



UNIVERSITÉ
DE LORRAINE



ADEME



Agence de l'Environnement
et de la Maîtrise de l'Énergie

Ecole Doctorale BioSE (Biologie-Santé-Environnement)

Thèse

Présentée et soutenue publiquement pour l'obtention du titre de

DOCTEUR DE L'UNIVERSITE DE LORRAINE

et de

L'ECOLE DES HAUTES ETUDES EN SANTE PUBLIQUE

Mention : « Sciences de la Vie et de la Santé »

présentée par

Cindy PADILLA

Inégalités sociales de santé et expositions environnementales. Une analyse spatio-temporelle du risque de mortalité infantile et néonatale dans 4 agglomérations françaises.

Date de la soutenance : 24 Octobre 2013

Membres du Jury :

Rapporteurs : **Jacqueline CLAVEL**, Directeur de Recherche, Epidémiologie environnementale des cancers, Université Paris-Sud
Béatrice FERVERS, Professeur associé, Coordinatrice Unité cancer et environnement, Université Lyon 1

Examineurs : **Denis ZMIROU-NAVIER**, Professeur, Université de Lorraine, EHESP, Directeur de thèse
Séverine DEGUEN, Professeur, Biostatistiques, EHESP, Co-directeur de thèse
Verónica.M VIEIRA, Professeur, Santé Environnement, Université Irvine, CA
Hélène DESQUEYROUX, Ingénieur de recherche, ADEME



UNIVERSITÉ
DE LORRAINE



ADEME



Agence de l'Environnement
et de la Maîtrise de l'Énergie

Ecole Doctorale BioSE (Biologie-Santé-Environnement)

Thèse

Présentée et soutenue publiquement pour l'obtention du titre de

DOCTEUR DE L'UNIVERSITE DE LORRAINE

et de

L'ECOLE DES HAUTES ETUDES EN SANTE PUBLIQUE

Mention : « Sciences de la Vie et de la Santé »

présentée par

Cindy PADILLA

Inégalités sociales de santé et expositions environnementales. Une analyse spatio-temporelle du risque de mortalité infantile et néonatale dans 4 agglomérations françaises.

Date de la soutenance : 24 Octobre 2013

Membres du Jury :

Rapporteurs : **Jacqueline CLAVEL**, Directeur de Recherche, Epidémiologie environnementale des cancers, Université Paris-Sud
Béatrice FERVERS, Professeur associé, Coordinatrice Unité cancer et environnement, Université Lyon 1

Examineurs : **Denis ZMIROU-NAVIER**, Professeur, Université de Lorraine, EHESP, Directeur de thèse
Séverine DEGUEN, Professeur, Biostatistiques, EHESP, Co-directeur de thèse
Verónica.M VIEIRA, Professeur, Santé Environnement, Université Irvine, CA
Hélène DESQUEYROUX, Ingénieur de recherche, ADEME

REMERCIEMENTS

Je souhaite remercier en premier lieu mon directeur de thèse, Denis Zmirou-Navier, Professeur et Directeur du département santé environnement travail & génie sanitaire de L'EHESP pour m'avoir accueilli au sein de son équipe et son département. Je lui suis reconnaissante pour ses qualités scientifiques et son aide précieuse dans la relecture et la correction des articles, des rapports ainsi que ma thèse. La justesse de ses critiques a été très constructive et utile.

J'aimerais aussi remercier mon co-directeur de thèse, Séverine Deguen, Professeur à l'EHESP pour la confiance qu'elle m'a témoignée du début à la fin de la thèse. J'ai été extrêmement sensible à ses qualités humaines d'écoute et de compréhension tout au long de ce travail doctoral. J'aimerais également lui dire à quel point j'ai apprécié sa grande disponibilité et ses multiples conseils avisés et experts sur cette recherche. J'ai pris un grand plaisir à travailler avec elle.

Je voudrais remercier les rapporteurs de cette thèse, Jacqueline Clavel, Directeur de recherches de l'équipe épidémiologie environnementale des cancers de l'Université Paris-Sud, et Béatrice Fervers, Professeur associé et coordinatrice de l'unité cancer environnement du centre Léon Bérard, pour l'intérêt qu'elles ont porté à mon travail et d'avoir accepté de consacrer une partie de leur temps à la lecture attentive de ce manuscrit et un grand merci aux autres membres de mon jury Hélène Desqueyroux Ingénieur à l'ADEME et Veronica Vieira.

Un grand merci aux membres de mon comité de thèse, Sylvaine Cordier Directeur de recherche à l'Université de Rennes I, Pierre Chauvin Directeur de l'équipe déterminants sociaux de la santé et du recours aux soins à l'université Paris 6, Denis Bard Professeur à l'EHESP et Hélène Desqueyroux ingénieur ADEME pour avoir contribué, par leurs critiques et leurs conseils, à l'aboutissement et à l'amélioration de ce travail.

Je tiens à témoigner toute ma gratitude aux financeurs du projet, l'Ecole des Hautes Etudes en Santé Publique et l'ADEME.

Je désire grandement remercier, Olivier Thomas et Marie-Aline Block ancien et nouveau directeur, Jeanine Pommier directeur-adjoint, Sarah Kitar et tout particulièrement Maud Subtil du Réseau Doctoral en santé publique de l'EHESP, de m'avoir donné l'opportunité de participer au Réseau Doctoral où j'ai bénéficié d'échanges, d'un stage enrichissant et de cours très formateurs.

Je tiens à exprimer toute ma reconnaissance à Veronica Vieira, Professeur à l'université d'Irvine, pour m'avoir accueilli chaleureusement au sein du département Santé Environnement de l'Ecole de santé publique de Boston lors d'un stage de 3 mois dans le cadre du réseau doctoral.

Je tiens à remercier l'ancien et nouveau directeur de l'Ecole Doctorale Biologie Santé Environnement, Jean-Yves Jouzeau et Patrick Menu ainsi que leur assistante Hélène Primerano pour son aide administrative constante et de sa sympathie.

Bien sûr, atteindre ces objectifs n'aurait pas été possible sans l'aide des membres et partenaires du projet Equit'Area qui ont participé au projet. Je souhaite notamment remercier les membres des AASQAs Florence Troude, Charles Beaugard, Géraldine LeNir, Philippe Rosselo pour leur disponibilité, nos échanges et leur travail. Je souhaite aussi exprimer ma gratitude au CepiDC et plus particulièrement Grégoire Rey.

Je désire en outre remercier toutes les personnes des services d'états civils des mairies pour leur sympathie, leur accueil chaleureux et plus particulièrement le personnel des agglomérations de Lille, et Lyon que j'ai rencontré lors de mon travail de terrain, merci.

Ces remerciements seraient incomplets si je n'en adressais pas à l'ensemble du personnel du département de l'Institut du Management dans lequel se trouve le bureau des doctorants. Pour leur bonjour, leur histoire, leur considération et leur sympathie, un grand merci !

J'exprime ma gratitude à tous les amis-chercheurs du bureau des doctorants pour leur soutien moral ainsi que pour la très bonne ambiance que j'ai toujours trouvée. Je les remercie plus particulièrement pour nos nombreuses discussions philosophiques lors desquelles on a refait le monde et la confiance qu'ils ont manifestés à l'égard de mon travail : Alis, Céline, François, Michael, Wahida, Jacques-Francois, Anne-Lise. Enfin, j'adresse particulièrement mille mercis à Wahida pour sa patience, ses multiples conseils et à François mon ami et compatriote pour son soutien affectif sans faille.

Je souhaite remercier spécialement Kevin pour son soutien et sa patience tout au long de la thèse.

Mes dernières pensées vont à ma famille, belle famille et mes amis pour avoir toujours cru en moi et mes projets les plus fous. J'associe à ces remerciements le club de sport Orange bleue et le tapis de course n°2, pour avoir partagé mes doutes, mes peines et mes déceptions et me donner le courage de persévérer.

Enfin, j'exprime toute ma reconnaissance aux deux relecteurs de ce manuscrit et amis : Florence et Olivier.

VALORISATIONS SCIENTIFIQUES

ARTICLES PUBLIES (5)

Padilla CM, Deguen S, Lalloue B, Zmirou-Navier D, Vieira VM. Cluster analysis of social and environment inequalities of infant mortality. A spatial study in small areas revealed by local disease mapping in France. *Sci Total Environ*. 2013 Jun 1;454-455:433-41.

