée, il semble important de considérer une approche multiparamétrique pour la détermination des facteurs pouvant engager le pronostic vital des cas de syndromes coronariens aigus.

Ces travaux ont permis de souligner que certaines caractéristiques territoriales avaient une influence sur le ratio d'incidence standardisé des maladies coronariennes. Le profil complexe n°1 rassemble globalement le plus de résultats pouvant influencer de manière favorable la survenue des maladies coronariennes : la population n'est pas ou peu défavorisée, les conditions socio-économiques sont bénéfiques et les pressions anthropiques environnementales y sont les plus faibles au sein de la MEL (à l'exception du bruit la nuit). L'établissement des profils complexes a permis de proposer une caractérisation du territoire au regard de la multidimensionnalité des déterminants qui ont été considérés et qui peuvent être en lien soit avec la résilience soit avec la vulnérabilité. Les MCV ont

une étiologie multifactorielle et requièrent de ce fait une approche holistique pour approfondir les connaissances sur les impacts des déterminants sur la survenue de ces pathologies mais également sur leurs possibles interactions les uns avec les autres. Bien que la défaveur socio-économique soit un facteur influent dans l'étiologie des MCV, les résultats de ces travaux soulignent que selon les unités spatiales, la sur-morbidité coronarienne ne serait être la résultante de ce seul facteur. Au sein des unités spatiales et selon les profils complexes, il existe une grande hétérogénéité en termes de résultats vis-à-vis soit des facteurs initiant et/ou aggravant des pathologies coronariennes, soit des facteurs protecteurs. Ces profils complexes pourraient permettre de contextualiser et d'adapter les politiques de santé publique en matière de maladies coronariennes. Par la suite, il est possible d'envisager de rassembler les données d'incidence des périodes 2008-2011 (*i.e.* 3 268 cas incidents) et 2012-2016 (*i.e.* 4 464 cas incidents). Les SIRs devront toutefois être harmonisés avant de pouvoir envisager de poursuivre les analyses. En effet, en l'absence de données populationnelles par tranche de 5 ans sur la période 2008-2011, le calcul a considéré trois strates d'âge (35-44 ans, 45-59 ans, 60-74 ans).

Discussion générale

Ces travaux de thèse ont permis de développer une nouvelle approche de caractérisation des inégalités territoriales en santé environnement en lien avec la vulnérabilité et la résilience. La pertinence de l'utilisation de l'*open data* et des outils SIG a été démontrée pour optimiser la caractérisation multidimensionnelle des déterminants territoriaux. Leur influence sur la survenue de l'Insuffisance Rénale Chronique Terminale et des maladies coronariennes a ensuite été analysée. Grâce à des études écologiques, la survenue de ces pathologies chroniques a été reliée au fait que les populations vivent dans un territoire où les sources de vulnérabilité sont multiples, mais aussi au fait que certaines caractéristiques en lien avec la résilience pouvaient contrebalancer un espace de vie dégradé et *de facto* influencer positivement la survenue de ces pathologies.

La littérature sur la problématique des inégalités territoriales en santé environnement est vaste et a pour but d'avancer dans la compréhension et la mesure de l'influence de l'environnement sur la construction des inégalités de santé (Benmarhnia, 2010; Maresca and Helmi, 2014). Il existe une grande diversité d'indicateurs permettant d'évaluer des critères d'inégalités de santé. Les indicateurs épidémiologiques aident à comparer l'état sanitaire des populations, à en surveiller l'évolution ou encore à identifier et quantifier les inégalités qui peuvent être observées entre les différents territoires (Saib, 2015). Depuis plusieurs décennies, l'épidémiologie environnementale s'est particulièrement focalisée sur le compartiment atmosphérique, comme le démontrent les nombreuses publications traitant de la problématique de la qualité de l'air et de ses impacts sur la santé humaine. La dynamique actuelle de publications sur ce lien est toujours très forte (Landrigan *et al.*, 2018; Markozannes *et al.*, 2022). Toutefois, ces indicateurs environnementaux qui permettent de retranscrire la qualité des milieux (*e.g.* concentrations en polluants atmosphériques) ne constituent qu'une seule facette du prisme au regard du contexte multifactoriel dans lequel s'inscrivent les populations et qui peut avoir un impact sur leur santé. Depuis plusieurs années, des déterminants environnementaux sont également proposés afin de pouvoir retranscrire la présence d'éléments remarquables au sein du cadre de vie des populations (*e.g.* présence de parcs et espaces verts). La mise en relation de ces multiples indicateurs est dès lors indispensable pour prendre en compte la pluralité des contextes (économique, social, culturel, environnemental, politique, ...) dans lesquels les populations vivent et qui peuvent être à l'origine de situations favorables ou défavorables à la santé (Rican *et al.*, 2011).

Le premier objectif de la thèse visait à répondre à cet objectif ainsi qu'à la définition de la santé environnementale de l'OMS. Le présent travail de thèse s'appuie de ce fait sur une démarche interdisciplinaire. Les indices composites en santé environnement sont plébiscités dans la littérature et sont généralement développés sous le prisme de la vulnérabilité. Bien qu'ils tendent à être pluridimensionnels, ils se focalisent principalement sur des aspects socio-économiques et environnementaux. Les indices de vulnérabilité que nous avons développés ont montré des corrélations modérées à relativement importantes avec d'autres indices de défaveur socio-économique classiquement utilisés dans les études de santé environnement en France. Les coefficients de corrélation calculés sur la région des Hauts-de-France étaient de 0,65 avec le FDep et 0,70 avec le French EDI, ce qui indique une convergence mais également une véritable complémentarité. Ces résultats soulignent que la vulnérabilité évaluée dans le cadre de ces travaux peut être expliquée en partie par la défaveur socio-économique mais qu'elle intègre également d'autres dimensions dont l'impact sanitaire est avéré ou fortement suspecté. Les indices composites de vulnérabilité qui ont été développés permettent en effet de prendre en compte une plus grande

variabilité des contextes de vie, notamment dans des territoires pour lesquels la vulnérabilité est peu ou prou en lien avec la défaveur socio-économique. Les indices de résilience ont été développés sur le même précepte que ceux de vulnérabilité, la capacité d'un territoire à faire face aux inégalités de santé étant considérée comme également multidimensionnelle. La corrélation modérée entre l'indice de résilience et l'IDH à l'échelle des communes de la région des HDF met en évidence que l'indice de résilience développé au cours de ces travaux de thèse apporte des résultats complémentaires aux aspects de niveau de vie, sanitaires et éducatifs par l'intégration de variables en lien avec l'environnement, la politique et les services. La méthodologie de sélection de variables qui a été appliquée a permis de minimiser la redondance des informations intégrées dans les indices composites tout en s'assurant que ceux-ci permettaient de proposer une caractérisation efficiente de la résilience et de la vulnérabilité sur les territoires considérés. Les coefficients de corrélation modérés entre les indices de résilience et de vulnérabilité ont permis de s'assurer que les indices développés mesuraient bien des phénomènes différents (*i.e.* -0,01 sur les communes des HDF, 0,177 sur les IRIS de la CUD, -0,199 sur les communes du NPC et -0,04 sur les IRIS de la MEL).

Au travers de la multi-dimensionnalité des déterminants intégrés à la caractérisation des territoires, la méthodologie qui a été proposée dans le Chapitre I permet de répondre aux enjeux essentiels que constituent les inégalités territoriales en santé environnement pour les gestionnaires de territoires. La valeur agrégée proposée par les indices composites de résilience et de vulnérabilité permet de hiérarchiser les unités spatiales les unes par rapport aux autres offrant aux gestionnaires territoriaux la possibilité de prioriser les actions de santé publique à mettre en œuvre. La classification et/ou la hiérarchisation des entités spatiales ainsi mises en œuvre permettent de pouvoir faciliter les comparaisons pour des gestionnaires territoriaux bien que cela soit parfois controversé en raison de la certaine rivalité qui peut être induite au regard des résultats (Remington *et al.*, 2015). Il apparaît de ce fait nécessaire aux chercheurs de s'impliquer dans l'accompagnement des gestionnaires territoriaux dans l'analyse et l'appropriation des résultats de la caractérisation des inégalités territoriales en santé environnement qui serviront de point d'ancrage lors du développement de politiques publiques. En analysant plus en détail les résultats de chaque variable intégrée aux indices composites, il est alors possible (i) d'approfondir les connaissances sur les raisons de la classification des unités spatiales en hautement vulnérables et/ou hautement résilientes qui peuvent différer d'une unité à une autre et (ii) de comprendre l'origine des inégalités de santé (Fernández and Wu, 2016). Pour des gestionnaires territoriaux, il est également fondamental de déterminer si des territoires possèdent des ressources à mettre en œuvre pour pallier les vulnérabilités auxquelles les populations sont ou pourraient être confrontées (*e.g.* pollution atmosphérique vs végétalisation). Les sources de vulnérabilité ou l'absence de réciprocité entre la vulnérabilité et la résilience permettent de distinguer des leviers d'action. Le développement de politiques de santé publique spécifiquement adaptées aux populations et notamment pour les populations qui en auraient le plus besoin, contribue à éviter que les inégalités de santé continuent de se creuser (Desenclos and Arwidson, 2016).

Le cadre méthodologique proposé dans ce travail de thèse est issu d'une analyse critique des différentes méthodes de développement d'indices composites retrouvées dans la littérature, à l'issue de laquelle l'ACP a été retenue. Le FDep synthétise les informations contenues parmi les quatre variables intégrées et correspond à l'inverse du score sur la première composante principale de l'ACP. Ce mode de calcul n'a pas pu être envisagé pour les indices de résilience et de vulnérabilité. En effet, le nombre de variables à intégrer aux indices composites était plus élevé ce qui a contribué à faire diminuer conséquemment la part de variance expliquée de la première composante principale. Pour chacun des indices qui ont été développés, deux à trois variables contribuaient majoritairement à chaque composante principale. Proposer un indice composite exclusivement à partir du score sur la

première composante de l'ACP aurait entraîné une image parcellaire et tronquée de la résilience et de la vulnérabilité. Des méthodes de statistique spatiale ont été envisagées pour pouvoir tenir compte de la structure spatiale des variables considérées pour le développement des indices composites. Les variables qui ont été intégrées présentaient toutefois différents profils de structure spatiale. La distribution spatiale des variables au sein des périmètres d'analyse pouvait être (i) aléatoire sans aucun effet de voisinage, (ii) régulière ou (iii) agrégée. Des phénomènes continus ont également pu être observés, avec des distributions spatiales relativement proches. Les données de la dimension environnementale tendaient à présenter des autocorrélations spatiales positives, les résultats pour chaque unité spatiale étant influencés par les valeurs des unités spatiales qui leur sont proches. *A contrario*, certaines autres variables présentaient une absence de structure spatiale. Ces résultats peuvent être expliqués par des décisions politiques de gestionnaires territoriaux et qui *de facto* sont propres à certaines unités spatiales. Les variables dans ce cas appartenaient plutôt à la dimension « Politique », bien que des variables d'autres dimensions peuvent également être impactées par des décisions de politiques publiques locales sans qu'il n'y ait un phénomène de continuité spatiale pour des unités contiguës. Les variables en lien avec les équipements de sports et loisirs considérés dans nos travaux comme faisant partie de la dimension des services sont des exemples de déterminants pour lesquels l'absence ou la faible autocorrélation spatiale peut être reliée à des spécificités locales. En outre, la prise en compte de l'espace dans l'ACP aurait impliqué de considérer une pondération *via* les liens de voisinage. Cette pondération des liens de voisinage ne correspondait pas à nos conditions de développement des indices composites (*i.e.* pas de présomption *a priori* de la prédominance d'un ou de plusieurs déterminants vis-à-vis des autres) et constitue la principale limite des méthodes multivariées incluant l'espace (Chessel *et al.*, 2003). De plus, une pondération ne correspondait pas aux critères d'entrée de la caractérisation des inégalités territoriales en santé environnement qui a été développée sans envisager de pathologie en particulier. Dans la littérature, l'introduction d'une pondération dans le développement d'un indice composite ne fait toujours pas consensus au regard de la subjectivité et du jugement de valeur qui sont introduits par (i) le choix de pondérer ou non et (ii) la sélection d'une méthode de pondération (European Commission *et al.*, 2008; Gan *et al.*, 2017). En outre, bien que le poids de certains facteurs sur la survenue des pathologies chroniques soit connu, les connaissances pour d'autres déterminants restent en effet encore parcellaires. Envisager une pondération des déterminants intégrés aux indices composites développés aurait artificiellement (i) accru l'influence des déterminants pour lesquels le lien avec la pathologie est avéré ou suspecté ou (ii) amoindri l'influence de ceux pour lesquels les connaissances restent encore à déterminer et à approfondir.

La méthodologie d'analyse des inégalités territoriales en santé environnement qui a été présentée dans le Chapitre I peut être déployée à l'identique dans d'autres régions de France. Les résultats qui seraient alors obtenus ne pourraient être directement comparés avec ceux de la région des Hauts-de-France. En effet, la méthodologie permet d'obtenir une classification relative des unités spatiales en fonction des niveaux de vulnérabilité et de résilience au sein d'un périmètre d'analyse, la valeur obtenue pour chaque unité ne constituant pas une mesure absolue. Si des comparaisons inter-régionales sont souhaitées, l'ensemble des unités spatiales regroupées en un seul territoire d'étude. Ces questions de l'adaptabilité au changement d'échelle et de la transposabilité pour d'autres contextes constituaient des aspects essentiels pour la proposition d'une méthodologie d'évaluation des inégalités territoriales en santé environnement au regard de l'absence de consensus sur une méthodologie unique. Les inégalités territoriales en santé environnement sont intrinsèquement des problématiques spatiales, la compréhension des phénomènes nécessite de définir l'échelle spatiale

la plus appropriée dans l'objectif de proposer des politiques de santé publique efficaces (Fernández and Wu, 2016).

Les travaux qui ont été menés sur différents périmètres à différentes échelles ont permis de souligner que la méthodologie pouvait être déployée sur d'autres territoires que celui pour lequel elle avait été initialement appliquée (*i.e.* les communes des Hauts-de-France). Les analyses à l'échelle des IRIS des inégalités territoriales (i) en santé environnement sur le territoire de la CUD et (ii) vis-à-vis de la santé cardiovasculaire sur la MEL ont mis en évidence que la méthodologie était adaptable au changement d'échelle. Différents critères de sélection ont été proposés pour pouvoir élaborer un set de données qui visent à caractériser les inégalités territoriales en santé environnement. Parmi ceux-ci, le verrou de la disponibilité d'une donnée a principalement été levé grâce à l'*open data*. Toutefois, les données pertinentes pour la caractérisation des inégalités territoriales en santé environnement peuvent ne pas être disponibles à certains niveaux d'agrégation (*e.g.* budget des communes à l'échelle des IRIS) et/ou sur des périmètres donnés (*e.g.* bruit dans l'environnement sur les Hauts-de-France). Cela peut dès lors contraindre le choix de l'unité spatiale mais également les variables à inclure dans le set de données utilisé pour le développement des indices composites. Sur des périmètres plus restreints, l'utilisation d'une unité spatiale plus petite permet de tendre vers une plus grande variance des informations qui seront agrégées dans les indices composites.

Outre les aspects statistiques en lien avec la variance, le choix d'une unité spatiale plus restreinte peut permettre également de réduire les corrélations entre les variables lors de l'étape de sélection. Dans l'analyse comparative qui a été menée sur le périmètre de la CUD, l'analyse bivariée réalisée sur le set de données à l'échelle des communes présentait un plus grand nombre de corrélations ce qui a contribué à réduire le nombre de variables pouvant être intégrées aux indices composites selon nos critères d'inclusion. Ce constat s'est avéré particulièrement vrai pour l'indice de vulnérabilité où 13 variables y avaient été intégrées lorsque la sélection avait été menée sur l'ensemble des communes des Hauts-de-France tandis que 7 ont été retenues lors de l'analyse spécifiquement menée sur les communes de la CUD. De plus, des variables qui présentent une distribution aberrante auront tendance à influencer la construction des indices composites par l'ACP, celle-ci étant sensible à ce type de variables (Tharrault *et al.*, 2007). Le choix d'une unité spatiale pour l'analyse des inégalités territoriales au sein d'un périmètre peut dès lors être conditionné par (i) la disponibilité et la qualité (*i.e.* exhaustivité) de données pertinentes pour l'objectif de l'analyse, (ii) des critères statistiques (*i.e.* variance, aberrance) et/ou (iii) du niveau décisionnel des gestionnaires et des parties prenantes qui pourraient être impliqués dans ces analyses afin d'identifier les leviers d'action pour déterminer les politiques de santé publique qui pourraient être implémentées, l'ensemble de ces critères pouvant être interconnectés. Les perspectives pour cet axe de recherche portent dans un premier temps sur le développement d'un outil d'aide à la décision permettant d'objectiver l'étape de sélection des variables. Dans un second temps, il s'agira de poursuivre sur la problématique de la transposabilité de la méthodologie à l'international. Une collaboration avec l'Institut Scientifique de Service Public¹³ (ISSEP) a été initiée afin de pouvoir éprouver la méthodologie face à un changement de contexte national. Une analyse comparative des conclusions entre les communes des Hauts-de-France et de Wallonie sera développée. Les travaux menés dans le cadre de cette collaboration pourront également venir étayer la discussion sur les problématiques en lien avec les échelles spatiales. En effet, bien que les unités spatiales qui seront considérées dans ce cadre portent le même nom, elles sont néanmoins différentes. Une commune en Belgique est en effet une division

¹³ L'ISSEP est soumis à l'autorité du Gouvernement wallon et est en charge de nombreux programmes d'études de caractérisation des milieux environnementaux (air, eaux, sol, ...) pour en proposer un diagnostic.

<https://www.issep.be/>

administrative qui correspond à (i) une ville avec des villages environnants, (ii) un regroupement de plusieurs villages ou (iii) un village important entouré de hameaux. Les 262 communes de Wallonie sont de fait plus grandes et plus peuplées : la superficie moyenne des communes des Hauts-de-France est de 8,37 km² (*sd* = 5,84) avec une population moyenne de 1 569 habitants (*sd* = 6 303) tandis que la superficie moyenne des communes wallonnes est de 64,52 km² (*sd* = 43,86), celles-ci comportant en moyenne 12 426 habitants (*sd* = 20 860) (l'Annexe 7 présente une comparaison des deux territoires). Ces travaux contribueront également à la recherche transfrontalière et étudieront la continuité géographique des phénomènes en lien avec la résilience et la vulnérabilité.

Le second objectif de la thèse portait sur la caractérisation de l'influence de l'environnement dans l'étiologie de l'IRCT et des maladies coronariennes, deux pathologies multifactorielles. Les travaux présentés dans le Chapitre II décrivent les démarches écologiques qui ont été mises en place afin d'explorer l'influence des caractéristiques territoriales qui pourraient être en lien avec la survenue de pathologies chroniques, et notamment des pathologies ayant un impact important sur certains territoires. Pour pouvoir formuler des hypothèses étiologiques et améliorer la connaissance sur les contributions des déterminants territoriaux sur la survenue d'une pathologie chronique, il est primordial d'avoir une exhaustivité des cas de celle-ci sur un territoire. Dans le cadre de ces travaux de thèse, la collaboration avec des registres de santé a permis d'obtenir des données exhaustives et robustes pour une pathologie donnée sur un périmètre donné (Jomier et Lassarade, 2021). Les pathologies chroniques qui ont été considérées dans le cadre de ces travaux ont été sélectionnées en fonction (i) de leur prédominance au sein de la région Hauts-de-France et (ii) de la disponibilité d'un registre de santé permettant d'avoir l'exhaustivité des cas sur un territoire considéré. Ainsi, les MCV ont été intégrées du fait de l'existence du registre MONICA et de données épidémiologiques telles que la part des décès imputables (24,5 % des décès ont été causés par les MCV et ont entraîné près d'une admission sur trois en ALD (OR2S, 2019)). Bien que l'IRCT ne figure pas parmi les principales causes de décès et d'admission en ALD dans les Hauts-de-France, les principales causes d'évolution des MRC vers l'IRCT (*i.e.* diabète, hypertension) présentent des résultats très défavorables au sein de la région. En effet, 22,3 % de la population des Hauts-de-France suivait en 2014 un traitement par des antihypertenseurs, ce qui correspond à la prévalence la plus élevée de France hors Mayotte (DREES and Santé Publique France, 2017). La surmortalité et les admissions en ALD pour le diabète de type 2 y sont plus importants par rapport aux résultats observés sur la France : la mortalité y est supérieure de 35 % sur la période 2006-2013 et les ALD sur la période 2007-2014 y sont supérieures de 34 % pour les femmes et de 23 % pour les hommes (OR2S, 2019).

Dans un premier temps, après ajustement sur des facteurs de confusion (*i.e.* âge, sexe, éloignement au centre lourd de dialyse le plus proche, DFG et FDep), une variabilité résiduelle de l'incidence d'IRCT a été observée. Cela a mis en évidence que ces paramètres ne suffisaient pas à expliquer à eux seuls l'hétérogénéité spatiale de l'incidence de l'IRCT observée sur les départements du Nord et du Pas de Calais. Des RR en lien avec une sous incidence ont été mis en évidence sur une zone étendue, de Calais à Auxi-le-Château. A l'inverse, les communes qui présentent des RR en lien avec une sur-incidence sont globalement localisées (i) au nord de l'ancien bassin minier et comprend une partie des communes de la MEL et (ii) à l'est du département du Nord, dans l'Avesnois. Les communes présentant des RR en lien avec une sur-incidence présentent des profils différents, certaines étant plutôt situées en milieu rural tandis que d'autres sont en milieu péri-urbain voire urbain. L'incidence résiduelle de l'IRCT a été mise en relation avec les indices de vulnérabilité et de résilience qui ont été spécifiquement calculés à partir de données décrivant des déterminants impliqués dans des liens avérés ou suspectés avec les MRC. Aucune association significative n'a été soulignée entre l'indice

de vulnérabilité et l'incidence de l'IRCT (modèle 4). Bien que la littérature soit équivoque sur l'influence de certains déterminants environnementaux avec les MRC et/ou l'IRCT, certains ont été associés à ces pathologies et constituent des facteurs de risque nouvellement identifiés (*e.g.* pollution atmosphérique et notamment les PM_{2,5}). Des pistes de réflexion peuvent être suggérées pour comprendre ces différences, (i) l'utilisation de moyennes sur une plage temporelle relativement étendue a pu contribuer à réduire la variabilité en lien avec des déterminants environnementaux (la pollution atmosphérique notamment), (ii) les méthodologies mises en œuvre pour analyser l'influence de certains déterminants environnementaux étaient plutôt épidémiologiques dans d'autres études (*e.g.* études de cohorte), (iii) les variables intégrées à l'indice de vulnérabilité n'ont pas présenté de pouvoir discriminant supplémentaire par rapport aux facteurs de confusion intégrés dans la détermination de la variabilité résiduelle de l'incidence de l'IRCT.

L'intégration de l'indice de résilience dans les modèles statistiques a permis de mettre en évidence qu'il contribuait à expliquer la variabilité observée de l'incidence de l'IRCT. Individuellement, le NDVI est le déterminant qui est apparu le plus associé à la variabilité résiduelle de l'incidence de l'IRCT. Dans la littérature, l'influence de la présence végétale a été peu mise en évidence pour les MRC et jamais pour l'IRCT, bien que cet indicateur soit régulièrement considéré dans l'analyse du lien entre la végétation et d'autres événements de santé (*e.g.* MCV, santé mentale). Dans le cadre de l'analyse des contributions territoriales sur la survenue des maladies coronariennes, des indices de vulnérabilité et de résilience ont été développés à l'échelle des IRIS de la MEL et ont été comparés au ratio d'incidence standardisé des maladies coronariennes sur la période 2012-2016. L'indice de vulnérabilité a montré une corrélation très modérée (*i.e.* 0,192, *p*-value < 0,001) tandis que l'indice de résilience a présenté une corrélation négative relativement élevée avec les SIRs lissés (-0,511, *p*-value < 0,001). Ce résultat souligne l'importance de l'influence des déterminants en lien avec la résilience sur la survenue de pathologies chroniques. Bien que la corrélation ne signifie pas causalité, ces résultats permettent d'avancer dans la connaissance des relations entre les déterminants de santé et la survenue de pathologies chroniques au sein d'un territoire.

Dans un second temps, des profils complexes des IRIS de la MEL ont été réalisés à partir des données ayant permis de développer les indices composites de résilience et de vulnérabilité afin d'étudier leur influence possible sur la survenue des maladies coronariennes. Les déterminants sont en effet susceptibles d'interagir les uns avec les autres et ont de ce fait été considérés simultanément dans l'établissement des profils territoriaux complexes. En outre, la démarche qui a été menée visait à considérer des déterminants en lien avec la résilience, les inégalités territoriales en santé environnement n'étant pas la seule résultante des nuisances et contraintes auxquelles sont soumises les populations (Lanier *et al.*, 2020). Dans le cadre de ces travaux, trois profils ont été déterminés et les résultats ont montré que les ratios d'incidence standardisés étaient significativement différents pour chacun des profils complexes, le profil 1 étant celui en sous-incidence avec des déterminants environnementaux présentant majoritairement une influence positive sur la survenue des maladies coronariennes, le profil 2 correspondait à une sur-incidence plutôt en lien avec la défaveur socio-économique tandis que le profil 3 présentait une sur-incidence en lien avec des indicateurs environnementaux défavorables. De la même manière que pour les indices composites, il est essentiel de revenir aux variables qui ont été intégrées dans la caractérisation des territoires pour déterminer les profils complexes afin de pouvoir identifier (i) les principaux contributeurs à une sous-incidence ou une sur-incidence et (ii) les leviers d'action pour les politiques de santé publique qui visent à réduire la charge des maladies cardiovasculaires. La méthodologie mise en œuvre pour élaborer des profils complexes des IRIS de la MEL a mis en exergue sa capacité à discriminer les territoires. L'analyse de ces profils complexes a permis de souligner que la défaveur socio-économique impactait significativement certains IRIS tandis que l'impact de la pollution

atmosphérique sur la santé cardiovasculaire était probablement plus important dans des IRIS où la population n'était pas ou peu défavorisée socio-économiquement. Ces résultats sont venus corroborer ceux qui avaient été précédemment mis en évidence sur le territoire. En effet, sur la période 2008-2011, la défaveur socio-économique et la pollution atmosphérique avaient été identifiées comme les principales causes de maladies coronariennes (Ocelli *et al.*, 2020).

L'analyse des profils complexes qui ont été déterminés pour étudier l'influence des contributions territoriales en lien avec la santé cardiovasculaire a souligné l'importance des déterminants en lien avec la végétation au sein de la MEL. Pour les deux profils en sur-incidence significative, ces indicateurs présentaient des résultats défavorables tandis que le profil de sous-incidence était caractérisé par les résultats les plus favorables sur le territoire de la MEL. Ces résultats concordent avec les données de la littérature où des relations inverses avaient été mises en évidence entre la végétation (*e.g.* espaces verts, NDVI) et les maladies cardiovasculaires en général (Aerts *et al.*, 2020; Astell-Burt *et al.*, 2021; Patino *et al.*, 2021; K. Wang *et al.*, 2019; Yitshak-Sade *et al.*, 2017). Les résultats pour les déterminants en lien avec la résilience ont mis en évidence des résultats contrastés au sein des profils complexes. La diversité et la richesse spécifique des catégories socio-professionnelles mais également des professions médicales et paramédicales ainsi que le fait d'utiliser des modes de transport actif pourraient également contribuer à la sous-incidence des maladies coronariennes au sein de la MEL. Dans le cadre de la mobilité quotidienne, l'utilisation des modes de transport actif¹⁴ encourage en effet l'activité physique ce qui peut concourir à l'amélioration de la santé cardiovasculaire ou qui limite sa dégradation. Les populations utilisant ce type de transport présenteraient des IMC plus faibles et ont été significativement identifiées comme moins susceptibles d'être en surpoids que les populations ayant recours à leur véhicule personnel (Laverty *et al.*, 2013; Patterson *et al.*, 2020). La diversité des professions médicales et paramédicales peut de plus contribuer à un état de santé meilleur des populations par une prise en charge précoce lorsqu'une ou des comorbidités sont identifiées. Toutefois, la question du recours, et surtout du non-recours au soin, au regard des difficultés socio-économiques, culturelles ou individuelles peut nuancer ces résultats. Cela peut être accentué par des problématiques d'accessibilité des populations (*e.g.* transport, délai pour l'obtention d'un rendez-vous). Il est à préciser également que des catégories socio-professionnelles élevées peuvent également présenter un non-recours au soin. Ces travaux ont permis de souligner que, quelle qu'ait été la méthodologie mise en œuvre pour analyser l'influence de la vulnérabilité et de la résilience sur l'apparition des pathologies chroniques considérées (*i.e.* indices composites et variabilité résiduelle de l'incidence après ajustement sur des facteurs confondants et profils complexes caractérisés par le ratio d'incidence standardisé), l'indice de résilience au travers de la caractérisation multidimensionnelle des déterminants contribuait de manière prégnante et inverse à l'incidence.

Certaines limites sont à prendre en considération notamment vis-à-vis des choix méthodologiques qui ont été opérés. Les modèles de régression écologique sont reconnus dans la littérature car ils permettent de prendre en compte l'autocorrélation spatiale des phénomènes considérés. Toutefois, ce type de modèle ne permet pas de prendre en compte les paramètres individuels. Un modèle de Poisson hiérarchique mixte aurait pu permettre d'intégrer ces facteurs dans l'analyse, leur limite principale portant sur la perte de l'information spatiale (les unités spatiales sont considérées indépendantes). Un modèle « mixte » aurait pu combiner les avantages de ces deux types de méthodologies. Toutefois, l'intégration des données individuelles dans des modèles spatiaux s'avère

¹⁴ Actifs de 15 ans ou plus allant principalement travailler à pied, en vélo ou en transports en commun

complexe et nécessite des données non disponibles pour les études menées dans le cadre de cette thèse. Par exemple, les méthodologies utilisées dans certaines publications sur l'IRCT s'apparentent à des études cas-témoin, où les « cas » sont des patients pour lesquels un déclin rapide du DFG est constaté, tandis que les « témoins » sont des patients pour lesquels ce déclin pourrait être qualifié de « normal » (Bowe *et al.*, 2017). Une telle approche ne pouvait être déployée car le périmètre de ces travaux de thèse inclut la seule population des incidents, et les paramètres individuels (*i.e.* pour l'IRCT : DFG, historique de cancer, diabète, IMC, statut tabagique et antécédents de MCV, BPCO ; pour les maladies coronariennes : IMC, statut tabagique, hypertension artérielle, dyslipidémie, diabète, obésité) sont recueillis au moment de l'enregistrement du patient au sein du registre. En outre, la part de données manquantes peut être relativement élevée selon les paramètres individuels et dans le cadre des maladies coronariennes, ils ne sont enregistrés que depuis 2013. Une réflexion est actuellement menée afin de pouvoir proposer ce type de démarche à l'avenir en adaptant la méthodologie développée pour la définition d'un eDFG qui a permis de pouvoir considérer la variabilité géographique de mise en place de la suppléance par la dialyse pour les patients atteints d'IRCT, ce paramètre étant un des facteurs de confusion contribuant à expliquer la variabilité spatiale de l'incidence de l'IRCT. L'estimation d'autres paramètres individuels (*e.g.* diabète et hypertension) ainsi que de facteurs comportementaux (*e.g.* statut tabagique) pourrait permettre de les considérer dans les modèles au regard de leur importance dans l'étiologie des maladies chroniques qui ont été étudiées dans le cadre de ces travaux.

La définition de l'exposition est une problématique complexe et lors de l'analyse de l'influence des caractéristiques territoriales sur la survenue de maladies chroniques, seul le lieu de résidence des patients intégrés dans les registres a été considéré en l'absence d'informations sur l'ensemble des lieux de vie fréquentés par les populations. Selon Aerts *et al.* (2020) et Tamura *et al.* (2019), l'exposition déterminée au lieu de résidence peut en effet différer de l'exposition réelle et quotidienne car elle ne permet pas de prendre en compte l'ensemble des lieux fréquentés par les populations pour le travail, les heures d'école ou encore les loisirs, ces déplacements pouvant de plus être variables selon les heures du jour ou encore selon les jours de la semaine. L'impossibilité de prendre en compte la mobilité des populations dans l'exposition considérée comme non résidentielle est identifiée dans la littérature sous le phénomène de « *Neighborhood Effect Averaging Problem* » (NEAP) (Kwan, 2018). Des travaux sur l'exposition individuelle à la pollution atmosphérique ont montré que considérer la mobilité des populations pouvaient entraîner des différences significatives avec l'estimation réalisée au lieu de vie de ces mêmes populations (Park and Kwan, 2017; Shafran-Nathan *et al.*, 2017). Des individus vivant dans des unités spatiales pour lesquelles les concentrations moyennes en polluants atmosphériques peuvent travailler dans des unités spatiales au sein desquelles les concentrations moyennes en polluants atmosphériques peuvent être inférieures ce qui contribue à surestimer leur exposition (Kwan, 2018). A l'inverse, l'exposition des populations vivant dans des zones géographiques où la pollution atmosphérique est plus faible peut être sous-estimée si dans le cadre de leur mobilité quotidienne ils sont amenés à fréquenter des lieux où celle-ci est plus élevée (Kwan, 2018). La problématique du NEAP n'est pas limitée au seul phénomène de pollution atmosphérique, les travaux de Kou *et al.* (2020) et Tao *et al.* (2020) ont montré que la mobilité quotidienne des individus pouvait également avoir un impact sur l'estimation de l'exposition chronique au bruit. Toutefois, ces travaux doivent encore être approfondis car les relations entre le bruit mesuré, le bruit perçu et les impacts sanitaires sont complexes et vont au-delà d'une mesure quantitative du bruit (Kou *et al.*, 2020).

Notre modélisation des relations entre les caractéristiques territoriales et la survenue d'une maladie a reposé sur l'utilisation de données agrégées (Anselin, 2000). Toutefois, les individus peuvent ne pas être bien représentés par des données populationnelles, ainsi, par exemple, un individu avec un

statut socio-économique élevé peut vivre dans un quartier où le statut socioéconomique moyen du quartier est plus faible (Marshall *et al.*, 2015). De plus, bien que les questions de recherche soient différentes, les associations mises en évidence sur la base de données agrégées peuvent être différentes de ce qui aurait pu être mis en évidence à un niveau individuel, ce phénomène étant connu sous le terme de l'« *ecological fallacy* » (Davide *et al.*, 2022). Une autre problématique méthodologique mentionnée comme le « *Uncertain Geographic Context Problem* » (UGCoP) repose sur le fait que les résultats peuvent différer selon les délimitations géographiques contextuelles des unités spatiales (*e.g.* les quartiers) (Kwan, 2012b). Les deux principales problématiques en lien avec l'UGCoP reposent sur les incertitudes en lien avec (i) la configuration spatiale des unités appropriées pour analyser l'influence des caractéristiques territoriales sur les inégalités de santé et (ii) sur les moments et les durées auxquelles les individus sont réellement exposés aux influences contextuelles (Kwan, 2012a). Ce problème tend à se rapprocher de celui du MAUP qui fait référence aux résultats qui peuvent être affectés par un choix d'une unité géographique plutôt qu'une autre et par l'inférence statistique (Kwan, 2012b). La littérature sur cette problématique des biais écologiques (*i.e.* *ecological fallacy*, NEAP, UGCoP, MAUP) est vaste mais ne propose pas de consensus sur des méthodes visant à les limiter. Ces biais reposent sur (i) l'inférence écologique qui agrège les données et (ii) la variabilité au sein des aires d'étude considérées pour déterminer l'exposition et des facteurs de confusion (Wakefield and Smith, 2016). Selon Shafran-Nathan *et al.* (2017), l'association entre les facteurs de risque environnementaux et des événements de santé pourrait être plus importante parce qu'une unité spatiale présente plus de cas. Cette association pourrait être fragilisée lorsque les résultats au sein des unités spatiales présentent une variabilité plus faible et /ou sont moins précis. Afin de pouvoir minimiser ces biais écologiques, plusieurs études ont montré que travailler avec des unités spatiales très petites permettait de diminuer les biais écologiques, sans la possibilité toutefois de les éliminer (Bryere *et al.*, 2017; Lalloué *et al.*, 2013; Maresca and Helmi, 2014; Vallée, 2019).

Les résultats des approches écologiques visant à analyser les relations entre des caractéristiques territoriales et la survenue de pathologies chroniques ne peuvent pas être transposés directement à l'échelle individuelle (Elliott and Wakefield, 2000, Greenland, 2001). La vision agrégée proposée par les approches écologiques ne permet pas de rendre compte la perception individuelle et négative de certaines caractéristiques territoriales qui peut engendrer un stress chronique (van Deurzen *et al.*, 2016; Walker and Crotty, 2015). Au-delà de ce stress chronique, ces différences de perception individuelle d'un déterminant avaient été soulignées dans l'analyse de la littérature qui avait été menée et dont les résultats ont été valorisés dans l'Article 1. Il avait été mis en évidence l'équivocité de certaines variables, et que selon les auteurs, certaines variables pouvaient avoir été considérées comme ayant un impact soit négatif soit positif sur les inégalités territoriales en santé environnement. Cela était notamment le cas pour la densité de population qui pouvait être considéré comme un *proxy* pour de multiples sources de nuisances (*e.g.* trafic routier, bruit) mais également pour différentes aménités (*e.g.* professions médicales, installations sportives). Selon Padilla (2013), plus l'unité spatiale est petite au sein d'une approche écologique (*e.g.* l'IRIS), plus les écarts entre les profils des personnes qui y résident et les caractéristiques agrégées à cette unité spatiale tendent à diminuer. En outre, des échelles spatiales larges (*e.g.* canton, région) agrègent les informations et peuvent, de ce fait, diluer des réalités territoriales potentiellement peu homogènes (Maresca and Helmi, 2014). Bien que la moyenne tende à rester sensiblement identique lors de l'augmentation de l'agrégation, la variance quant à elle diminue ce qui entraîne un lissage et une perte de l'hétérogénéité spatiale des informations agrégées (Jelinski and Wu, 1996). Les études écologiques permettent toutefois d'améliorer les connaissances sur les mécanismes qui façonnent les inégalités de santé à l'échelle individuelle mais également populationnelle en étudiant l'influence potentielle de l'environnement compris au sens élargi (*e.g.* physique, social) sur la santé humaine et sur certains

comportements en lien avec la santé (Macintyre and Ellaway, 2000). Ce type d'étude répond aux besoins de l'épidémiologie qui cherche à expliquer la distribution spatiale des pathologies influencées par des caractéristiques individuelles qui peuvent interagir également avec le contexte dans lequel les populations vivent (Arcaya *et al.*, 2016).

Les hypothèses étiologiques qui ont été formulées devront faire l'objet de recherches au niveau individuel afin d'enrichir le faisceau de preuves sur les liens existants entre ces caractéristiques et l'IRCT et les maladies coronariennes. De plus, l'utilisation des registres de santé devrait être développée pour mieux évaluer le lien entre l'apparition d'une pathologie et des facteurs environnementaux suspectés d'avoir un impact sur la pathologie (Jomier and Lassarade, 2021). Pour pouvoir mettre en place un traitement personnalisé qui serait adapté à chaque profil de patients, des études de cohortes s'avèreront nécessaires pour avancer dans la compréhension des liens à l'interface des aspects socio-économiques, environnementaux, culturels, familiaux ou encore génétiques et de pathologies chroniques (INSERM, 2017).

Les résultats issus des travaux présentés dans le Chapitre II ont permis de distinguer plusieurs perspectives de recherche. Dans le cadre de l'IRCT, à court terme, une caractérisation de l'occupation des sols des communes pour lesquelles des sur- ou des sous-incidences ont été mises en évidence pourra renseigner d'une possible différenciation entre territoires artificialisés et territoires agricoles de ces zones géographiques. A moyen/long terme, il s'agira d'élargir la caractérisation de l'environnement, notamment vis-à-vis de l'indice de vulnérabilité afin d'améliorer la part de la variabilité expliquée par des contaminations environnementales en lien avec les compartiments autres qu'atmosphérique qui ont été prises en compte dans les indices de vulnérabilité. En effet, l'intégration de données en lien avec les émissions dans l'air des HAPs fortement corrélées aux ETMs n'ont pas permis de venir étayer les résultats de la littérature sur les liens entre contamination métallique et insuffisance rénale. Des campagnes de biosurveillance végétale et fongique et des prélèvements des sols pourront permettre sur certaines zones géographiques d'améliorer la connaissance sur l'imprégnation des milieux vis-à-vis de certains contaminants. En parallèle de la contamination des sols, des données en lien avec l'EDCH pourraient être intégrées également au regard des études de la littérature traitant de l'impact des contaminations en lien avec ce milieu. En effet, les ETMs font partie des contaminants pour lesquels les expositions sont multiples, les populations pouvant y être exposées *via* l'ensemble des milieux (eau, air, sol). L'intégration des données en lien avec l'EDCH comportent toutefois des limites au regard de (i) la plage temporelle pour laquelle des données sont accessibles (2016-2020) et (ii) des ETMs qui présentent des concentrations supérieures aux limites de quantification (*i.e.* antimoine, chrome, cuivre, fer, nickel, plomb).

Concernant les maladies coronariennes, il s'agira de poursuivre les analyses statistiques sur les déterminants qui présentaient des résultats contrastés selon les profils complexes. Des facteurs de risque et des facteurs protecteurs pourraient être identifiés et viendraient confirmer ou infirmer les hypothèses étiologiques qui ont été formulées dans cet axe de recherche. Les campagnes de prélèvements des sols pourraient permettre améliorer la caractérisation du territoire de la MEL en élargissant la caractérisation de l'environnement. De plus, les données d'incidence pour les deux périodes pourraient être rassemblées afin de pouvoir ce qui permettrait d'obtenir un total de 7 732 cas incidents sur 2008-2016 concourant à améliorer la robustesse des analyses qui seraient menées.

Pour l'ensemble des travaux de thèse, la question de la temporalité constitue une perspective de recherche qui permettrait d'avoir une vision plus précise des inégalités territoriales en santé

environnement. Les bases de données disponibles en *open data* ne proposent pas nécessairement des mises à jour annuelles. Cette limite a été particulièrement constatée pour les données environnementales, qu'elles soient en lien avec la vulnérabilité ou la résilience. La méthodologie développée pour l'analyse de l'influence des indices de résilience et de vulnérabilité sur l'incidence de l'IRCT a permis de considérer en partie cette question et de caractériser la pollution chronique sur une plage temporelle étendue. Des méthodes statistiques devront être mises en place afin de considérer le double aspect de la variabilité spatiale et temporelle pouvant être constatée aussi bien pour les données d'incidence des pathologies chroniques que pour les données permettant de caractériser le territoire, et notamment des données environnementales.

Articles dans des revues avec comité de lecture

Brousmiche D, Lanier C, Cuny D, Frevent C, Genin M, Blanc-Garin C, Amouyel P, Deram A, Occelli F, Meirhaeghe A (2022) How do territorial characteristics affect spatial inequalities in the risk of coronary heart disease? *Environmental Health Perspectives (in prep., relecture interne)*

Brousmiche D, Ahmed-Salem M, Genin M, Lanier C, Gomis S, Cuny D, Occelli F, Glowacki F (2022) Spatial variability of the incidence of End-Stage Renal Disease: what impacts of confounding factors, vulnerability and resilience? *Journal of the American Society of Nephrology (in prep.)*

Brousmiche D, Lanier C, Occelli F, Pochet L, Degezelle S, Genin M, Deram A, Cuny D (2022) Spatial scale and spatial effects: shedding light on how a change of scale can affect the characterization of environmental health inequalities. *Environmental Science & Policy (submit)*

Brousmiche D, Genin M, Occelli F, Frank L, Deram A, Cuny D, Lanier C (2021) Data for the assessment of vulnerability and resilience in the field of environmental health in the north of France. *Data In Brief*. Doi: <https://doi.org/10.1016/j.dib.2021.107220>

Brousmiche D, Genin M, Occelli F, Frank L, Deram A, Cuny D, Lanier C (2020) How to analyse environmental health resilience and vulnerability? A joint analysis with composite indices applied to the north of France. *Science of the Total Environment*, October 2020;142983. Doi: [10.1016/j.scitotenv.2020.142983](https://doi.org/10.1016/j.scitotenv.2020.142983)

Lanier C, **Brousmiche D**, Deram A, Frank L, Genin M, Occelli F, Cuny D (2020) Caractérisation de l'hétérogénéité spatiale des inégalités environnementales et sociales de santé par une approche spatialisée de la balance résilience/vulnérabilité. *Environnement, Risques & Santé*, 19(4):250-256. Doi: [10.1684/ers.2020.1451](https://doi.org/10.1684/ers.2020.1451)

Brousmiche D, Occelli F, Genin M, Cuny D, Deram A, Lanier C (2020) Spatialized composite indices to evaluate environmental health inequalities: Meeting the challenge of selecting relevant variables. *Ecological Indicators* April 2020 111;106023. <https://doi.org/10.1016/j.ecolind.2019.106023>

Communication dans des colloques internationaux

Brousmiche D, Lanier C, Occelli F, Cuny D, Meirhaeghe A (2022) Influence of territorial characteristics on incidence of myocardial infarction, a study at fine geographic scale. *ISES Annual Meeting 2022, International Society of Exposure Sciences*. Lisbon (Portugal), 25-29 September 2022

Brousmiche D, Lanier C, Cuny D, Occelli F, Montaye M, Amouyel P, Meirhaeghe A, Genin M (2021) Influence of vulnerability and resilience on incidence of coronary heart disease, a study at fine geographic scale. *4TH World academic conference on environmental health "The real situation in environmental health – challenges and solutions"*, session: urban health. International Federation of Environmental Health. Tartu (Estonia), 04-06 May 2021

Communications dans des colloques nationaux

Brousmiche D, Pochet L, Degezelle S, Lanier C, Cuny D (2022) Potentiels de vulnérabilité, de résilience et analyse des inégalités environnementales et sociales de santé au sein de la Communauté

Urbaine de Dunkerque. *Journée d'étude SIGLES : Analyses spatiales en santé environnement*. Lille, 20 Mai 2022

Brousmiche D, Lanier C, Deram A, Occelli F, Genin M, Meirhaeghe A, Cuny D (2021) Développement d'indices composites spatialisés pour l'étude du lien entre environnement au sens élargi et survenue de pathologies chroniques : apports et limites de l'open data. *Journée d'Etude sur l'accès et l'utilisation des données en libre accès en géographie de la population, de la santé et des mobilités*. Paris, 02 Décembre 2021

Brousmiche D, Pochet L, Degezelle S, Lanier C, Cuny D (2021) Potentiels de vulnérabilité, de résilience et analyse des inégalités environnementales et sociales de santé grâce au développement d'indices composites à l'échelle des IRIS au sein de la Communauté Urbaine de Dunkerque. *Les 2^{èmes} rencontres de Géographie de la Santé*. Rennes, 04-05 Février 2021.

Brousmiche D, Degezelle S, Pochet L, Occelli F, Deram A, Lanier C, Cuny D (2020) Intégration des multi-expositions dans la caractérisation des territoires, de la recherche à la décision publique : exemple d'une démarche collaborative menée sur le territoire de la Communauté Urbaine de Dunkerque (CUD). *11^{ème} congrès de la Société Française de Santé Environnement*. Lille, 18-20 novembre 2020.

Brousmiche D, Occelli F, Genin M, Cuny D, Deram A, Lanier C (2019) Développement d'un indice spatialisé de vulnérabilité pour une approche globale de la santé-environnement en Hauts-de-France. Communication affichée. *Quatrièmes Rencontres nationales de la Recherche sur les Sites et Sols Pollués*. Le Beffroi de Montrouge (Portes de Paris), 26 et 27 novembre 2019.

Brousmiche D, Occelli F, Genin M, Cuny D, Deram A, Lanier C (2019) Développement d'un indice spatialisé de vulnérabilité pour une approche globale de la santé-environnement en Hauts-de-France. Communication orale. *Quatrièmes Rencontres nationales de la Recherche sur les Sites et Sols Pollués*. Le Beffroi de Montrouge (Portes de Paris), 26 et 27 novembre 2019.

Lanier C, **Brousmiche D**, Occelli F, Cuny D, Deram A, Génin M (2019) Approche bénéfiques/risques et IESS : développement d'indices de vulnérabilité et de résilience adaptés au territoire français. Communication orale. *Congrès de la Société Francophone de Santé et Environnement (SFSE) 2019*. Paris, 27 et 28 Novembre 2019.

Brousmiche D, Occelli F, Genin M, Cuny D, Deram A, Lanier C (2019) Développement d'indices spatialisés de vulnérabilité et de résilience en santé environnement. Communication orale. *SIG 2019, La Conférence Francophone Esri*. Les Docks de Paris, 16-17 Octobre 2019.

Activités d'encadrement

Année universitaire	Prénom & Nom du stagiaire	Durée du stage	Formation	Intitulé du stage
2021-2022	Camélia Bousbaine	8 semaines	L2 – Sciences Pour la Santé (ILIS, Université de Lille)	Développement d'indicateurs en lien avec des habitudes de vie
	Vincent Herbeau	10 semaines	M1 – Information Communication Sciences du Document (UFR DECCID, Université de Lille)	Création d'un outil de communication pour les élus de la CUD en lien avec l'analyse des inégalités territoriales de santé
	Julie-Kim Ngo	8 semaines	L3 – Sciences Pour la Santé (ILIS, Université de Lille)	Défaveur socio-économique et incidence des Accidents Vasculaires Cérébraux sur Lille
2020-2021	Léa Cœugniet	8 semaines	L3 – Sciences Pour la Santé (ILIS, Université de Lille)	Développement d'indicateurs à l'IRIS en lien avec la résilience et la vulnérabilité à l'échelle des IRIS de MEL
	Louis Pochet	5 mois	M1 – Qualité Environnement Santé Toxicologie (ILIS, Université de Lille)	Inégalités territoriales de santé en lien avec la résilience et la vulnérabilité à l'échelle des IRIS de la CUD

Activités d'enseignement

Année universitaire	Composante	Nombre d'heures	Intitulé
2021-2022	UFR3S – Pharmacie (Lille)	26,5 HTD	2A : Module d'accompagnement pour l'acquisition des méthodes d'apprentissage dans le supérieur (BCC1 : maîtriser les gestes fondamentaux pour les travaux pratiques, communiquer, connaître son futur environnement professionnel) ED : 7 * 1,5 H – TP : 8 * 2 H
	Composante FST – Département Biologie (Lille)	2 HTD	M2 : Master ECOREMID (Biodiversité Ecologie Evolution parcours ECOREMID - Ecologie et Restauration des Milieux Dégradés). UEC Diagnostic biologique des pollutions
	Composante UFR3S – ILIS (Lille)	15 HTD	M1 : Master Management Sectoriel, BCC3 Identifier et analyser les enjeux en management et de santé publique, UE 2 amélioration de la qualité et gestion des risques, module « Santé au travail » 10 H CM
	Paris Nanterre	2 HTD	M2 : Master Territoires, Villes et Santé. Cours séminaire « Changement climatique et la santé »
2020-2021	UFR3S – Pharmacie (Lille)	23 HTD	2A : Module d'accompagnement pour l'acquisition des méthodes d'apprentissage dans le supérieur (BCC1 : maîtriser les gestes fondamentaux pour les travaux pratiques, communiquer, connaître son futur environnement professionnel) ED : 6 * 1,5 H – TP : 7 * 2 H
	Composante FST – Département Biologie (Lille)	2 HTD	M2 : Master ECOREMID (Biodiversité Ecologie Evolution parcours ECOREMID - Ecologie et Restauration des Milieux Dégradés). UEC Diagnostic biologique des pollutions
		70 HTD	

Implication dans des sociétés savantes

Depuis 2018 :

Membre de l'axe « Territoires et Santé » du Collège International des Sciences Territoriales (CIST).
<https://cist.cnrs.fr/axes-structurants/sante-territoires-et-sante/>

Depuis 2019 :

Membre de la commission « Les Géographies de la Santé » du Comité National Français de Géographie (CNFG). <https://www.cnfg.fr/commissions/geographie-de-la-sante/>

2021 – 2022 :

Représentante élue des doctorants de l'Ecole Doctorale Biologie Santé de Lille – ED446 (scrutin avril 2021). Participation aux réunions du conseil de l'ED et aux concours pour l'attribution des contrats doctoraux en tant qu'observatrice

Depuis 2022 :

Membre de l'International Society of Exposure Science. <https://intlexposurescience.org/>

Membre de la Fédération Hospitalo-Universitaire (FHU) RESPIRE, Work-Packages « Environnement ». <https://www.chu-amiens.fr/chercheurs/fhu-respire/>

Organisations de colloques

2021 : « Accès et l'utilisation des données en libre accès en géographie de la population, de la santé et de la mobilité » 2 Décembre 2021, Paris

Membre du comité d'organisation d'une journée d'étude sur l'accès et l'utilisation des données en libre accès en géographie de la population, de la santé et de la mobilité (Comité National Français de Géographie CNFG, Collège International des Sciences Territoriales CIST, axes Territoires et santé, Information territoriale locale et Mobilités et territoires).

2022 : « Journée d'étude SIGLES : analyses spatiales en santé environnement » 20 Mai 2022, Lille

Membre du comité d'organisation d'une journée d'étude visant à rassembler les équipes travaillant dans les champs de l'analyse spatiale en santé environnement et sur la caractérisation écologique d'un territoire, sur les données et les méthodes mobilisées et sur la mesure du poids de l'environnement dans l'étiologie des maladies multifactorielles.

Bibliographie

- ABM, 2015. Rapport annuel 2015 : Réseau, Épidémiologie, Information, Néphrologie. Registre français des traitements de suppléance de l'insuffisance rénale chronique. Agence de la Bio-Médecine.
- Aerts, R., Nemery, B., Bauwelinck, M., Trabelsi, S., Deboosere, P., Van Nieuwenhuysse, A., Nawrot, T.S., Casas, L., 2020. Residential green space, air pollution, socioeconomic deprivation and cardiovascular medication sales in Belgium: A nationwide ecological study. *Science of The Total Environment* 712, 136426. <https://doi.org/10.1016/j.scitotenv.2019.136426>
- Afsar, B., Elsurer Afsar, R., Kanbay, A., Covic, A., Ortiz, A., Kanbay, M., 2019. Air pollution and kidney disease: review of current evidence. *Clinical Kidney Journal* 12, 19–32. <https://doi.org/10.1093/ckj/sfy111>
- AGUR, 2019. SCOT Région Flandre-Dunkerque, Diagnostic thématique - Description du territoire. Agence d'Urbanisme et de développement de la région Flandre-Dunkerque.
- AGUR, 2015. Plan Local D'urbanisme intercommunal Habitat Déplacements - Etat des lieux. Agence d'Urbanisme Flandre Dunkerque.
- Alla, F., 2016. 3. Les déterminants de la santé, in: *Traité de santé publique*. Lavoisier, p. 15. <https://doi.org/10.3917/lav.bourd.2016.01.0040>
- Almaguer, M., Herrera, R., Orrantes, 2014. Chronic Kidney Disease of Unknown Etiology in Agricultural Communities. *MEDICC Review* 16, 9–15.
- Amann, M., Organizaci??n Mundial de la Salud, Regional Office for Europe, 2008. Health risks of ozone from long-range transboundary air pollution. World Health Organization, Regional Office for Europe, Copenhagen.
- Alvarez, C.H., Evans, C.R., 2021. Intersectional environmental justice and population health inequalities: A novel approach. *Social Science & Medicine* 269, 113559. <https://doi.org/10.1016/j.socscimed.2020.113559>
- Anctil, H., Bédard, Lucie, Québec (Province), Ministère de la santé et des services sociaux, 2012. La santé et ses déterminants : mieux comprendre pour mieux agir. Ministère de la santé et des services sociaux, Québec.
- Andersson, E.M., Ögren, M., Molnár, P., Segersson, D., Rosengren, A., Stockfelt, L., 2020. Road traffic noise, air pollution and cardiovascular events in a Swedish cohort. *Environmental Research* 185, 109446. <https://doi.org/10.1016/j.envres.2020.109446>
- Angeon, V., Bates, S., 2015. Reviewing Composite Vulnerability and Resilience Indexes: A Sustainable Approach and Application. *World Development* 72, 140–162. <https://doi.org/10.1016/j.worlddev.2015.02.011>
- Anselin, L., 2000. The Alchemy of Statistics, or Creating Data Where No Data Exist. *Annals of the Association of American Geographers* 90, 586–592. <https://doi.org/10.1111/0004-5608.00212>
- Arcaya, M., Brewster, M., Zigler, C.M., Subramanian, S.V., 2012. Area variations in health: A spatial multilevel modeling approach. *Health & Place* 18, 824–831. <https://doi.org/10.1016/j.healthplace.2012.03.010>
- Arcaya, M.C., Arcaya, A.L., Subramanian, S.V., 2015. Inequalities in health: definitions, concepts, and theories. *Global Health Action* 8, 27106. <https://doi.org/10.3402/gha.v8.27106>
- Arcaya, M.C., Tucker-Seeley, R.D., Kim, R., Schnake-Mahl, A., So, M., Subramanian, S.V., 2016. Research on neighborhood effects on health in the United States: A systematic review of study characteristics. *Social Science & Medicine* 168, 16–29. <https://doi.org/10.1016/j.socscimed.2016.08.047>

- Argacha, J.F., Bourdrel, T., van de Borne, P., 2018. Ecology of the cardiovascular system: A focus on air-related environmental factors. *Trends in Cardiovascular Medicine* 28, 112–126. <https://doi.org/10.1016/j.tcm.2017.07.013>
- Arnett, D.K., Blumenthal, R.S., Albert, M.A., Buroker, A.B., Goldberger, Z.D., Hahn, E.J., Himmelfarb, C.D., Khera, A., Lloyd-Jones, D., McEvoy, J.W., Michos, E.D., Miedema, M.D., Muñoz, D., Smith, S.C., Virani, S.S., Williams, K.A., Yeboah, J., Ziaeian, B., 2019. 2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Circulation* 140. <https://doi.org/10.1161/CIR.0000000000000678>
- ARS, 2018. Projt régional de santé des Hauts-de-France : programme régional d'accès à la prévention et aux soins des peersonnes les plus démunies 2018-2013., Cadre d'orientation stratégique schéma régional de santé PRAPS. Agence Régionale de Santé, Hauts-de-France.
- ARS, 2017. Priorités de santé du territoire de démocratie sanitaire du Pas-de-Calais - Diagnostic partagé du Conseil Territorial de Santé. Agence Régionale de Santé Hauts-de-France.
- Astell-Burt, T., Navakatikyan, M.A., Walsan, R., Davis, W., Figtree, G., Arnolda, L., Feng, X., 2021. Green space and cardiovascular health in people with type 2 diabetes. *Health & Place* 69, 102554. <https://doi.org/10.1016/j.healthplace.2021.102554>
- AUDIAR, 2016. Indicateurs de bien-être - Etat des lieux. Agence d'Urbanisme et de Développement Intercommunal de l'Agglomération Rennaise.
- Avdic, D., 2016. Improving efficiency or impairing access? Health care consolidation and quality of care: Evidence from emergency hospital closures in Sweden. *Journal of Health Economics* 48, 44–60. <https://doi.org/10.1016/j.jhealeco.2016.02.002>
- Badaloni, C., Cesaroni, G., Cerza, F., Davoli, M., Brunekreef, B., Forastiere, F., 2017. Effects of long-term exposure to particulate matter and metal components on mortality in the Rome longitudinal study. *Environment International* 109, 146–154. <https://doi.org/10.1016/j.envint.2017.09.005>
- Bailleul, Y., Lefèvre, L., 2019. Un jeune sur huit non scolarisé et non diplômé - Les jeunes en Hauts-de-France. Insee Analyses Hauts-de-France.
- Bambra, C., Cairns, J.M., Kasim, A., Smith, J., Robertson, S., Copeland, A., Johnson, K., 2015. This divided land: An examination of regional inequalities in exposure to brownfield land and the association with morbidity and mortality in England. *Health & Place* 34, 257–269. <https://doi.org/10.1016/j.healthplace.2015.05.010>
- Barzegar, N., Tohidi, M., Ghodssi-Ghassemabadi, R., Amiri, P., Azizi, F., Hadaegh, F., 2021. Impact of educational level on incident chronic kidney disease during 13 years of follow-up: a prospective cohort study. *Public Health* 195, 98–104. <https://doi.org/10.1016/j.puhe.2021.04.006>
- Bauwelinck, M., Casas, L., Nawrot, T.S., Nemery, B., Trabelsi, S., Thomas, I., Aerts, R., Lefebvre, W., Vanpoucke, C., Van Nieuwenhuysse, A., Deboosere, P., Vandenheede, H., 2021. Residing in urban areas with higher green space is associated with lower mortality risk: A census-based cohort study with ten years of follow-up. *Environment International* 148, 106365. <https://doi.org/10.1016/j.envint.2020.106365>
- Beale, L., Abellan, J.J., Hodgson, S., Jarup, L., 2008. Methodologic Issues and Approaches to Spatial Epidemiology. *Environmental Health Perspectives* 116, 1105–1110. <https://doi.org/10.1289/ehp.10816>
- Beccari, B., 2016. A Comparative Analysis of Disaster Risk, Vulnerability and Resilience Composite Indicators. *PLoS Currents*. <https://doi.org/10.1371/currents.dis.453df025e34b682e9737f95070f9b970>
- Becker, W., Saisana, M., Paruolo, P., Vandecasteele, I., 2017. Weights and importance in composite indicators: Closing the gap. *Ecological Indicators* 80, 12–22. <https://doi.org/10.1016/j.ecolind.2017.03.056>
- Becuwe, B., Lecomte, M., 2015. Des trajectoires sociales différenciées entre les territoires du Nord-Pas de Calais. INSEE Analyses Nord-Pas-de-Calais 4.

- Begou, P., Kassomenos, P., Kelessis, A., 2020. Effects of road traffic noise on the prevalence of cardiovascular diseases: The case of Thessaloniki, Greece. *Science of The Total Environment* 703, 134477. <https://doi.org/10.1016/j.scitotenv.2019.134477>
- Belhakem, N., Lermechin, H., 2018. Une pauvreté très élevée, mais moins intense. *INSEE Analyses Hauts de France* 4.
- Benmarhnia, T., 2010. Les inégalités environnementales et leurs rôles dans les inégalités sociales de santé : état de l'art et réflexions sur le développement d'indicateurs de défaveur environnementale (Mémoire). Ecole des Hautes Etudes en Santé Publique, Rennes.
- Bentley, R.J., Pevalin, D., Baker, E., Mason, K., Reeves, A., Beer, A., 2016. Housing affordability, tenure and mental health in Australia and the United Kingdom: a comparative panel analysis. *Housing Studies* 31, 208–222. <https://doi.org/10.1080/02673037.2015.1070796>
- Besag, J., York, J., Mollié, A., 1991. Bayesian image restoration, with two applications in spatial statistics. *Ann Inst Stat Math* 43, 1–20. <https://doi.org/10.1007/BF00116466>
- Bevan, G., Pandey, A., Griggs, S., Dalton, J.E., Zidar, D., Patel, S., Khan, S.U., Nasir, K., Rajagopalan, S., Al-Kindi, S., 2022. Neighborhood-Level Social Vulnerability and Prevalence of Cardiovascular Risk Factors and Coronary Heart Disease. *Current Problems in Cardiology* 101182. <https://doi.org/10.1016/j.cpcardiol.2022.101182>
- Bhatnagar, A., 2016. Cardiovascular Perspective of the Promises and Perils of E-Cigarettes. *Circ Res* 118, 1872–1875. <https://doi.org/10.1161/CIRCRESAHA.116.308723>
- Biggeri, M., Mauro, V., 2018. Towards a more 'Sustainable' Human Development Index: Integrating the environment and freedom. *Ecological Indicators* 91, 220–231. <https://doi.org/10.1016/j.ecolind.2018.03.045>
- Bilgel, F., 2019. Spatial distribution of inequalities in end-stage renal disease in the United States. *Spatial and Spatio-temporal Epidemiology* 30, 100282. <https://doi.org/10.1016/j.sste.2019.100282>
- Bixby, H., Hodgson, S., Fortunato, L., Hansell, A., Fecht, D., 2015. Associations between Green Space and Health in English Cities: An Ecological, Cross-Sectional Study. *PLoS ONE* 10, e0119495. <https://doi.org/10.1371/journal.pone.0119495>
- Blangiardo, M., Cameletti, M., 2015. *Spatial and spatio-temporal Bayesian models with R-INLA*. John Wiley and Sons, Inc, Chichester, West Sussex.
- Böhringer, C., Jochem, P.E.P., 2007. Measuring the immeasurable — A survey of sustainability indices. *Ecological Economics* 63, 1–8. <https://doi.org/10.1016/j.ecolecon.2007.03.008>
- Bourret, C., 2016. Relever le défi des Inégalités Sociales et Territoriales en Santé (ISTS) par la rencontre des organisations d'interface et de la télésanté dans une approche d'« humanisme numérique » en santé ? *rfsic*. <https://doi.org/10.4000/rfsic.2013>
- Bowe, B., Xie, Y., Li, T., Yan, Y., Xian, H., Al-Aly, Z., 2019. Estimates of the 2016 global burden of kidney disease attributable to ambient fine particulate matter air pollution. *BMJ Open* 9, e022450. <https://doi.org/10.1136/bmjopen-2018-022450>
- Bowe, B., Xie, Y., Li, T., Yan, Y., Xian, H., Al-Aly, Z., 2018. Particulate Matter Air Pollution and the Risk of Incident CKD and Progression to ESRD. *JASN* 29, 218–230. <https://doi.org/10.1681/ASN.2017030253>
- Bowe, B., Xie, Y., Xian, H., Lian, M., Al-Aly, Z., 2017. Geographic Variation and US County Characteristics Associated With Rapid Kidney Function Decline. *Kidney International Reports* 2, 5–17. <https://doi.org/10.1016/j.ekir.2016.08.016>
- Braconnier, P., Caire, G., Dupuy, R., Jahan, E., Le Masne, P., Nivoix, S., 2011. Une évaluation du bien-être au sein des régions françaises dans une approche de développement durable. *Innovations* 34, 93. <https://doi.org/10.3917/inno.034.0093>
- Branchu, M., Costemalle, V., Fontaine, M., 2018. Données carroyées et confidentialité (13è Journées de méthodologie statistique de l'INSEE (JMS)). INSEE, Paris.
- Bravo, G., 2014. The Human Sustainable Development Index: New calculations and a first critical analysis. *Ecological Indicators* 37, 145–150. <https://doi.org/10.1016/j.ecolind.2013.10.020>

- Briggs, A.C., Black, A.W., Lucas, F.L., Siewers, A.E., Fairfield, K.M., 2019. Association between the food and physical activity environment, obesity, and cardiovascular health across Maine counties. *BMC Public Health* 19, 374. <https://doi.org/10.1186/s12889-019-6684-6>
- Briggs, D.J., 2008. A framework for integrated environmental health impact assessment of systemic risks. *Environ Health* 7, 61. <https://doi.org/10.1186/1476-069X-7-61>
- Brousmiche, D., Genin, M., Occelli, F., Frank, L., Deram, A., Cuny, D., Lanier, C., 2020a. How can we analyze environmental health resilience and vulnerability? A joint analysis with composite indices applied to the north of France. *Science of The Total Environment* 142983. <https://doi.org/10.1016/j.scitotenv.2020.142983>
- Brousmiche, D., Occelli, F., Genin, M., Cuny, D., Deram, A., Lanier, C., 2020b. Spatialized composite indices to evaluate environmental health inequalities: Meeting the challenge of selecting relevant variables. *Ecological Indicators* 111, 106023. <https://doi.org/10.1016/j.ecolind.2019.106023>
- Brousmiche, D., Genin, M., Occelli, F., Frank, L., Deram, A., Cuny, D., Lanier, C., 2021. Data for the assessment of vulnerability and resilience in the field of environmental health in the north of France. Data in Brief 107220. <https://doi.org/10.1016/j.dib.2021.107220>
- Brousmiche, D., Lanier, C., Occelli, F., Pochet, L., Degezelle, S., Genin, M., Deram, A., Cuny, D., *Under Review*. Spatial scale and spatial effects: shedding light on how a change of scale can affect the characterization of environmental inequalities. *Environmental Sciences & Policy*
- Bryere, J., Pernet, C., Copin, N., Launay, L., Gusto, G., Grosclaude, P., Delpierre, C., Lang, T., Lantieri, O., Dejardin, O., Launoy, G., 2017. Assessment of the ecological bias of seven aggregate social deprivation indices. *BMC Public Health* 17, 86. <https://doi.org/10.1186/s12889-016-4007-8>
- Burgass, M.J., Halpern, B.S., Nicholson, E., Milner-Gulland, E.J., 2017. Navigating uncertainty in environmental composite indicators. *Ecological Indicators* 75, 268–278. <https://doi.org/10.1016/j.ecolind.2016.12.034>
- Buzzelli, M., 2020. Modifiable Areal Unit Problem, in: *International Encyclopedia of Human Geography*. pp. 169–173
- Cai, Q., Dekker, L.H., Bakker, S.J.L., de Borst, M.H., Navis, G., 2019. Intraregional differences in renal function in the Northern Netherlands: The Lifelines Cohort Study. *PLoS ONE* 14, e0223908. <https://doi.org/10.1371/journal.pone.0223908>
- Cai, Y., Hodgson, S., Blangiardo, M., Gulliver, J., Morley, D., Fecht, D., Vienneau, D., de Hoogh, K., Key, T., Hveem, K., Elliott, P., Hansell, A.L., 2018. Road traffic noise, air pollution and incident cardiovascular disease: A joint analysis of the HUNT, EPIC-Oxford and UK Biobank cohorts. *Environment International* 114, 191–201. <https://doi.org/10.1016/j.envint.2018.02.048>
- Cambois, E., Ghosn, W., Rey, G., 2021. État des lieux des inégalités sociales de santé: Actualité et dossier en santé publique N° 113, 10–16. <https://doi.org/10.3917/aedesp.113.0010>
- Carter, A.R., Gill, D., Davies, N.M., Taylor, A.E., Tillmann, T., Vaucher, J., Wootton, R.E., Munafò, M.R., Hemani, G., Malik, R., Seshadri, S., Woo, D., Burgess, S., Davey Smith, G., Holmes, M.V., Tzoulaki, I., Howe, L.D., Dehghan, A., 2019. Understanding the consequences of education inequality on cardiovascular disease: mendelian randomisation study. *BMJ* l1855. <https://doi.org/10.1136/bmj.l1855>
- Cartone, A., Postiglione, P., 2021. Principal component analysis for geographical data: the role of spatial effects in the definition of composite indicators. *Spatial Economic Analysis* 16, 126–147. <https://doi.org/10.1080/17421772.2020.1775876>
- Caudeville, J., Rican, S., 2016. Étude de l'inégalité socio-environnementale en France : associations spatiales entre désavantage social des populations et proximité à un site potentiellement dangereux. *Env. Risque Sante* 15, 39–47. <https://doi.org/10.1684/ers.2015.0828>
- Cebrecos, A., Domínguez-Berjón, M.F., Duque, I., Franco, M., Escobar, F., 2018. Geographic and statistic stability of deprivation aggregated measures at different spatial units in health research. *Applied Geography* 95, 9–18. <https://doi.org/10.1016/j.apgeog.2018.04.001>

- Ceccarelli, C., Guandalini, A., Martini, A., Pontecorvo, M.E., 2020. Accuracy evaluation of LFS-BES Indicators: A regional assessment. *Soc Indic Res.* <https://doi.org/10.1007/s11205-020-02532-3>
- Chang, H.H., 2016. Data assimilation for environmental pollution fields, in: *Handbook of Spatial Epidemiology*, Chapman & Hall/CRC Handbooks of Modern Statistical Methods. CRC Press/Taylor & Francis, Boca Raton, p. 684.
- Chang, K., Wu, I.-W., Huang, B.-R., Juang, J.-G., Wu, J.-C., Chang, S.-W., Chang, C., 2018. Associations between Water Quality Measures and Chronic Kidney Disease Prevalence in Taiwan. *IJERPH* 15, 2726. <https://doi.org/10.3390/ijerph15122726>
- Charlson, M.E., Pompei, P., Ales, K.L., MacKenzie, C.R., 1987. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *Journal of Chronic Diseases* 40, 373–383. [https://doi.org/10.1016/0021-9681\(87\)90171-8](https://doi.org/10.1016/0021-9681(87)90171-8)
- Chauvin, P., Estecahandy, P., 2010. Inégalités sociales de santé et précarité. *Actualité et Dossier en Santé Publique (ADSP)* 17–18.
- Chen, K., Breitner, S., Wolf, K., Hampel, R., Meisinger, C., Heier, Margit, von Scheidt, W., Kuch, B., Peters, Annette, Schneider, A., for the KORA Study Group, Peters, A., Schulz, H., Schwettmann, L., Leidl, R., Heier, M., Strauch, K., 2019. Temporal variations in the triggering of myocardial infarction by air temperature in Augsburg, Germany, 1987–2014. *European Heart Journal* 40, 1600–1608. <https://doi.org/10.1093/eurheartj/ehz116>
- Cheng, Y.-Y., Chang, Y.-T., Cheng, H.-L., Shen, K.-H., Sung, J.-M., Guo, H.-R., 2018. Associations between arsenic in drinking water and occurrence of end-stage renal disease with modifications by comorbidities: A nationwide population-based study in Taiwan. *Science of The Total Environment* 626, 581–591. <https://doi.org/10.1016/j.scitotenv.2018.01.043>
- Chessel, D., Ollier, S., Dray, S., 2003. Ordination sous contrainte spatiale (Fiche de Biostatistique No. 8), *Biométrie et Biologie Evolutive*. Université Lyon 1.
- Cheung, A.K., Chang, T.I., Cushman, W.C., Furth, S.L., Hou, F.F., Ix, J.H., Knoll, G.A., Muntner, P., Pecoits-Filho, R., Sarnak, M.J., Tobe, S.W., Tomson, C.R.V., Lytvyn, L., Craig, J.C., Tunnicliffe, D.J., Howell, M., Tonelli, M., Cheung, M., Earley, A., Mann, J.F.E., 2021. Executive summary of the KDIGO 2021 Clinical Practice Guideline for the Management of Blood Pressure in Chronic Kidney Disease. *Kidney International* 99, 559–569. <https://doi.org/10.1016/j.kint.2020.10.026>
- Chevillard, G., Lucas-Gabrielli, V., Mousques, J., 2018. « Déserts médicaux » en France : état des lieux et perspectives de recherches: *L'Espace géographique* Tome 47, 362–380. <https://doi.org/10.3917/eg.474.0362>
- Chien, J.-W., Wu, C., Chan, C.-C., 2021. Open space and adult's chronic kidney disease, overweight and diabetes in the metropolitan area of New Taipei City. *Journal of the Formosan Medical Association* S0929664621005015. <https://doi.org/10.1016/j.jfma.2021.11.004>
- Chignard, S., 2012. *Open data: comprendre l'ouverture des données publiques*, Collection Entreprendre. Éditions fyp, Paris.
- Cincinelli, A., Katsoyiannis, A., 2017. Atmospheric pollution in city centres and urban environments. The impact of scientific, regulatory and industrial progress. *Science of The Total Environment* 579, 1057–1058. <https://doi.org/10.1016/j.scitotenv.2016.11.057>
- Clayton, D., Kaldor, J., 1987. Empirical Bayes Estimates of Age-Standardized Relative Risks for Use in Disease Mapping. *Biometrics* 43, 671. <https://doi.org/10.2307/2532003>
- Cliff, A.D., Ord, J.K., 1981. *Spatial processes: Models and Applications*. Pion Limited, London.
- Collister, D., Ferguson, T., Komenda, P., Tangri, N., 2016. The Patterns, Risk Factors, and Prediction of Progression in Chronic Kidney Disease: A Narrative Review. *Seminars in Nephrology* 36, 273–282. <https://doi.org/10.1016/j.semnephrol.2016.05.004>
- Colonna, M., Sauleau, E.-A., 2013. How to interpret and choose a Bayesian spatial model and a Poisson regression model in the context of describing small area cancer risks variations. *Revue d'Épidémiologie et de Santé Publique* 61, 559–567. <https://doi.org/10.1016/j.respe.2013.07.686>