Padilla CM, Lalloue B, Pie C, Lucas E, Zmirou-Navier D, Deguen S. An ecological study to identify census blocks supporting a higher burden of disease - Infant mortality in the Lille Metropolitan Area, France. *Matern Child Health J*. 2013 Mar 12

Padilla C, Lalloué B, Zmirou-Navier D, Deguen S. Infant mortality, deprivation and proximity to polluting industrial facilities – A small-scale spatial analysis with census block data (Lille Metropolitan Area, France). *Environnement Risques et Santé*. Vol 10, n°3 ; 216-21 – Déc2010.

Lalloué B, Monnez JM, **Padilla C**, LeMeur N, Zmirou-Navier D, Deguen S. A statistical procedure to create a neighborhood socioeconomic index for health inequalities analysis. *Int J Equity Health*. 2013 Mar 28;12:21

Lalloué B, Deguen S, Monnez JM, **Padilla C**, Kihal W, Zmirou-Navier D, LeMeur N. *SesIndexCreator*: An R Package for Socioeconomic Indices Computation and Visualization. *Journal of Statistical Software*.

ARTICLES SOUMIS (2)

Padilla CM, Kihal W, Vieira VM, Rosselo P, LeNir G, Zmirou-Navier D, Deguen S. Air quality and social deprivation in four French metropolitan areas– A spatio-temporal environmental inequality analysis conducted at a small geographical level. Submitted in *Environmental Research journal*.

Lalloué B, Monnez JM, **Padilla C**, Kihal W, Zmirou-Navier D, Deguen S. Data Analysis technics, A tool for Cumulative Exposure Assessment. Submitted in Exposure Science and Environmental Epidemiology journal

ARTICLES EN REVISION (2)

Kihal W, **Padilla CM**, Lalloué B, Gelormini M, Rey G, Zmirou-Navier D, Deguen S. Green space and Social inequalities among a cohort of neonatal mortality in France. In revision : BMC Pregnancy and child births journal

Kihal W, **Padilla CM**, Lalloué B, Rougier C, Defrance J, Zmirou-Navier D, Deguen S. An exploratory spatial analysis to assess the relationship between deprivation, noise and infant mortality. In revision : Environmental Health journal.

COMMUNICATIONS SCIENTIFIQUES

1. “Environmental Health and Social Vulnerabilities” – European Congress of Environment and Health – Namur, Belgium, October 21 & 22, 2010.

Poster presentation: Mortality, deprivation and proximity to industrial plants. A spatial study in Lille, France.

2. “Territoires et Santé des populations” – 12^{ème} Congrès national des Observatoires Régionaux de la Santé– Lyon, France, Novembre 9 & 10, 2010.

Poster presentation: Mortality, deprivation and proximity to industrial plants. A spatial study in Lille, France.

3. “SFSE -Genes, environment and social inequalities” – 1^{er} Congrès national en Santé Environnement– Rennes, France, December 14 & 15, 2010.

Oral presentation: Mortality, deprivation and proximity to industrial plants. A spatial study in Lille, France.

4. “Rencontres de l’Hôtel Dieu“ – Congrès national du réseau doctoral en santé publique – May 12 & 13 2011 – Paris

Oral presentation: Association of proximity to pollution facilities, deprivation and infant mortality. A spatial analysis using census data _Lille in France

5. “ISEE – International Society for Environmental Epidemiology – September, 13-16 2011 – Barcelona

Oral presentation: *Proximity to Industrial plants, Neighbourhood Socio-economic Characteristics and the Risk of Infant Mortality - A Bayesian analysis in Lyon, France*

6. “UEP – International Congress: Urban Environmental Pollution – June, 17-20 2012 – Amsterdam

Poster presentation: Spatial analysis and mapping of the effect of the socioeconomic deprivation on the association between ambient air NO₂ and infant mortality in Lille , France.

7. “Health and Space”– 2nd national congress of health geography and spatial epidemiology– September 19-21 2012 – Marseille

Oral presentation: Spatial analysis and mapping of the effect of the socioeconomic deprivation on the association between ambient air NO₂ and infant mortality in Lille , France

8 “Science and Policy to protect future generations” – International Congress in Environmental health –March, 03-06 2013 – Boston

Oral presentation: Environmental inequalities in France A spatio-temporal analysis conducted at a small geographical level in four French metropolitan areas

9. “ Rencontres nationales scientifiques du réseau doctoral en santé publique“ –Avril 15 &16 2013 – Paris

Oral presentation: Environmental inequalities in France – A spatio-temporal analysis conducted at a small geographical level in four French metropolitan areas.

10 “ISEE, ISES, ISEA”_ International congress in Environment and Health _ Aout, 19-23 2013 _ Bale

Oral Presentation: Neonatal mortality, outdoor air pollution and social inequalities_ a spatial analysis conducted at a small geographical level in three French metropolitan areas.

TABLE DES MATIERES

Remerciements	1
Valorisations Scientifiques.....	5
Table des Matières	9
Liste des Tableaux	13
Liste des Figures	15
Liste des Abréviations	17
Introduction générale.....	19
Chapitre 1 : Contexte.....	21
I. Les inégalités de santé liées à la mortalité infantile et néonatale	21
A. Choix de la mortalité infantile et néonatale	21
B. Le constat : Situation actuelle et évolution entre 2000 et 2009	24
II. Les inégalités sociales de santé liées à la mortalité infantile et néonatale.....	28
A. Des déterminants socio-économiques nombreux.....	29
B. Des déterminants contextuels	30
III. L'environnement, un déterminant suspecté	32
A. Mécanismes par lesquels peuvent opérer les expositions environnementales dans les ISS.....	32
IV. Le projet de recherche Equit' Aréa.....	35
V. Objectifs et structure de la thèse	36
Schéma conceptuel de la thèse	38
Références : Chapitre 1	39
Chapitre 2 : Matériels et Méthodes	47
I. Le cadre d'étude.....	47
A. Les zones d'étude.....	47
B. Le type d'étude : étude écologique	48
C. L'unité géographique : l'IRIS	49
D. Description du cadre de l'étude	50
II. Données sanitaires	52
A. Définition	52
B. Recueil des données sanitaires.....	53

C.	Les difficultés rencontrées	56
D.	Description des données sanitaires	58
III.	Données socio-économiques.....	62
A.	Choix d'un indice de défaveur.....	62
B.	Construction de l'indice de défaveur	63
C.	Description de l'indice de défaveur	64
IV.	Données environnementales : le dioxyde d'azote.....	66
A.	Présentation du dioxyde d'azote	66
B.	Choix du dioxyde d'azote	68
C.	Modélisation du dioxyde d'azote.....	69
D.	Distribution des concentrations du NO ₂	71
V.	Des méthodes statistiques adaptées : l'épidémiologie spatiale en santé publique.....	73
A.	Choix de ces modèles.....	73
B.	Approche Bayésienne	74
C.	Approche Spatiale : GAM	76
D.	Le système d'information géographique	79
	Références : Chapitre 2	80

Chapitre 3 : Exposition au dioxyde d'azote et inégalités socio-spatiales au regard de la mortalité néonatale et infantile

I.	Etude des inégalités socio-spatiales de santé.....	87
A.	Contexte	87
B.	Publication et résultats	88
C.	Discussion	103
II.	Etude du rôle du NO ₂ dans les inégalités socio-spatiales de santé.....	105
A.	Contexte	105
B.	Publication et résultats	106
C.	Discussion	122
III.	Etude spatio-temporelle des inégalités environnementales	125
A.	Contexte	125
B.	Publication	126
C.	Discussion	153
IV.	Etude de l'évolution temporelle du rôle du NO ₂ dans les inégalités socio-spatiales de santé	155
A.	Contexte	155
B.	Résultats.....	157
C.	Discussion	166
	Références : Chapitre 3	167

Chapitre 4 : Discussion générale et perspectives	171
I. Conclusion générale.....	171
A. Les mérites de l'étude	171
B. Les limites de l'étude	172
C. Apports scientifiques	175
II. Perspectives et autres contributions de la recherche.....	176
A. Autres contributions de recherche	176
B. Perspectives en santé publique.....	178
Références : Chapitre 4	179
ANNEXES	181