- Comber, A., Brunsdon, C., Green, E., 2008. Using a GIS-based network analysis to determine urban greenspace accessibility for different ethnic and religious groups. *Landscape and Urban Planning* 86, 103–114. <https://doi.org/10.1016/j.landurbplan.2008.01.002>
- Commission on Social Determinants of Health, 2008. Closing the gap in a generation : health equity through action on the social determinants of health. (Final report of the commission on social determinants of health). World Health Organization, Geneva.
- Concialdi, P., 2014. Pauvreté, précarité et chômage. *Spécificités* 6, 30. <https://doi.org/10.3917/spec.006.0030>
- Costanza, R., d'Arge, R., de Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'Neill, R.V., Paruelo, J., Raskin, R.G., Sutton, P., van den Belt, M., 1998. The value of the world's ecosystem services and natural capital. *Ecological Economics* 25, 3–15. [https://doi.org/10.1016/S0921-8009\(98\)00020-2](https://doi.org/10.1016/S0921-8009(98)00020-2)
- Couchoud, C., Guihenneuc, C., Bayer, F., Lemaitre, V., Brunet, P., Stengel, B., On behalf of the REIN Registry, 2012. Medical practice patterns and socio-economic factors may explain geographical variation of end-stage renal disease incidence. *Nephrology Dialysis Transplantation* 27, 2312–2322. <https://doi.org/10.1093/ndt/gfr639>
- Cuchère, E., Terra, S., 2017. Atouts et défis de la région Hauts-de-France. *INSEE Analyses Hauts de France* 4.
- Cuny, D., Rzepka, M., Van Haluwyn, C., 2008. Introduction à la biosurveillance végétale et fongique de la qualité de l'air, concept et intérêt. *Spectra Analyse* 31–35.
- Cuny, D., 2012. La biosurveillance végétale et fongique de la pollution atmosphérique : concepts et applications. *Annales Pharmaceutiques Françaises* 70, 182–187. <https://doi.org/10.1016/j.pharma.2012.05.003>
- Cutter, S.L., Burton, C.G., Emrich, C.T., 2010. Disaster Resilience Indicators for Benchmarking Baseline Conditions. *Journal of Homeland Security and Emergency Management* 7. <https://doi.org/10.2202/1547-7355.1732>
- D2DPE, 2010. Fiche signalétique CU Dunkerque - Grand Littoral, Mission Observation Régionale et Analyse Spatiale. Direction du Développement Durable, de la Prospective et de l'Évaluation, Hauts-de-France.
- Dahlgren, G., Whitehead, M., 2007. European strategies for tackling social inequalities in health: Levelling up Part 2 (No. WHOLIS E89384). World Health Organization, Regional Office for Europe.
- Dahlgren, G., Whitehead, M., 1991. Policies and strategies to promote social equity in health. Background document to WHO - strategy paper for Europe. Institute for futures studies.
- Daiber, A., Lelieveld, J., Steven, S., Oelze, M., Kröller-Schön, S., Sørensen, M., Münzel, T., 2019. The “exposome” concept – how environmental risk factors influence cardiovascular health. *Acta Biochim Pol.* https://doi.org/10.18388/abp.2019_2853
- Danet, S., 2010. Les caractéristiques des inégalités sociales de santé. *Actualité et Dossier en Santé Publique (ADSP), les inégalités sociales de santé* 8–14.
- Dark, S.J., Bram, D., 2007. The modifiable areal unit problem (MAUP) in physical geography. *Progress in Physical Geography: Earth and Environment* 31, 471–479. <https://doi.org/10.1177/0309133307083294>
- Daoui, D., 2018. Indicateur de Développement Humain “IDH” : vers un indicateur d'attractivité. *Journal of the Geopolitics and Geostrategic Intelligence* 1, 39–44.
- Davide, D.F., Alessandra, F., Roberto, P., 2022. Distributive justice in environmental health hazards from industrial contamination: A systematic review of national and near-national assessments of social inequalities. *Social Science & Medicine* 297, 114834. <https://doi.org/10.1016/j.socscimed.2022.114834>
- Davoodabadi, Z., Soleimani, A., Pourmoghaddas, A., Hosseini, S.M., Jafari-Koshki, T., Rahimi, M., Shishehforoush, M., Lahijan-zadeh, A., Sadeghian, B., Moazam, E., Mohebi, M.B., Ezatian, V., Rabiei, K., Sarrafzadegan, N., 2019. Correlation between air pollution and hospitalization due

- to myocardial infarction. *ARYA Atherosclerosis* 15. <https://doi.org/10.22122/arya.v15i4.1834>
- de Groot, R., Brander, L., van der Ploeg, S., Costanza, R., Bernard, F., Braat, L., Christie, M., Crossman, N., Ghermandi, A., Hein, L., Hussain, S., Kumar, P., McVittie, A., Portela, R., Rodriguez, L.C., ten Brink, P., van Beukering, P., 2012. Global estimates of the value of ecosystems and their services in monetary units. *Ecosystem Services* 1, 50–61. <https://doi.org/10.1016/j.ecoser.2012.07.005>
- de Lapasse, B., Garnier, M., Haran, L., Kaldi, M., Lecomte, C., Toutin, G., 2018. Regards sur les territoires: population, conditions de vie, éducation, emploi transfrontalier Rapport 2017 (No. ISBN 978-2-11-139268-7), Regards sur les territoires. Commissariat général à l'égalité des territoires, Saint-Denis.
- de Vries, S., Verheij, R.A., Groenewegen, P.P., Spreeuwenberg, P., 2003. Natural Environments—Healthy Environments? An Exploratory Analysis of the Relationship between Greenspace and Health. *Environ Plan A* 35, 1717–1731. <https://doi.org/10.1068/a35111>
- Dekkers, M., Loutas, N., De Keyzer, M., Goedertier, S., 2013. La qualité des données et métadonnées ouvertes. Module de formation 2.2.
- Desenclos, J.C., Arwidson, P., 2016. Inégalités sociales et territoriales de santé : des connaissances et des faits probants pour l'action ! *Bulletin Epidémiologique Hebdomadaire Editorial*, 274–275.
- Devilleine, V., 2016. Une forte surmortalité par maladies cardiovasculaires.
- Dhindsa, D.S., Khambhati, J., Schultz, W.M., Tahhan, A.S., Quyyumi, A.A., 2020. Marital status and outcomes in patients with cardiovascular disease. *Trends in Cardiovascular Medicine* 30, 215–220. <https://doi.org/10.1016/j.tcm.2019.05.012>
- Dialga, I., 2016. L'Indice du Développement Humain ou l'indice monétaire de développement ? Evidences par l'Analyse en Composantes Principales. HAL open science.
- Dollé, M., 2008. La transmission intergénérationnelle de la pauvreté. Regards croisés sur l'économie 4, 97. <https://doi.org/10.3917/rce.004.0097>
- Dominski, F.H., Lorenzetti Branco, J.H., Buonanno, G., Stabile, L., Gameiro da Silva, M., Andrade, A., 2021. Effects of air pollution on health: A mapping review of systematic reviews and meta-analyses. *Environmental Research* 201, 111487. <https://doi.org/10.1016/j.envres.2021.111487>
- Donovan, G.H., Michael, Y.L., Gatzolis, D., Prestemon, J.P., Whitsel, E.A., 2015. Is tree loss associated with cardiovascular-disease risk in the Women's Health Initiative? A natural experiment. *Health & Place* 36, 1–7. <https://doi.org/10.1016/j.healthplace.2015.08.007>
- Dourlens, C., 2008. Les usages de la fragilité dans le champ de la santé. *Alter* 2, 156–178. <https://doi.org/10.1016/j.alter.2008.03.004>
- DREAL, 2020. Plan de Protection de l'Atmosphère du Nord-Pas de Calais - Rapport d'évaluation. Direction Régionale de l'Environnement, de l'Aménagement et du Logement.
- DREES, Santé Publique France, 2017. L'état de santé de la population en France - Rapport 2017 (No. ISBN: 978-2-11-151590-1). Direction de la Recherche, des Etudes, de l'Evaluation et des Statistiques (DREES).
- Ducimetière, P., Haas, B., Ruidavets, J.-B., Montaye, M., Wagner, A., Ferrières, J., Dallongeville, J., Bingham, A., 2011. Fréquence et mortalité à 28 jours des divers épisodes d'insuffisance coronaire aiguë dans trois régions françaises en 2006. *Bulletin Epidémiologique Hebdomadaire* 419–22.
- Ducros, D., Nicoules, V., Chehoud, H., Bayle, A., Souche, A., Tanguy, M., Valière, J.-P., Cayla, F., Grosclaude, P., 2015. Les bases médico-administratives pour mesurer les inégalités sociales de santé. *Santé Publique* 27, 383. <https://doi.org/10.3917/spub.153.0383>
- Dulong de Rosnay, M., 2017. Données ouvertes (Open Data) in Marie Cornu, Judith Rochfeld et Fabienne Orsi (dir.), *Dictionnaire des biens communs*, Presses Universitaires de France.

- Dunn, J.R., 2002. Housing and inequalities in health: a study of socioeconomic dimensions of housing and self reported health from a survey of Vancouver residents. *Journal of Epidemiology & Community Health* 56, 671–681. <https://doi.org/10.1136/jech.56.9.671>
- Dunn, J.R., Hayes, M.V., 2000. Social inequality, population health, and housing: a study of two Vancouver neighborhoods. *Social Science & Medicine* 51, 563–587. [https://doi.org/10.1016/S0277-9536\(99\)00496-7](https://doi.org/10.1016/S0277-9536(99)00496-7)
- Eckersley, R., 2015. Beyond inequality: Acknowledging the complexity of social determinants of health. *Social Science & Medicine* 147, 121–125. <https://doi.org/10.1016/j.socscimed.2015.10.052>
- EGR, 2013. Les états généraux du rein - le rapport.
- Estabrooks, P.A., Lee, R.E., Gyurcsik, N.C., 2003. Resources for physical activity participation: Does availability and accessibility differ by neighborhood socioeconomic status? *Annals of Behavioral Medicine* 25, 100–104. https://doi.org/10.1207/S15324796ABM2502_05
- Estoque, R.C., Murayama, Y., 2017. A worldwide country-based assessment of social-ecological status (c. 2010) using the social-ecological status index. *Ecological Indicators* 72, 605–614. <https://doi.org/10.1016/j.ecolind.2016.08.047>
- Estoque, R.C., Murayama, Y., 2014. Social–ecological status index: A preliminary study of its structural composition and application. *Ecological Indicators* 43, 183–194. <https://doi.org/10.1016/j.ecolind.2014.02.031>
- Estoque, R.C., Murayama, Y., 2013. Landscape pattern and ecosystem service value changes: Implications for environmental sustainability planning for the rapidly urbanizing summer capital of the Philippines. *Landscape and Urban Planning* 116, 60–72. <https://doi.org/10.1016/j.landurbplan.2013.04.008>
- Etsy, D.C., Levy, M., Srebotnjak, T., de Sherbinin, A., 2005. Environmental Sustainability Index: Benchmarking Nationale Environmental Stewardship. Yale Center for Environmental Law and Policy, New Haven.
- European Commission, Organisation for Economic Co-operation and Development, SourceOECD (Online service) (Eds.), 2008. Handbook on constructing composite indicators: methodology and user guide. OECD, Paris.
- Exner, A., Politti, E., Schriefl, E., Erker, S., Stangl, R., Baud, S., Warmuth, H., Matzenberger, J., Kranzl, L., Paulesich, R., Windhaber, M., Supper, S., Stöglehner, G., 2016. Measuring regional resilience towards fossil fuel supply constraints. Adaptability and vulnerability in socio-ecological Transformations-the case of Austria. *Energy Policy* 91, 128–137. <https://doi.org/10.1016/j.enpol.2015.12.031>
- Eze, I.C., Schaffner, E., Fischer, E., Schikowski, T., Adam, M., Imboden, M., Tsai, M., Carballo, D., von Eckardstein, A., Künzli, N., Schindler, C., Probst-Hensch, N., 2014. Long-term air pollution exposure and diabetes in a population-based Swiss cohort. *Environment International* 70, 95–105. <https://doi.org/10.1016/j.envint.2014.05.014>
- Faburel, G., 2008. Les inégalités environnementales comme inégalités de moyens des habitants et des acteurs territoriaux: Pour que l’environnement soit un facteur réel de cohésion urbaine. *eps* 111–126. <https://doi.org/10.4000/eps.2430>
- Fang, M., Hu, L., Chen, D., Guo, Y., Liu, J., Lan, C., Gong, J., Wang, B., 2021. Exposome in human health: Utopia or wonderland? *The Innovation* 2, 100172. <https://doi.org/10.1016/j.xinn.2021.100172>
- Fekete, A., 2018. Societal resilience indicator assessment using demographic and infrastructure data at the case of Germany in context to multiple disaster risks. *International Journal of Disaster Risk Reduction* 31, 203–211. <https://doi.org/10.1016/j.ijdrr.2018.05.004>
- Feng, Y., Jones, M.R., Chu, N.M., Segev, D.L., McAdams-DeMarco, M., 2021. Ambient Air Pollution and Mortality among Older Patients Initiating Maintenance Dialysis. *Am J Nephrol* 52, 217–227. <https://doi.org/10.1159/000514233>

- Fernández, I.C., Wu, J., 2016. Assessing environmental inequalities in the city of Santiago (Chile) with a hierarchical multiscale approach. *Applied Geography* 74, 160–169. <https://doi.org/10.1016/j.apgeog.2016.07.012>
- Fischer, R.S.B., Unrine, J.M., Vangala, C., Sanderson, W.T., Mandayam, S., Murray, K.O., 2020. Evidence of nickel and other trace elements and their relationship to clinical findings in acute Mesoamerican Nephropathy: A case-control analysis. *PLoS ONE* 15, e0240988. <https://doi.org/10.1371/journal.pone.0240988>
- Flacke, J., 2015. Spatial urban health equity indicators – a framework-based approach supporting spatial decision making, in: Köckler, H. (Ed.), . Presented at the SUSTAINABLE DEVELOPMENT AND PLANNING 2015, Istanbul, Turkey, pp. 365–376. <https://doi.org/10.2495/SDP150311>
- Flacke, J., Schüle, S., Köckler, H., Bolte, G., 2016. Mapping Environmental Inequalities Relevant for Health for Informing Urban Planning Interventions—A Case Study in the City of Dortmund, Germany. *International Journal of Environmental Research and Public Health* 13, 711. <https://doi.org/10.3390/ijerph13070711>
- Fleury, D., Peytavin, J.F., Alam, T., Godillon, S., Saint-Gérand, T., Medjkane, M., Millot, M., 2010. Inégalité sociale et risque routier. L'apport d'une approche territorialisée. (No. hal-00850864). *Les cahiers scientifiques du transport*.
- Fong, K.C., Hart, J.E., James, P., 2018. A Review of Epidemiologic Studies on Greenness and Health: Updated Literature Through 2017. *Curr Envir Health Rpt* 5, 77–87. <https://doi.org/10.1007/s40572-018-0179-y>
- Freudenberg, M., 2003. Composite Indicators of Country Performance (OECD Science, Technology and Industry Working Papers No. 2003/16). <https://doi.org/10.1787/405566708255>
- Frigerio, I., De Amicis, M., 2016. Mapping social vulnerability to natural hazards in Italy: A suitable tool for risk mitigation strategies. *Environmental Science & Policy* 63, 187–196. <https://doi.org/10.1016/j.envsci.2016.06.001>
- Fromageot, A., Coppieters, Y., Parent, F., Lagasse, R., 2005. Epidémiologie et géographie : une interdisciplinarité à développer pour l'analyse des relations entre santé et environnement. *Environnement Risques Santé* 4, 1–9.
- Gadrey, J., Jany-Catrice, F., 2016. II. Le PIB et la croissance en question, in: *Repères*. La Découverte, Paris, pp. 15–24.
- Galli, A., Wiedmann, T., Ercin, E., Knoblauch, D., Ewing, B., Giljum, S., 2012. Integrating Ecological, Carbon and Water footprint into a “Footprint Family” of indicators: Definition and role in tracking human pressure on the planet. *Ecological Indicators* 16, 100–112. <https://doi.org/10.1016/j.ecolind.2011.06.017>
- Gan, X., Fernandez, I.C., Guo, J., Wilson, M., Zhao, Y., Zhou, B., Wu, J., 2017. When to use what: Methods for weighting and aggregating sustainability indicators. *Ecological Indicators* 81, 491–502. <https://doi.org/10.1016/j.ecolind.2017.05.068>
- García-Sánchez, I.-M., Almeida, T.A. das N., Camara, R.P. de B., 2015. A proposal for a Composite Index of Environmental Performance (CIEP) for countries. *Ecological Indicators* 48, 171–188. <https://doi.org/10.1016/j.ecolind.2014.08.004>
- Gehlke, C.E., Biehl, K., 1934. Certain Effects of Grouping upon the Size of the Correlation Coefficient in Census Tract Material. *Journal of the American Statistical Association* 29, 169–170. <https://doi.org/10.1080/01621459.1934.10506247>
- Getis, A., 2007. Reflections on spatial autocorrelation. *Regional Science and Urban Economics* 37, 491–496. <https://doi.org/10.1016/j.regsciurbeco.2007.04.005>
- Ghosn, W., 2018a. Indicateurs écologiques synthétiques du niveau socio-économique pour la recherche en Santé (Note méthodologique). Institut national de la santé et de la recherche médicale (INSERM), Centre d'épidémiologie sur les causes médicales de décès CépiDc.
- Ghosn, W., 2018b. Note méthodologique : indicateurs écologiques synthétiques du niveau socio-économique pour la recherche en santé. CépiDc-Inserm.
- Ginot, L., Peyr, C., 2010. Habitat dégradé et santé perçue : une étude à partir des demandes de logement social. *Santé Publique* 22, 493. <https://doi.org/10.3917/spub.105.0493>

- Gkiouleka, A., Huijts, T., Beckfield, J., Bamba, C., 2018. Understanding the micro and macro politics of health: Inequalities, intersectionality & institutions - A research agenda. *Social Science & Medicine* 200, 92–98. <https://doi.org/10.1016/j.socscimed.2018.01.025>
- Glaser, J., Lemery, J., Rajagopalan, B., Diaz, H.F., García-Trabanino, R., Taduri, G., Madero, M., Amarasinghe, M., Abraham, G., Anutrakulchai, S., Jha, V., Stenvinkel, P., Roncal-Jimenez, C., Lanaspá, M.A., Correa-Rotter, R., Sheikh-Hamad, D., Burdmann, E.A., Andres-Hernando, A., Milagres, T., Weiss, I., Kanbay, M., Wesseling, C., Sánchez-Lozada, L.G., Johnson, R.J., 2016. Climate Change and the Emergent Epidemic of CKD from Heat Stress in Rural Communities: The Case for Heat Stress Nephropathy. *CJASN* 11, 1472–1483. <https://doi.org/10.2215/CJN.13841215>
- Glasser, A.M., Onnen, N., Craigmile, P.F., Schwartz, E., Roberts, M.E., 2022. Associations between disparities in tobacco retailer density and disparities in tobacco use. *Preventive Medicine* 154, 106910. <https://doi.org/10.1016/j.ypmed.2021.106910>
- Goldberg, M., Quantin, C., Guéguen, A., Zins, M., 2008. Bases de données médico-administratives et épidémiologie : intérêts et limites. *Courrier des statistiques* 124, 59–70.
- Golden, S.D., Kuo, T.-M., Kong, A.Y., Baggett, C.D., Henriksen, L., Ribisl, K.M., 2020. County-level associations between tobacco retailer density and smoking prevalence in the USA, 2012. *Preventive Medicine Reports* 17, 101005. <https://doi.org/10.1016/j.pmedr.2019.101005>
- Goodchild, M.F., 1986. Spatial autocorrelation, CATMOG. Geo Books, Norwich.
- Greco, S., Ishizaka, A., Tasiou, M., Torrisi, G., 2018. On the Methodological Framework of Composite Indices: A Review of the Issues of Weighting, Aggregation, and Robustness. *Social Indicators Research*. <https://doi.org/10.1007/s11205-017-1832-9>
- Guerrien, M., 2003. L'intérêt de l'analyse en composantes principales (ACP) pour la recherche en sciences sociales: Présentation à partir d'une étude sur le Mexique. *cal* 181–192. <https://doi.org/10.4000/cal.7364>
- Guo, L.-C., Lv, Z., Ma, W., Xiao, J., Lin, H., He, G., Li, X., Zeng, W., Hu, J., Zhou, Y., Li, M., Yu, S., Xu, Y., Zhang, J., Zhang, H., Liu, T., 2022. Contribution of heavy metals in PM_{2.5} to cardiovascular disease mortality risk, a case study in Guangzhou, China. *Chemosphere* 297, 134102. <https://doi.org/10.1016/j.chemosphere.2022.134102>
- Gutiérrez, O.M., 2015. Contextual Poverty, Nutrition, and Chronic Kidney Disease. *Advances in Chronic Kidney Disease* 22, 31–38. <https://doi.org/10.1053/j.ackd.2014.05.005>
- Haines, A., Ebi, K., 2019. The Imperative for Climate Action to Protect Health. *N Engl J Med* 380, 263–273. <https://doi.org/10.1056/NEJMra1807873>
- Halonen, J.I., Stenholm, S., Kivimäki, M., Pentti, J., Subramanian, S.V., Kawachi, I., Vahtera, J., 2015. Is change in availability of sports facilities associated with change in physical activity? A prospective cohort study. *Preventive Medicine* 73, 10–14. <https://doi.org/10.1016/j.ypmed.2015.01.012>
- Hamman, P., 2011. Les échelles spatiales et temporelles de la « ville durable ». *Espaces et sociétés* 144–145, 213. <https://doi.org/10.3917/esp.144.0213>
- Hartig, T., Mitchell, R., de Vries, S., Frumkin, H., 2014. Nature and Health. *Annu. Rev. Public Health* 35, 207–228. <https://doi.org/10.1146/annurev-publhealth-032013-182443>
- Harwell, L.C., Vivian, D.N., McLaughlin, M.D., Hafner, S.F., 2019. Scientific Data Management in the Age of Big Data: An Approach Supporting a Resilience Index Development Effort. *Front. Environ. Sci.* 7, 72. <https://doi.org/10.3389/fenvs.2019.00072>
- HAS, 2021. Guide du parcours de soins - Maladie rénale chronique de l'adulte (MRC°, Recommander les bonnes pratiques. Haute Autorité de Santé.
- HAS, 2010a. Analyse critique du développement d'indicateurs composites : le cas de l'infarctus du myocarde à la sortie de l'établissement de santé, Projet COMPAQH. INSERM.
- HAS, 2010b. Recommandations en santé publique : les conditions de mise en oeuvre de la télémédecine en unité de dialyse médicalisée - Argumentaire. Collège de la Haute Autorité de Santé.

- Hassan Bhat, T., Jiawen, G., Farzaneh, H., 2021. Air Pollution Health Risk Assessment (AP-HRA), Principles and Applications. *IJERPH* 18, 1935. <https://doi.org/10.3390/ijerph18041935>
- Hat, K., Stöglehner, G., 2019. How Resilient is Growth? Resilience Assessment of Austrian Municipalities on the Basis of Census Data from 1971 to 2011. *Sustainability* 11, 1818. <https://doi.org/10.3390/su11061818>
- Havard, S., 2008. Contribution de la pollution atmosphérique aux inégalités socio-spatiales de santé : analyse écologique du risque d'infarctus du myocarde dans l'agglomération de Strasbourg. Université de Rennes 1, Ecole des Hautes Etudes en Santé Publique.
- HCSP, 2021. Rapport relatif aux indicateurs composites en santé-environnement. Haut Conseil de la Santé Publique.
- He, Y.X., Jiao, Z., Yang, J., 2018. Comprehensive evaluation of global clean energy development index based on the improved entropy method. *Ecological Indicators* 88, 305–321. <https://doi.org/10.1016/j.ecolind.2017.12.013>
- Hellström, L., Elinder, C.-G., Dahlberg, B., Lundberg, M., Järup, L., Persson, B., Axelson, O., 2001. Cadmium exposure and end-stage renal disease. *American Journal of Kidney Diseases* 38, 1001–1008. <https://doi.org/10.1053/ajkd.2001.28589>
- Henaff, N., Lange, M.-F., Martin, J.-Y., 2009. Revisiter les relations entre pauvreté et éducation. *Revue Française de Socio-Économie* 3, 187. <https://doi.org/10.3917/rfse.003.0187>
- Héritier, H., Vienneau, D., Foraster, M., Eze, I.C., Schaffner, E., Thiesse, L., Ruzdik, F., Habermacher, M., Köpfl, M., Pieren, R., Schmidt-Trucksäss, A., Brink, M., Cajochen, C., Wunderli, J.M., Probst-Hensch, N., Röösli, M., 2018. Diurnal variability of transportation noise exposure and cardiovascular mortality: A nationwide cohort study from Switzerland. *International Journal of Hygiene and Environmental Health* 221, 556–563. <https://doi.org/10.1016/j.ijheh.2018.02.005>
- Host, S., 2013. Exposition à la pollution atmosphérique liée au trafic routier et risques sanitaires. *Vertigo*. <https://doi.org/10.4000/vertigo.12816>
- Houlden, V., Porto de Albuquerque, J., Weich, S., Jarvis, S., 2019. A spatial analysis of proximate greenspace and mental wellbeing in London. *Applied Geography* 109, 102036. <https://doi.org/10.1016/j.apgeog.2019.102036>
- Howard, B.V., Comuzzie, A., Devereux, R.B., Ebbesson, S.O.E., Fabsitz, R.R., Howard, Wm.J., Laston, S., MacCluer, J.W., Silverman, A., Umans, J.G., Wang, H., Weissman, N.J., Wenger, C.R., 2010. Cardiovascular disease prevalence and its relation to risk factors in Alaska Eskimos. *Nutrition, Metabolism and Cardiovascular Diseases* 20, 350–358. <https://doi.org/10.1016/j.numecd.2009.04.010>
- Hsu, A., Johnson, L.A., Lloyd, A., 2013. *Measuring Progress: A Practical Guide from the Developers of the Environmental Performance Index (EPI)*, Yale Center for Environmental Law & Policy. New Haven.
- Hsu, L.-I., Hsieh, F.-I., Wang, Y.-H., Lai, T.-S., Wu, M.-M., Chen, C.-J., Chiou, H.-Y., Hsu, K.-H., 2017. Arsenic Exposure From Drinking Water and the Incidence of CKD in Low to Moderate Exposed Areas of Taiwan: A 14-Year Prospective Study. *American Journal of Kidney Diseases* 70, 787–797. <https://doi.org/10.1053/j.ajkd.2017.06.012>
- Hu, H.-B., Hou, Z.-H., Huang, C.-H., LaMonte, M.J., Wang, M., Lu, B., 2022. Associations of exposure to residential green space and neighborhood walkability with coronary atherosclerosis in Chinese adults. *Environmental Pollution* 292, 118347. <https://doi.org/10.1016/j.envpol.2021.118347>
- Hu, L., Liu, B., Li, Y., 2020. Ranking sociodemographic, health behavior, prevention, and environmental factors in predicting neighborhood cardiovascular health: A Bayesian machine learning approach. *Preventive Medicine* 141, 106240. <https://doi.org/10.1016/j.ypmed.2020.106240>
- Huang, G., London, J.K., 2012. Cumulative Environmental Vulnerability and Environmental Justice in California's San Joaquin Valley. *International Journal of Environmental Research and Public Health* 9, 1593–1608. <https://doi.org/10.3390/ijerph9051593>

- Humpel, N., 2002. Environmental factors associated with adults' participation in physical activity A review. *American Journal of Preventive Medicine* 22, 188–199. [https://doi.org/10.1016/S0749-3797\(01\)00426-3](https://doi.org/10.1016/S0749-3797(01)00426-3)
- Hwang, S.Y., Jeong, S., Choi, S., Kim, D.H., Kim, S.R., Lee, G., Son, J.S., Park, S.M., 2021. Association of Air Pollutants with Incident Chronic Kidney Disease in a Nationally Representative Cohort of Korean Adults. *IJERPH* 18, 3775. <https://doi.org/10.3390/ijerph18073775>
- Hystad, P., Davies, H.W., Frank, L., Van Loon, J., Gehring, U., Tamburic, L., Brauer, M., 2014. Residential Greenness and Birth Outcomes: Evaluating the Influence of Spatially Correlated Built-Environment Factors. *Environmental Health Perspectives* 122, 1095–1102. <https://doi.org/10.1289/ehp.1308049>
- IARC, 2013. La pollution atmosphérique une des premières causes environnementales de décès par cancer, selon le CIRC (Communiqué de presse No. 221). International Agency for Research on Cancer, Lyon, France.
- INERIS, 2014. Les inégalités environnementales sur les territoires - Exposition environnementale et vulnérabilité des populations (INERIS références). Institut National de l'Environnement industriel et des RISques.
- INSEE, 2021a. L'essentiel sur... la pauvreté, L'essentiel sur... Institut National de la Statistique et des Etudes Economiques.
- INSEE, 2021b. Fichier Localisé Social et Fiscal (Filosofi) - Sources et méthodes. Institut National de la Statistique et des Etudes Economiques.
- INSEE, 2015. Sphère présentielle et sphère productive. Institut National de la Statistique et des Etudes Economiques.
- INSERM, 2021. La recherche en santé publique. Institut National de la Santé et de la REcherche Médicale.
- INSERM, 2017. Insuffisance rénale - Décrypter les mécanismes de destruction du rein. Institut National de la Santé et de la REcherche Médicale.
- IRDES, ABM, 2018. Atlas de l'insuffisance rénale chronique terminale en France. Institut de recherche et documentation en économie de la santé (France) et Agence de la Biomédecine (France).
- Jayasumana, M.A.C.S., Paranagama, P.A., Amarasinghe, M.D., Wijewardane, K.M.R.C., Dahanayake, K.S., Fonseka, S.I., Rajakaruna, K.D.L.M.P., Mahamithawa, A.M.P., Samarasinghe, U.D., Senanayake, V.K., 2013. Possible link of Chronic arsenic toxicity with Chronic Kidney Disease of unknown etiology in Sri Lanka. *Journal of Natural Sciences Research* 3, 1–10.
- Jelinski, D.E., Wu, J., 1996. The modifiable areal unit problem and implications for landscape ecology. *Landscape Ecol* 11, 129–140. <https://doi.org/10.1007/BF02447512>
- Jenks, G.F., 1967. The Data Model Concept in Statistical Mapping, Frenzel K. Rand McNally&Co. ed, *International Yearbook of Cartography*. Kirschbaum Verlag, Bonn, Germany.
- Johnson, R.J., Sánchez-Lozada, L.G., Newman, L.S., Lanaspá, M.A., Diaz, H.F., Lemery, J., Rodriguez-Iturbe, B., Tolan, D.R., Butler-Dawson, J., Sato, Y., Garcia, G., Hernando, A.A., Roncal-Jimenez, C.A., 2019. Climate Change and the Kidney. *Ann Nutr Metab* 74, 38–44. <https://doi.org/10.1159/000500344>
- Jomier, B., Lassarade, F., 2021. Les orientations et la gouvernance de la politique de santé environnementale. Mieux connaître les effets de l'environnement sur la santé de la population par le recueil de données et en informer le public (Rapports d'information No. 479 (2020-2021)), Travaux parlementaires. SENAT, Commission des affaires sociales.
- JORF, 2014. Lois et Décrets, Journal Officiel de la République Française.
- Josselin, D., Louvet, R., 2019. Impact of the Scale on Several Metrics Used in Geographical Object-Based Image Analysis: Does GEOBIA Mitigate the Modifiable Areal Unit Problem (MAUP)? *IJGI* 8, 156. <https://doi.org/10.3390/ijgi8030156>
- Journard, R., 2016. Mise au point d'indicateurs de développement durable. Presented at the 20è Colloque international en évaluation environnementale, Antananarivo, Madagascar.
- Jouve, N., 2013. Open Data : synthèse bibliographique. CEREMA, Direction territoriale Nord-Picardie.

- Joyce Fan, Z., Lackland, D.T., Lipsitz, S.R., Nicholas, J.S., Egan, B.M., Tim Garvey, W., Hutchison, F.N., 2007. Geographical patterns of end-stage renal disease incidence and risk factors in rural and urban areas of South Carolina. *Health & Place* 13, 179–187. <https://doi.org/10.1016/j.healthplace.2005.12.002>
- Juhász, A., Nagy, C., Páldy, A., Beale, L., 2010. Development of a Deprivation Index and its relation to premature mortality due to diseases of the circulatory system in Hungary, 1998–2004. *Social Science & Medicine* 70, 1342–1349. <https://doi.org/10.1016/j.socscimed.2010.01.024>
- Jungers, P., Joly, D., Man, N.K., Legendre, C., 2011. *L'insuffisance rénale chronique - Prévention et traitement*, Médecine Sciences Publications. ed. Lavoisier, Paris.
- Kaiser, H.F., 1958. The varimax criterion for analytic rotation in factor analysis. *Psychometrika* 23, 187–200. <https://doi.org/10.1007/BF02289233>
- Kamel, E.G., El-Minshawy, O., 2010. Environmental factors incriminated in the development of end stage renal disease in El-Minia Governorate, Upper Egypt. *Int J Nephrol Urol* 2, 431–437.
- Kastorini, C.-M., Milionis, H.J., Georgousopoulou, E., Kalantzi, K., Nikolaou, V., Vemmos, K.N., Goudevenos, J.A., Panagiotakos, D.B., 2015. Defining the Path Between Social and Economic Factors, Clinical and Lifestyle Determinants, and Cardiovascular Disease. *gh* 10, 255. <https://doi.org/10.1016/j.gh.2015.01.002>
- Kataria, A., Trasande, L., Trachtman, H., 2015. The effects of environmental chemicals on renal function. *Nat Rev Nephrol* 11, 610–625. <https://doi.org/10.1038/nrneph.2015.94>
- Katz, S.F., Rodriguez, F., Knowles, J.W., 2021. Health disparities in cardiometabolic risk among Black and Hispanic youth in the United States. *American Journal of Preventive Cardiology* 6, 100175. <https://doi.org/10.1016/j.ajpc.2021.100175>
- KDIGO, 2013. Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease. Official Journal of the International Society of Nephrology, *Kidney International Supplements* 3, 1–150.
- Kerényi, Á., 2011. The Better Life Index of the Organisation for Economic Co-operation and Development. *Public Finance Quarterly, Journal of economic literature* LVI, 518–538.
- Khomenko, S., Nieuwenhuijsen, M., Ambròs, A., Wegener, S., Mueller, N., 2020. Is a liveable city a healthy city? Health impacts of urban and transport planning in Vienna, Austria. *Environmental Research* 183, 109238. <https://doi.org/10.1016/j.envres.2020.109238>
- Khreis, H., Warsow, K.M., Verlinghieri, E., Guzman, A., Pellecuer, L., Ferreira, A., Jones, I., Heinen, E., Rojas-Rueda, D., Mueller, N., Schepers, P., Lucas, K., Nieuwenhuijsen, M., 2016. The health impacts of traffic-related exposures in urban areas: Understanding real effects, underlying driving forces and co-producing future directions. *Journal of Transport & Health* 3, 249–267. <https://doi.org/10.1016/j.jth.2016.07.002>
- Kihal-Talantikite, W., Padilla, C.M., Lalloue, B., Rougier, C., Defrance, J., Zmirou-Navier, D., Deguen, S., 2013. An exploratory spatial analysis to assess the relationship between deprivation, noise and infant mortality: an ecological study. *Environmental Health* 12. <https://doi.org/10.1186/1476-069X-12-109>
- Kim, M.H., Lee, K.M., Kim, S.J., Kim, S.Y., Park, J.S., Cho, Y.R., Park, K., Park, T.H., Kim, Y.D., Serebruany, V., 2018. Impact of marital status on outcomes following ST-segment elevation myocardial infarction. *Cardiovascular Revascularization Medicine* 19, 237–240. <https://doi.org/10.1016/j.carrev.2017.07.013>
- Kim, N.H., Hyun, Y.Y., Lee, K.-B., Chang, Y., Rhu, S., Oh, K.-H., Ahn, C., 2015. Environmental Heavy Metal Exposure and Chronic Kidney Disease in the General Population. *J Korean Med Sci* 30, 272. <https://doi.org/10.3346/jkms.2015.30.3.272>
- Kim, J.H., 2019. Multicollinearity and misleading statistical results. *Korean J Anesthesiol* 72, 558–569. <https://doi.org/10.4097/kja.19087>
- King, J.R., Jackson, D.A., 1999. Variable selection in large environmental data sets using Principal Components Analysis. *Environmetrics* 10, 67–77.