LISTE DES TABLEAUX¹

Tableau 1. Mortalité infantile et néonatale dans les agglomérations de Lille, Lyon, Marseille et Paris-ville, 2000-2009.	58
Tableau 2. Caractéristiques des données d’entrées des modèles.....	70
Tableau 3. Statistiques descriptives des concentrations ambiantes du dioxyde d’azote par IRIS dans les 4 agglomérations pour la période 2002-2009.	71
Tableau 4: Valeurs moyennes des caractéristiques socio-économiques de l’IRIS selon l’appartenance à la classe de défaveur.....	100
Tableau 5 : Risques relatifs (RR) et intervalle de crédibilité à 95% de la mortalité infantile de plusieurs modèles selon l’ajustement à la défaveur socio-économique.	101
Tableau 6: Risques relatifs (RR) et intervalle de crédibilité à 95% de la mortalité néonatale de plusieurs modèles selon l’ajustement à la défaveur socio-économique.	102
Tableaux 7 AB: Caractéristiques de chaque modèle spatial d’estimation du risque, A) de mortalité infantile pour la ville de Paris, B) de mortalité néonatale de l’agglomération de Lille, Lyon et Paris-ville. P value = valeur de la déviance du test global dont hypothèse Ho est l’homogénéité spatiale du risque de mortalité infantile ou néonatale.	118
Tableau 8 : Evolution du taux de mortalité infantile et néonatale entre les périodes 2002-2005 et 2006-2009 pour les agglomérations de Lille et Lyon et les périodes 2004-2006 et 2007- 2009 pour la ville de Paris.....	159
Tableau 9 : Evolution du taux de mortalité infantile et néonatale entre les périodes 2002-2005 et 2006-2009 selon la défaveur socio-économique du quartier de résidence pour les 3 agglomérations,159	
Tableau 10 : Caractéristiques de chaque modèle spatial d’estimation du risque A) de mortalité infantile pour les agglomérations de Lille, Lyon et Paris-ville B) de mortalité néonatale pour les agglomérations de Lille, Lyon et Paris-ville. P valeur de la déviance du test global dont hypothèse Ho est l’homogénéité spatiale du risque de mortalité infantile ou néonatale.....	160

¹ Les tableaux joints aux articles intégrés dans la thèse ne sont pas indiqués dans la liste ci-dessous.

LISTE DES FIGURES ²

Figure 1 : Taux de mortalité avant 1 an, 1 mois, 1 semaine en France	25
Figure 2 : Évolution du rang occupé par la France dans le classement européen des plus bas taux de mortalité infantile dans 30 pays européens - 1999-2009	26
Figure 3 : Carte des taux de mortalité infantile et néonatale en 2000 et 2009 par région en France.....	27
Figure 4 : Evolution de la mortalité infantile et néonatale entre 2000-2009 pour les 4 agglomérations de l'étude.	28
Figure 5 : Cartographies des quatre agglomérations du projet Equit'Area.	51
Figure 6 : Cartographies des rapports de mortalités standardisés (SMR) de mortalité infantile à A) Lille, B) Lyon, C) Paris sur la période 2000 à 2009.....	60
Figure 7 : Cartographies des rapports de mortalités standardisés (SMR) de mortalité néonatale à A) Lille, B) Lyon, C) Paris sur la période 2000 à 2009.....	61
Figure 8 : Cartographies de l'indice de défaveur en quintiles par IRIS à A) Lille, B) Lyon, C) Paris, D) Marseille.....	65
Figure 9 : Cartographies des concentrations moyennes par IRIS du dioxyde d'azote en quintiles pour la période 2002-2009 à A) Lille, B) Lyon, C) Paris, D) Marseille.....	72
Figure 10 : Prévalence de la mortalité infantile estimée par le modèle GAM pour l'agglomération de Paris, issu du modèle Brut A), ajusté sur l'exposition au NO ₂ B), ajusté sur la défaveur sociale C) et sur le modèle complet D). La ligne noire et solide identifie les IRIS à risque significativement plus élevés de mortalité infantile.	119
Figure 11 : Prévalence de la mortalité néonatale estimée par le modèle GAM pour l'agglomération de Lille, issu du modèle Brut A), ajusté sur l'exposition au NO ₂ B), ajusté sur la défaveur sociale C) modèle complet D). La ligne noire et solide identifie les IRIS à risque significativement plus élevés de mortalité néonatale.	120

² Les figures jointes aux articles intégrés dans la thèse ne sont pas indiquées dans la liste ci-dessous.

Figure 12 : Prévalence de la mortalité néonatale estimée par le modèle GAM pour l'agglomération de Lyon, issu du modèle Brut A), ajusté sur l'exposition au NO ₂ B), ajusté sur la défaveur sociale C) modèle complet D). La ligne noire et solide identifie les IRIS à risque significativement plus élevés de mortalité néonatale.	121
Figure 13 : Relation entre la défaveur socio-économique et la concentration en dioxyde d'azote pour les agglomérations de A) Lille, B) Lyon et C) la ville de Paris sur la période 2002-2009.	124
Figure 14 : Prévalence de la mortalité néonatale estimée par le modèle GAM pour l'agglomération de Lille, issu du modèle Brut A), ajusté sur l'exposition au NO ₂ B), ajusté sur la défaveur sociale C) sur les deux périodes d'études. La ligne noire et solide identifie les IRIS à risque significativement plus élevés de mortalité néonatale.	161
Figure 15 : Prévalence de la mortalité infantile estimée par le modèle GAM pour l'agglomération de Lille, issu du modèle Brut A), ajusté sur l'exposition au NO ₂ B), ajusté sur la défaveur sociale C) sur les deux périodes d'études. La ligne noire et solide identifie les IRIS à risque significativement plus élevés de mortalité infantile.	162
Figure 16 : Prévalence de la mortalité infantile estimée par le modèle GAM pour l'agglomération de Paris, issu du modèle Brut A), ajusté sur l'exposition au NO ₂ B), ajusté sur la défaveur sociale C) sur les deux périodes d'études. La ligne noire et solide identifie les IRIS à risque significativement plus élevés de mortalité infantile.	163
Figure 17 : Prévalence de la mortalité infantile estimée par le modèle GAM pour l'agglomération de Lyon, issu du modèle Brut A), ajusté sur l'exposition au NO ₂ B), ajusté sur la défaveur sociale C) sur les deux périodes d'études. La ligne noire et solide identifie les IRIS à risque significativement plus élevés de mortalité infantile.	164
Figure 18 : Prévalence de la mortalité néonatale estimée par le modèle GAM pour l'agglomération de Lyon, issu du modèle Brut A), ajusté sur l'exposition au NO ₂ B), ajusté sur la défaveur sociale C) sur les deux périodes d'études. La ligne noire et solide identifie les IRIS à risque significativement plus élevés de mortalité néonatale.	165

LISTE DES ABREVIATIONS³

AASQA :	Association agréée de surveillance de la qualité de l'air
ACP :	Analyse en Composantes Principales
ADMS :	Atmospheric Dispersion. Modelling System
AIC :	Akaike information criterion
API :	Allocation parents isolés
ARS :	Agence régionale de la santé
BYM :	Besag, York and Mollié
CNIL :	Commission nationale de l'informatique et des libertés
CO/ CO ₂ :	Monoxyde de carbone/ dioxyde de carbone
DIC :	Deviance Information Criteria
DoHaD :	Developmental origins of health disease
DOM :	Département d'outre-mer
EHESP :	Ecole des Hautes Études en Santé Publiques
EPER :	European Pollutant Emission Register
GAM :	Generalized Additive Models
HCSP :	Haut conseil de la santé publique
HLM :	Habitation à loyer modéré
IC :	Intervalle de Crédibilité
IGAS :	Inspections générales des Affaires Sociales
IGN :	Institut géographique national
INSEE :	Institut national de la statistique et des études économiques.
INSERM :	Institut national de la santé et de la recherche médicale

³ Les abréviations mentionnées dans les articles intégrés dans la thèse ne sont pas indiquées dans la liste ci-dessous.

INVS :	Institut national de veille sanitaire
IRIS :	Ilots Regroupés pour l'Information Statistiques
ISS :	Inégalités sociales de santé
LOESS :	Locally weighted regression smoothers
NO/ NO ₂ :	Monoxyde d'azote/ dioxyde d'azote
O ₂ :	Ozone
OCDE :	Organisation de coopération et de développement économiques
OMS :	L'organisation Mondiale de la Santé
OMD :	Objectifs du millénaire pour le développement
ONU :	Organisation des Nations Unis
PACA :	Provence Alpes Côtes d'Azur
PM :	Particulate matter en µm de diamètre
PNSE :	Plan National Santé Environnement
RMI :	Revenu minimum d'insertion
RSA :	Revenu de solidarité active
RR :	Risque Relatif
SIDA :	Syndrome d'immunodéficience acquise
SIG :	Système d'Information Géographique
SMR :	Standardized Mortality/Morbidity Ratios,
SO ₂ :	Dioxyde de soufre
UNICEF :	United Nations International Children's Emergency Fund
UN-IGME :	United Nations Inter-Agency Group
ZUS :	Zones Urbaines Sensibles

INTRODUCTION GENERALE

« La place des individus dans l'échelle sociale influe sur les conditions dans lesquelles ils grandissent, apprennent, vivent, travaillent et vieillissent et sur leur vulnérabilité face à leur environnement, à la maladie et à ses conséquences. »

Déterminants sociaux de la santé -OMS(1)

L'objectif d'améliorer l'état de santé de la population dans son ensemble et de chacun en particulier, quelle que soit son origine ou son appartenance sociale, est largement partagé et par ailleurs inscrit dans les textes fondateurs de notre république. Le premier de ces objectifs a été atteint : l'espérance de vie s'est constamment améliorée. L'autre ne l'a pas été, puisque les progrès n'ont pas profité de la même manière à tous et que les inégalités sociales de santé ont eu tendance à se creuser (2).

En effet, malgré l'amélioration des conditions de vie et les progrès des prises en charge médicales, des disparités sociales significatives demeurent en matière de santé en France. Un important travail de l'Institut national de la santé et de la recherche médicale (INSERM) (3), un rapport de l'Inspection générale des affaires sociales (IGAS) (2), ainsi qu'un rapport du Haut conseil de la santé publique (HCSP) (4), leur ont été consacrées. Ce dernier soulignait que les inégalités sociales de santé (morbidité, mortalité, santé perçue, santé fonctionnelle) demeurent dans notre pays plus importantes que dans la plupart des autres pays européens. En France, réduire les inégalités sociales de santé (ISS) est devenu un objectif explicite des politiques de santé ces dernières années.

En dépit de nombreux facteurs de risques déjà identifiés, une meilleure connaissance des déterminants économiques, sociaux et culturels de telles disparités manque encore afin d'améliorer l'interprétation des observations et de leur évolution. Cette connaissance est aussi nécessaire en vue d'engager des politiques et des actions de prévention pertinentes. Une part des inégalités sociales de santé (quel que soit l'évènement sanitaire étudié) demeure à ce jour inexplicée. Si divers travaux ont déjà été engagés dans ce sens, l'influence des expositions environnementales sur les inégalités sociales de santé reste en France un terrain encore mal

connu. Déterminer comment ces expositions environnementales peuvent pour partie expliquer les inégalités de santé constitue donc aujourd'hui un objectif de recherche en santé publique.

Ce travail de thèse s'inscrit dans le cadre du programme Equit'Area, recherche initiée en 2008 à l'École des Hautes Études en Santé Publique (EHESP) qui investigate la contribution des expositions environnementales aux inégalités sociales de la santé. Dans ce cadre, ma thèse porte plus particulièrement sur le rôle de la pollution atmosphérique sur les ISS en lien avec la mortalité infantile et néonatale. Ce manuscrit est l'aboutissement de quatre années de travail au sein du département Santé Environnement et Travail & Génie Sanitaire et présente les travaux accomplis.

CHAPITRE 1 :

CONTEXTE

I. LES INÉGALITÉS DE SANTÉ LIÉES À LA MORTALITÉ INFANTILE ET NÉONATALE

La France peut se prévaloir d'un taux de natalité qui la place actuellement au deuxième rang après l'Irlande au sein des pays européens. En 2010, en France, les femmes ont eu en moyenne 2,01 enfants - soit près de 833 000 naissances - contre 1,6 en moyenne en Europe. Toutefois, selon l'éditorial d'une publication de l'Institut national de veille sanitaire, « *ces données rassurantes ne doivent pas cacher la réalité beaucoup plus péjorative de la mortalité périnatale* ».

A. Choix de la mortalité infantile et néonatale

1. De l'histoire à l'indicateur reconnu de l'état de santé de la population.

La mortalité infantile se définit par le rapport du nombre de décès d'enfants de moins de un an au nombre de naissances vivantes de la période choisie. La mortalité néonatale est le nombre de décès survenant de 0 à 28 jours révolus rapporté au nombre de naissances vivantes. On distingue la mortalité néonatale précoce pour les décès survenus dans les 6 premiers jours après la naissance et la mortalité néonatale tardive définissant un décès survenus entre 7 et 28 jours révolus.

A partir des années 1870, de nets progrès améliorent les conditions de vie (ou plutôt de survie) des nouveau-nés. D'abord, les accouchements à l'hôpital constituent un progrès

décisif. Sous la III^{ème} République, une série de lois de protection de l'enfance est votée, notamment la loi Roussel de 1874 qui organise au niveau national la surveillance des jeunes enfants placés en nourrice. Ce sont surtout les progrès de l'ère pastoriennne, notamment en matière d'obstétrique, de prophylaxie, de stérilisation et d'asepsie puerpérale inspirés par les travaux de Semmelweis, et de nutrition infantile, qui contribuent à la baisse de la mortalité infantile dans le dernier tiers du XIX^{ème} siècle. Le vaccin contre la variole, mis au point par l'Anglais Jenner dès 1796, se diffuse peu à peu. " Les enfants furent les premiers bénéficiaires d'une découverte extraordinaire, celle de la vaccine ", qui réduit les ravages de la variole (5). La mort des enfants en bas âge devient dès lors évitable et la mortalité infantile décroît rapidement.

A la veille de la Première Guerre mondiale, elle se stabilise à 11% environ. En dépit de ces succès, la situation reste préoccupante. Par exemple, dans le dernier tiers du XIX^{ème} siècle, la vaccination suscite encore une certaine méfiance en France. En outre, elle n'est pas obligatoire : il faudra attendre 1902 (pour la variole), 1938 (pour la diphtérie) et 1940 (pour le tétanos) pour que tous les enfants soient indistinctement vaccinés (5).

C'est seulement en 1990 que l'Organisation des Nations Unies (ONU) inclut la mortalité infantile dans les « Objectifs du millénaire pour le développement » (OMD), Millennium Development Goals en anglais. Huit objectifs sont adoptés à New York (États-Unis) par 193 États membres de l'ONU et au moins 23 organisations internationales, le but étant de les atteindre pour 2015. Ces objectifs recouvrent de grands enjeux humanitaires : la réduction de l'extrême pauvreté et de la mortalité infantile, la lutte contre plusieurs épidémies dont le SIDA, l'accès à l'éducation, l'égalité des sexes, et l'application du développement durable. La réduction de la mortalité infantile et plus précisément, « Réduire de 2/3 le taux de mortalité des enfants de moins de 5 ans » est l'objectif n°4 (6). En 2004, l'Organisation Mondiale de la Santé (OMS) ainsi que les membres de l'UNICEF (United Nations International Children's Emergency Fund), la Banque Mondiale et les Nations Unies se réunissent sous le nom de « United Nations Inter-agency Group » (UN-IGME) dans le but de faire avancer les recherches et réaliser les objectifs de l'OMD (7).

2. Intérêt en santé publique

Aujourd'hui, la mortalité infantile et la mortalité périnatale sont reconnues de longue date comme des indicateurs pertinents pour caractériser l'état de santé général d'une population (8). Il reflète le lien logique entre les facteurs tels que le développement économique, les conditions de vie, le bien-être social, les taux de mortalité, la qualité de l'environnement et l'état de santé générale de la population (8–10). De plus, il permet de comparer les systèmes de santé et mettre en évidence les populations qui requièrent une attention particulière de la part des services de santé. (11)

Le nouveau-né est particulièrement sensible à la qualité de son environnement (12). Cette plus grande vulnérabilité aux agressions de son environnement va se prolonger durant toute la première année de sa vie, avec une fenêtre critique pendant la première semaine de vie. Son état de santé durant cette période pourra influencer sa santé à l'âge adulte, selon le concept de l' « Origine Développementale de la Santé et des Maladies » ou DOHaD, proposé il y a vingt ans par David Barker qui mettait en lien le poids de naissance et le risque ultérieur de pathologie coronarienne (13). En effet, l'extrême fragilité du nourrisson le rend entièrement dépendant des conditions de qualité de son environnement immédiat en matière d'infrastructures d'hygiène, de santé, de développement socio-économique, d'éducation, de gestion mais aussi en fonction de l'enclavement géographique plus ou moins prononcé et de la situation politique du pays. Dans la mesure où ces conditions influencent intimement le déroulement de la grossesse et les caractéristiques de santé et sociales des femmes, on peut les considérer comme des marqueurs de la santé des femmes, marqueurs de leur « possibilité » d'avoir une grossesse qui se termine par la naissance à terme d'un enfant vivant, en bonne santé et de poids normal (7,14,15).