- Kleinschmidt, I., Hills, M., Elliott, P., 1995. Smoking behaviour can be predicted by neighbourhood deprivation measures. *Journal of Epidemiology & Community Health* 49, S72–S77. https://doi.org/10.1136/jech.49.Suppl_2.S72
- Klompaker, J.O., Hoek, G., Bloemasma, L.D., Gehring, U., Strak, M., Wijga, A.H., van den Brink, C., Brunekreef, B., Lebret, E., Janssen, N.A.H., 2018. Green space definition affects associations of green space with overweight and physical activity. *Environmental Research* 160, 531–540. <https://doi.org/10.1016/j.envres.2017.10.027>
- Klugman, J., Rodríguez, F., Choi, H.J., 2011a. The HDI 2010: new controversies, old critiques. *The Journal of Economic Inequality* 9, 249–288. <https://doi.org/10.1007/s10888-011-9178-z>
- Klugman, J., Rodríguez, F., Choi, H.J., 2011b. Human Development Research Paper 2011/01 The HDI 2010: New Controversies, Old Critiques (Research Paper), Human Development Reports. United Nations Development Programme.
- Knol, A.B., Briggs, D.J., Lebret, E., 2010. Assessment of complex environmental health problems: Framing the structures and structuring the frameworks. *Science of The Total Environment* 408, 2785–2794. <https://doi.org/10.1016/j.scitotenv.2010.03.021>
- Kolenikov, S., Angeles, G., 2004. The Use of Discrete Data in PCA: Theory, Simulations, and Applications to Socioeconomic Indices (Working paper No. WP-04-85). University of North Carolina.
- Kou, L., Tao, Y., Kwan, M.-P., Chai, Y., 2020. Understanding the relationships among individual-based momentary measured noise, perceived noise, and psychological stress: A geographic ecological momentary assessment (GEMA) approach. *Health & Place* 64, 102285. <https://doi.org/10.1016/j.healthplace.2020.102285>
- Kovesdy, C.P., 2022. Epidemiology of chronic kidney disease: an update 2022. *Kidney International Supplements* 12, 7–11. <https://doi.org/10.1016/j.kisu.2021.11.003>
- Kramer, A., Pippias, M., Noordzij, M., Stel, V.S., Afentakis, N., Ambühl, P.M., Andrusev, A.M., Fuster, E.A., Arribas Monzón, F.E., Åsberg, A., Barbullushi, M., Bonthuis, M., Caskey, F.J., Castro de la Nuez, P., Cernevskis, H., des Grottes, J.-M., Garneata, L., Golan, E., Hemmeler, M.H., Ioannou, K., Jarraya, F., Kolesnyk, M., Komissarov, K., Lassalle, M., Macario, F., Mahillo-Duran, B., Martín de Francisco, A.L., Palsson, R., Pechter, Ü., Resic, H., Rutkowski, B., Santiuste de Pablos, C., Seyahi, N., Simic Ogrizovic, S., Slon Roblero, M.F., Spustova, V., Stojceva-Taneva, O., Traynor, J., Massy, Z.A., Jager, K.J., 2018. The European Renal Association – European Dialysis and Transplant Association (ERA-EDTA) Registry Annual Report 2015: a summary. *Clinical Kidney Journal* 11, 108–122. <https://doi.org/10.1093/ckj/sfx149>
- Kruize, H., Droomers, M., van Kamp, I., Ruijsbroek, A., 2014. What Causes Environmental Inequalities and Related Health Effects? An Analysis of Evolving Concepts. *International Journal of Environmental Research and Public Health* 11, 5807–5827. <https://doi.org/10.3390/ijerph110605807>
- Kwan, M.-P., 2018. The Neighborhood Effect Averaging Problem (NEAP): An Elusive Confounder of the Neighborhood Effect. *IJERPH* 15, 1841. <https://doi.org/10.3390/ijerph15091841>
- Kwan, M.-P., 2012a. The Uncertain Geographic Context Problem. *Annals of the Association of American Geographers* 102, 958–968.
- Kwan, M.-P., 2012b. How GIS can help address the uncertain geographic context problem in social science research. *Annals of GIS* 18, 245–255. <https://doi.org/10.1080/19475683.2012.727867>
- Labib, S.M., Lindley, S., Huck, J.J., 2020. Spatial dimensions of the influence of urban green-blue spaces on human health: A systematic review. *Environmental Research* 180, 108869. <https://doi.org/10.1016/j.envres.2019.108869>
- Lahelma, E., Martikainen, P., Laaksonen, M., Aittomäki, A., 2004. Pathways between socioeconomic determinants of health. *Journal of Epidemiology & Community Health* 58, 327–332. <https://doi.org/10.1136/jech.2003.011148>

- Lalloué, B., Monnez, J.-M., Padilla, C., Kihal, W., Le Meur, N., Zmirou-Navier, D., Deguen, S., 2013. A statistical procedure to create a neighborhood socioeconomic index for health inequalities analysis. *Int J Equity Health* 12, 21. <https://doi.org/10.1186/1475-9276-12-21>
- Lan, Y., Roberts, H., Kwan, M.-P., Helbich, M., 2022. Daily space-time activities, multiple environmental exposures, and anxiety symptoms: A cross-sectional mobile phone-based sensing study. *Science of The Total Environment* 155276. <https://doi.org/10.1016/j.scitotenv.2022.155276>
- Landrigan, P.J., Fuller, R., Acosta, N.J.R., Adeyi, O., Arnold, R., Basu, N. (Nil), Baldé, A.B., Bertollini, R., Bose-O'Reilly, S., Boufford, J.I., Breyse, P.N., Chiles, T., Mahidol, C., Coll-Seck, A.M., Cropper, M.L., Fobil, J., Fuster, V., Greenstone, M., Haines, A., Hanrahan, D., Hunter, D., Khare, M., Krupnick, A., Lanphear, B., Lohani, B., Martin, K., Mathiasen, K.V., McTeer, M.A., Murray, C.J.L., Ndahimananjara, J.D., Perera, F., Potočník, J., Preker, A.S., Ramesh, J., Rockström, J., Salinas, C., Samson, L.D., Sandilya, K., Sly, P.D., Smith, K.R., Steiner, A., Stewart, R.B., Suk, W.A., Schayck, O.C.P. van, Yadama, G.N., Yumkella, K., Zhong, M., 2018. The Lancet Commission on pollution and health. *The Lancet* 391, 462–512. [https://doi.org/10.1016/S0140-6736\(17\)32345-0](https://doi.org/10.1016/S0140-6736(17)32345-0)
- Lang, T., 2011. Numéro thématique - Inégalités sociales de santé. Editorial. *Bulletin Epidémiologique Hebdomadaire* 73–74.
- Lang, T., Ulrich, V., France, Ministère de l'emploi et de la solidarité, Direction de la recherche, des études, de l'évaluation et des statistiques, 2017. Les inégalités sociales de santé: actes du séminaire de recherche de la DREES 2015-2016. Editions Dicom.
- Lanier, C., Brousmiche, D., Deram, A., Frank, L., Genin, M., Occelli, F., Cuny, D., 2020. Characterization of the spatial heterogeneity of environmental and social inequalities in health: a spatialized approach to the resilience-vulnerability balance. *Env. Risque Santé* 19, 250–256. <https://doi.org/10.1684/ers.2020.1451>
- Lanier, C., Deram, A., Cuny, M.-A., Cuny, D., Occelli, F., 2019. Spatial analysis of environmental inequalities caused by multiple air pollutants: A cumulative impact screening method, applied to the north of France. *Ecological Indicators* 99, 91–100. <https://doi.org/10.1016/j.ecolind.2018.12.011>
- Larroche, V., Dymytrova, V., 2018. L'apport des SIC aux projets Open Data : pour une meilleure prise en compte des réutilisateurs dans la conception des outils de traitement des data. *Communication organisation* 69–80. <https://doi.org/10.4000/communicationorganisation.6818>
- Lasso de la Vega, M.C., Urrutia, A.M., 2001. HDPI: a framework for pollution-sensitive Human Development Indicators. *Environment, Development and Sustainability* 3, 199–215. <https://doi.org/10.1023/A:1012738731198>
- Last, J.M., International Epidemiological Association (Eds.), 2001. A dictionary of epidemiology, 4th ed. ed. Oxford University Press, New York.
- Laurent, E., 2013. Vers l'égalité des territoires : Dynamique, mesures, politiques. Ministère de l'Égalité des territoires et du Logement.
- Laverty, A.A., Mindell, J.S., Webb, E.A., Millett, C., 2013. Active Travel to Work and Cardiovascular Risk Factors in the United Kingdom. *American Journal of Preventive Medicine* 45, 282–288. <https://doi.org/10.1016/j.amepre.2013.04.012>
- Lebov, J.F., Engel, L.S., Richardson, D., Hogan, S.L., Hoppin, J.A., Sandler, D.P., 2016. Pesticide use and risk of end-stage renal disease among licensed pesticide applicators in the Agricultural Health Study. *Occup Environ Med* 73, 3–12. <https://doi.org/10.1136/oemed-2014-102615>
- Lee, D., Robertson, C., Ramsay, C., Pyper, K., 2020. Quantifying the impact of the modifiable areal unit problem when estimating the health effects of air pollution. *Environmetrics* 31. <https://doi.org/10.1002/env.2643>
- Lee, J., Shin, A., Cho, S., Choi, J.-Y., Kang, D., Lee, J.-K., 2020. Marital status and the prevalence of obesity in a Korean population. *Obesity Research & Clinical Practice* 14, 217–224. <https://doi.org/10.1016/j.orcp.2020.04.003>

- Lee, W., Prifti, K., Kim, H., Kim, E., Yang, J., Min, J., Park, J.Y., Kim, Y.C., Lee, J.P., Bell, M.L., 2022. Short-term Exposure to Air Pollution and Attributable Risk of Kidney Diseases: A Nationwide Time-series Study. *Epidemiology* 33, 17–24. <https://doi.org/10.1097/EDE.0000000000001430>
- Leppo, K. (Ed.), 2013. Health in all policies: seizing opportunities, implementing policies. Ministry of Social Affairs and Health, Finland, s.l.
- Lejeune, Z., Chevau, T., Teller, J., 2012. La qualité du logement comme variable environnementale : l'exemple de la région urbaine de Liège (Wallonie). *Flux* N° 89-90, 30. <https://doi.org/10.3917/flux.089.0030>
- Leroux, L., Riem, B., 2020. Bilan démographique 2019 : la région quitte le trio de tête. *Insee Flash Hauts-de-France*.
- Li, Y.-L., Chuang, T.-W., Chang, P., Lin, L.-Y., Su, C.-T., Chien, L.-N., Chiou, H.-Y., 2021. Long-term exposure to ozone and sulfur dioxide increases the incidence of type 2 diabetes mellitus among aged 30 to 50 adult population. *Environmental Research* 194, 110624. <https://doi.org/10.1016/j.envres.2020.110624>
- Liang, Z., Wang, W., Yang, C., Wang, Y., Shen, J., Li, P., Ma, L., Wei, F., Chen, R., Liang, C., Li, S., Zhang, L., 2022. Residential greenness and prevalence of chronic kidney disease: Findings from the China National Survey of Chronic Kidney Disease. *Science of The Total Environment* 806, 150628. <https://doi.org/10.1016/j.scitotenv.2021.150628>
- Lin, S.-Y., Ju, S.-W., Lin, C.L., Hsu, W.-H., Lin, C.-C., Ting, I.-W., Kao, C.-H., 2020. Air pollutants and subsequent risk of chronic kidney disease and end-stage renal disease: A population-based cohort study. *Environmental Pollution* 261, 114154. <https://doi.org/10.1016/j.envpol.2020.114154>
- Linares, C., Díaz, J., Negev, M., Martínez, G.S., Debono, R., Paz, S., 2020. Impacts of climate change on the public health of the Mediterranean Basin population - Current situation, projections, preparedness and adaptation. *Environmental Research* 182, 109107. <https://doi.org/10.1016/j.envres.2019.109107>
- Lindström, M., Rosvall, M., 2012. Marital status, social capital, economic stress, and mental health: A population-based study. *The Social Science Journal* 49, 339–342. <https://doi.org/10.1016/j.soscij.2012.03.004>
- Liu, X.-X., Ma, X.-L., Huang, W.-Z., Luo, Y.-N., He, C.-J., Zhong, X.-M., Dadvand, P., Browning, M.H.E.M., Li, L., Zou, X.-G., Dong, G.-H., Yang, B.-Y., 2022. Green space and cardiovascular disease: A systematic review with meta-analysis. *Environmental Pollution* 301, 118990. <https://doi.org/10.1016/j.envpol.2022.118990>
- Liu, Y., Yuan, Y., Xiao, Y., Li, Y., Yu, Y., Mo, T., Jiang, H., Li, X., Yang, H., Xu, C., He, M., Guo, H., Pan, A., Wu, T., 2020. Associations of plasma metal concentrations with the decline in kidney function: A longitudinal study of Chinese adults. *Ecotoxicology and Environmental Safety* 189, 110006. <https://doi.org/10.1016/j.ecoenv.2019.110006>
- Loomis, D., Grosse, Y., Lauby-Secretan, B., Ghissassi, F.E., Bouvard, V., Benbrahim-Tallaa, L., Guha, N., Baan, R., Mattock, H., Straif, K., 2013. The carcinogenicity of outdoor air pollution. *The Lancet Oncology* 14, 1262–1263. [https://doi.org/10.1016/S1470-2045\(13\)70487-X](https://doi.org/10.1016/S1470-2045(13)70487-X)
- Lue, S.-H., Wellenius, G.A., Wilker, E.H., Mostofsky, E., Mittleman, M.A., 2013. Residential proximity to major roadways and renal function. *J Epidemiol Community Health* 67, 629–634. <https://doi.org/10.1136/jech-2012-202307>
- Luo, K., Li, R., Wang, Z., Zhang, R., Xu, Q., 2017. Effect modification of the association between temperature variability and daily cardiovascular mortality by air pollutants in three Chinese cities. *Environmental Pollution* 230, 989–999. <https://doi.org/10.1016/j.envpol.2017.07.045>
- Maantay, J., 2007. Asthma and air pollution in the Bronx: Methodological and data considerations in using GIS for environmental justice and health research. *Health & Place* 13, 32–56. <https://doi.org/10.1016/j.healthplace.2005.09.009>
- Macintyre, S., Ellaway, A., 2000. Ecological approaches: Rediscovering the role of the physical and social environment, in: Berkman, L. and Kawachi, I. (Eds.) *Social Epidemiology*. Oxford University Press, Oxford, pp. 332–348.

- Mackenbach, J.P., Stirbu, I., Roskam, A.-J.R., Schaap, M.M., Menvielle, G., Leinsalu, M., Kunst, A.E., 2008. Socioeconomic Inequalities in Health in 22 European Countries. *N Engl J Med* 358, 2468–2481. <https://doi.org/10.1056/NEJMsa0707519>
- Maclachlan, J.C., Jerrett, M., Abernathy, T., Sears, M., Bunch, M.J., 2007. Mapping health on the Internet: A new tool for environmental justice and public health research. *Health & Place* 13, 72–86. <https://doi.org/10.1016/j.healthplace.2005.09.012>
- Maday, C., 2016. Qualité des données et Open data. *I2D - Information, données & documents* 53, 58. <https://doi.org/10.3917/i2d.164.0058>
- Madika, A.-L., Mounier-Vehier, C., 2017. Tabac et pression artérielle : une relation complexe à mieux connaître. *La Presse Médicale* 46, 697–702. <https://doi.org/10.1016/j.lpm.2017.07.001>
- Maggino, F., 2017. Developing Indicators and Managing the Complexity, in: Maggino, F. (Ed.), *Complexity in Society: From Indicators Construction to Their Synthesis*. Springer International Publishing, Cham, pp. 87–114. https://doi.org/10.1007/978-3-319-60595-1_4
- Mallamaci, F., Pisano, A., Tripepi, G., 2020. Physical activity in chronic kidney disease and the EXerCise Introduction To Enhance trial. *Nephrology Dialysis Transplantation* 35, ii18–ii22. <https://doi.org/10.1093/ndt/gfaa012>
- Manisalidis, I., Stavropoulou, E., Stavropoulos, A., Bezirtzoglou, E., 2020. Environmental and Health Impacts of Air Pollution: A Review. *Front. Public Health* 8, 14. <https://doi.org/10.3389/fpubh.2020.00014>
- Maresca, B., Helmi, S., 2014. Les inégalités territoriales de santé : une approche des bassins de santé (Cahier de recherche No. 320). Centre de recherche pour l'étude et l'observation des conditions de vie (CREDOC).
- Markozannes, G., Pantavou, K., Rizos, E.C., Sindosi, O.A., Tagkas, C., Seyfried, M., Saldanha, I.J., Hatzianastassiou, N., Nikolopoulos, G.K., Ntzani, E., 2022. Outdoor air quality and human health: An overview of reviews of observational studies. *Environmental Pollution* 306, 119309. <https://doi.org/10.1016/j.envpol.2022.119309>
- Marlier, G., 2009. Note D2DPE n°24 - Un indicateur de développement humain à l'échelle des territoires du Nord-Pas de Calais. *Observation Régionale et Analyse Spatiale, Direction du Développement Durable, de la Prospective et de l'Évaluation, Région Nord Pas de Calais*.
- Marlier, G., Dallery, T., Chusseau, N., 2015. Le Nord-Pas-de-Calais, entre mutations économiques et développement humain. *Revue de l'OFCE* 143, 225. <https://doi.org/10.3917/reof.143.0225>
- Marmot, M., Friel, S., Bell, R., Houweling, T.A., Taylor, S., 2008. Closing the gap in a generation: health equity through action on the social determinants of health. *The Lancet* 372, 1661–1669. [https://doi.org/10.1016/S0140-6736\(08\)61690-6](https://doi.org/10.1016/S0140-6736(08)61690-6)
- Marot, G., 2019. Le bassin minier : un territoire densément peuplé confronté à diverses fragilités sociales (No. 74). *INSEE FLASH HAUTS-DE-FRANCE*.
- Marshall, I.J., Wang, Y., Crichton, S., McKeivitt, C., Rudd, A.G., Wolfe, C.D.A., 2015. The effects of socioeconomic status on stroke risk and outcomes. *The Lancet Neurology* 14, 1206–1218. [https://doi.org/10.1016/S1474-4422\(15\)00200-8](https://doi.org/10.1016/S1474-4422(15)00200-8)
- Marzi, S., Mysiak, J., Santato, S., 2018. Comparing adaptive capacity index across scales: The case of Italy. *Journal of Environmental Management* 223, 1023–1036. <https://doi.org/10.1016/j.jenvman.2018.06.060>
- Mauro, V., Giusti, C., Marchetti, S., Pratesi, M., 2021. Does uncertainty in single indicators affect the reliability of composite indexes? An application to the measurement of environmental performances of Italian regions. *Ecological Indicators* 127, 107740. <https://doi.org/10.1016/j.ecolind.2021.107740>
- McCartney, G., Popham, F., McMaster, R., Cumbers, A., 2019. Defining health and health inequalities. *Public Health* 172, 22–30. <https://doi.org/10.1016/j.puhe.2019.03.023>
- McGeehin, M.A., Mirabelli, M., 2001. The potential impacts of climate variability and change on temperature-related morbidity and mortality in the United States. *Environmental Health Perspectives* 109, 185–189.

- Meijer, M., Röhl, J., Bloomfield, K., Grittner, U., 2012. Do neighborhoods affect individual mortality? A systematic review and meta-analysis of multilevel studies. *Social Science & Medicine* 74, 1204–1212. <https://doi.org/10.1016/j.socscimed.2011.11.034>
- Menvielle, G., Lang, T., 2021. Les inégalités sociales de santé : vingt ans d'évolution: Actualité et dossier en santé publique N° 113, 2–2. <https://doi.org/10.3917/aedesp.113.0002>
- Mercier, L., 2005. La pauvreté : phénomène complexe et multidimensionnel. ss 44, 7–27. <https://doi.org/10.7202/706704ar>
- Mericksay, B., 2021. Les effets de l'open data et du Gig Data dans la fabrique des cartes de l'actuib publiques : entre repositionnement des acteurs et démonopolisation de l'expertise autour des données territoriales, *Les cartes de l'action publique : Pouvoirs, territoires, résistances*.
- Mericksay, B., 2018. Open data et crowdsourcing : Quelle place pour le géomaticien ?, in: *Géoséminaire*. Presented at the *La géomatique dépassée par ses usages : quelles bonnes pratiques pour demain ?*, Université de Rennes 2.
- Miao, J., Wu, X., 2016. Urbanization, socioeconomic status and health disparity in China. *Health & Place* 42, 87–95. <https://doi.org/10.1016/j.healthplace.2016.09.008>
- Miller, K.K., Johnson, A., Dabson, B., 2017. *Research Methodology of Resilience and Vulnerability Indexes*. Institute of Public Policy Harry S Truman School of Public Affairs, University of Missouri.
- Ministère des Solidarités et de la Santé, 2022. *Santé Environnement : 3è plan national 2015 > 2019*.
- Modica, M., Reggiani, A., 2015. Spatial Economic Resilience: Overview and Perspectives. *Networks and Spatial Economics* 15, 211–233. <https://doi.org/10.1007/s11067-014-9261-7>
- Moine, A., 2006. Le territoire comme un système complexe : un concept opératoire pour l'aménagement et la géographie. *Espace géographique* 35, 115. <https://doi.org/10.3917/eg.352.0115>
- Moleux, M., Schaetzel, F., Scotton, C., 2011. *Les inégalités sociales de santé : déterminants sociaux et modèles d'action (No. RM2011- 061P)*. Inspection Générale des Affaires Sociales.
- Monteiro, A., Carvalho, V., Velho, S., Sousa, C., 2012. Assessing and monitoring urban resilience using COPD in Porto. *Science of The Total Environment* 414, 113–119. <https://doi.org/10.1016/j.scitotenv.2011.11.009>
- Moran, P.A.P., 1948. The Interpretation of Statistical Maps. *Journal of the Royal Statistical Society: Series B (Methodological)* 10, 243–251. <https://doi.org/10.1111/j.2517-6161.1948.tb00012.x>
- MTEs, 2019. *L'environnement en France : 2019 - rapport de synthèse*, Ministère de la Transition Ecologique et Solidaire. ed, Services de la Donnée et des Etudes Statistiques (SDES). La Documentation française.
- Mukherjee, D., Safraj, S., Tayyab, M., Shivashankar, R., Patel, S.A., Narayanan, G., Ajay, V.S., Ali, M.K., Narayan, K.V., Tandon, N., Prabhakaran, D., 2017. Park availability and major depression in individuals with chronic conditions: Is there an association in urban India? *Health & Place* 47, 54–62. <https://doi.org/10.1016/j.healthplace.2017.07.004>
- Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., 2005. *Tools for Composite Indicators Building (No. EUR 61682)*, Joint Research Centre European Commission. Institute for the Protection and Security of the Citizen Econometrics and Statistical Support to Antifraud Unit, Ispra Italy.
- Nascimento, I., Carrage, S., 2007. *Indicateurs Stratégiques de Développement Durable - Un Indice de Qualité de Vie et Bien-Être pour le Région Île de France*.
- Nash, D., McClure, G., Mastracci, T.M., Anand, S.S., 2021. Social Deprivation and Peripheral Artery Disease. *Canadian Journal of Cardiology* S0828282X2100965X. <https://doi.org/10.1016/j.cjca.2021.12.011>
- Neves Almeida, T.A. das, García-Sánchez, I.-M., 2016. A comparative analysis between composite indexes of environmental performance: An analysis on the CIEP and EPI. *Environmental Science & Policy* 64, 59–74. <https://doi.org/10.1016/j.envsci.2016.06.011>

- Nicholas, S.B., Kalantar-Zadeh, K., Norris, K.C., 2015. Socioeconomic Disparities in Chronic Kidney Disease. *Advances in Chronic Kidney Disease* 22, 6–15. <https://doi.org/10.1053/j.ackd.2014.07.002>
- Nicolas, J.-P., Vanco, F., Verry, D., 2012. Mobilité quotidienne et vulnérabilité des ménages: Revue d'Économie Régionale & Urbaine février, 19–44. <https://doi.org/10.3917/reru.121.0019>
- Nieuwenhuijsen, M.J., 2020. Urban and transport planning pathways to carbon neutral, liveable and healthy cities; A review of the current evidence. *Environment International* 140, 105661. <https://doi.org/10.1016/j.envint.2020.105661>
- Nieuwenhuijsen, M.J., 2016. Urban and transport planning, environmental exposures and health—new concepts, methods and tools to improve health in cities. *Environ Health* 15, S38. <https://doi.org/10.1186/s12940-016-0108-1>
- Nieuwenhuijsen, M.J., Khreis, H., Verlinghieri, E., Mueller, N., Rojas-Rueda, D., 2017. Participatory quantitative health impact assessment of urban and transport planning in cities: A review and research needs. *Environment International* 103, 61–72. <https://doi.org/10.1016/j.envint.2017.03.022>
- Nieuwenhuijsen, M., Khreis, H., 2019. Urban and Transport Planning, *Environment and Health*, in: Nieuwenhuijsen, M., Khreis, H. (Eds.), *Integrating Human Health into Urban and Transport Planning: A Framework*. Springer International Publishing, Cham, pp. 3–16. https://doi.org/10.1007/978-3-319-74983-9_1
- Nortier, J., Plennevaux, V., Nolf, A., Hougardy, J.M., Petit, A.D., 2017. Le trajet de soins de l'insuffisance rénale chronique : développement et perspectives. *Rev. Med. Brux.* 38, 271–278.
- Nourry, M., 2008. Measuring sustainable development: Some empirical evidence for France from eight alternative indicators. *Ecological Economics* 67, 441–456. <https://doi.org/10.1016/j.ecolecon.2007.12.019>
- Occelli, F., Bavdek, R., Deram, A., Hellequin, A.-P., Cuny, M.-A., Zwarterook, I., Cuny, D., 2016. Using lichen biomonitors to assess environmental justice at a neighbourhood level in an industrial area of Northern France. *Ecological Indicators* 60, 781–788. <https://doi.org/10.1016/j.ecolind.2015.08.026>
- Occelli, F., Cuny, M.-A., Devred, I., Deram, A., Quarré, S., Cuny, D., 2014a. Étude de l'imprégnation de l'environnement de trois bassins de vie de la région Nord-Pas-de-Calais par les éléments traces métalliques. *Pollution atmosphérique*. <https://doi.org/10.4267/pollution-atmospherique.2497>
- Occelli, F., Deram, A., Génin, M., Noël, C., Cuny, D., Glowacki, F., on behalf of the Néphronor Network, 2014b. Mapping End-Stage Renal Disease (ESRD): Spatial Variations on Small Area Level in Northern France, and Association with Deprivation. *PLoS ONE* 9, e110132. <https://doi.org/10.1371/journal.pone.0110132>
- Occelli, F., Lanier, C., Cuny, D., Deram, A., Dumont, J., Amouyel, P., Montaye, M., Dauchet, L., Dallongeville, J., Genin, M., 2020. Exposure to multiple air pollutants and the incidence of coronary heart disease: A fine-scale geographic analysis. *Science of The Total Environment* 714, 136608. <https://doi.org/10.1016/j.scitotenv.2020.136608>
- OR2S, 2018. Indicateurs socio-sanitaires en région (données locales), Mortalité tous âges [WWW Document]. Plateforme sanitaire et sociale Hauts-de-France, open data. URL <https://calcul2.or2s.fr/openData/#>
- OR2S, 2019. Diagnostic territorialisé des Hauts-de-France, territoires de proximité.
- Orfeuill, J.P., 2006. Déplacements et inégalités, Actes de la journée d'étude "Faire société en France et en Europe au début du XXI^e siècle." Centre d'analyse stratégique, Palais du Luxembourg.
- Padilla, C., 2013. Inégalités sociales de santé et expositions environnementales. Une analyse spatio-temporelle du risque de mortalité infantile et néonatale dans 4 agglomérations françaises. Université de Lorraine, Nancy.
- Pampalon, R., Raymond, G., 2000. A Deprivation Index for Health and Welfare Planning in Quebec. *Chronic Diseases in Canada* 21, 104–113.

- Panagos, P., Van Liedekerke, M., Jones, A., Montanarella, L., 2012. European Soil Data Centre: Response to European policy support and public data requirements. *Land Use Policy* 29, 329–338. <https://doi.org/10.1016/j.landusepol.2011.07.003>
- Parini, P., Frikke-Schmidt, R., Tselepis, A.D., Moulin, P., von Eckardstein, A., Binder, C.J., Catapano, A.L., Ray, K.K., Tokgözoğlu, L., 2021. Taking action: European Atherosclerosis Society targets the United Nations Sustainable Development Goals 2030 agenda to fight atherosclerotic cardiovascular disease in Europe. *Atherosclerosis* 322, 77–81. <https://doi.org/10.1016/j.atherosclerosis.2021.02.007>
- Park, Y.M., Kwan, M.-P., 2017. Individual exposure estimates may be erroneous when spatiotemporal variability of air pollution and human mobility are ignored. *Health & Place* 43, 85–94. <https://doi.org/10.1016/j.healthplace.2016.10.002>
- Parodi, A.L., 2021. Le concept « One Health », une seule santé : réalité et perspectives. *Bulletin de l'Académie Nationale de Médecine* 205, 659–661. <https://doi.org/10.1016/j.banm.2021.05.001>
- Patino, J.E., Hong, A., Duque, J.C., Rahimi, K., Zapata, S., Lopera, V.M., 2021. Built environment and mortality risk from cardiovascular disease and diabetes in Medellín, Colombia: An ecological study. *Landscape and Urban Planning* 213, 104126. <https://doi.org/10.1016/j.landurbplan.2021.104126>
- Patterson, R., Panter, J., Vamos, E.P., Cummins, S., Millett, C., Laverty, A.A., 2020. Associations between commute mode and cardiovascular disease, cancer, and all-cause mortality, and cancer incidence, using linked Census data over 25 years in England and Wales: a cohort study. *The Lancet Planetary Health* 4, e186–e194. [https://doi.org/10.1016/S2542-5196\(20\)30079-6](https://doi.org/10.1016/S2542-5196(20)30079-6)
- Pearce, J.R., Richardson, E.A., Mitchell, R.J., Shortt, N.K., 2010. Environmental justice and health: the implications of the socio-spatial distribution of multiple environmental deprivation for health inequalities in the United Kingdom: Environmental justice and health. *Transactions of the Institute of British Geographers* 35, 522–539. <https://doi.org/10.1111/j.1475-5661.2010.00399.x>
- Pearson, A.L., Pearce, J., Kingham, S., 2013. Deprived yet healthy: Neighbourhood-level resilience in New Zealand. *Social Science & Medicine* 91, 238–245. <https://doi.org/10.1016/j.socscimed.2012.09.046>
- Pedersen, A., Mikkelsen, E., Cronin-Fenton, D., Kristensen, N., Pham, T.M., Pedersen, L., Petersen, I., 2017. Missing data and multiple imputation in clinical epidemiological research. *Clinical Epidemiology Volume 9*, 157–166. <https://doi.org/10.2147/CLEP.S129785>
- Petroni, M., Howard, S., Howell, I.B., Collins, M.B., 2021. NYenviroScreen: An open-source data driven method for identifying potential environmental justice communities in New York State. *Environmental Science & Policy* 124, 348–358. <https://doi.org/10.1016/j.envsci.2021.07.004>
- Petrovic, D., de Mestral, C., Bochud, M., Bartley, M., Kivimäki, M., Vineis, P., Mackenbach, J., Stringhini, S., 2018. The contribution of health behaviors to socioeconomic inequalities in health: A systematic review. *Preventive Medicine* 113, 15–31. <https://doi.org/10.1016/j.ypmed.2018.05.003>
- Pierret, R., 2013. Qu'est-ce que la précarité ? *socio* 307–330. <https://doi.org/10.4000/socio.511>
- Pillon, F., Tan, K., Jouty, P., Frullani, Y., 2014. Diabète de type 2 : l'essentiel sur la pathologie. *Actualités Pharmaceutiques* 53, 18–22. <https://doi.org/10.1016/j.actpha.2014.10.004>
- PNSE4, 2021. Un environnement, Un santé - 4è Plan National Santé Environnement (Synthèse 4è PNSE). Ministère de la transition écologique.
- Porcherie, M., Vaillant, Z., Faure, E., Rican, S., Simos, J., Cantoreggi, N.L., Heritage, Z., Le Gall, A.R., Cambon, L., Diallo, T.A., Vidales, E., Pommier, J., 2017. The GREENH-City interventional research protocol on health in all policies. *BMC Public Health* 17, 820. <https://doi.org/10.1186/s12889-017-4812-8>
- Pornet, C., Delpierre, C., Dejardin, O., Grosclaude, P., Launay, L., Guittet, L., Lang, T., Launoy, G., 2012. Construction of an adaptable European transnational ecological deprivation index: the

- French version. *Journal of Epidemiology and Community Health* 66, 982–989. <https://doi.org/10.1136/jech-2011-200311>
- Porter, M.E., Stern, S., Artavia Loría, R., 2013. *Social Progress Index 2013. The Social Progress Imperative*, Washington D.C.
- Potthoff, R.F., Whittinghill, M., 1966. Testing for Homogeneity: I. The Binomial and Multinomial Distributions. *Biometrika* 53, 167. <https://doi.org/10.2307/2334062>
- Potvin, L., Moquet, M.-J., Jones, C.M., 2010. Réduire les inégalités sociales en santé, Dossiers santé en action. INPES, Saint-Denis.
- Pretty, J., Griffin, M., Peacock, J., Hine, R., Sellens, M., South, N., 2005. A countryside for health and wellbeing: the physical and mental health benefits of green exercise - executive summary. *Countryside Recreation* 13, 42.
- Prüss-Üstün, A., Wolf, J., Corvalán, C., Bos, R., Neira, M., 2016. Preventing disease through healthy environments: a global assessment of the burden of disease from environmental risks, Second edition. ed. World Health Organization, Geneva, Switzerland.
- Raimondeau, J., Carmona, É., 2020. Chapitre 3. Méthodes épidémiologiques; in: *Manuel de Santé Publique*. Presses de l'EHESP, pp. 73–99. <https://doi.org/10.3917/ehesp.raimo.2020.01.0073>
- Rajagopalan, S., Landrigan, P.J., 2021. Pollution and the Heart. *N Engl J Med* 385, 1881–1892. <https://doi.org/10.1056/NEJMra2030281>
- Ramirez-Rubio, O., Daher, C., Fanjul, G., Gascon, M., Mueller, N., Pajín, L., Plasencia, A., Rojas-Rueda, D., Thondoo, M., Nieuwenhuijsen, M.J., 2019. Urban health: an example of a “health in all policies” approach in the context of SDGs implementation. *Global Health* 15, 87. <https://doi.org/10.1186/s12992-019-0529-z>
- Raulet-Croset, N., 2014. Le territoire comme objet de gestion: *Annales des Mines - Responsabilité et environnement* N° 74, 48–52. <https://doi.org/10.3917/re.074.0048>
- Rebouillat, V., 2020. Entrepôts de données de recherche : mesurer l'impact de l'Open Science à l'aune de la consultation des jeux de données déposés. Presented at the 7^e conférence Document numériques & Société - Humains et données : création, médiation, décision, narration, Nancy, p. 19.
- Regmi, K., Mudyarabikwa, O., 2020. A systematic review of the factors - barriers and enablers - affecting the implementation of clinical commissioning policy to reduce health inequalities in the National Health Service (NHS), UK. *Public Health* 186, 271–282. <https://doi.org/10.1016/j.puhe.2020.07.027>
- Remington, P.L., Catlin, B.B., Gennuso, K.P., 2015. The County Health Rankings: rationale and methods. *Popul Health Metrics* 13, 11. <https://doi.org/10.1186/s12963-015-0044-2>
- Rey, G., Jouglu, E., Fouillet, A., Hémon, D., 2009. 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* 9. <https://doi.org/10.1186/1471-2458-9-33>
- Rey, G., Rican, S., Jouglu, E., 2011. Mesure des inégalités de mortalité par cause de décès. Approche écologique à l'aide d'un indice de désavantage social. *Bulletin Epidémiologique Hebdomadaire* 8–9, 87–90.
- Reynard, R., Vialette, P., Gass, C., 2015. Trente ans de mutations fonctionnelles de l'emploi dans les territoires (No. 1538), INSEE Première. Institut National de la Statistique et des Etudes Economiques.
- Reynaud, M., Foucaud, D., Natali, J.P., Bonnevide, P., 2012. Rapport annuel 2012. Haut Conseil pour l'avenir de l'assurance maladie (HCAAM).
- Rican, S., Rey, G., Lucas-Gabrielli, V., Bard, D., Zeitlin, J., Charreire, H., Jouglu, E., Salem, G., Vaillant, Z., Combier, E., Oppert, J.M., Hercberg, S., Castetbon, K., Méjéan, C., Pampalon, R., 2011. Désavantages locaux et santé : construction d'indices pour l'analyse des inégalités sociales et territoriales de santé en France et leurs évolutions. *Env. Risque Sante* 10, 211–215. <https://doi.org/10.1684/ers.2011.0450>