Ces indicateurs mesurent aussi la qualité du système de soins ainsi que l'efficacité des soins préventifs et l'attention accordée à la santé de la mère et de l'enfant (2,4). En effet, les principales causes de décès du nourrisson en France aujourd'hui, suite à la réduction des décès attribuables à la mort subite des années 90 (16) sont l'immaturité et les anomalies congénitales (17,18,14).

B. Le constat : Situation actuelle et évolution entre 2000 et 2009

En dépit d'un taux globalement bas en France métropolitaine et dans la plupart des pays développés, on démontrera que des disparités spatiales persistent pour la mortalité infantile et néonatale en termes de niveau atteint et d'évolution récente.

1. Etat des lieux

Selon les statistiques de l'état civil, le taux de mortalité infantile est ainsi passé de 51,9 décès pour 1 000 enfants nés vivants en 1950 à 4,9 et 3,8 pour 1 000 en 1995 et 2005. La mortalité néonatale est passée de 3,0 en 1995 à 2,5 pour 1 000 enfants nés vivants en 2005. La diminution observée au cours des années 80 concerneraient principalement les décès survenant au cours de la première semaine de vie, dite période néonatale précoce (19). Le décès du nouveau-né âgé de moins d'une semaine serait principalement influencé par les complications pendant la grossesse ou lors de l'accouchement, et par les malformations congénitales (20). À la fin des années 1990, c'est grâce à la forte réduction des décès entre le premier mois, et la première année, c'est-à-dire mortalité post-néonatale, que la mortalité infantile s'abaisse à 4 décès pour 1 000 naissances vivantes. Les campagnes de prévention préconisant de coucher les bébés sur le dos ont en effet fait chuter le nombre de morts subites du nourrisson ces années-là, en France comme dans les autres pays européens (16). Entre 2005 et 2008, la mortalité infantile et néonatale stagne en France métropolitaine à 3,8 et 2,5 respectivement décès pour 1 000 naissances vivantes. En 2008, elles augmentent très légèrement, atteignent 3,9 et 2,6 décès pour 1 000 naissances vivantes et diminuent en 2010 pour se situer depuis sur un nouveau plancher, autour de 3,6 enfants décédés à moins d'un an et 2,5 enfants décédés de moins de 1 mois pour 1 000 naissances vivantes (Figure 1).

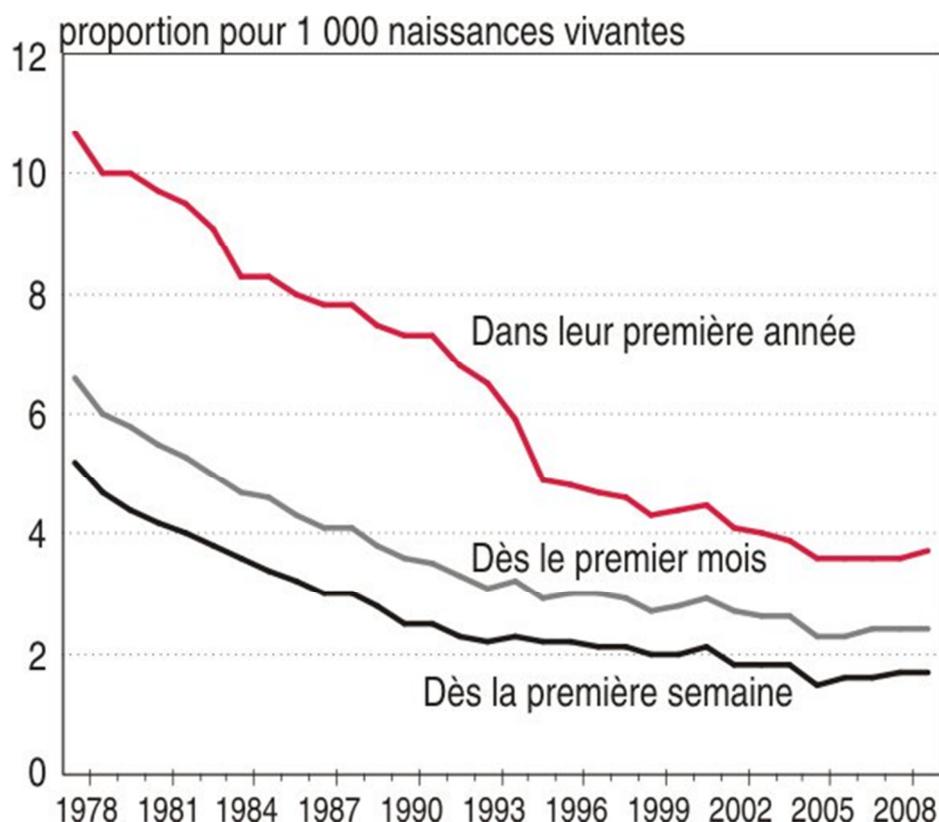


Figure 1 : Taux de mortalité avant 1 an, 1 mois, 1 semaine en France

Champ : France métropolitaine.

Source : Insee, statistiques de l'état civil.

2. Des disparités territoriales

Parmi les pays européens

Sur la période où la mortalité infantile a stagné en France, elle a continué de se réduire dans d'autres pays européens. Dans le classement européen des plus faibles taux de mortalité infantile, la France (France métropolitaine + DOM) occupait le 7^e rang en 1999 juste derrière les pays nordiques ; elle se situe en 2009 à la 20^e place d'après les statistiques d'Eurostat portant sur les 27 pays de l'union européenne, auxquels nous avons ajouté l'Islande, la Norvège et la Suisse. La France métropolitaine, sans les DOM, occupait la 6^{ème} place en 1999 et est passée à la 17^{ème} en 2009. La plupart des pays de l'Union européenne ont un taux de mortalité infantile très faible, compris entre 3 et 4 pour 1000 naissances vivantes. La Suède et la Finlande ont des taux inférieurs avec respectivement 2,5 et 2,6‰. Parmi les pays les plus riches, le Royaume-Uni est dans la plus mauvaise situation, avec un taux atteignant 4,7‰.

Avec 3,8‰, la France ne figure pas parmi les pays les mieux lotis. À l’opposé, les taux de mortalité infantile étaient environ quatre fois supérieurs en Roumanie (10,1‰) et en Bulgarie (9,0‰) (Figure 2). Pour rester à la 7ème place, il aurait fallu que la France affiche en 2009 un taux de mortalité infantile de 3,0 décès pour 1000 naissances vivantes, soit environ 500 décès de moins. Si la France avait le taux de mortalité de l’Islande (1,8 en 2009, soit deux fois moins qu’en France), il y aurait 1 500 décès d’enfants de moins d’un an en moins chaque année.

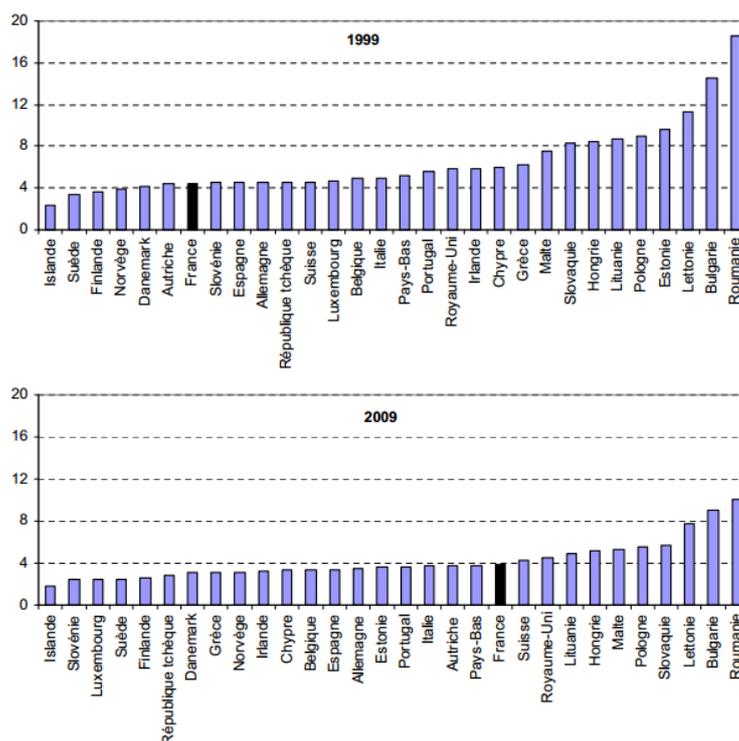


Figure 2 : Évolution du rang occupé par la France dans le classement européen des plus bas taux de mortalité infantile dans 30 pays européens - 1999-2009

Champ pour la France : France métropolitaine+DOM

Source : Eurostat

En France

Au sein même du pays, la France est caractérisée par des disparités territoriales de mortalité infantile et néonatale. En 2000, à l’échelle de la région, on compte 3,3 décès d’enfants de moins de 1 an dont 2,1 moins de 1 mois pour 1000 nés vivants en Basse Normandie, 5,4 moins de 1 an dont 3,6 moins d’1 mois pour la Lorraine. En 2009, on compte 1,5 moins de 1 mois en Limousin et 5,0 de moins de 1 an dont 3,3 de moins de 1 mois en Alsace (Figure 3).