- Rican, S., Vaillant, Z., Bochaton, A., Salem, G., 2014. Inégalités géographiques de santé en France. Les Tribunes de la santé n°43, 39. <https://doi.org/10.3917/seve.043.0039>
- Richardson, E.A., Mitchell, R., Shortt, N.K., Pearce, J., Dawson, T.P., 2010. Developing Summary Measures of Health-Related Multiple Physical Environmental Deprivation for Epidemiological Research. *Environ Plan A* 42, 1650–1668. <https://doi.org/10.1068/a42459>
- Rojas-Rueda, D., de Nazelle, A., Teixido, O., Nieuwenhuijsen, M.J., 2013. Health impact assessment of increasing public transport and cycling use in Barcelona: A morbidity and burden of disease approach. *Preventive Medicine* 57, 573–579. <https://doi.org/10.1016/j.ypmed.2013.07.021>
- Rojas-Rueda, D., Nieuwenhuijsen, M.J., Gascon, M., Perez-Leon, D., Mudu, P., 2019. Green spaces and mortality: a systematic review and meta-analysis of cohort studies. *The Lancet Planetary Health* 3, e469–e477. [https://doi.org/10.1016/S2542-5196\(19\)30215-3](https://doi.org/10.1016/S2542-5196(19)30215-3)
- Rollet, P., Thomas, P., 2021. Combattre les inégalités et la pauvreté en Hauts-de-France - premier rapport 2021, Eclairer l'avenir. Conseil Economique, Social et Environnemental Régional (CESER), Hauts-de-France.
- Rothenberg, R., Stauber, C., Weaver, S., Dai, D., Prasad, A., Kano, M., 2015. Urban health indicators and indices—current status. *BMC Public Health* 15, 494. <https://doi.org/10.1186/s12889-015-1827-x>
- Rousse, B., 2017. Les enjeux éthiques de “l'éducation” thérapeutique du patient : l'exemple de la personne en dialyse. Université de Paris-Saclay, Université de Paris XI.
- Rydin, Y., Bleahu, A., Davies, M., Dávila, J.D., Friel, S., De Grandis, G., Groce, N., Hallal, P.C., Hamilton, I., Howden-Chapman, P., Lai, K.-M., Lim, C., Martins, J., Osrin, D., Ridley, I., Scott, I., Taylor, M., Wilkinson, P., Wilson, J., 2012. Shaping cities for health: complexity and the planning of urban environments in the 21st century. *The Lancet* 379, 2079–2108. [https://doi.org/10.1016/S0140-6736\(12\)60435-8](https://doi.org/10.1016/S0140-6736(12)60435-8)
- Sadd, J.L., Pastor, M., Morello-Frosch, R., Scoggins, J., Jesdale, B., 2011. Playing It Safe: Assessing Cumulative Impact and Social Vulnerability through an Environmental Justice Screening Method in the South Coast Air Basin, California. *International Journal of Environmental Research and Public Health* 8, 1441–1459. <https://doi.org/10.3390/ijerph8051441>
- Saib, M.S., 2015. Construction d'une méthode de caractérisation des inégalités territoriales, environnementales et sociales de santé - Développement d'un outil d'intégration des phénomènes spatiaux pour la construction d'indicateurs composites à l'échelle régionale (Spécialité Biologie Santé). Université de Picardie Jules Verne, Amiens.
- Saib, M.-S., Caudeville, J., Beauchamp, M., Carré, F., Ganry, O., Trugeon, A., Cicoella, A., 2015. Building spatial composite indicators to analyze environmental health inequalities on a regional scale. *Environmental Health* 14, 68. <https://doi.org/10.1186/s12940-015-0054-3>
- Saisana, M., Cartwright, F., 2007. Measuring Lifelong Learning and its impact on Happiness - The Canadian paradigm. Presented at the International Conference on Policies for Happiness, Canadian Council on Learning, Monitoring and Reporting - Ottawa, Siena, p. 21.
- Saghapour, T., Giles-Corti, B., Jafari, A., Qaisrani, M.A., Turrell, G., 2021. Supporting pandemic disease preparedness: Development of a composite index of area vulnerability. *Health & Place* 70, 102629. <https://doi.org/10.1016/j.healthplace.2021.102629>
- Salem, G., 1995. Géographie de la santé, santé de la géographie. *Espace, populations, sociétés* 13, 25–30. <https://doi.org/10.3406/espos.1995.1667>
- Salvati, L., Carlucci, M., 2014. A composite index of sustainable development at the local scale: Italy as a case study. *Ecological Indicators* 43, 162–171. <https://doi.org/10.1016/j.ecolind.2014.02.021>
- Santé Publique France, 2021. Les inégalités sociales et territoriales de santé (Dossier thématique). Santé Publique France, Saint-Maurice.
- Sarkar, C., Webster, C., 2017. Urban environments and human health: current trends and future directions. *Current Opinion in Environmental Sustainability* 25, 33–44. <https://doi.org/10.1016/j.cosust.2017.06.001>

- Scarwell, H.-J., Roussel, I., Charles, L., 2013. Environnement et santé : quels enjeux, quels acteurs, quelle intelligibilité ? Développement durable & territoires, Santé et environnement 4. <https://doi.org/DOI : 10.4000/developpementdurable.9848>
- Schold, J.D., Flechner, S.M., Poggio, E.D., Augustine, J.J., Goldfarb, D.A., Sedor, J.R., Buccini, L.D., 2018. Residential Area Life Expectancy: Association With Outcomes and Processes of Care for Patients With ESRD in the United States. *American Journal of Kidney Diseases* 72, 19–29. <https://doi.org/10.1053/j.ajkd.2017.12.014>
- Schultz, W.M., Kelli, H.M., Lisko, J.C., Varghese, T., Shen, J., Sandesara, P., Quyyumi, A.A., Taylor, H.A., Gulati, M., Harold, J.G., Mieres, J.H., Ferdinand, K.C., Mensah, G.A., Sperling, L.S., 2018. Socioeconomic Status and Cardiovascular Outcomes: Challenges and Interventions. *Circulation* 137, 2166–2178. <https://doi.org/10.1161/CIRCULATIONAHA.117.029652>
- Schuurman, N., Bell, N., Dunn, J.R., Oliver, L., 2007. Deprivation Indices, Population Health and Geography: An Evaluation of the Spatial Effectiveness of Indices at Multiple Scales. *J Urban Health* 84, 591–603. <https://doi.org/10.1007/s11524-007-9193-3>
- Sen, A., 1999. Development as Freedom, in: *The Globalization and Development Reader - Perspectives on Development and Global Change*. Second Edition. Edited by J. Timmons Roberts, Amy Bellone Hite, and Nitsan Chorev. John Wiley & Sons, Ltd, pp. 525–548.
- Shafran-Nathan, R., Yuval, Levy, I., Broday, D.M., 2017. Exposure estimation errors to nitrogen oxides on a population scale due to daytime activity away from home. *Science of The Total Environment* 580, 1401–1409. <https://doi.org/10.1016/j.scitotenv.2016.12.105>
- Shannon, C.E., 1948. A Mathematical Theory of Communication. *Bell System Technical Journal* 27, 379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Sharifi, A., 2016. A critical review of selected tools for assessing community resilience. *Ecological Indicators* 69, 629–647. <https://doi.org/10.1016/j.ecolind.2016.05.023>
- Sherrieb, K., Norris, F.H., Galea, S., 2010. Measuring Capacities for Community Resilience. *Social Indicators Research* 99, 227–247. <https://doi.org/10.1007/s11205-010-9576-9>
- Shrestha, R., Flacke, J., Martinez, J., van Maarseveen, M., 2016. Environmental Health Related Socio-Spatial Inequalities: Identifying “Hotspots” of Environmental Burdens and Social Vulnerability. *International Journal of Environmental Research and Public Health* 13, 691. <https://doi.org/10.3390/ijerph13070691>
- Silventoinen, K., Korhonen, K., Martikainen, P., 2022. Changing associations of coronary heart disease incidence with current partnership status and marital history over three decades. *SSM - Population Health* 18, 101080. <https://doi.org/10.1016/j.ssmph.2022.101080>
- Simpson, D., Rue, H., Riebler, A., Martins, T.G., Sørbye, S.H., 2017. Penalising Model Component Complexity: A Principled, Practical Approach to Constructing Priors. *Statist. Sci.* 32. <https://doi.org/10.1214/16-STS576>
- Sirven, N., Debrand, T., 2012. Social capital and health of older Europeans: Causal pathways and health inequalities. *Social Science & Medicine* 75, 1288–1295. <https://doi.org/10.1016/j.socscimed.2012.05.009>
- Song, X., Wang, S., Hu, Y., Yue, M., Zhang, T., Liu, Y., Tian, J., Shang, K., 2017. Impact of ambient temperature on morbidity and mortality: An overview of reviews. *Science of The Total Environment* 586, 241–254. <https://doi.org/10.1016/j.scitotenv.2017.01.212>
- Sorensen, G., 2000. INTRODUCTION TO SOCIAL DETERMINANTS OF HEALTH, in: *Women and Health*. Elsevier, pp. 523–527. <https://doi.org/10.1016/B978-012288145-9/50048-6>
- Steptoe, A., 2015. Coronary Heart Disease: Psychosocial Aspects, in: *International Encyclopedia of the Social & Behavioral Sciences*. Elsevier, pp. 917–922. <https://doi.org/10.1016/B978-0-08-097086-8.14023-1>
- Stewart, I.T., Bacon, C.M., Burke, W.D., 2014. The uneven distribution of environmental burdens and benefits in Silicon Valley’s backyard. *Applied Geography* 55, 266–277. <https://doi.org/10.1016/j.apgeog.2014.09.016>
- Stiglitz, J.E., Sen, A., Fitoussi, J.-P., 2009. Rapport de la Commission sur la mesure des performances économiques et du progrès social.

- Strezov, V., Evans, A., Evans, T.J., 2017. Assessment of the Economic, Social and Environmental Dimensions of the Indicators for Sustainable Development: Economic, social and environmental dimensions of sustainability. *Sustainable Development* 25, 242–253. <https://doi.org/10.1002/sd.1649>
- Su, F., Zeeshan, M., Xiong, L.-H., Lv, J.-Y., Wu, Y., Tang, X.-J., Zhou, Y., Ou, Y.-Q., Huang, W.-Z., Feng, W.-R., Zeng, X.-W., Dong, G.-H., 2022. Co-exposure to perfluoroalkyl acids and heavy metals mixtures associated with impaired kidney function in adults: A community-based population study in China. *Science of The Total Environment* 839, 156299. <https://doi.org/10.1016/j.scitotenv.2022.156299>
- Su, J.G., Morello-Frosch, R., Jesdale, B.M., Kyle, A.D., Shamasunder, B., Jerrett, M., 2009. An Index for Assessing Demographic Inequalities in Cumulative Environmental Hazards with Application to Los Angeles, California. *Environmental Science & Technology* 43, 7626–7634. <https://doi.org/10.1021/es901041p>
- Tagne Fotso, R., 2016. Imprégnation aux métaux et métalloïdes en population générale du Nord-Pas-de-Calais : niveaux, déterminants et liens avec le débit de filtration glomérulaire, à partir de l'enquête IMePoGe (Santé publique, recherche clinique, technologies biomédicales). Université de Lille.
- Tamura, K., Langerman, S.D., Ceasar, J.N., Andrews, M.R., Agrawal, M., Powell-Wiley, T.M., 2019. Neighborhood Social Environment and Cardiovascular Disease Risk. *Curr Cardiovasc Risk Rep* 13, 7. <https://doi.org/10.1007/s12170-019-0601-5>
- Tanner, R.M., Gutiérrez, O.M., Judd, S., McClellan, W., Bowling, C.B., Bradbury, B.D., Safford, M.M., Cushman, M., Warnock, D., Muntner, P., 2013. Geographic Variation in CKD Prevalence and ESRD Incidence in the United States: Results From the Reasons for Geographic and Racial Differences in Stroke (REGARDS) Study. *American Journal of Kidney Diseases* 61, 395–403. <https://doi.org/10.1053/j.ajkd.2012.10.018>
- Tao, Y., Chai, Y., Kou, L., Kwan, M.-P., 2020. Understanding noise exposure, noise annoyance, and psychological stress: Incorporating individual mobility and the temporality of the exposure-effect relationship. *Applied Geography* 125, 102283. <https://doi.org/10.1016/j.apgeog.2020.102283>
- Tešić, M., Hermans, E., Lipovac, K., Pešić, D., 2018. Identifying the most significant indicators of the total road safety performance index. *Accident Analysis & Prevention* 113, 263–278. <https://doi.org/10.1016/j.aap.2018.02.003>
- Tharrault, Y., Mourot, G., Ragot, J., Maquin, D., 2007. Détection et isolation de défauts par analyse en composantes principales robuste. 2èmes Journées Doctorales / Journées Nationales MACS, Reims.
- Thatipelli, S., Kershaw, K.N., Colangelo, L.A., Gordon-Larsen, P., Jacobs, D.R., Dransfield, M.T., Meza, D., Rosenberg, S.R., Washko, G.R., Parekh, T.M., Carnethon, M.R., Kalhan, R., 2022. Neighborhood Socioeconomic Deprivation in Young Adulthood and Future Respiratory Health: The CARDIA Lung Study. *The American Journal of Medicine* 135, 211-218.e1. <https://doi.org/10.1016/j.amjmed.2021.07.048>
- Tibuakuu, M., Michos, E.D., Navas-Acien, A., Jones, M.R., 2018. Air Pollution and Cardiovascular Disease: a Focus on Vulnerable Populations Worldwide. *Curr Epidemiol Rep* 5, 370–378. <https://doi.org/10.1007/s40471-018-0166-8>
- Tobías, A., Recio, A., Díaz, J., Linares, C., 2015. Health impact assessment of traffic noise in Madrid (Spain). *Environmental Research* 137, 136–140. <https://doi.org/10.1016/j.envres.2014.12.011>
- Tobler, W.R., 1970. A computer movie simulating urban growth in the Detroit region. *Econ. Geogr.* 46, 234–240
- Togtokh, C., Gaffney, O., 2010. 2010 Human Sustainable Development Index. *The Our World web magazine*.
- Tsai, C.-C., Wu, C.-L., Kor, C.-T., Lian, I.-B., Chang, C.-H., Chang, T.-H., Chang, C.-C., Chiu, P.-F., 2018. Prospective associations between environmental heavy metal exposure and renal outcomes

- in adults with chronic kidney disease: Heavy metals in soil and CKD progression. *Nephrology* 23, 830–836. <https://doi.org/10.1111/nep.13089>
- Twohig-Bennett, C., Jones, A., 2018. The health benefits of the great outdoors: A systematic review and meta-analysis of greenspace exposure and health outcomes. *Environmental Research* 166, 628–637. <https://doi.org/10.1016/j.envres.2018.06.030>
- Tzoulas, K., Korpela, K., Venn, S., Yli-Pelkonen, V., Kaźmierczak, A., Niemela, J., James, P., 2007. Promoting ecosystem and human health in urban areas using Green Infrastructure: A literature review. *Landscape and Urban Planning* 81, 167–178. <https://doi.org/10.1016/j.landurbplan.2007.02.001>
- UNDP (Ed.), 1990. Human development report 1990. Published for the United Nations (New York) Development Programme, United Nations Development Programme. ed. Oxford Univ. Pr, New York.
- Vaillant, Z., Bardes, J., Rican, S., 2021. De la discrimination positive à la discrimination territoriale : les quartiers en politique de la ville, inégaux face à la santé: Les cahiers de la LCD N° 12, 67–91. <https://doi.org/10.3917/clcd.012.0067>
- Vallée, J., 2019. Les effets de lieu au quotidien (Rapport Habilitation à Diriger des Recherches). Université de Paris I - Panthéon - Sorbonne, Paris.
- Vallée, J., 2015. Santé et échelles territoriales. La santé en action, Les échelles territoriales pour diagnostiquer, comprendre et agir face aux inégalités de santé sont-elles les mêmes ? Dossier Urbanisme et aménagements favorables à la santé, 30–31.
- van den Berg, P., Kemperman, A., de Kleijn, B., Borgers, A., 2016. Ageing and loneliness: The role of mobility and the built environment. *Travel Behaviour and Society* 5, 48–55. <https://doi.org/10.1016/j.tbs.2015.03.001>
- van den Bosch, M., Ode Sang, Å., 2017. Urban natural environments as nature-based solutions for improved public health – A systematic review of reviews. *Environmental Research* 158, 373–384. <https://doi.org/10.1016/j.envres.2017.05.040>
- van der Ploeg, S., de Groot, D., Wang, Y., 2010. The TEEB Valuation Database: overview of structure, data and results - Final report. Foundation for Sustainable Development, Wageningen, the Netherlands.
- van Deurzen, I., Rod, N.H., Christensen, U., Hansen, Å.M., Lund, R., Dich, N., 2016. Neighborhood perceptions and allostatic load: Evidence from Denmark. *Health & Place* 40, 1–8. <https://doi.org/10.1016/j.healthplace.2016.04.010>
- Vasan, R.S., Zachariah, J.P., Xanthakis, V., 2020. Life Course Developmental Approach to Cardiovascular Health and Cardiovascular Disease Prevention. *Journal of the American College of Cardiology* 76, 2708–2711. <https://doi.org/10.1016/j.jacc.2020.10.011>
- Vigneron, E., 2013. Inégalités de santé, inégalités de soins dans les territoires français. *Les Tribunes de la santé* n° 38, 41. <https://doi.org/10.3917/seve.038.0041>
- Villanueva, K., Badland, H., Hooper, P., Koohsari, M.J., Mavoa, S., Davern, M., Roberts, R., Goldfeld, S., Giles-Corti, B., 2015. Developing indicators of public open space to promote health and wellbeing in communities. *Applied Geography* 57, 112–119. <https://doi.org/10.1016/j.apgeog.2014.12.003>
- Vineis, P., Robinson, O., Chadeau-Hyam, M., Dehghan, A., Mudway, I., Dagnino, S., 2020. What is new in the exposome? *Environment International* 143, 105887. <https://doi.org/10.1016/j.envint.2020.105887>
- Vorstius, C., Rowan, J.S., Brown, I., Frogbrook, Z., Palarea-Albaladejo, J., 2019. Large-scale risk screening of raw water quality in the context of drinking water catchments and integrated response strategies. *Environmental Science & Policy* 100, 84–93. <https://doi.org/10.1016/j.envsci.2019.05.012>
- Wagstaff, A., 2002. Poverty and health sector inequalities. *Bull. World Health Organ.* 80, 97–105.
- Wakefield, J.C., Smith, T.R., 2016. Ecological modeling: general issues, in: *Handbook of Spatial Epidemiology*. CRC Press/Taylor & Francis group, Boca Raton.

- Walker, K.E., Crotty, S.M., 2015. Classifying high-prevalence neighborhoods for cardiovascular disease in Texas. *Applied Geography* 57, 22–31. <https://doi.org/10.1016/j.apgeog.2014.11.011>
- Waller, L.A., Gotway, C.A., 2004. *Applied spatial statistics for public health data*, Wiley series in probability and statistics. John Wiley & Sons, Hoboken, N.J.
- Wang, K., Lombard, J., Rundek, T., Dong, C., Gutierrez, C.M., Byrne, M.M., Toro, M., Nardi, M.I., Kardys, J., Yi, L., Szapocznik, J., Brown, S.C., 2019. Relationship of Neighborhood Greenness to Heart Disease in 249 405 US Medicare Beneficiaries. *JAHA* 8. <https://doi.org/10.1161/JAHA.118.010258>
- Wang, S., Cheng, C., Tan, S., 2019. Housing Determinants of Health in Urban China: A Structural Equation Modeling Analysis. *Soc Indic Res* 143, 1245–1270. <https://doi.org/10.1007/s11205-018-2022-0>
- Ward, J.H., 1963. Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association* 58, 236–244. <https://doi.org/10.1080/01621459.1963.10500845>
- Ward, M.M., 2008. Socioeconomic Status and the Incidence of ESRD. *American Journal of Kidney Diseases* 51, 563–572. <https://doi.org/10.1053/j.ajkd.2007.11.023>
- Watts, N., Amann, M., Ayeb-Karlsson, S., Belesova, K., Bouley, T., Boykoff, M., Byass, P., Cai, W., Campbell-Lendrum, D., Chambers, J., Cox, P.M., Daly, M., Dasandi, N., Davies, M., Depledge, M., Depoux, A., Dominguez-Salas, P., Drummond, P., Ekins, P., Flahault, A., Frumkin, H., Georgeson, L., Ghanei, M., Grace, D., Graham, H., Grojsman, R., Haines, A., Hamilton, I., Hartinger, S., Johnson, A., Kelman, I., Kiesewetter, G., Kniveton, D., Liang, L., Lott, M., Lowe, R., Mace, G., Odhiambo Sewe, M., Maslin, M., Mikhaylov, S., Milner, J., Latifi, A.M., Moradi-Lakeh, M., Morrissey, K., Murray, K., Neville, T., Nilsson, M., Oreszczyn, T., Owfi, F., Pencheon, D., Pye, S., Rabhaniha, M., Robinson, E., Rocklöv, J., Schütte, S., Shumake-Guillemot, J., Steinbach, R., Tabatabaei, M., Wheeler, N., Wilkinson, P., Gong, P., Montgomery, H., Costello, A., 2018. The Lancet Countdown on health and climate change: from 25 years of inaction to a global transformation for public health. *The Lancet* 391, 581–630. [https://doi.org/10.1016/S0140-6736\(17\)32464-9](https://doi.org/10.1016/S0140-6736(17)32464-9)
- Whitehead, M., 1992. The Concepts and Principles of Equity and Health. *Int J Health Serv* 22, 429–445. <https://doi.org/10.2190/986L-LHQ6-2VTE-YRRN>
- Whitehead, M., Dahlgren, G., 2006. Concept and principles for tackling social inequalities in health: Levelling up Part 1 (Studies on social and economic determinants of populations health No. 2). World Health Organization, Regional Office for Europe, University of Liverpool.
- WHO, 2022. Almost everyone now breathing polluted air, warns WHO, UN News, Global perspective Human stories. World Health Organization.
- WHO, 2021a. Ambient (outdoor) air pollution. World Health Organization.
- WHO, 2021b. Lignes directrices OMS relatives à la qualité de l'air : particules (PM_{2,5} et PM₁₀), ozone, dioxyde d'azote, dioxyde de soufre et monoxyde de carbone : résumé d'orientation. Organisation mondiale de la Santé, Genève.
- WHO, 2021c. Noncommunicable diseases - Key facts. World Health Organization.
- WHO, 2021d. Cardiovascular diseases (CVDs) - Key facts.
- WHO (Ed.), 2019. Environmental health inequalities in Europe: second assessment report. World Health Organization, Regional Office for Europe, Copenhagen.
- WHO, 2018a. Santé, environnement et changement climatique - Rapport du Directeur général (Soixante et onzième Assemblée Mondiale de la Santé No. A71/10). World Health Organization.
- WHO, 2018b. Environmental noise guidelines for the European Region.
- WHO, 2016. Urban green spaces and health - A review of evidence, European Environment and Health Process. World Health Organization Regional Office for Europe, Copenhagen.
- WHO, 2012a. Cadre politique et stratégie Santé 2020 (Comité régional de l'Europe No. soixante-deuxième session). World Health Organization, Malte.

- WHO (Ed.), 2012b. Environmental health inequalities in Europe: assessment report. World Health Organization, Regional Office for Europe, Copenhagen.
- WHO (Ed.), 2011. Burden of disease from environmental noise: quantification of healthy life years lost in Europe. World Health Organization, Regional Office for Europe, Copenhagen.
- WHO, 2010. La face cachée des villes mettre au jour et vaincre: les inégalités en santé en milieu urbain. Organisation mondiale de la Santé ; ONU-HABITAT, Genève; Nairobi.
- WHO, 1999. Environmental Health Indicators: Framework and Methodologies (No. WHO/SDE/OEH/99.10). World Health Organization Sustainable Development and Healthy Environments, Geneva.
- Wilke, R.A., Qamar, M., Lupu, R.A., Gu, S., Zhao, J., 2019. Chronic Kidney Disease in Agricultural Communities. *The American Journal of Medicine* 132, e727–e732. <https://doi.org/10.1016/j.amjmed.2019.03.036>
- Williams, L., Zhang, R., Packard, K.C., 2017. Factors affecting the physical and mental health of older adults in China: The importance of marital status, child proximity, and gender. *SSM - Population Health* 3, 20–36. <https://doi.org/10.1016/j.ssmph.2016.11.005>
- Wilund, K.R., Thompson, S., Viana, J.L., Wang, A.Y.-M., 2021. Physical Activity and Health in Chronic Kidney Disease, in: Silva Junior, G.B., Ferreiro Fuentes, A., Nangaku, M., Remuzzi, G., Ronco, C. (Eds.), *Contributions to Nephrology*. S. Karger AG, pp. 1–13. <https://doi.org/10.1159/000517696>
- Windenberger, F., Rican, S., Jouglu, E., Rey, G., 2012. Spatiotemporal association between deprivation and mortality: trends in France during the nineties. *European Journal of Public Health* 22, 347–353. <https://doi.org/10.1093/eurpub/ckr029>
- Wiréhn, L., Danielsson, Å., Neset, T.-S.S., 2015. Assessment of composite index methods for agricultural vulnerability to climate change. *Journal of Environmental Management* 156, 70–80. <https://doi.org/10.1016/j.jenvman.2015.03.020>
- Wolf, K., Hoffmann, B., Andersen, Z.J., Atkinson, R.W., Bauwelinck, M., Bellander, T., Brandt, J., Brunekreef, B., Cesaroni, G., Chen, J., de Faire, U., de Hoogh, K., Fecht, D., Forastiere, F., Gulliver, J., Hertel, O., Hvidtfeldt, U.A., Janssen, N.A.H., Jørgensen, J.T., Katsouyanni, K., Ketzel, M., Klompmaker, J.O., Lager, A., Liu, S., MacDonald, C.J., Magnusson, P.K.E., Mehta, A.J., Nagel, G., Oftedal, B., Pedersen, N.L., Pershagen, G., Raaschou-Nielsen, O., Renzi, M., Rizzuto, D., Rodopoulou, S., Samoli, E., van der Schouw, Y.T., Schramm, S., Schwarze, P., Sigsgaard, T., Sørensen, M., Stafoggia, M., Strak, M., Tjønneland, A., Verschuren, W.M.M., Vienneau, D., Weinmayr, G., Hoek, G., Peters, A., Ljungman, P.L.S., 2021. Long-term exposure to low-level ambient air pollution and incidence of stroke and coronary heart disease: a pooled analysis of six European cohorts within the ELAPSE project. *The Lancet Planetary Health* 5, e620–e632. [https://doi.org/10.1016/S2542-5196\(21\)00195-9](https://doi.org/10.1016/S2542-5196(21)00195-9)
- Wong, C.W., Kwok, C.S., Narain, A., Gulati, M., Mihalidou, A.S., Wu, P., Alasnag, M., Myint, P.K., Mamas, M.A., 2018. Marital status and risk of cardiovascular diseases: a systematic review and meta-analysis. *Heart* 104, 1937–1948. <https://doi.org/10.1136/heartjnl-2018-313005>
- Wong, D.W.S., 2004. The Modifiable Areal Unit Problem (MAUP), in: Janelle, D.G., Warf, B., Hansen, K. (Eds.), *WorldMinds: Geographical Perspectives on 100 Problems*. Springer Netherlands, Dordrecht, pp. 571–575. https://doi.org/10.1007/978-1-4020-2352-1_93
- Wresinski, J., 1987. Grande pauvreté et précarité économique et sociale (Avis et rapports du Conseil Economique et Social No. 6), *Journal Officiel de la République Française*.
- Wu, C.-D., Chern, Y.-R., Pan, W.-C., Lung, S.-C.C., Yao, T.-C., Tsai, H.-J., Spengler, J.D., 2020. Effects of surrounding environment on incidence of end stage renal disease. *Science of The Total Environment* 723, 137915. <https://doi.org/10.1016/j.scitotenv.2020.137915>
- Wu, Y.-H., Wu, C.-D., Chung, M.-C., Chen, C.-H., Wu, L.-Y., Chung, C.-J., Hsu, H.-T., 2022. Long-Term Exposure to Fine Particulate Matter and the Deterioration of Estimated Glomerular Filtration Rate: A Cohort Study in Patients With Pre-End-Stage Renal Disease. *Front. Public Health* 10, 858655. <https://doi.org/10.3389/fpubh.2022.858655>

- Xiao, Q., Berrigan, D., Matthews, C.E., 2017. A prospective investigation of neighborhood socioeconomic deprivation and self-rated health in a large US cohort. *Health & Place* 44, 70–76. <https://doi.org/10.1016/j.healthplace.2017.01.003>
- Yao, L., Liu, J., Wang, R., Yin, K., Han, B., 2014. Effective green equivalent—A measure of public green spaces for cities. *Ecological Indicators* 47, 123–127. <https://doi.org/10.1016/j.ecolind.2014.07.009>
- Yeager, R., Riggs, D.W., DeJarnett, N., Srivastava, Shweta, Lorkiewicz, P., Xie, Z., Krivokhizhina, T., Keith, R.J., Srivastava, Sanjay, Browning, M.H.E.M., Zafar, N., Krishnasamy, S., DeFilippis, A., Turner, J., Rai, S.N., Bhatnagar, A., 2020. Association between residential greenness and exposure to volatile organic compounds. *Science of The Total Environment* 707, 135435. <https://doi.org/10.1016/j.scitotenv.2019.135435>
- Yeager, R.A., Smith, T.R., Bhatnagar, A., 2019. Green environments and cardiovascular health. *Trends in Cardiovascular Medicine*. <https://doi.org/10.1016/j.tcm.2019.06.005>
- Yitshak-Sade, M., Kloog, I., Novack, V., 2017. Do air pollution and neighborhood greenness exposures improve the predicted cardiovascular risk? *Environment International* 107, 147–153. <https://doi.org/10.1016/j.envint.2017.07.011>
- Yu, Y., Jerrett, M., Paul, K.C., Su, J., Shih, I.-F., Wu, J., Lee, E., Inoue, K., Haan, M., Ritz, B., 2021. Ozone Exposure, Outdoor Physical Activity, and Incident Type 2 Diabetes in the SALSA Cohort of Older Mexican Americans. *Environ Health Perspect* 129, 097004. <https://doi.org/10.1289/EHP8620>
- Zelem, M.C., 2013. Mobilités choisies (vs) mobilités contraintes (No. Note 11), *Débat National sur la Transition Energétique*.
- Zhang, L., Niu, M., Zhang, Haiyang, Wang, Y., Zhang, Haiqing, Mao, Z., Zhang, X., He, M., Wu, T., Wang, Z., Wang, C., 2022. Nonlaboratory-based risk assessment model for coronary heart disease screening: Model development and validation. *International Journal of Medical Informatics* 162, 104746. <https://doi.org/10.1016/j.ijmedinf.2022.104746>
- Zhang, Q.-L., Rothenbacher, D., 2008. Prevalence of chronic kidney disease in population-based studies: Systematic review. *BMC Public Health* 8, 117. <https://doi.org/10.1186/1471-2458-8-117>
- Zhang, S., Wei, Z., Liu, W., Yao, L., Suo, W., Xing, J., Huang, B., Jin, D., Wang, J., 2015. Indicators for Environment Health Risk Assessment in the Jiangsu Province of China. *International Journal of Environmental Research and Public Health* 12, 11012–11024. <https://doi.org/10.3390/ijerph120911012>
- Zhang, X., Zhao, D., Xie, X., Wei, Z., Wang, W., Sun, J., Wang, M., Liu, Jun, Qi, Y., Guo, M., Zheng, J., Liu, Jing, 2011. Marital status in relation to case fatality after myocardial infarction — A population based study. *International Journal of Cardiology* 152, S54. <https://doi.org/10.1016/j.ijcard.2011.08.646>
- Zhao, C., Zhou, Y., Jiang, J., Xiao, P., Wu, H., 2021. Spatial characteristics of cultivated land quality accounting for ecological environment condition: A case study in hilly area of northern Hubei province, China. *Science of The Total Environment* 774, 145765. <https://doi.org/10.1016/j.scitotenv.2021.145765>
- Zhou, T., Hu, B., Meng, X., Sun, L., Li, H., Xu, P., Cheng, B., Sheng, J., Tao, F., Yang, L., Wu, Q., 2021. The associations between urinary metals and metal mixtures and kidney function in Chinese community-dwelling older adults with diabetes mellitus. *Ecotoxicology and Environmental Safety* 226, 112829. <https://doi.org/10.1016/j.ecoenv.2021.112829>
- Zuindeau, B., Lipovac, J.-C., 2009. Elaboration d'un indicateur participatif de bien-être (IPBE) : l'exemple du territoire d'Hénin-Carvin (Nord Pas de Calais). *Revue Critique d'Ecologie Politique*.

Annexes

Table des annexes

Annexe 1 : Supplementary material and results – Article 3.....	ii
Annexe 2 : Supplementary materials and results – Article 4.....	xiv
Annexe 3 : Liste des centres de dialyse (Hauts-de-France).....	xliv
Annexe 4 : Modèles épidémiologiques développés pour déterminer la variabilité résiduelle de l'IRCT.....	xliv
Annexe 5 : Supplementary materials and results – Article 5.....	xlvii
Annexe 6 : Maladies coronariennes, évènements « fatal » : ratio d'incidence standardisé, par standardisation indirecte (2012-2016)	lxvii
Annexe 7 : Comparatif Hauts-de-France Wallonie.....	lxviii

Annexe 1

Supplementary material and results – Article 3

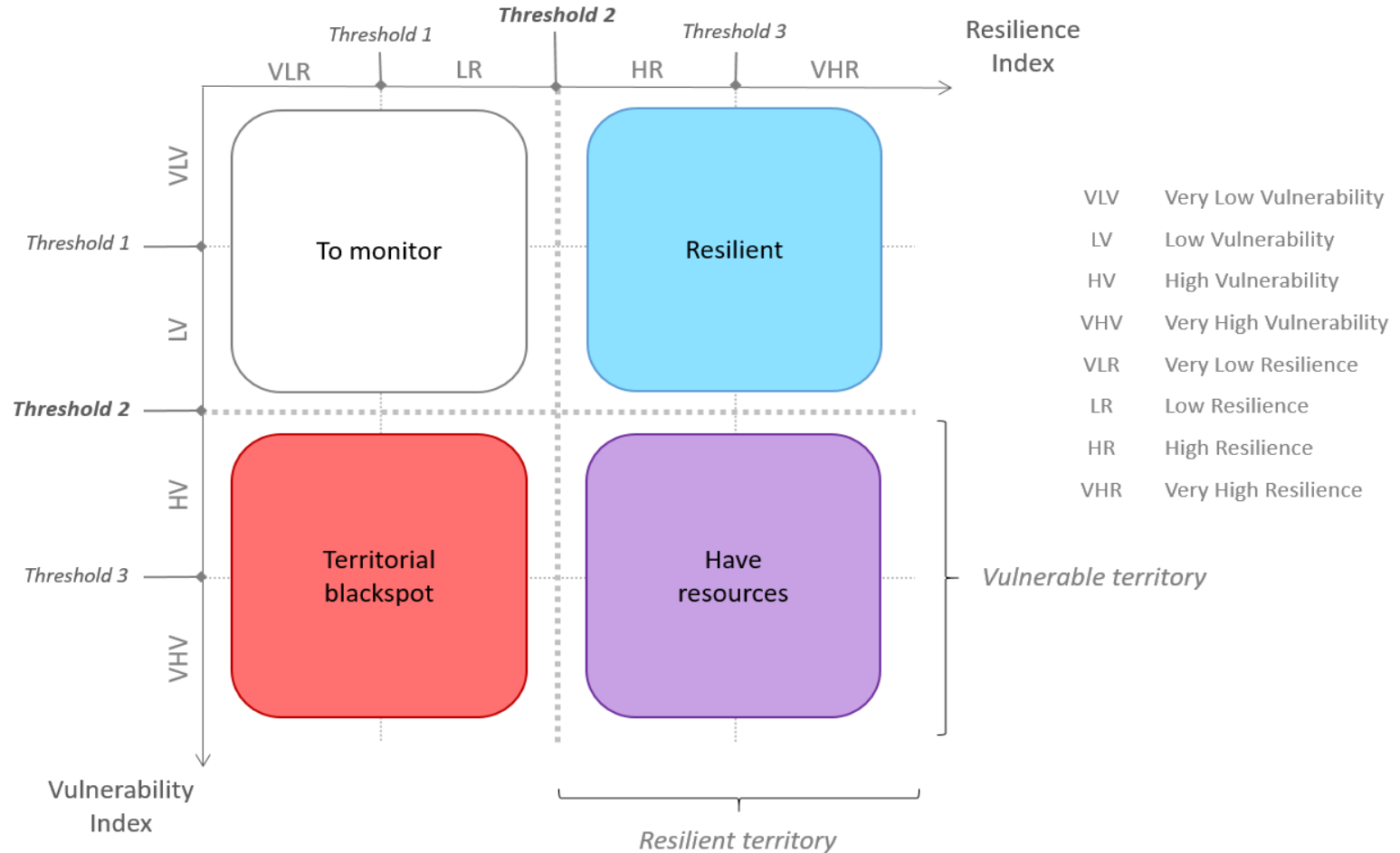
Supplementary material

1. Cross interpretation of resilience and vulnerability indices

Cross-interpretation of resilience and vulnerability indices

Distribution in four territory profiles according to the levels of resilience and vulnerability (very low, low, high, very high).

The territorial blackspots are characterized by a high (or very high) vulnerability paired with a low (or very low) resilience

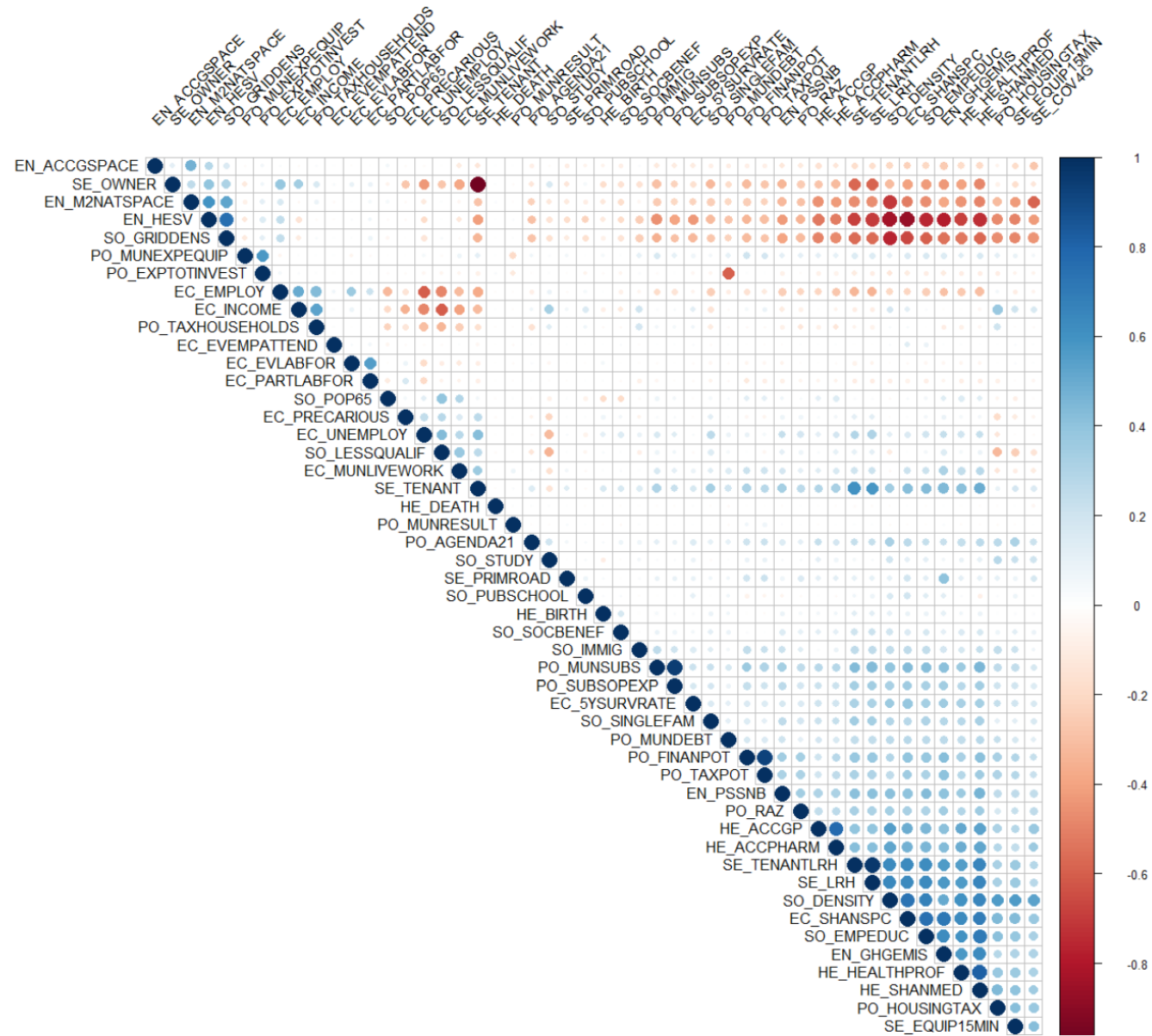


Supplementary results – summary

1. Acronym and variable
2. Step (i): “Collection and analysis of data”, Bivariate analysis - correlation matrix (1/2)
3. Step (i): “Collection and analysis of data”, Bivariate analysis - correlation matrix (2/2)
4. Step (i): “Collection and analysis of data”, Correlogram
5. Step (i): “Collection and analysis of data”, Final use of the collected data
6. Step (i): “Collection and analysis of data”, PCA and graphical representation of the 50 variables
7. Step (ii): “Calculation of the resilience and vulnerability indices”, graphical representation of variables included in Resilience Index
8. Step (ii): “Calculation of the resilience and vulnerability indices”, graphical representation of variables included in Vulnerability Index

Acronym	Data	Acronym	Data
EC_5YSURVRATE	The 5-year survival rate of companies	PO_TAXHOUSEHOLDS	Share of tax households taxed
EC_EMPLOY	Employment rate	PO_TAXPOT	Tax potential per inhabitant
EC_EVEMPATTEND	Employment evolution for attendance activities	HE_ACCGP	Potential spatial accessibility to general practitioners
EC_EVLABFOR	Evolution of the labour force	HE_ACCPHARM	Potential spatial accessibility to pharmacist
EC_INCOME	Average income per consumption unit	HE_BIRTH	Proportion of births
EC_MUNLIVWORK	Share of the labour force living and working in the same municipality	HE_DEATH	Proportion of deaths
EC_PARTLABFOR	Evolution of the share of the labour force in the total population	HE_HEALTHPROF	Number of health professionals, per 1,000 inhabitants
EC_PRECARIOUS	Precarious employment rate	HE_SHANMED	Shannon-Weaver index on the representation of medical and paramedical professions
EC_SHANSPC	Shannon-Weaver index on the representation of socio-professional categories	SO_DENSITY	Population density
EC_UNEMPLOY	Unemployment rate	SO_EMPEDUC	Employment in education and training, per 1,000 inhabitants
EN_ACCGSPACE	Share of population with access to green space within 400 meters of their homes	SO_GRIDDENS	Communal density grid
EN_GHGEMIS	Greenhouse gas (GHG) emissions	SO_IMMIG	Share of immigrants in the total population
EN_HESV	Monetisation of ecosystem services per capita	SO_LESSQUALIF	Share of working-age population with few or no qualifications
EN_M2NATSPACE	Area (m ²) of natural spaces per inhabitant	SO_POP65	Share of the population aged over 65 years
EN_PSSNB	Number of contaminated sites (polluted sites and soils)	SO_PUBSCHOOL	Number of state schools, per 1,000 inhabitants
PO_AGENDA21	Municipality covered by an Agenda 21	SO_SINGLEFAM	Share of single-parent families
PO_EXPTOTINVEST	Share of expenditure for municipality equipment compared to the total investment uses	SO_SOCBENEF	Share of persons covered by a social benefit
PO_FINANPOT	Financial potential per inhabitant	SO_STUDY	Study continuation rate for 18-24 years old
PO_HOUSINGTAX	Outcome of the housing tax, over the number of primary and secondary residences	SE_4GCOV	Part of area with 4G cover provided by the better operator
PO_MUNDEBT	Proportion of medium- and long-term debt of the municipality, per inhabitant	SE_EQUIP15MIN	Share of population with average access to 12- mid-range municipal equipments in 15 minutes or less
PO_MUNEXPEQUIP	Expenditure for municipal equipment, per inhabitant	SE_LRH	Low-rent housing (LRH) per 1,000 inhabitants
PO_MUNRESULT	Overall financial result of the municipality, per inhabitant	SE_OWNER	Proportion of people owning their main residence
PO_MUNSUBS	Amount of subsidies paid by municipalities, per inhabitant	SE_PRIMROAD	Length of primary roads per capita
PO_RAZ	Eligible areas for regional aid zoning	SE_TENANT	Share of population who are tenants of their main residence
PO_SUBSOPEXP	Share of subsidies paid within operating expenses	SE_TENANTLRH	Share of tenants who rent their main residence in Low-Rent Housing

Step (i): "Collection and analysis of data", Correlogram



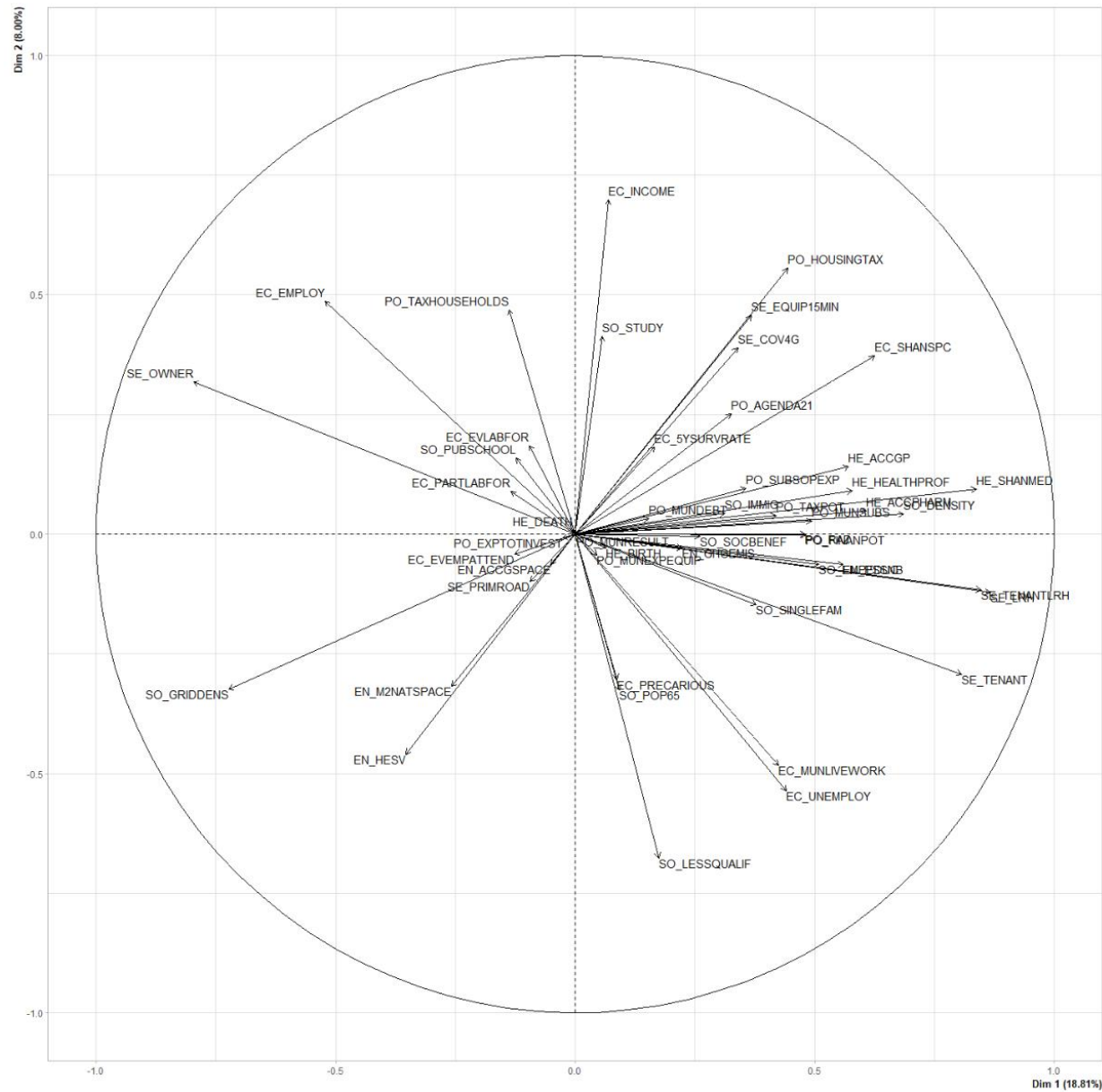
Step (i): “Collection and analysis of data”, final use of the collected data

If in final use there is “ $cor > |0,6|$ ” or “ $|0,2| < cor$ ”, the corresponding variable was not considered for the development of composite indices

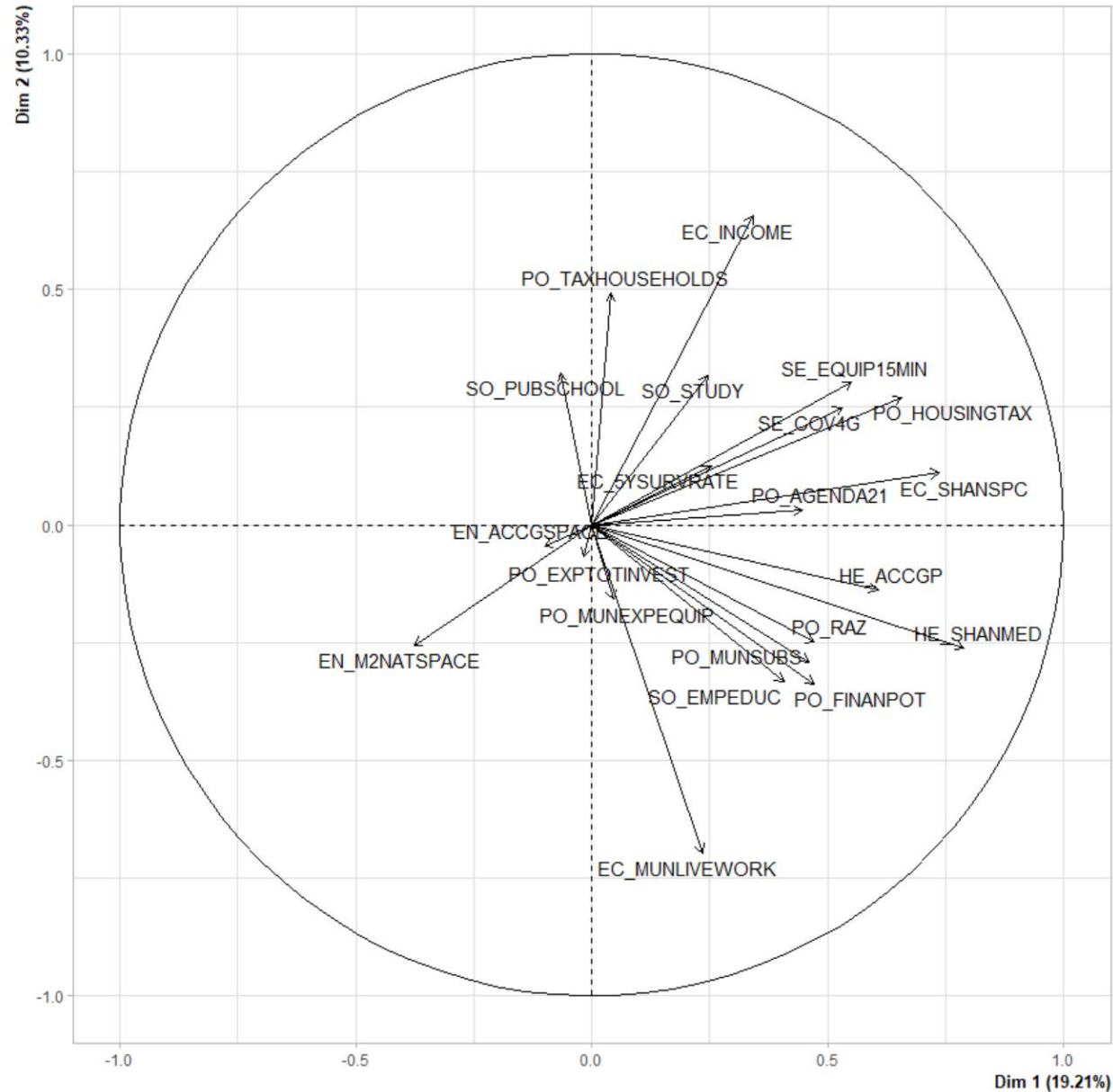
Acronym	Final use
EC_5YSURVRATE	Resilience Index
EC_EMPLOY	$cor > 0,6 $
EC_EVEMPATTEND	$ 0,2 < cor$
EC_EVLABFOR	$ 0,2 < cor$
EC_INCOME	Resilience Index
EC_MUNLIVEWORK	Resilience Index
EC_PARTLABFOR	$ 0,2 < cor$
EC_PRECARIOUS	Vulnerability Index
EC_SHANSPC	Resilience Index
EC_UNEMPLOY	Vulnerability Index
EN_ACCGSPACE	Resilience Index
EN_GHGEMIS	Vulnerability Index
EN_HESV	$ 0,2 < cor$
EN_M2NATSPACE	Resilience Index
EN_PSSNB	Vulnerability Index
PO_AGENDA21	Resilience Index
PO_EXPTOTINVEST	Resilience Index
PO_FINANPOT	Resilience Index
PO_HOUSINGTAX	Resilience Index
PO_MUNDEBT	Vulnerability Index
PO_MUNEXPEQUIP	Resilience Index
PO_MUNRESULT	$ 0,2 < cor$
PO_MUNSUBS	Resilience Index
PO_RAZ	Resilience Index
PO_SUBSOPEXP	$cor > 0,6 $

Acronym	Final use
PO_TAXHOUSEHOLDS	Resilience Index
PO_TAXPOT	$cor > 0,6 $
HE_ACCGP	Resilience Index
HE_ACCPHARM	$cor > 0,6 $
HE_BIRTH	$ 0,2 < cor$
HE_DEATH	Vulnerability Index
HE_HEALTHPROF	$cor > 0,6 $
HE_SHANMED	Resilience Index
SO_DENSITY	Vulnerability Index
SO_EMPEDUC	Resilience Index
SO_GRIDDENS	$cor > 0,6 $
SO_IMMIG	Vulnerability Index
SO_LESSQUALIF	Vulnerability Index
SO_POP65	Vulnerability Index
SO_PUBSCHOOL	Resilience Index
SO_SINGLEFAM	Vulnerability Index
SO_SOCBENEF	$ 0,2 < cor$
SO_STUDY	Resilience Index
SE_4GCOV	Resilience Index
SE_EQUIP15MIN	Resilience Index
SE_LRH	$cor > 0,6 $
SE_OWNER	$cor > 0,6 $
SE_PRIMROAD	Vulnerability Index
SE_TENANT	Vulnerability Index
SE_TENANTLRH	$cor > 0,6 $

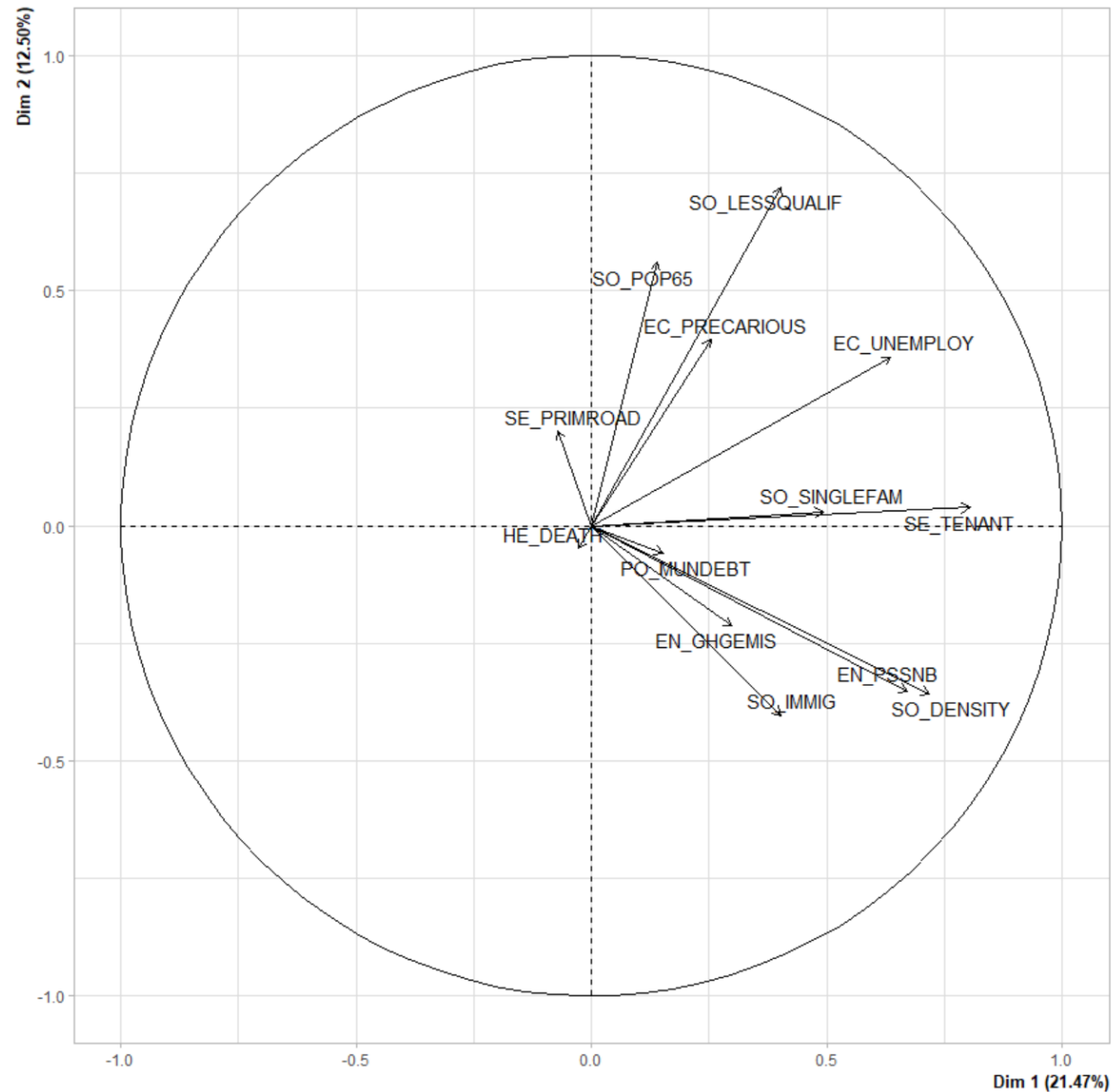
Step (i): "Collection and analysis of data", PCA and graphical representation of the 50 variables



Step (ii): "Calculation of the resilience and vulnerability indices", graphical representation of variables included in Resilience Index (prior the Varimax rotation)



Step (ii): "Calculation of the resilience and vulnerability indices", graphical representation of variables included in Vulnerability Index (prior the Varimax rotation)



Annexe 2

Supplementary materials and results – Article 4

Supplementary materials

1. Descriptive measures for the population and cover areas
 - Page 2
2. List of official national organizations from which data were collected
 - Page 3
3. Meta-data concerning dataset used to develop composite indices at the IRIS scale (IRIS: the smallest aggregated unit for spatial information):
 - Page 4: meta-data related to “Economics”
 - Page 5-6: meta-data related to “Environment”
 - Page 7: meta-data related to “Health”
 - Page 8: meta-data related to “Social factors”
 - Page 9: meta-data related to “Services”
 - Page 10: meta-data related to “Policy”

Descriptive measures for the population and cover areas

		Municipalities (17)	IRIS (93)
Population (inhab.)	Median	5,320	2,108
	Q1-Q3	2,437 - 8,022	1,739 - 2,675
Cover area (ha)	Median	950.062	69.401
	Q1-Q3	773.450 - 3,060.781	34.480 - 223.266

ATMO HDF: Region air observatory in Hauts-de-France region, belonging to the French Federation of Certified Associations for the air quality Monitoring

CNAF: National Health Insurance Fund

IGN: National Institute for Geographic and forest information

INSEE: French National Institute of Statistics and Economic Studies

MEDDE: Ministry of Ecological and Solidarity Transition

Open Data DUC: open data portal of the Dunkerque Urban Community

Meta-data related to “Economics”

Data	Acronym	Specifications	Unit	Source(s)	Time period
Average income per consumption unit	DEC_MED17	The disposable income of a household comprises earned income (net of social contributions), income from assets, transfers from other households and social benefits (including pensions and unemployment benefits), and net of direct taxes. The consumption unit (CU) is an OECD unit of measurement that weights individuals in a household according to their age, allowing comparison between households. CU are generally calculated in the following manner: 1 CU is allocated to the first adult in the household, 0.5 CU to other persons aged 14 or over and 0.3 CU to children under 14	€/CU	INSEE	2017
Employment rate	TX_EMP	Proportion of economically active persons aged 15-64 (in employment) in the total population aged 15-64	%	INSEE	2016
Evolution of the labour force	V_EMP_1116	Calculation of a change in the number of people aged 15-64 in the labour force between 2 non-consecutive years	%	INSEE	2011-2016
Evolution of the share of the labour force in the total population	EV_POP_ACT_TOT_1116	Calculation of the evolution of the share of the labour force aged 15 and over within the total population between 2 non-consecutive years.	%	INSEE	2011-2016
Precarious employment rate	T_PREC	Share of the labour force aged 15-64 in fixed-term contracts, temporary employment, assisted employment and apprenticeships among the labour force aged 15-64	%	INSEE	2016
Shannon-Weaver index on the representation of socio-professional categories	SHANNON_CSP	Use of the Shannon-Weaver index to describe a population in terms of the number of categories (category richness) and their relative abundance within those categories (species equitability). 15 socio-professional categories were considered for the calculation, identified at the place of residence	Index	INSEE	2016
Share of the labour force living and working in the same municipality	P_AC_OCC_MMCOM	The labour force considered is aged 15 and over	%	INSEE	2016
Unemployment rate	T_CHOM	Share of people aged 15-64 who are unemployed and looking for work	%	INSEE	2016

Meta-data related to “Environment”

Data	Acronym	Specifications	Unit	Source(s)	Time period
SO ₂	SO2_2013	Annual mean concentrations	µg/m ³	ATMO HDF	2013
NO ₂	NO2_2013	Annual mean concentrations	µg/m ³	ATMO HDF	2013
Particulate matters (PM ₁₀)	PM10_2013	Annual mean concentrations	µg/m ³	ATMO HDF	2013
Dust deposition	CUD_pouss	Biomonitoring with poplars was used. Dust deposition was quantified by measuring the accumulation of total dust on poplar leaves during the flowering season. Dust quantity was determined in relation to leaf mass	µg/g	Data from our laboratory	2018
Eutrophication	Eutro_2017	Biomonitoring with lichen was used. The lichen diversity and abundance were used to determine the Biological Index of Epiphytic Lichens which indicates the ecological responses of epiphytic lichens as integrators of environmental conditions over a ten-year period. This index was weighted by the Wirth index to obtain an eutrophic ratio, quantifying the influence of environmental nitrophilia on epiphytic communities	Index	Data from our laboratory	2017
Length of greenway network and bicycle paths per ha	LNGR_HA	A greenway is reserved for non-motorized travel, such as pedestrians and bicycles. Greenways are developed with a view to integrated planning that enhances the environment, heritage, quality of life and conviviality	m/ha	Open data DUC, INSEE	2016, 2019
Mean Impregnation Ratio	RIM_2018	Biomonitoring with lichen was used. The bioaccumulation of 18 trace elements (Al, Sb, As, Be, Cd, Co, Cr, Cu, Mn, Hg, Ni, Pb, Pd, Pt, Rh, Ti, V, Zn) was determined in the thalli of the foliose lichen Xanthoria parietina. The multi-metallic burden is generated by the calculation of a composite index	Index	Data from our laboratory	2018
Monetisation of ecosystem services per capita	H_ESV	Use of the Corine Land Cover database to determine 4 land-use categories (grassland, urban, cultivated and/or arable land, forests) and The Economics of Ecosystems and Biodiversity valuation database to assign a monetary value to the ecosystem services provided by these 4 categories	USD/ha/year	INSEE, MEDDE	2016-2018
Number of facilities classified for environmental protection per km ²	NB_ICPE_Superf	Any industrial or agricultural facility likely to create risks or cause pollution or nuisances, in particular for the safety and health of local residents, is a facility classified for environmental protection. A facility is subject to numerous environmental risk prevention regulations.	Number/km ²	MEDDE	2019

Meta-data related to “Environment”

Data	Acronym	Specifications	Unit	Source(s)	Time period
Share of population exposed to noise throughout the day	PART_EXPO_B_Lden70	The L_{DEN} is the noise indicator for overall annoyance which was used (level day evening night). The considered sources of noise were roads, railways and industrial facilities. Only areas with levels above 70 dB(A) were kept for the calculation. The population data are those of the INSEE’s 200-meter grid	%	Open data DUC, INSEE	2015, 2019
Share of population exposed to noise at night	PART_EXPO_B_Ln60	The L_N is the noise indicator for overall annoyance which was used (level night, between 22:00 and 6:00). The considered sources of noise were roads, railways and industrial facilities. Only areas with levels above 60dB(A) were kept for the calculation. The population data are those of the INSEE’s 200-meter grid	%	Open data DUC, INSEE	2015, 2019
Share of population with access to green space within 400meters of their homes	Pt_pop_Acc400_parc	Green spaces of less than 0.5ha have been excluded. The criterion included parks, leisure centers and natural areas. The population data are those of the INSEE’s 200-meter grid	%	Open data DUC, INSEE	2015, 2019

Meta-data related to “Health”

Data	Acronym	Specifications	Unit	Source(s)	Time period
Accessibility to general practitioners	ACC_MG_HC	Theoretical calculation performed in off-peak hours, between the nearest general practitioners and 200-metre gridded data of the French National Institute of Economics and Studies	Minute	Open data DUC, INSEE	2015, 2019
Accessibility to medical biology laboratory	ACC_LABO_H	Theoretical calculation performed in off-peak hours, between the nearest medical biology laboratory and 200-metre gridded data of the French National Institute of Economics and Studies	Minute	Open data DUC, INSEE	2015, 2019
Accessibility to pharmacy	ACC_PHARM_	Theoretical calculation performed in off-peak hours, between the nearest pharmacy and 200-metre gridded data of the French National Institute of Economics and Studies	Minute	Open data DUC, INSEE	2015, 2019
Number of health equipment per 1,000 inhabitants	NB_EQUIP_SANTE_1000HAB	Number of health equipment (medical biology laboratory, pharmacy, hospital centre, mental health centre, retirement homes, dental practitioners, nurses and midwives, general practitioners, counted at the IRIS-code of the place of activity, per 1,000 inhabitants.	Number per 1,000 inhabitants	INSEE	2016
Number of health professionals per 1,000 inhabitants	P_SANTE_1000H	Number of health professions, all specialities combined, counted at the IRIS-code of the place of activity, per 1,000 inhabitants.	Number per 1,000 inhabitants	INSEE	2016
Shannon-Weaver index on the representation of medical and paramedical professions	SHAN_MED	Use of the Shannon-Weaver index to describe a population in terms of the number of categories (category richness) and their relative abundance within those categories (category equitability). 25 medical and paramedical professions were considered for the calculation.	Index	INSEE	2016

Meta-data related to “Social factors”

Data	Acronym	Specifications	Unit	Source(s)	Time period
Study continuation rate for 18-24 years old	TXSCO_1824	Share of the population aged 18 to 24 pursuing higher education	%	INSEE	2016
Share of working-age population with few or no qualifications	PEU_DIPL_1564	Ratio between the number of out-of-school persons aged 15 years or over with no diploma (or at most a BEPC, a French certificate of general education, awarded by schools at the end of the first four years of general secondary education) and the total number of out-of-school persons aged 15 or over	%	INSEE	2016
Share of single-parent families	P_MONOP	A single-parent family includes a lone parent and one or more unmarried (childless) children	%	INSEE	2016
Population density	Densite_pop	Number of inhabitants per km ²	Inhab./km ²	INSEE	2016
Share of population under 11 years old	P_PERS_moins11	Percentage of the population under age 11 in the total population	%	INSEE	2016
Share of population aged 65 and over	P_PERS_plus65	Percentage of the population aged over 65 years in the total population	%	INSEE	2016
Share of immigrants in the total population	P_IMM_15	An immigrant is a person born abroad and residing in France. A distinction is made between foreign and immigrant populations: an immigrant is not necessarily a foreigner and conversely, some foreigners were born in France. Immigrant status is permanent: an individual continues to belong to the immigrant population even if he/she becomes French by acquisition	%	INSEE	2016
Number of state schools per 1,000 inhabitants	NB_EC_PUB_1000H	This correspond to the sum of the number of state nursery schools, state elementary schools and state elementary schools with a pre-elementary class	Number per 1,000 inhabitants	Open data DUC, INSEE	2016
Share of persons covered by a social benefit	P_COUV_AL	Percentage of persons covered by a benefit from the family branch, such as family benefits, early childhood benefits, family supplement, education allowance for a disabled child, family support allowance, back-to-school allowance, family housing allowance, social housing allowance, etc.	%	CNAF, INSEE	2016, 2018
Share of poor households	MEN_PAUV	Ratio between the number of poor households and the total number of households. The poverty line is set at 60% of the median standard of living	%	INSEE	2015
Number of social housing per 1,000 inhabitants	LOG_SOC_1000HAB	A social housing is housing intended, as a result of a public or private initiative, for people with modest incomes who could have difficulty to find housing on the private market	Number per 1,000 inhabitants	INSEE	2015

Meta-data related to “Services”

Data	Acronym	Specifications	Unit	Source(s)	Time period
Low-rent housing (LRH) per 1,000 inhabitants	HLM	Low-rent housing is housing units managed by a public or private low-rent housing organization, which benefits from partial, direct (subsidy) or indirect public funding (various advantages including credits, tax exemptions, etc.)	Number per 1,000 inhabitants	INSEE	2016
Proportion of people owning their main residence	P_PROP	Ownership status applies to owner, co-owner and first-time homeowner households	%	INSEE	2016
Share of population who are tenants of their main residence	P_LOCA	The rental stock comprises the public and the private sectors	%	INSEE	2016
Share of tenants who rent their main residence in Low-Rent Housing (LRH)	P_LOCA_HLM	LRH dwellings offer rents well below those of the private rental sector	%	INSEE	2016
Share of main residences built before 1919	Pt_RP_ACH19	It does not take into account the possible renovation of dwellings	%	INSEE	2016
Share of main residences built between 1919-1945	Pt_RP_ACH45	It does not take into account the possible renovation of dwellings	%	INSEE	2016
Share of main residences built between 1946-1970	Pt_RP_ACH70	It does not take into account the possible renovation of dwellings	%	INSEE	2016
Share of main residences built between 1971-1990	Pt_RP_ACH90	It does not take into account the possible renovation of dwellings	%	INSEE	2016
Share of main residences built between 1991-2005	Pt_RP_ACH05	It does not take into account the possible renovation of dwellings	%	INSEE	2016
Share of main residences built between 2006-2013	Pt_RP_ACH13	It does not take into account the possible renovation of dwellings	%	INSEE	2016
Number of sports clubs, associations and cultural facilities per 1,000 inhabitants	NB_EQP_SLC_1000H	The associations and cultural facilities are those supported by the DUC. Data on sports clubs were extracted from a national database which registers all establishments in France, all sectors of activity included (SIRENE)	Number per 1,000 inhabitants	Open data DUC, INSEE	2016
Length of primary roads per capita	LNG_RT_H	The primary road is defined as a portion of a roadway intended for pedestrian or cycles and includes only sections of major road networks	m/inhabitant	IGN, INSEE	2016
Length of bus line per inhabitant	LNG_BUS_HAB	This data corresponds to the entire bus network implemented in September 2018. It includes “chrono”, direct, connection and city center shuttle lines	m/inhabitant	Open data DUC, INSEE	2018

Meta-data related to “Policy”