La mortalité infantile régionale a diminué en 10 ans mais les disparités régionales persistent. Les régions qui ont les taux les plus élevés entre 2000 et 2009 sont la Champagne –Ardenne, Lorraine et l’Alsace. Les « bons élèves » sont les régions Poitou-Charentes, Pays de la Loire, Auvergne et Bourgogne.

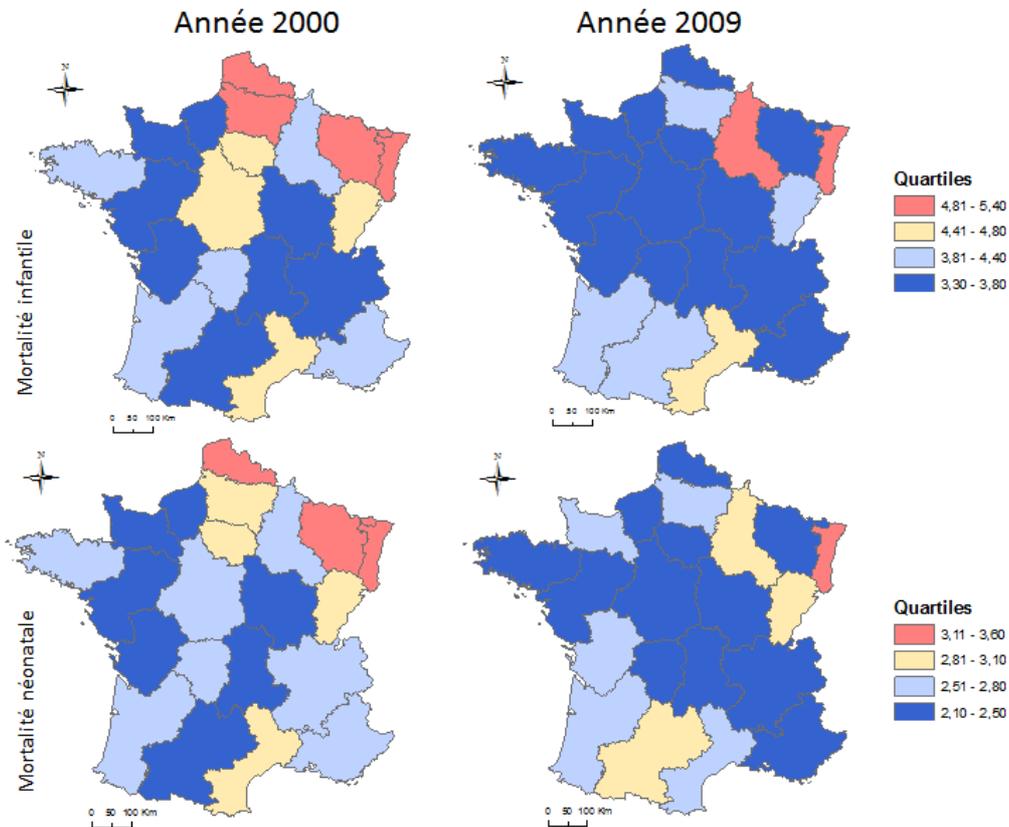


Figure 3: Carte des taux de mortalité infantile et néonatale en 2000 et 2009 par région en France

Source : www.ecosante.fr

Données : [Insee](http://www.insee.fr)

En terme d'évolution, le graphique de la figure 4 montre bien que la région Ile de France et Nord pas de Calais ont vu leur taux de mortalité infantile et néonatale diminuer sur la période alors que les régions Rhône Alpes et Provence Alpes Côte d'Azur montrent une légère augmentation des taux depuis l'année 2005. En termes d'évolution sur la France entière, les régions qui n'ont pas baissé leur taux sur la période, voire augmenter sont la Champagne-Ardenne, Basse Normandie, Midi-Pyrénées, à l'inverse les réductions les plus importantes sont pour les régions Centre, Nord Pas de Calais, Lorraine et le Limousin.

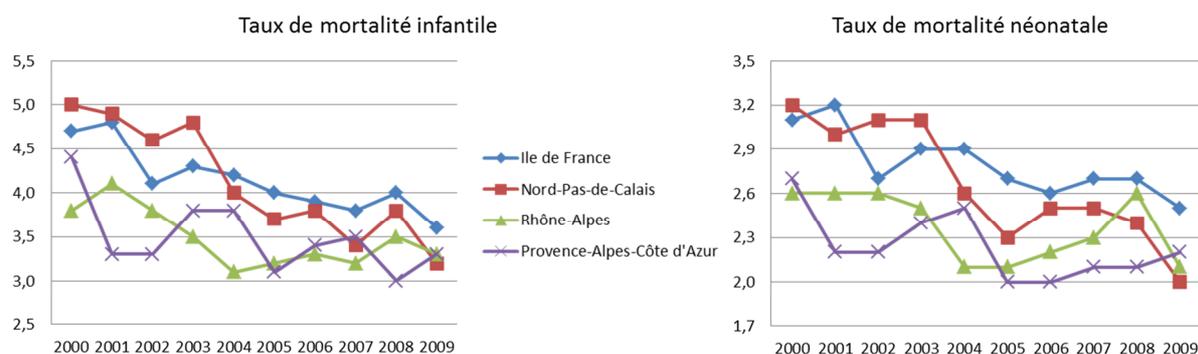


Figure 4: Evolution de la mortalité infantile et néonatale entre 2000-2009 pour les 4 agglomérations de l'étude.

Source : Eco-Santé Régions & Départements 2013
Données : INSEE

II. LES INÉGALITÉS SOCIALES DE SANTÉ LIÉES À LA MORTALITÉ INFANTILE ET NÉONATALE

En dépit du faible taux de mortalité infantile dans les pays développés, les inégalités sociales liées à la mortalité infantile et néonatale persistent. Les inégalités sociales de santé constituent une réalité objectivée par des données socio-épidémiologiques robustes bien connues de la littérature.

Plusieurs facteurs, identifiés depuis longtemps, influencent la mortalité infantile et néonatale, parmi lesquels des caractéristiques liées à l'enfant comme le poids à la naissance, l'âge gestationnel, la multiplicité ou le genre. Il existe une surmortalité masculine à la naissance. (21–24). Les autres facteurs de risques sont liés aux caractéristiques des parents (situation socio-économique, ou style de vie), à leur environnement de vie (défaveur sociale contextuelle et exposition environnementale, par exemple) et au système de soins (14,25,26). Néanmoins, les déterminants de la mortalité infantile et néonatale apparaissent différents. Alors que la première est plus sensible aux déterminants de la mortalité post-néonatale, décrite comme un indicateur des conditions de vie et de la qualité de l'environnement immédiat du nouveau-né, la deuxième est plus sensible au bon déroulement de la grossesse, à la qualité de la vie et de l'environnement qui entoure la femme enceinte (27).

A. Des déterminants socio-économiques nombreux

En termes de facteurs socio-économiques, les déterminants les plus souvent rapportés sont : l'éducation, la profession, le revenu et le chômage. L'éducation reflète une position sociale, environnementale et culturelle (21). L'âge de la mère est un déterminant important à prendre en compte avec un risque plus élevé de décès pour l'enfant né d'une mère de plus de 35 ans ou au contraire pendant l'adolescence (28,26). De plus, les mères en âge de procréer sont susceptibles d'avoir fini leurs études même si elles n'ont pas d'emploi stable. Des importantes revues de la littérature ont mis en évidence des taux de mortalité infantile, néonatale et postnéonatale plus élevés pour les femmes avec un niveau d'éducation faible dans les pays du Nord : Danemark, Suède, Finlande (29,22). La profession est également un facteur important à considérer ; un travail pénible combiné avec une exposition chimique ou biologique peut par exemple altérer la reproduction et avoir un impact sur les issues de grossesse (30,25,14). Le chômage est aussi un facteur de risque connu pour produire une situation de précarité et de stress, néfaste pour la mère et l'enfant (31). Le revenu est largement étudié dans la littérature, il est considéré comme une mesure de la défaveur matérielle. Les ressources économiques et matérielles peuvent influencer la santé d'une personne dans le sens où cela affecterait la qualité de la nourriture, la qualité du logement, le repos instauré par des vacances, la proximité des espaces verts, une mauvaise assiduité aux examens de santé (32).

Le lien par lequel les déterminants sociaux influent la santé du nouveau-né peut découler d'un facteur moins conventionnel mais bien réel qui est le stress. La littérature décrit deux types de stress de la mère, le stress nutritionnel et le stress psychosocial (33). Le stress peut interagir en complément d'un autre facteur ou découler d'un mal causé par d'autres facteurs (34,35). Ce facteur psychosocial peut être émotionnel (anxiété, déprime) ou/et lié au cadre de vie (lié à la profession, au revenu, au rythme de vie des grandes villes), et peut engendrer des dommages et des issues de grossesse défavorables de par la relation étroite qui relie la mère au fœtus (36–41).

L'accès et la qualité des soins primaires sont reconnus également comme des déterminants du décès du nouveau-né. De la littérature émerge deux mécanismes à travers lesquels la qualité des soins primaires pendant la grossesse peuvent conduire à des issues de grossesses

défavorable(15). Le premier mécanisme est un accès aux systèmes de santé approprié et un suivi rigoureux tout au long de la grossesse (42) afin d'identifier à temps les infections et de les traiter rapidement. De plus, un accès tardif ou non rigoureux aux soins prénataux peut en interaction avec d'autres facteurs conduire à un faible poids de naissance, une naissance prématurée, eux-mêmes principales causes de mortalité infantile et néonatale (43–46). Le cadre de vie des femmes enceintes peut en partie expliquer un accès tardif au système de soins. En effet, les causes peuvent être un faible nombre de médecins généralistes ou gynécologues proches du ménage ; un temps d'attente trop important peut également décourager la femme enceinte à se déplacer (47,32). L'enquête nationale périnatale en France en 2010, a démontré que la surveillance prénatale et l'issue de la grossesse étaient différentes selon l'accessibilité économique (défavorable pour les femmes ayant des ressources provenant d'une allocation chômage, de l'API (allocation parents isolés), du RMI (revenu minimum d'insertion), ou du RSA (revenu solidarité active), ou encore n'ayant aucune ressource (32)). Elles ont moins de consultations prénatales que les autres femmes, mais sont plus souvent hospitalisées pendant la grossesse, ce qui suggère des problèmes de santé plus fréquents (32).

Le deuxième mécanisme lié aux soins primaires est l'état de santé et certains comportements non hygiéniques de la mère lors de la grossesse (48,49,43). Les problèmes de santé de la mère combinés à certains facteurs de risques individuels augmentent le taux de mortalité infantile ; il s'agit des habitudes tabagiques (50) de la consommation d'alcool (51) ou de drogue (52) mais également le fait ne pas manger équilibré ou perdre trop de poids pendant la grossesse (33,53,14). Les principales maladies de la mère constituant des facteurs à risque pour l'enfant sont l'obésité, le diabète, les maladies sexuellement transmissibles (54,15).

B. Des déterminants contextuels

Une variabilité importante de la répartition des catégories sociales au sein d'un territoire est constatée quelle que soit l'échelle géographique considérée : régionale, départementale, communale ou à l'échelle fine du quartier. L'implantation différentielle des populations selon leur profil socio-économique résulte de processus complexes qui ont contribué au façonnage des territoires. Lorsque l'analyse porte sur des unités spatiales de petite taille, telles que les

IRIS (Ilots Regroupés pour l'Information Statistiques), les caractéristiques socio-économiques du territoire sont proches des profils des personnes qui y résident (55). La caractérisation et la typologie des quartiers en site urbain deviennent ainsi un enjeu primordial pour identifier (surveillance) et comprendre (recherche) les inégalités sociales de santé. Des études écologiques de type géographique se sont intéressées à répondre à la question suivante : quels sont les déterminants qui peuvent influencer sur le fait que certaines zones du territoire ont des taux élevés de mortalité infantile et d'autres non ?

Deux voies sont empruntées pour investiguer l'analyse des relations entre la mortalité infantile/ néonatale et la défaveur socio-économique du quartier. Tout d'abord un indicateur global de défaveur, ou indicateur composite intégrant de multiples domaines de la défaveur comme par exemple le revenu, la profession, l'éducation ou la qualité du logement. De nombreux travaux ont mis en évidence une augmentation du risque avec le degré de défaveur du quartier (56–58). Les indices composite de défaveur les plus couramment utilisés sont ceux de Townsend et de Carstairs qui relatent une défaveur matérielle (56). D'autres indices plus récents et plus complets en termes de nombre de dimensions socio-économiques pris en compte ont également démontré une association avec la mortalité infantile (59,60,58). Plusieurs études dans des pays industrialisés ont confirmé l'existence d'un gradient social : plus la défaveur du quartier est élevée et plus le risque de mortalité infantile sera élevé (24,59,61–63) La deuxième voie est l'utilisation d'indicateur dit de « pauvreté du quartier », déterminants important des taux de mortalité ; citons l'index de Gini et le taux de pauvreté. L'index de Gini est principalement utilisé dans les pays moyennement développés jusqu'aux plus riches, qui mesure la distribution non équitable des revenus. Les pays européens mesurent la pauvreté monétaire de manière relative (le seuil de 60% du niveau de vie médian) alors que d'autres pays (comme les États-Unis ou l'Australie) ont une approche absolue (la proportion de la population qui vit avec moins de 1 dollar par jour). Les études montrent qu'actuellement dans les pays moyennement développés, l'index de Gini explique en grande partie la variabilité des taux à l'échelle macro (entre pays) alors que pour les pays les plus développés d'autres critères sont nécessaires (64,65). Des études plus récentes ont démontré que les pays avec un fort taux de mortalité infantile avaient de fortes inégalités en termes de revenu (64,66,67).

De plus, des études ont illustré le fait que l'assiduité à se faire soigner et l'adhésion aux mesures préventives et hygiéno-diététiques majeures dépendent du niveau de la défaveur de la zone de résidence et des caractéristiques culturelles (68,15). De même, certains critères sont des bons proxy de l'accès aux systèmes de soin comme l'éloignement du domicile de la mère à la maternité la plus proche (47), vivre dans une zone urbaine ou rural (69) ou plus communément, la taille de la commune (62).

III. L'ENVIRONNEMENT, UN DETERMINANT SUSPECTÉ

Le terme « environnement » est utilisé dans son acception matérielle (la qualité chimique et physique des milieux de vie), et cet environnement « physique » est très immergé dans son milieu social. L'estimation de la contribution des expositions environnementales aux inégalités sociales de santé représente à l'heure actuelle un enjeu de recherche majeur en santé publique, en raison du caractère évitable et injuste de cette situation.

A. Mécanismes par lesquels peuvent opérer les expositions environnementales dans les ISS

De nombreuses recherches visant à démontrer des liens entre environnement et inégalités sociales liées à une maladie se sont développées ces dernières décennies. Elles ont, entre autres, mis en évidence deux mécanismes par lesquels les expositions environnementales pourraient contribuer à générer des inégalités sociales de santé (70): le **différentiel d'exposition** et le **différentiel de susceptibilité**. Ces deux mécanismes peuvent opérer de manière indépendante ou synergique. Ainsi les populations socio-économiquement défavorisées pourraient être à la fois davantage exposées à un plus grand nombre de nuisances et/ou à des niveaux d'expositions plus élevés (**différentiel d'exposition**) et/ou les populations plus défavorisées pourraient être plus vulnérables aux effets de l'environnement (**différentiel de vulnérabilité/susceptibilité**) en raison d'un état de santé plus fragile que les populations plus favorisées. L'hypothèse sous-jacente serait que les populations défavorisées pourraient être exposées à un plus grand nombre de nuisances et/ou à des niveaux d'exposition plus élevés. Dans les paragraphes qui vont suivre nous présenterons de manière très synthétique ces deux mécanismes.