Data	Acronym	Specifications	Unit	Source(s)	Time period
Share of areas to be urbanized	Pt_AU	This includes certain areas identified in the Local Urbanism Plan (LUP) for an urbanization, validated or subject to a modification of the LUP	%	Open data DUC	2019
Share of mixed urbanized areas	Pt_XY	This includes certain areas identified in restructuring or urban reconversion	%	Open data DUC	2019
Share of priority neighborhoods	QRT_PRIO	The priority neighborhoods are areas of intervention of the Ministry of the City. This public policy aims to counterbalance for differences in living standards with the rest of the territory. These neighborhoods are therefore those with the lowest incomes.	%	Open data DUC	2019

Supplementary results

1. Case 1: Included data in Resilience Index (page 12), in Vulnerability Index (page 13)
2. Case 2:
 - Bivariate analysis – correlation matrix: pages 14-15
 - Included data in Resilience Index (page 16), in Vulnerability Index (page 17)
 - Graphical representation of variables included in composite indices (prior the Varimax rotation): Resilience Index (page 18), Vulnerability Index (page 19)
3. Case 3:
 - Bivariate analysis – correlation matrix (Case 3): Pages 20-21
 - Included data in Resilience Index (page 22), in Vulnerability Index (page 23)
 - Variance Inflation Factor: page 24
 - Graphical representation of variables included in composite indices (prior the Varimax rotation): Resilience Index (page 25), Vulnerability Index (page 26)
4. Comparison of the three cases:
 - Included data in “Economics”, “Environment”, “Health” dimensions: page 27
 - Included data in “Policy”, “Services”, “Social” dimensions: page 28
 - Comparison of spatial autocorrelation: page 29

Included data in case 1 – Resilience Index

DIMENSION	DATA
ECONOMICS	Average income per consumption unit
ECONOMICS	Shannon-Weaver index on the representation of socio professional categories
ECONOMICS	Share of the labor force living and working in the same municipality
ECONOMICS	The 5-year survival rate of companies
ENVIRONMENT	Area (m ²) of natural spaces per inhabitant
ENVIRONMENT	Share of population with access to green space within their 400m of their homes
HEALTH	Potential spatial accessibility to general practitioners
HEALTH	Shannon-Weaver index of the representation of medical and paramedical professions
POLICY	Amount of subsidies paid by municipalities, per inhabitant
POLICY	Eligible areas for regional aid zoning
POLICY	Expenditure for municipal equipment, per inhabitant
POLICY	Financial potential per inhabitant
POLICY	Housing tax proceeds compared to the number of primary and secondary residences
POLICY	Municipality covered by an Agenda 21
POLICY	Share of taxable tax households
POLICY	Share of expenditure allocated to municipal facilities and services compared to the total investment
SERVICES	Part of area with 4G cover provided by the most efficient operator
SERVICES	Share of population with average access to 12-mid-range municipal facilities and services in 15 minutes or less
SOCIAL FACTORS	Employment in education and training, per 1,000 inhabitants
SOCIAL FACTORS	Number of state schools, per 1,000 inhabitants
SOCIAL FACTORS	Study continuation rate for 18-24 years old

Included data in case 1 – Vulnerability Index

DIMENSION	DATA
ECONOMICS	Precarious employment rate
ECONOMICS	Unemployment rate
ENVIRONMENT	Greenhouse gas emissions
ENVIRONMENT	Number of contaminated sites (polluted sites and soils)
HEALTH	Proportion of deaths
POLICY	Proportion of medium- and long-term debt of the municipality, per inhabitant
SERVICES	Length of primary roads per capita
SERVICES	Share of population who are tenants of their main residence
SOCIAL FACTORS	Population density
SOCIAL FACTORS	Share of immigrants in the total population
SOCIAL FACTORS	Share of single-parent families
SOCIAL FACTORS	Share of the population aged over 65 years
SOCIAL FACTORS	Share of working-age population with few or no qualifications

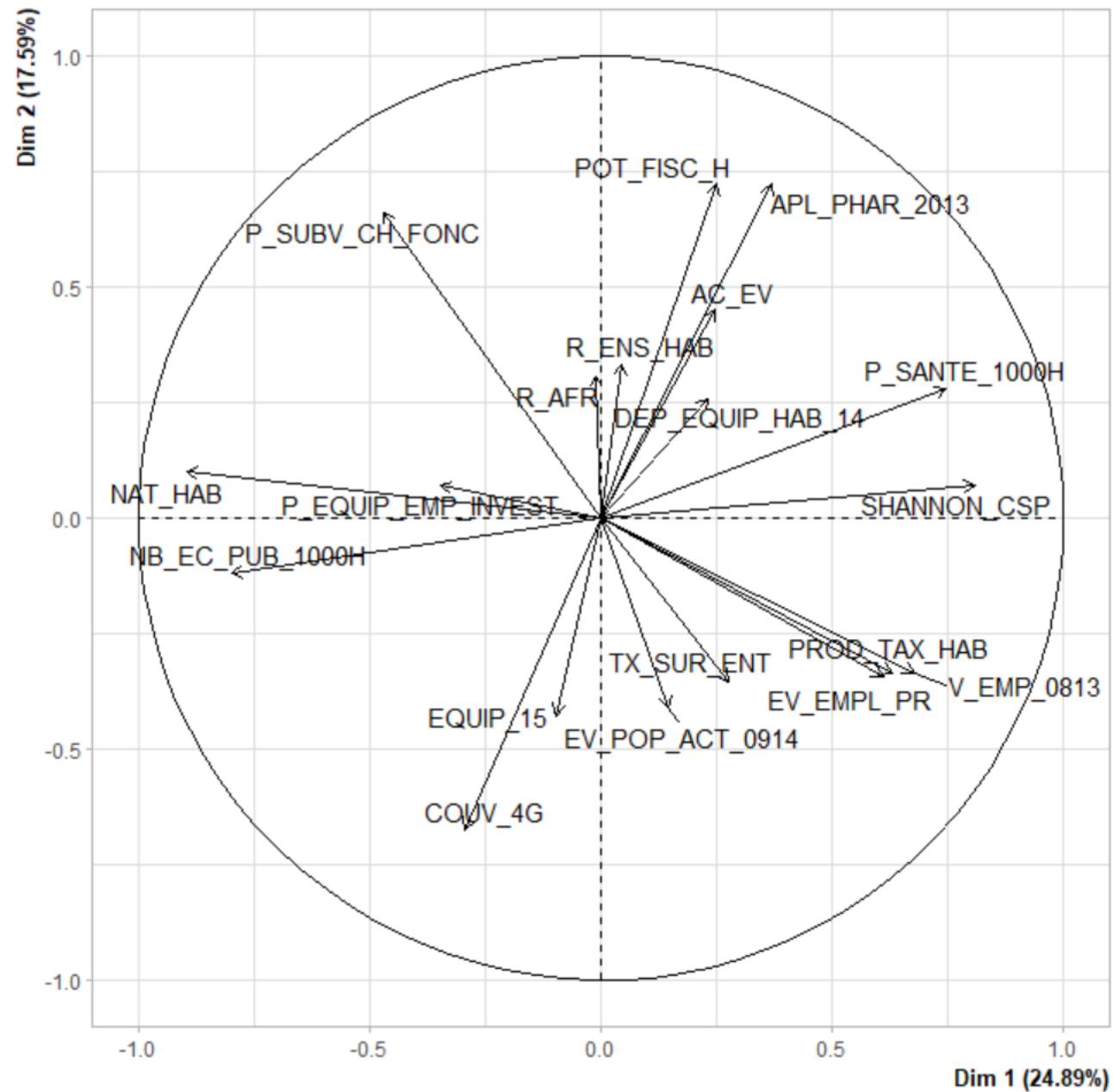
Included data in case 2 – Resilience Index

DIMENSION	DATA
ECONOMICS	Employment evolution for attendance activities
ECONOMICS	Evolution of the labor force
ECONOMICS	Shannon-Weaver index on the representation of socio professional categories
ECONOMICS	The 5-year survival rate of companies
ECONOMICS	Evolution of the share of the labour force in the total population
ENVIRONMENT	Area (m ²) of natural spaces per inhabitant
ENVIRONMENT	Share of population with access to green space within their 400m of their homes
HEALTH	Number of health professionals, per 1,000 inhabitants
HEALTH	Potential spatial accessibility to pharmacist
POLICY	Expenditure for municipal equipment, per inhabitant
POLICY	Share of expenditure allocated to municipal facilities and services compared to the total investment
POLICY	Housing tax proceeds compared to the number of primary and secondary residences
POLICY	Overall financial result of the municipality, per inhabitant
POLICY	Eligible areas for regional aid zoning
POLICY	Share of subsidies paid within operating expenses
POLICY	Tax potential per inhabitant
SERVICES	Part of area with 4G cover provided by the most efficient operator
SERVICES	Share of population with average access to 12-mid-range municipal facilities and services in 15 minutes or less
SOCIAL FACTORS	Number of state schools, per 1,000 inhabitants

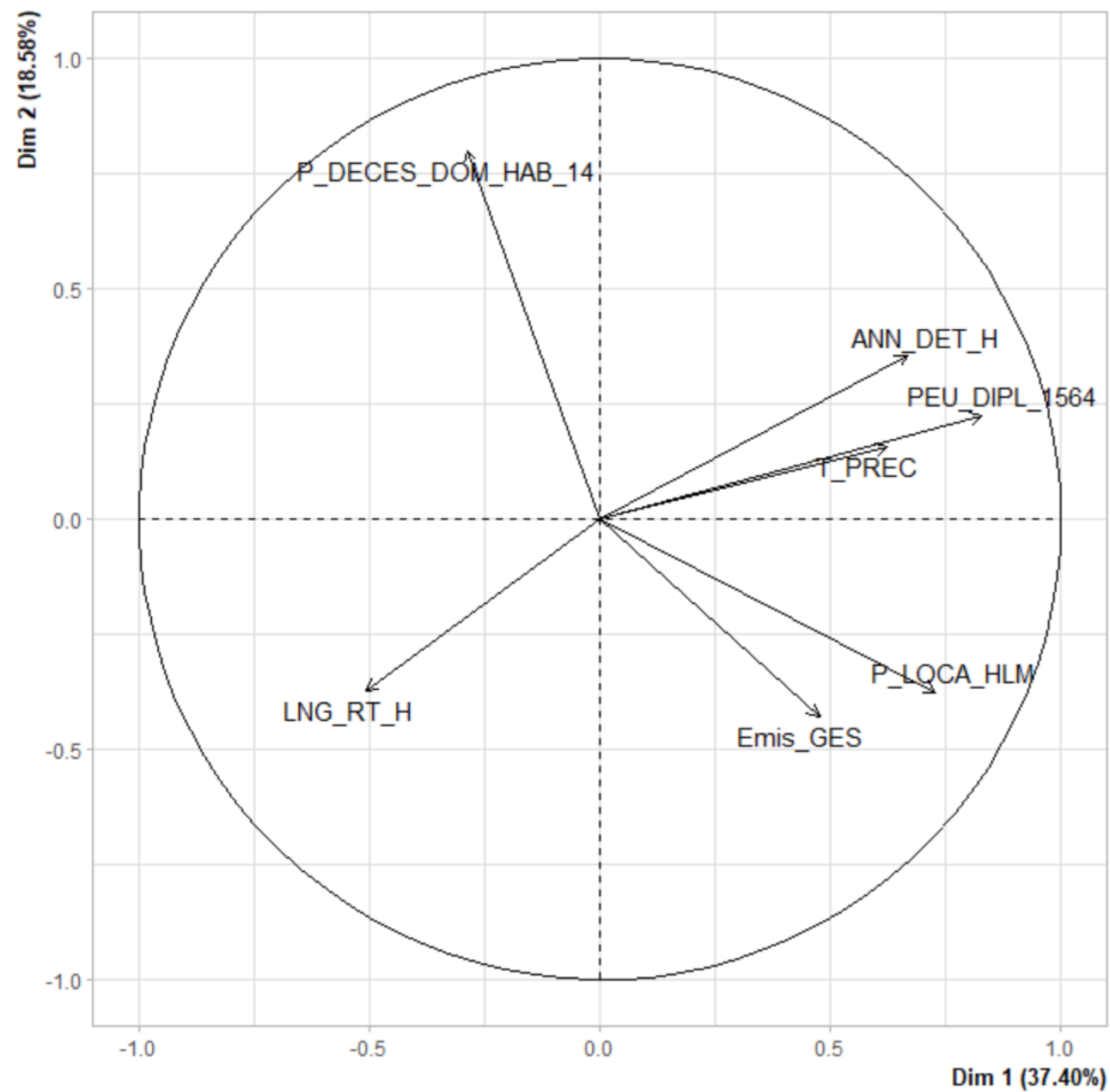
Included data in case 2 – Vulnerability Index

DIMENSION	DATA
ECONOMICS	Precarious employment rate
ENVIRONMENT	Greenhouse gas emissions
HEALTH	Proportion of deaths
POLICY	Proportion of medium- and long-term debt of the municipality, per inhabitant
SERVICES	Length of primary roads per capita
SERVICES	Share of tenants who rent their main residence in Low-Rent Housing
SOCIAL FACTORS	Share of working-age population with few or no qualifications

Graphical representation of variables included in Resilience Index (prior the Varimax rotation) (Case 2)



Graphical representation of variables included in Vulnerability Index (prior the Varimax rotation) (Case 2)



Included data in case 3 – Resilience Index

DIMENSION	DATA
ECONOMICS	Average income per consumption unit
ECONOMICS	Employment in education and training, per 1,000 inhabitants
ECONOMICS	Shannon-Weaver index on the representation of socio professional categories
ECONOMICS	Share of the labor force living and working in the same municipality
ECONOMICS	The 5-year survival rate of companies
ENVIRONMENT	Area (m ²) of natural spaces per inhabitant
ENVIRONMENT	Share of population with access to green space within their 400m of their homes
HEALTH	Potential spatial accessibility to general practitioners
HEALTH	Shannon-Weaver index of the representation of medical and paramedical professions
POLICY	Amount of subsidies paid by municipalities, per inhabitant
POLICY	Eligible areas for regional aid zoning
POLICY	Expenditure for municipal equipment, per inhabitant
POLICY	Financial potential per inhabitant
POLICY	Housing tax proceeds compared to the number of primary and secondary residences
POLICY	Municipality covered by an Agenda 21
SERVICES	Part of area with 4G cover provided by the most efficient operator
SERVICES	Share of population with average access to 12-mid-range municipal facilities and services in 15 minutes or less
SOCIAL FACTORS	Number of state schools, per 1,000 inhabitants
SOCIAL FACTORS	Study continuation rate for 18-24 years old

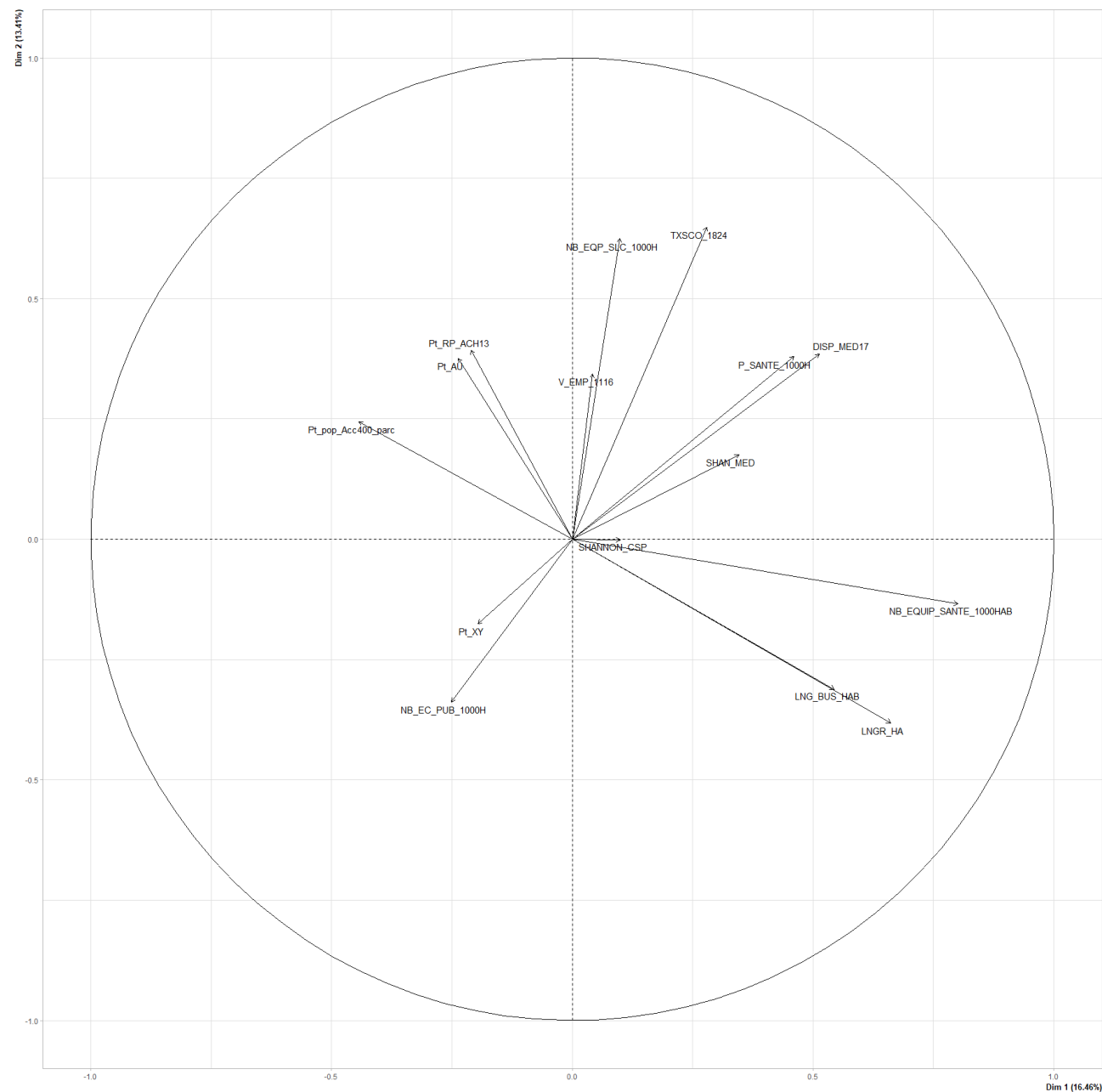
Included data in case 3 – Vulnerability Index

DIMENSION	DATA
ECONOMICS	Precarious employment rate
ECONOMICS	Unemployment rate
ENVIRONMENT	Greenhouse gas emissions
ENVIRONMENT	Number of contaminated sites (polluted sites and soils)
HEALTH	Proportion of deaths
POLICY	Proportion of medium- and long-term debt of the municipality, per inhabitant
POLICY	Share of expenditure allocated to municipal facilities and services compared to the total investment
POLICY	Share of taxable households
SERVICES	Length of primary roads per capita
SERVICES	Share of population who are tenants of their main residence
SOCIAL FACTORS	Population density
SOCIAL FACTORS	Share of immigrants in the total population
SOCIAL FACTORS	Share of single-parent families
SOCIAL FACTORS	Share of the population aged over 65 years
SOCIAL FACTORS	Share of working-age population with few or no qualifications

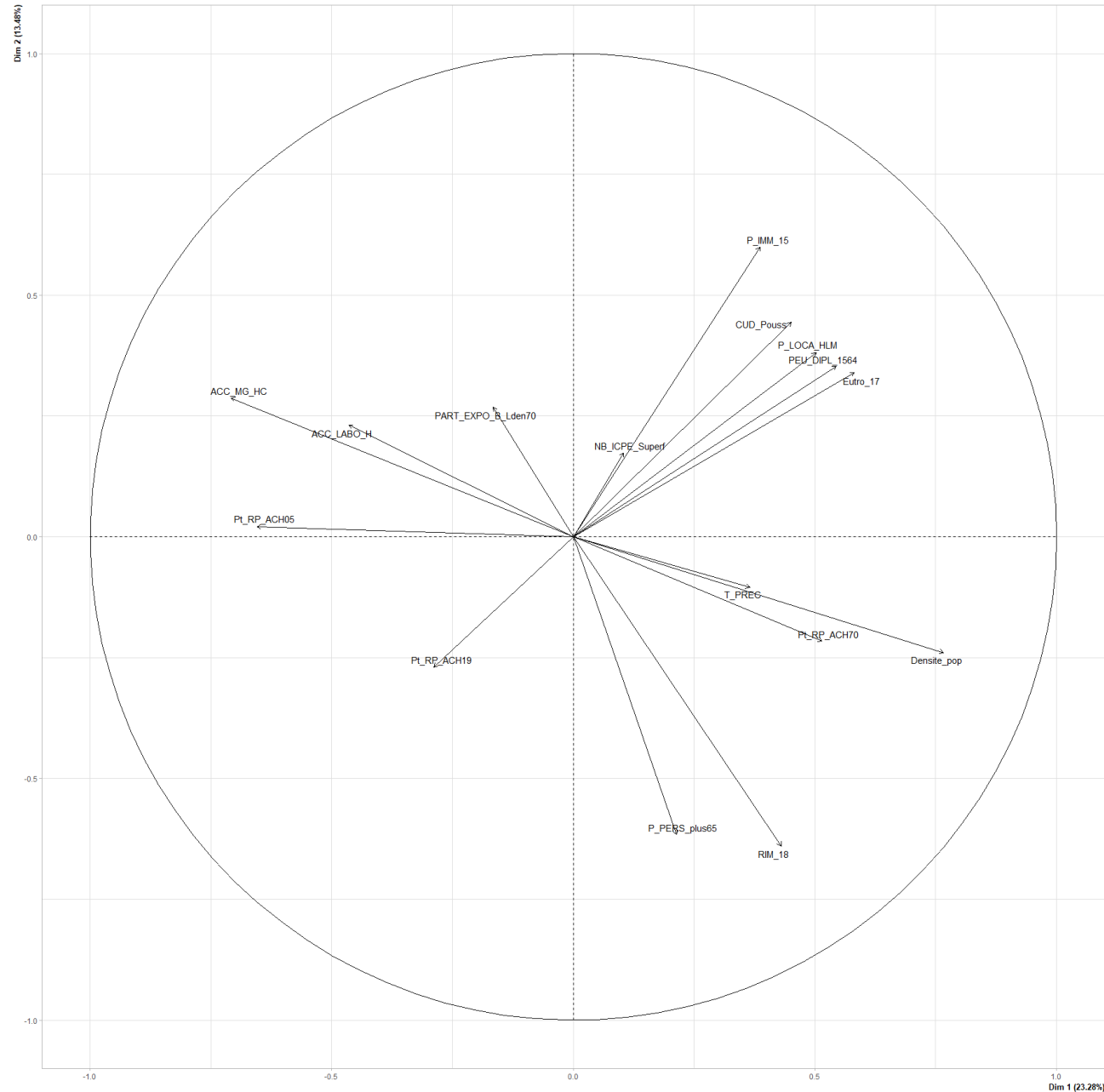
Variation Inflation Factor (Case 3)

ACRONYM OF THE VARIABLE	R ²	VIF
V_EMP_1116	0.3026	1.43
T_PREC	0.4688	1.88
DISP_MED17	0.5916	2.45
SHANNON_CSP	0.2033	1.26
Pt_AU	0.2709	1.37
Pt_XY	0.4876	1.95
TXSCO_1824	0.6632	2.97
PEU_DIPL_1564	0.8116	5.31
Densite_pop	0.7794	4.53
P_PERS_plus65	0.7068	3.41
P_IMM_15	0.6265	2.68
P_LOCA_HLM	0.7002	3.34
Pt_RP_ACH19	0.6469	2.83
Pt_RP_ACH70	0.3443	1.53
Pt_RP_ACH05	0.6026	2.52
Pt_RP_ACH13	0.4934	1.97
NB_EQP_SLC_1000H	0.4336	1.77
LNG_BUS_HAB	0.4241	1.74
P_SANTE_1000H	0.7124	3.48
SHAN_MED	0.5282	2.12
NB_EQUIP_SANTE_1000HAB	0.6557	2.90
ACC_LABO_H	0.5273	2.12
ACC_MG_HC	0.6898	3.22
Eutro_17	0.5748	2.35
PART_EXPO_B_Lden70	0.4635	1.86
Pt_pop_Acc400_parc	0.4493	1.82
NB_ICPE_Superf	0.4729	1.90
CUD_Pouss	0.5832	2.40
LNGR_HA	0.6512	2.87

Graphical representation of variables included in Resilience Index (prior the Varimax rotation) (Case 3)



Graphical representation of variables included in Vulnerability Index (prior the Varimax rotation) (Case 3)



Comparison of included data for each of the three cases

DIMENSION	INDEX	DATA	CASE 1	CASE 2	CASE 3
ECONOMICS	RESILIENCE	Average income per consumption unit	X		X
ECONOMICS	RESILIENCE	Employment evolution for attendance activities		X	
ECONOMICS	RESILIENCE	Evolution of the labor force		X	X
ECONOMICS	VULNERABILITY	Precarious employment rate	X	X	X
ECONOMICS	RESILIENCE	Shannon-Weaver index on the representation of socio professional categories	X	X	X
ECONOMICS	RESILIENCE	Share of the labor force living and working in the same municipality	X		
ECONOMICS	RESILIENCE	The 5-year survival rate of companies	X	X	
ECONOMICS	VULNERABILITY	Unemployment rate	X		
ENVIRONMENT	RESILIENCE	Area (m ²) of natural spaces per inhabitant	X	X	
ENVIRONMENT	VULNERABILITY	Dust deposition			X
ENVIRONMENT	VULNERABILITY	Eutrophication			X
ENVIRONMENT	VULNERABILITY	Greenhouse gas emissions	X	X	
ENVIRONMENT	RESILIENCE	Length of greenway network and bicycle paths per ha			X
ENVIRONMENT	VULNERABILITY	Mean Impregnation Ratio			X
ENVIRONMENT	VULNERABILITY	Number of contaminated sites (polluted sites and soils)	X		
ENVIRONMENT	VULNERABILITY	Number of facilities classified for environmental protection per km ²			X
ENVIRONMENT	VULNERABILITY	Share of population exposed to noise throughout the day			X
ENVIRONMENT	RESILIENCE	Share of population with access to green space within their 400m of their homes	X	X	X
HEALTH	VULNERABILITY	Accessibility to general practitioners			X
HEALTH	VULNERABILITY	Accessibility to medical biology laboratory			X
HEALTH	RESILIENCE	Number of health equipment per 1,000 inhabitants			X
HEALTH	RESILIENCE	Number of health professionals, per 1,000 inhabitants		X	X
HEALTH	RESILIENCE	Potential spatial accessibility to general practitioners	X		
HEALTH	RESILIENCE	Potential spatial accessibility to pharmacist		X	
HEALTH	VULNERABILITY	Proportion of deaths	X	X	
HEALTH	RESILIENCE	Shannon-Weaver index of the representation of medical and paramedical professions	X		X

Comparison of included data for each of the three cases

DIMENSION	INDEX	DATA	CASE 1	CASE 2	CASE 3
POLICY	RESILIENCE	Amount of subsidies paid by municipalities, per inhabitant	X		
POLICY	RESILIENCE	Eligible areas for regional aid zoning	X		
POLICY	RESILIENCE	Expenditure for municipal equipment, per inhabitant	X	X	
POLICY	RESILIENCE	Financial potential per inhabitant	X		
POLICY	RESILIENCE	Housing tax proceeds compared to the number of primary and secondary residences	X	X	
POLICY	RESILIENCE	Municipality covered by an Agenda 21	X		
POLICY	VULNERABILITY	Proportion of medium- and long-term debt of the municipality, per inhabitant	X	X	
POLICY	RESILIENCE	Share of areas to be urbanized			X
POLICY	VULNERABILITY	Share of expenditure allocated to municipal facilities and services compared to the total investment	X	X	
POLICY	RESILIENCE	Share of mixed urbanized areas			X
POLICY	RESILIENCE	Share of subsidies paid within operating expenses		X	
POLICY	VULNERABILITY	Share of taxable tax households	X		
POLICY	VULNERABILITY	Tax potential per inhabitant		X	
SERVICES	RESILIENCE	Length of bus line per inhabitant			X
SERVICES	VULNERABILITY	Length of primary roads per capita	X	X	
SERVICES	RESILIENCE	Number of sports clubs, associations and cultural facilities per 1,000 inhabitants			X
SERVICES	RESILIENCE	Part of area with 4G cover provided by the most efficient operator	X	X	
SERVICES	VULNERABILITY	Share of main residences built before 1919			X
SERVICES	VULNERABILITY	Share of main residences built between 1946-1970			X
SERVICES	VULNERABILITY	Share of main residences built between 1991-2005			X
SERVICES	RESILIENCE	Share of main residences built between 2006-2013			X
SERVICES	VULNERABILITY	Share of population who are tenants of their main residence	X	X	
SERVICES	RESILIENCE	Share of population with average access to 12-mid-range municipal facilities and services in 15 minutes or less	X	X	
SERVICES	VULNERABILITY	Share of tenants who rent their main residence in Low-Rent Housing (LRH)			X
SOCIAL FACTORS	RESILIENCE	Number of state schools, per 1,000 inhabitants	X	X	X
SOCIAL FACTORS	RESILIENCE	Employment in education and training, per 1,000 inhabitants	X		
SOCIAL FACTORS	VULNERABILITY	Population density	X		X
SOCIAL FACTORS	VULNERABILITY	Share of immigrants in the total population	X		X
SOCIAL FACTORS	VULNERABILITY	Share of single-parent families	X		
SOCIAL FACTORS	VULNERABILITY	Share of the population aged over 65 years	X		X
SOCIAL FACTORS	VULNERABILITY	Share of working-age population with few or no qualifications	X	X	X
SOCIAL FACTORS	RESILIENCE	Study continuation rate for 18-24 years old	X		X

Comparison of spatial autocorrelation

	Index	Global Moran's I statistic	p-value
Case 1	RI-1	-0.030	$p > 0.05$
	VI-1	0.312	$p < 0.05$
Case 2	RI-2	-0.313	$p > 0.05$
	VI-2	0.193	$p > 0.05$
Case 3	RI-3	-0.007	$p > 0.05$
	VI-3	0.115	$p < 0.001$

Annexe 3

Liste des centres de dialyse (Hauts-de-France)

REGION	COMMUNE	LIBELLE	CODE ETABLISSEMENT
NORD PAS DE CALAIS	ARRAS	CH ARRAS	E62ARRA009041
	BETHUNE	CH BETHUNE	E62BETH007119
	BOIS-BERNARD	HOPITAL PRIVE DE BOIS BERNARD	E62BOIS001148
	BOULOGNE SUR MER	CH BOULOGNE SUR MER	E62BOUL012160
	CALAIS	CH CALAIS	E62CALA009193
	CAMBRAI	CH CAMBRAI	E59CAMB008122
	DECHY	CH DOUAI DECHY	E59DECH001170
	DUNKERQUE	CH DUNKERQUE	E59DUNK012183
	FOURMIES	CH FOURMIES	E59FOUR003249
	HELFAUT	NEPHROCARE HELFAUT	E62HELFO02423
	LENS	CH LENS	E62LENS004498
	LILLE	HOP CLAUDE HURIEZ CHR LILLE	E59LILL030350
		HOPITAL PRIVE LE BOIS	E59LILL037350
		HOPITAL PRIVE LA LOUVIERE	E59LILL039350
		HOP JEANNE DE FLANDRE CHR LILLE	E59LILL025350
	MAUBEUGE	CH SAMBRE AVESNOIS MAUBEUGE	E59MAUB005392
		NEPHROCARE MAUBEUGE	E59MAUB008392
	RANG-DU-FLIERS	CH ARRONDISSEMENT DE MONTREUIL	E62RANG002688
ROUBAIX	HOPITAL VICTOR PROVO	E59ROUB015512	
VALENCIENNES	POLYCLINIQUE VAUBAN	E59VALE017606	
	CH VALENCIENNES	E59VALE011606	
PICARDIE	ABBEVILLE	SA CLINIQUE SAINTE ISABELLE	E80ABBE010001
	BEAUVAIS	CH BEAUVAIS	E60BEAU007057
	COMPIEGNE	POLYCLINIQUE SAINT COME	E60COMP012159
	CREIL	CH GHPSO CREIL	E60CREI007175
	LAON	CH LAON	E02LAON004408
	SAINT-QUENTIN	CH SAINT QUENTIN	E02SAIN012691
	SALOUEL	CHU AMIENS SALOUEL	E80AMIE017021
	SOISSONS	CH SOISSONS	E02SOIS007722

Annexe 4

Modèles épidémiologiques développés pour déterminer la variabilité résiduelle de l'IRCT

We describe the ESRD cases counts O_i in each spatial unit i as a Poisson random variable

$$O_i \sim \text{Poisson}(\theta_i E_i), i = 1, \dots, N$$

where N is the total number of communes in northern of France, E_i is the expected number of ESRD cases calculated using an indirect standardization on age and gender, and θ_i the relative risk associated with neighborhood i .

We consider the following five log-linear models hierarchical Bayesian model described by Besag et al. (1991) with re-parametrisation of the Besag-York-Mollié conditional autoregressive prior distribution described by Simpson et al. (Simpson et al., 2017):

Model 1: adjusted for age and sex

$$\log(\theta_i) = \beta_0 + b_i$$

Model 2: adjusted for age, sex, and health confounders (eGFR and Time travel (TT))

$$\log(\theta_i) = \beta_0 + \beta_1 eGFR_i + \beta_2 TT_i + b_i$$

Model 3: adjusted for age, sex, health confounders and FDep index

$$\log(\theta_i) = \beta_0 + \beta_1 eGFR_i + \beta_2 TT_i + \beta_3 FDep_i + b_i$$

Model 4: with resilience index (RI) and adjusted for age, sex, and confounders (health confounders and FDep index)

$$\log(\theta_i) = \beta_0 + \beta_1 eGFR_i + \beta_2 TT_i + \beta_3 FDep_i + \beta_4 RI_i + b_i$$

Model 5: with vulnerability index (VI) and adjusted for age, sex, and all confounders

$$\log(\theta_i) = \beta_0 + \beta_1 eGFR_i + \beta_2 TT_i + \beta_3 FDep_i + \beta_4 VI_i + b_i$$

Where β_0 corresponds to the overall level of relative risk across the northern of France and β_j the corresponding effect of the j -th covariate. b_i is a random effect accounting for spatial auto-correlation. In line with references Simpson et al. (2017) and Riebler et al. (2016), b_i is defined as follows:

$$b_i = \frac{1}{\tau} (\sqrt{1 - \varphi} v_i + \sqrt{\varphi} u_i)$$

where v_i is a i.i.d. random effect, u_i is a spatially structured random effect, τ is the marginal precision, and φ is a mixing parameter (corresponding to the proportion of marginal variance explained by the spatial random effect u_i).

The β_j values have a non-informative normal prior distribution $N(0,100)$, as suggested by Lawson (2013). The model was fitted using an integrated nested Laplace approximation (Blangiardo & Cameletti, 2015).

To evaluate the proportion of the spatial variation of the ESRD incidence rate explained after adding new covariate, we consider the following ratio:

$$R = (\sigma_b - \sigma_a) / \sigma_b$$

where σ_b and σ_a are standard deviation of the random effect b_i before and after adding new covariate, respectively. It should be noted that by the definition of the random effect b_i , the standard deviation b_i is $\frac{1}{\tau}$.

Besag, J., York, J., & Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*, 43(1), 1-20. <https://doi.org/10.1007/BF00116466>

Blangiardo, M., & Cameletti, M. (2015). *Spatial and Spatio-temporal Bayesian Models with R - INLA* (1er édition). Wiley.

Riebler, A., Sørbye, S. H., Simpson, D., & Rue, H. (2016). An intuitive Bayesian spatial model for disease mapping that accounts for scaling. *Statistical Methods in Medical Research*, 25(4), 1145-1165. <https://doi.org/10.1177/0962280216660421>

Simpson, D., Rue, H., Riebler, A., Martins, T. G., & Sørbye, S. H. (2017). Penalising Model Component Complexity : A Principled, Practical Approach to Constructing Priors. *Statistical Science*, 32(1), 1-28. <https://doi.org/10.1214/16-STS576>

Annexe 5

Supplementary materials and results – Article 5

Supplementary materials

1. Descriptive measures for the population and cover areas (Table SM-I)
 1. Page 2
2. List of official national organizations from which data were collected
 - Page 3
3. Meta-data concerning dataset used to develop composite indices at the IRIS scale (IRIS: the smallest aggregated unit for spatial information):
 - Page 4: meta-data related to “Economics” (Table SM-II)
 - Page 5-6: meta-data related to “Environment” (Table SM-III)
 - Page 7: meta-data related to “Health” (Table SM-IV)
 - Page 8: meta-data related to “Social factors” (Table SM-V)
 - Page 9-10: meta-data related to “Services” (Table SM-VI)

Descriptive measures for the population and cover areas (Table SM-I)

Studied perimeter: European Lille Metropolis			Municipalities (85)	IRIS (472)
Population (inhab.)	1,126,558	Median	4,833	2,287
		Q1 - Q3	1,986 - 13,326	1,869 - 2,837
Cover area (ha)	61,289.820	Median	543.032	42.819
		Q1 - Q3	317.527 - 1,012.778	24.832 - 96.295
Density (inhab./ha)	18.381	Median	10.554	51.484
		Q1 - Q3	4.573 - 4.127	22.562 - 88,672

List of official national organizations

Copernicus: European Union's Earth observation programme, offering information services that draw from satellite Earth Observation and *in-situ* (non-space) data

Data.gouv.fr: open, collaborative and community platform for French public data

IGN: National Institute for Geographic and forest information

INERIS: French National Institute responsible for environmental risks, under the direction of the Ministry of Environment

INSEE: French National Institute of Statistics and Economic Studies

MEDDE: Ministry of Ecological and Solidarity Transition

ODELM: Open Data portal of the European Lille Metropolis

ODS: Open Data Soft

OSM : Open Street Map

Meta-data related to “Economics” (Table SM-II)

If in final use there is “ $\rho > |0,6|$ ” or “ $|0,2| < \rho$ ”, the corresponding variable was not considered for PCA – HCPC

Data	Acronym	Sub-dimension	Specifications	Unit	Source(s)	Time period	Median [Q1;Q3]	Final use
Deprivation Index	FDEP18	Deprivation (vulnerability)	An index aggregating four socio-economic variables (median household income, percentage of high school graduates in the population aged 15 and over, percentage of blue-collar workers in the active population, unemployment rate)	Index	ODS	2018	-0.04 [-1.18;1.24]	PCA – CAH
Proportion of poor households	P15_PMEN_PAU	Deprivation (vulnerability)	Ratio between the number of poor households and the total number of households. The poverty line is set at 60% of the median standard of living	%	INSEE	2015	0.34 [0.13;11.45]	PCA – CAH
Evolution of the share of the labor force in the total population	P1217_EV_PARACT	Economic drive (resilience)	Calculation of the evolution of the share of the labor force aged 15 and over within the total population between 2 non-consecutive years	%	INSEE	2012-2017	0.08 [-1.49;1.97]	PCA – CAH
Evolution of the labor force in the total population	P1217_EVPA	Economic drive (resilience)	Calculation of the evolution of the labor force aged 15 and between 2 non-consecutive years	%	INSEE	2012-2017	0.06 [-6.55;8.06]	$\rho > 0,6 $
The share of the labor force in the total population	P17_PARACT	Employability (resilience)	The share of the labor force aged 15 and over within the total population	%	INSEE	2017	46.86 [43.01;49.60]	PCA – CAH
Precarious employment rate	P17_EMPLAI	Deprivation (vulnerability)	Share of the labor force aged 15-64 in fixed-term contracts, temporary employment, assisted employment and apprenticeships among the labour force aged 15-64	%	INSEE	2016	11.40 [8.90;14.92]	$\rho > 0,6 $
Shannon-Weaver index on the representation of socio-professional categories	SHAN_CSP_17	Economic diversity (resilience)	Use of the Shannon-Weaver index to describe a population in terms of the number of categories (category richness) and their relative abundance within those categories (species equitability). 6 socio-professional categories were considered for the calculation, identified at the place of residence	Index	INSEE	2017	0.62 [0.59;0.64]	PCA – CAH
Share of the labor force living and working in the same municipality	P17_OCCRES	Economic influence (resilience)	The labor force considered is aged 15 and over	%	INSEE	2017	23.42 [16.40;39.51]	PCA – CAH

Meta-data related to “Environment” (Table SM-III)

If in final use there is “ $\rho > |0,6|$ ” or “ $|0,2| < \rho$ ”, the corresponding variable was not considered for PCA – HCPC

Data	Acronym	Sub-dimension	Specifications	Unit	Source(s)	Time period	Median [Q1;Q3]	Final use
Share of population with access to green space within 400meters of their homes	AC_EV_EB_F	Natural resources (resilience)	Green spaces of less than 0.5ha have been excluded. The criterion included parks, green spaces, wooded areas, forests. The population data are those of the INSEE’s 200-meter grid	%	ODELM INSEE	2015	79.07 [42.52;100]	PCA – CAH
Monetisation of ecosystem services per capita	H_ESV_BRUTE	Natural resources (resilience)	Use of the Corine Land Cover database to determine 4 land-use categories (grassland, urban, cultivated and/or arable land, forests) and The Economics of Ecosystems and Biodiversity valuation database to assign a monetary value to the ecosystem services provided by these 4 categories	USD/ha/year	INSEE MEDDE	2016-2018	79.55 [44.61; 137.87]	PCA – CAH
Mean Impregnation Ratio	L_RIM	Air quality (vulnerability)	Biomonitoring with lichen was used. The bioaccumulation of 18 trace elements (Al, Sb, As, Be, Cd, Co, Cr, Cu, Mn, Hg, Ni, Pb, Pd, Pt, Rh, Ti, V, Zn) was determined in the thalli of the foliose lichen Xanthoria parietina. The multi-metallic burden is generated by the calculation of a composite index	Index	Data from our laboratory	2009	1.71 [1.43;2.01]	PCA – CAH
NO ₂	NO2_COMPLETE	Air quality (vulnerability)	Annual mean concentrations. Data from Atmospheric Dispersion Modelling System, completed by data from INERIS in case of missing data	µg/m ³	INERIS	2009	33.16 [29.18;37.20]	PCA – CAH
Particulate matters (PM ₁₀)	PM10_COMPLETE	Air quality (vulnerability)	Annual mean concentrations. Data from Atmospheric Dispersion Modelling System, completed by data from INERIS in case of missing data	µg/m ³	INERIS	2009	33.12 [31.28;35.35]	$\rho > 0,6 $
Particulate matters (PM _{2,5})	MOY_PM25_2009_2016	Air quality (vulnerability)	Annual mean concentrations. PM _{2,5} data were not available before year 2009	µg/m ³	INERIS	2009-2016	17.69 [17.29;17.86]	$\rho > 0,6 $
O ₃	MOY_O3_2008_2016	Air quality (vulnerability)	Annual mean concentrations	µg/m ³	INERIS	2008-2016	40.30 [39.42;41.66]	$\rho > 0,6 $
Normalized Difference Vegetation Index	NDVI_2017	Natural resources (resilience)	The NDVI raster provides an indicator of the greenness of the biomes at 3003 spatial resolution. Even though it is not a physical property of the vegetation cover, its very simple formulation $NDVI = (REF_nir - REF_red)/(REF_nir + REF_red)$ where REF_nir and REF_red are the spectral reflectances measured in the near infrared and red wavebands respectively, makes it widely used for ecosystems monitoring	Index	Copernicus	2017	0.43 [0.36;0.49]	PCA – CAH

Meta-data related to “Environment” (Table SM-III, final section)

If in final use there is “ $\rho > |0,6|$ ” or “ $|0,2| < \rho$ ”, the corresponding variable was not considered for PCA – HCPC

Data	Acronym	Sub-dimension	Specifications	Unit	Source(s)	Time period	Median [Q1;Q3]	Final use
Share of population exposed to noise throughout the day	P_EXP_60DB_LDEN	Environmental risks (vulnerability)	The L _{DEN} is the noise indicator for overall annoyance which was used (level day evening night). The considered sources of noise were roads, railways, aircraft and industrial facilities. Only areas with levels above 60 dB(A) were kept for the calculation. The population data are those of the INSEE’s 200-meter grid	%	ODELM INSEE	2018 2015	21.73 [10.63;40.07]	PCA – CAH
Share of population exposed to noise at night	P_POP_LN45	Environmental risks (vulnerability)	The L _N is the noise indicator for overall annoyance which was used (level night, between 22:00 and 6:00). The considered sources of noise were roads, railways, aircraft and industrial facilities. Only areas with levels above 45 dB(A) were kept for the calculation. The population data are those of the INSEE’s 200-meter grid	%	ODELM INSEE	2018 2015	62.88 [42.19;82.07]	PCA – CAH
Dust deposition	Poussieres	Air quality (vulnerability)	Biomonitoring with poplars was used. Dust deposition was quantified by measuring the accumulation of total dust on poplar leaves during the flowering season. Dust quantity was determined in relation to leaf mass	µg/g	Data from our laboratory	2011	1,959.11 [1,814.88; 2,090.68]	PCA – CAH
Tree Cover Density	TCD	Natural resources (resilience)	The TCD raster product provides information on the proportional crown coverage per pixel at 10m spatial resolution and ranges from 0% (all non-tree covered areas) to 100%, whereby Tree Cover Density is defined as the „vertical projection of tree crowns to a horizontal earth’s surface“	%	Copernicus	2018	7.03 [4.39;11.02]	$\rho > 0,6 $

Meta-data related to “Health” (Table SM-IV)

If in final use there is “ $\rho > |0,6|$ ” or “ $|0,2| < \rho$ ”, the corresponding variable was not considered for PCA – HCPC

Data	Acronym	Sub-dimension	Specifications	Unit	Source(s)	Time period	Median [Q1;Q3]	Final use
Travel time to general practitioners	ACC_MG_HCHP	Health care facilities (vulnerability)	Theoretical calculation performed in off-peak hours, between the nearest general practitioners and 200-metre gridded data of the French National Institute of Economics and Studies	Minute	INSEE	2017	1.6 [1.31;2.00]	PCA – CAH
Travel time to emergency services	ACCES_URG	Health care facilities (vulnerability)	Mean of theoretical calculation performed in off-peak hours and peak hours, between the nearest emergency service and 200-metre gridded data of the French National Institute of Economics and Studies	Minute	INSEE	2017	9.58 [6.49;12.02]	PCA – CAH
Distance to emergency services	URG_KM	Health care facilities (vulnerability)	Theoretical distance between the nearest general practitioners and 200-metre gridded data of the French National Institute of Economics and Studies	Km	INSEE	2017	3.58 [2.19;5.39]	$\rho > 0,6 $
Number of health professionals per 1,000 inhabitants	P16_PSAN/HAB	Health care facilities (resilience)	Number of health professions, all specialities combined, counted at the IRIS-code of the place of activity, per 1,000 inhabitants	Number per 1,000 inhabitants	INSEE	2016	5.72 [2.46;10.67]	$\rho > 0,6 $
Shannon-Weaver index on the representation of medical and paramedical professions	SHAN_MED_19	Health care facilities (resilience)	Use of the Shannon-Weaver index to describe a population in terms of the number of categories (category richness) and their relative abundance within those categories (category equitability). 25 medical and paramedical professions were considered for the calculation. Information provided at the professional activity location	Index	INSEE	2019	0.62 [0.35;0.77]	PCA – CAH

Meta-data related to “Social factors” (Table SM-V)

If in final use there is “ $\rho > |0,6|$ ” or “ $|0,2| < \rho$ ”, the corresponding variable was not considered for PCA – HCPC

Data	Acronym	Sub-dimension	Specifications	Unit	Source(s)	Time period	Median [Q1;Q3]	Final use
Population density	P17_DENS	Urbanization (vulnerability)	Number of inhabitants per hectare	Inhab./ha	INSEE	2017	51.48 [22.56;88.67]	$\rho > 0,6 $
Study continuation rate for 18-24 years old	P17_PETU18-24	Education level	Share of the population aged 18 to 24 pursuing higher education	%	INSEE	2017	56.37 [46.72;66.74]	$\rho > 0,6 $
Share of working-age population with few or no qualifications	P17_PEUDIPL15-64	Education level (vulnerability)	Ratio between the number of out-of-school persons aged 15 to 64 with no diploma (or at most a BEPC, a French certificate of general education, awarded by schools at the end of the first four years of general secondary education) and the total number of out-of-school persons aged 15 to 64	%	INSEE	2017	4.55 [3.69;5.45]	PCA – CAH
Share of single-parent families	P17_PFAMONO	Social resources (vulnerability)	A single-parent family includes a lone parent and one or more unmarried (childless) children	%	INSEE	2017	11.17 [8.29;14.77]	$\rho > 0,6 $
Share of immigrants in the total population	P17_PIMM	Integration and anomie (vulnerability)	An immigrant is a person born abroad and residing in France. A distinction is made between foreign and immigrant populations: an immigrant is not necessarily a foreigner and conversely, some foreigners were born in France. Immigrant status is permanent: an individual continues to belong to the immigrant population even if he/she becomes French by acquisition	%	INSEE	2017	7.67 [4.66;12.90]	$\rho > 0,6 $
Proportion of single households	P17_PMENASEUL	Social resources (vulnerability)	Ratio between the number of single households and the total number of households	%	INSEE	2017	32.82 [26.37;40.78]	PCA – CAH
Marital status	S_MARITAL	Social resources (resilience)	Ratio between the population aged 15 and over married and the total population aged 15 and over	%	INSEE	2017	39.99 [33.75;47.09]	$\rho > 0,6 $

Meta-data related to “Services” (Table SM-VI)

If in final use there is “ $\rho > |0,6|$ ” or “ $|0,2| < \rho$ ”, the corresponding variable was not considered for PCA – HCPC

Data	Acronym	Sub-dimension	Specifications	Unit	Source(s)	Time period	Median [Q1;Q3]	Final use
Distance to sport and leisure facilities	EQUIP_SPORT_CO URT	Land use planning (vulnerability)	Theoretical distance by road between the nearest sport and leisure facilities and 200-metre gridded data of the French National Institute of Economics and Studies Exclusion of certain leisure facilities: cinema, museum, conservatory and library Exclusion of certain sport facilities: bowling alley, bowling, marina/anchorage, theater/street arts/circus center	Meters	INSEE IGN	2017	387.83 [281.80; 531.18]	PCA – CAH
Distance to fast food	FAST_FOOD_CO RT	Land use planning (resilience)	Theoretical distance by road between the nearest fast-food restaurant and 200-metre gridded data of the French National Institute of Economics and Studies Use of restaurant chains Mac Donald's, Burger King, KFC, Domino's pizza, So Good and Subway. Data includes also French fry vendor, kabob, burger, sandwiches and tacos	Meters	INSEE IGN OSM	2017	705.59 [442.43; 1,279.10]	$\rho > 0,6 $
Share of bicycle paths in the total public road	P_CYCL_VOI	Mobility (resilience)	All bicycle paths: two-way, one-way, on bus lanes, green lanes	%	Data.gouv INSEE	2017	78.95 [69.38;85.84]	PCA – CAH
Proportion of social housing	P15_PLOG_SOC	Land use planning (resilience)	The share of social housing in the total housing stock	%	INSEE	2015	19.43 [10.00;38.59]	$\rho > 0,6 $
Share of physically active commute modes	P17_DPT_ACTIFS	Mobility (resilience)	Physically active commute modes considered were walking, bike, public transport (bus, metro, tram)	%	INSEE	2017	25.04 [15.09;37.54]	PCA – CAH
Share of main residences built after 2006	P17_PRESPC06-14	Habitat characteristics (resilience)	It does not take into account the possible renovation of dwellings	%	INSEE	2017	5.81 [1.92;11.78]	PCA – CAH
Share of main residences built before 2005	P17_PRESPC19-05	Habitat characteristics (vulnerability)	It does not take into account the possible renovation of dwellings	%	INSEE	2017	92.80 [86.15;96.73]	$\rho > 0,6 $
Share of main residences built before 1970	P17_PRESPC19-70	Habitat characteristics (vulnerability)	It does not take into account the possible renovation of dwellings	%	INSEE	2016	53.74 [39.75;69.18]	PCA – CAH
Share of population who are tenants of their main residence	P17_PRESPOL	Access to the homeownership market (vulnerability)	The rental stock comprises the public and the private sectors	%	INSEE	2017	43.18 [27.69;66.73]	$\rho > 0,6 $

Meta-data related to “Services” (Table SM-VI, final section)

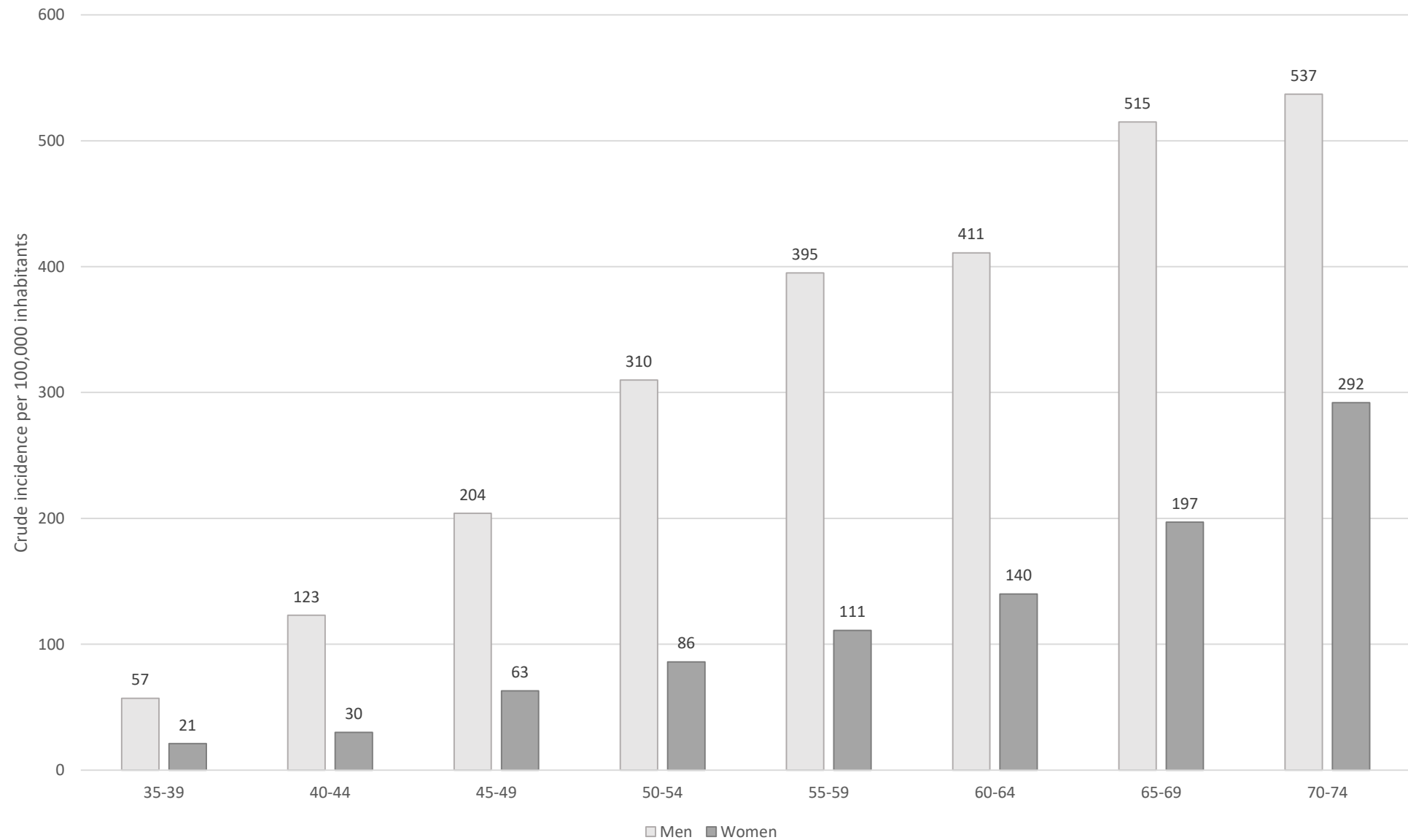
If in final use there is “ $\rho > |0,6|$ ” or “ $|0,2| < \rho$ ”, the corresponding variable was not considered for PCA – HCPC

Data	Acronym	Sub-dimension	Specifications	Unit	Source(s)	Time period	Median [Q1;Q3]	Final use
Proportion of people owning their main residence	P17_PRESPOP	Access to the homeownership market (resilience)	Ownership status applies to owner, co-owner and first-time homeowner households	%	INSEE	2017	55.31 [32.00;71.05]	$\rho > 0,6 $
Length of bicycle paths per inhabitant	P17_VOI_CY_H	Mobility (resilience)	Use of all bicycle paths (two-way, one-way, on bus lanes, green lanes)	m/inhab.	Data.gouv INSEE	2017	3.08 [2.00;5.18]	$\rho > 0,6 $
Length of public transport per inhabitant	TEC_H	Mobility (resilience)	Use of bus, metro, tram	m/inhab.	ODELM INSEE	2017	2.25 [1.07;5.13]	$\rho > 0,6 $

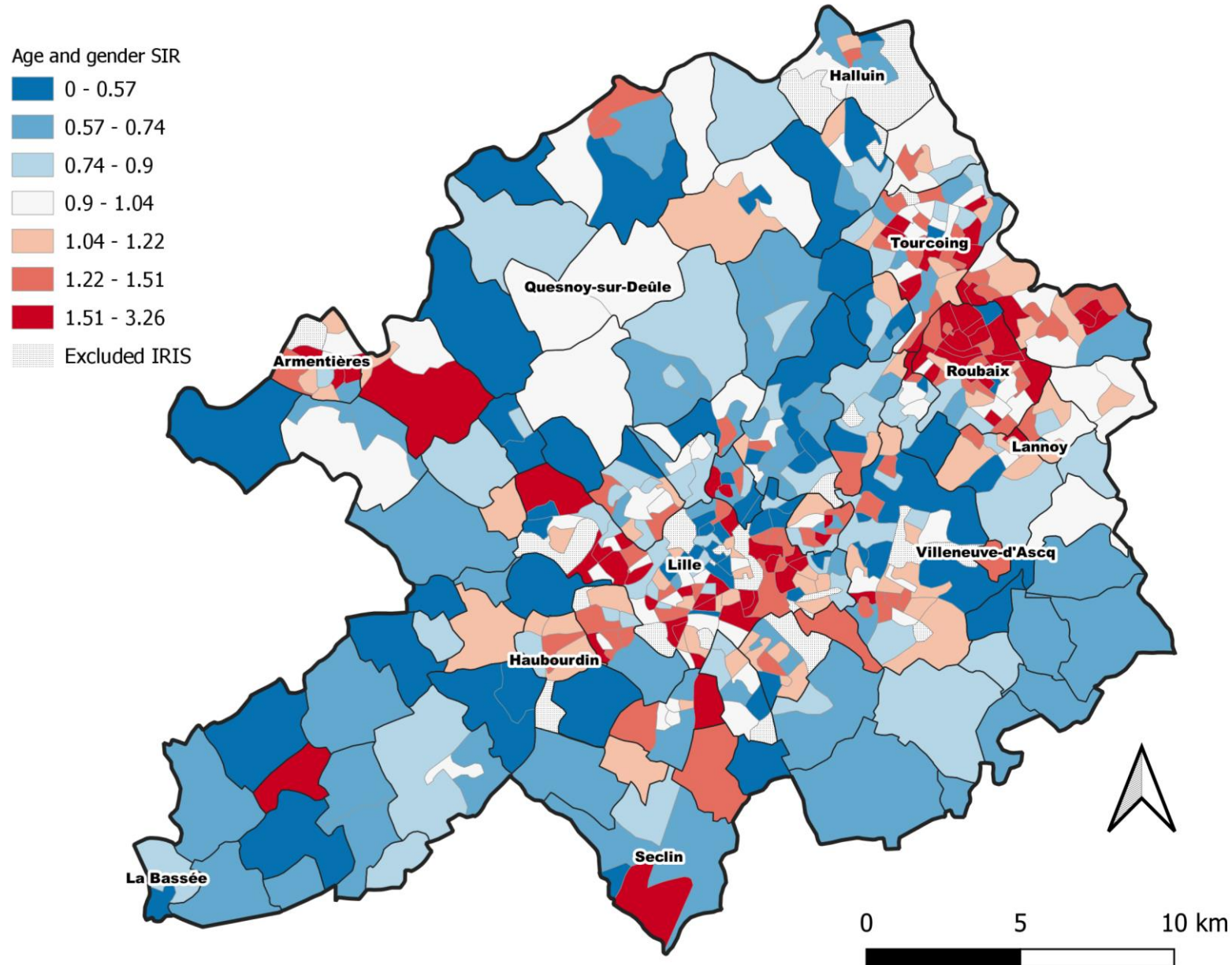
Supplementary results

1. Crude incidence of Coronary Heart Diseases (2012-2016) (Figure SM-1)
 - Page 12
2. Standardized Incidence Ratios before Bayesian Poisson regression model (Figure SM-2)
 - Page 13
3. Bivariate analysis – correlation matrix (Table SM-VII)
 - Pages 14-15
4. PCA - graph of variables (Figure SM-3)
 - Page 16
5. PCA – \cos^2 and cumulative % of variance (Table SM-VIII)
 - Page 17
6. SIRs: Tukey multiple comparisons of means (Table SM-IX, Figure SM-4)
 - Page 18
7. Population density in each class (Table SM-X)
 - Page 19

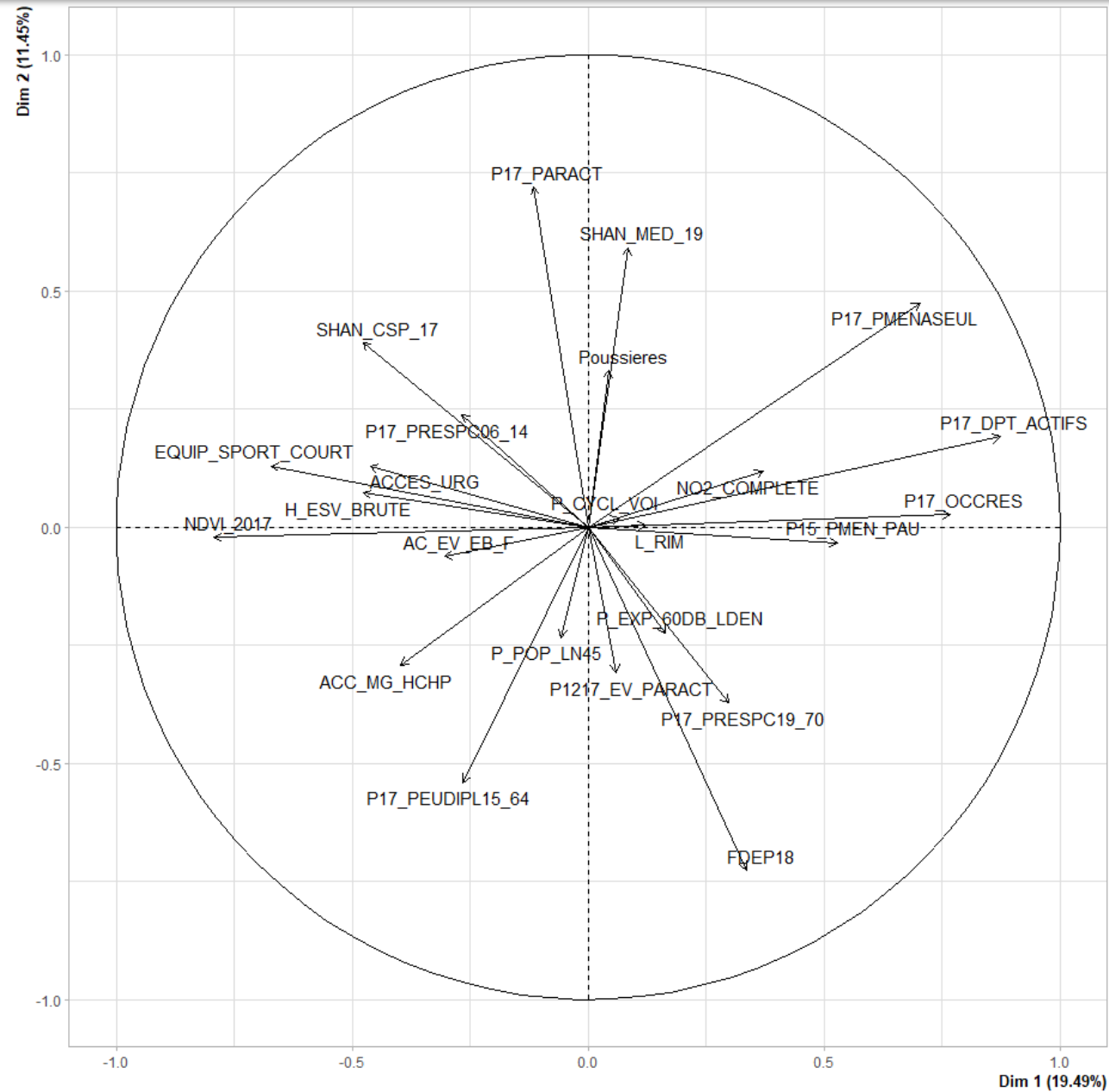
Crude incidence of Coronary Heart Diseases (2012-2016) (Figure SM-1)



Crude incidence of Coronary Heart Diseases (2012-2016) (Figure SM-2)



PCA graph of 24 variables (Figure SM-3)



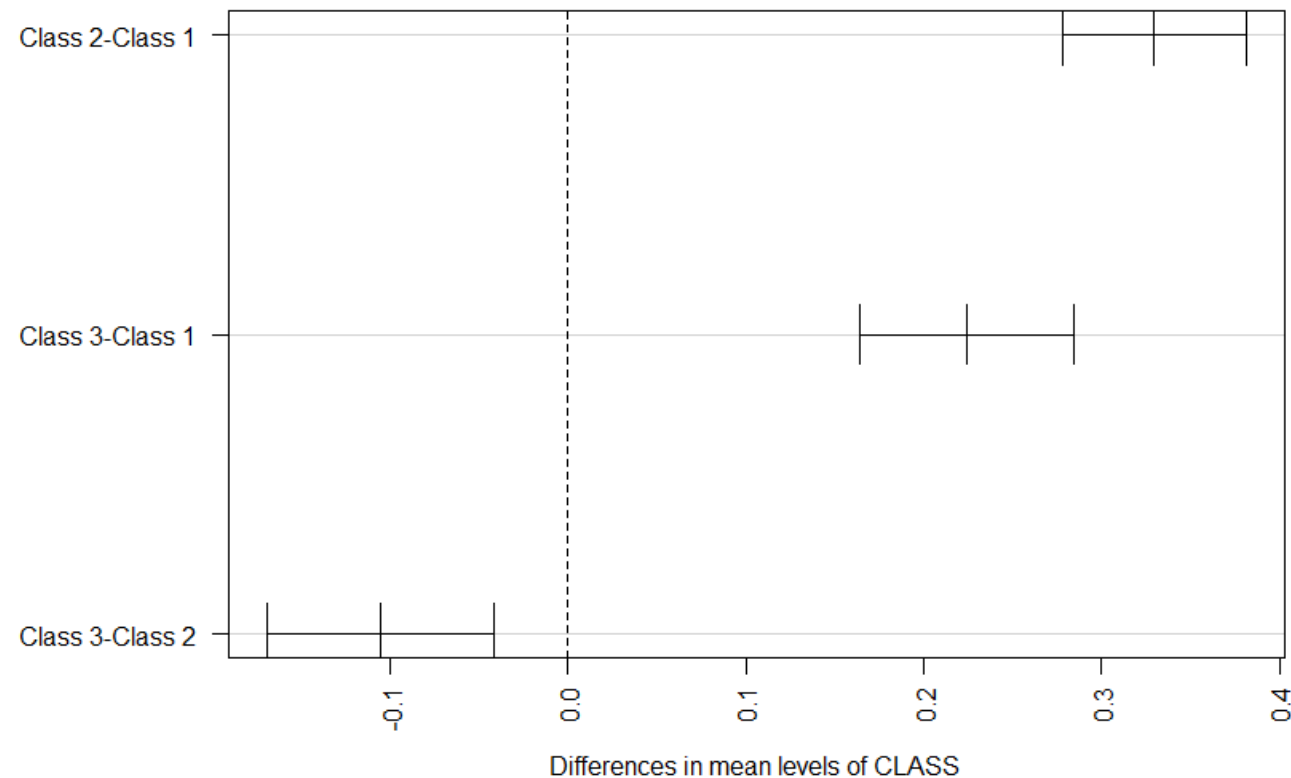
PCA cos² (cos² < 0.2 are omitted) (Table SM-VIII)

	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5	Dim.6	Dim.7	Dim.8
FDEP18		0.5277						
P17_PARACT		0.5162						
P1217_EV_PARACT					0.2578	0.2031		
P17_OCCRES	0.5892							
P17_PEUDIPL15_64		0.2924						
P17_PMENASEUL	0.4945	0.2247						
P15_PMEN_PAU	0.2799							
AC_EV_EB_F			0.1466					
P_EXP_60DB_LDEN			0.5295					
P_POP_LN45			0.5441					
NDVI_2017	0.6276							
NO2_COMPLETE			0.3371					
Poussieres				0.3391				
L_RIM						0.2964		
H_ESV_BRUTE	0.2261							
P17_PRESPC19_70				0.1741				
P17_PRESPC06_14						0.0933		
SHAN_CSP_17	0.2277							
P_CYCL_VOI			0.2040				0.2316	0.2292
EQUIP_SPORT_COURT	0.4501							
SHAN_MED_19		0.3491			0.2716			
ACCES_URG	0.2131			0.2014				
ACC_MG_HCHP					0.2406			
P17_DPT_ACTIFS	0.7626							
Cumulative % of variance	19.485	30.932	39.752	45.509	51.25	56.489	61.258	65.519

SIRs: Tukey multiple comparisons of means (Table SM-IX, Figure SM-4)

	Difference	Lower	Upper	p adj
Class 2 - Class 1	0.3296	0.2779	0.3812	0.0000
Class 3 - Class 1	0.2239	0.1636	0.2843	0.0000
Class 3 - Class 2	-0.1056	-0.1694	-0.0419	0.0003

95% family-wise confidence level

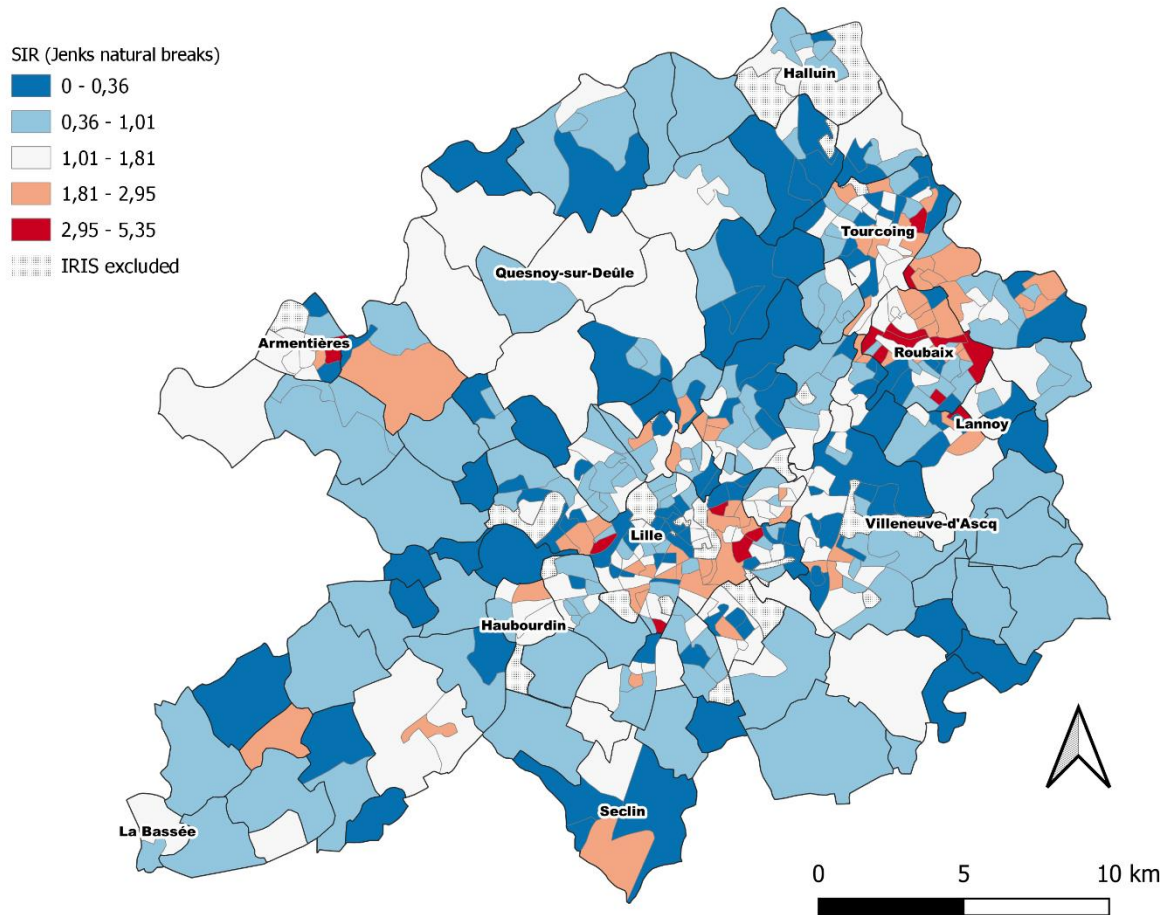


Population density in each class (Table SM-X)

	Mean	CI 95%
Class 1	28.399	[25.039 – 31.758]
Class 2	78.977	[72.534 – 85.420]
Class 3	108.547	[99.024 – 118.070]
ELM	61.800	[57.400 – 66.202]

Annexe 6

Maladies coronariennes, évènements « fatal » : ratio d'incidence standardisé, par standardisation indirecte (2012-2016)



Annexe 7

Comparaison Hauts-de-France Wallonie



		Hauts de France (3 817)	Wallonie (262)
Superficie (km ²)	Etendue	0,17 à 97,98	1,004 à 99,8
	Moyenne	8,35	37,30
	Médiane	6,91	33,92
Habitants	Etendue	de 5 à 231 491	1 415 à 206 214
	Moyenne	1 569	12 426
	Médiane	423	7 150