1. Différentiel d'exposition

L'existence de disparités territoriales des expositions environnementales selon le niveau socioéconomique serait contraire au principe de « justice environnementale » ou d'« équité environnementale », qui stipule qu'aucun groupe de population ne doit supporter une part disproportionnée des expositions. Ces préoccupations qui ont d'abord émergé aux Etats-Unis et au Canada (71,72) sous la pression des minorités ethniques, dans le cadre du vaste mouvement politique d'émancipation des noirs et autres minorités, et se développent progressivement en Europe depuis les années 2000 (73–77) plutôt sous l'angle de la précarité sociale et des expositions aux risques environnementaux. Selon l'OCDE (Organisation de coopération et de développement économiques) (78), la problématique des inégalités environnementales et de ses conséquences reste insuffisamment documentée en Europe.

Un grand nombre d'études aux Etats-Unis et au Canada ont démontré que les personnes les plus défavorisées sont fréquemment plus exposées à la pollution atmosphérique liée au trafic (axes routiers avec un important flux de voitures) ou aux émissions industrielles (complexes industriels, usines de traitement des déchets, décharges (72,77,79–87). C'est seulement récemment que la question de l'inégale répartition de la pollution parmi les populations de statut socio-économique différent a été abordée en Europe, pour exemple aux Pays-Bas (88), en Finlande (89) , en Suede (76), en Allemagne (90), au Royaume-Uni (91,92), en 'Italie (93) et en France (94,95).

La majorité des études Américaines démontrent que les plus défavorisés se retrouvent en situation de proximité ou d'exposition avec des agents physiques dangereux ou potentiellement dangereux. En Europe, les premiers résultats sont davantage controversés (79), notamment pour la qualité de l'air. En effet, alors que certaines études rapportent que les populations à faible statut socio-économique sont plus exposés aux polluants atmosphériques (88,96), d'autres conduites dans de grandes agglomérations marquées par un intense trafic automobile comme Rome, trouvent des résultats inverses (93). A Strasbourg en France, une étude de Havard et al a démontré que les classes moyennes étaient les plus exposées à la pollution issue du trafic routier (95), situation que l'on retrouve également dans

l'agglomération de Lyon (60). Ces contrastes tiennent à l'histoire de la construction urbaine, propre à chaque territoire.

La réduction considérable des pollutions industrielles historiques au cours des trente dernières années s'est accompagnée d'une croissance des émissions liées au trafic automobile, habituellement plus dense dans les hyper-centres où résident souvent, dorénavant, des catégories sociales plus aisées. Au sein d'un même pays, la géographie urbaine, fruit de l'histoire de chaque ville, se traduit également par des situations contrastées. La partie 4 du chapitre 3, sera spécifiquement dédiée aux inégalités environnementales en France.

2. Différentiel de vulnérabilité

Contrairement aux résultats des études s'intéressant aux inégalités environnementales, celles qui ont exploré l'existence d'un différentiel de vulnérabilité s'accordent à dire que, même lorsque les sujets de milieux sociaux défavorisés résident dans des secteurs urbains moins affectés par la pollution, les conséquences sanitaires de cette exposition y sont plus marquées que pour des milieux plus aisés (94). Cela résulte de divers processus qui conjuguent leurs effets : une plus grande vulnérabilité (état de santé précaire, comportements peu « hygiéniques », moindre recours aux soins), la possibilité plus réduite d'échapper aux effluves des pots d'échappement (le temps d'un week-end ou des congés), la qualité souvent plus dégradée de l'habitat et/ou moins isolé pour faire face aux polluants extérieurs et des conditions de travail plus difficiles des personnes d'origine modeste.

La littérature s'est encore peu intéressée au différentiel de vulnérabilité et le nombre d'études ayant testé cette hypothèse reste relativement limitée. Les travaux existants provenant du Canada (97), de la Norvège (98), du Brésil (99) ou des Etats-Unis ont constaté que le statut socio-économique modifiait la sensibilité de la population face aux effets de la pollution, avec des problèmes de santé plus fréquents dans des populations résidant dans les quartiers les plus défavorisés même s'ils ne sont pas plus exposés à la pollution (100). Cela résulterait d'une plus grande vulnérabilité liée notamment à un état de santé souvent plus dégradé, à des conditions de vie plus difficiles et à un plus faible recours au soin.

En France, à Strasbourg par exemple, ce sont les populations résidant dans les Iris les plus défavorisés qui présentent un risque d'infarctus du myocarde accru, alors que les concentrations en PM₁₀ n'y sont pas les plus élevées (101). À Bordeaux, Filleul *et al.* (102) ont montré une association significative entre la mortalité, dans une population âgée de plus de 65 ans, et l'exposition aux fumées noires chez les ouvriers uniquement. Quelques études françaises ne mettent cependant pas en évidence de modification de l'effet sanitaire lié à l'exposition à la pollution atmosphérique par les caractéristiques socioéconomiques des populations (103,104). Ceci montre bien toute la complexité des mécanismes sous-jacents à la modification de l'effet de la pollution atmosphérique, et sans d'autre d'autres facteurs environnementaux, par le niveau socioéconomique, et la nécessité de conduire davantage d'études sur ce sujet.

IV. **LE PROJET DE RECHERCHE EQUIT'ARÉA**

C'est dans ce contexte que le projet Equit'Area a été mis en place. Son objectif général est d'étudier la contribution des expositions environnementales et d'un cumul de celles-ci aux inégalités sociales de santé sur le territoire national. Il fait converger deux champs de recherche habituellement explorés de manière indépendante: l'épidémiologie environnementale et l'épidémiologie sociale. Le programme de recherche est structuré autour de trois étapes complémentaires : (i) la caractérisation des inégalités environnementales, (ii) la caractérisation des inégalités sociales de santé et (iii) la contribution des expositions environnementales aux inégalités sociales de santé.

Ce projet s'inscrit dans une double perspective. L'hypothèse d'une distribution inégale du « fardeau » environnemental selon le profil socio-économique des unités territoriales est en soi original puisque les travaux en France sur le sujet sont encore peu nombreux. La loi de santé publique de 2004, le PNSE2 ainsi que le Grenelle de l'environnement en font un axe prioritaire de recherche et de politique publique. Plus novateur encore est l'objectif de donner à voir ce phénomène dans une gamme de grandes agglomérations urbaines, pour en faire un outil d'analyse des politiques d'aménagement du territoire à la disposition des différentes parties intéressées. La seconde perspective est une vision de santé publique. Ce projet constitue une étape pour explorer l'impact sanitaire des inégalités d'expositions cumulées à des nuisances environnementales, exploration qui pourrait se déployer, après cette phase

pilote, à une échelle nationale grâce au formidable développement des bases d'information statistique socio-démographiques, économiques, environnementales et sanitaires, et aux outils nouveaux d'analyse géographique. Il s'agit ici d'une perspective épidémiologique, comme moyen d'accès à la connaissance des risques environnementaux et d'aide à l'élaboration des politiques publiques de santé et d'environnement

V. **OBJECTIFS ET STRUCTURE DE LA THÈSE**

Dans le cadre de ce projet, mon travail de thèse a, plus précisément, pour objectif d'étudier la contribution du dioxyde d'azote sur les inégalités sociales de santé liées à la mortalité infantile et néonatale sur 4 agglomérations françaises Lille, Lyon, Paris et Marseille. L'originalité de mon étude dans ce projet de recherche concerne la compréhension des inégalités sociales de santé par une approche spatiale et temporelle et une meilleure compréhension des mécanismes par lequel l'exposition au dioxyde d'azote contribue aux inégalités sociales de santé.

Dans cette perspective, les principaux objectifs de ma thèse sont les suivants

- (1) décrire les disparités socio-territoriales d'exposition au dioxyde d'azote et des évènements sanitaires (mortalité infantile et néonatale).
- (2) étudier les inégalités socio-spatiales de la mortalité infantile et néonatale. L'objectif de cette étape est d'examiner la répartition du risque de mortalité selon la défaveur du quartier.
- (3) étudier les inégalités environnementales afin d'analyser la distribution spatiale de l'exposition au dioxyde d'azote et la situation socio-économique du quartier de résidence (différentiel d'exposition).
- (4) étudier la relation entre ces évènements de santé, l'exposition environnementale et la défaveur socio-économique. L'objectif est tout d'abord analyser si à exposition au dioxyde d'azote égale, on retrouve une

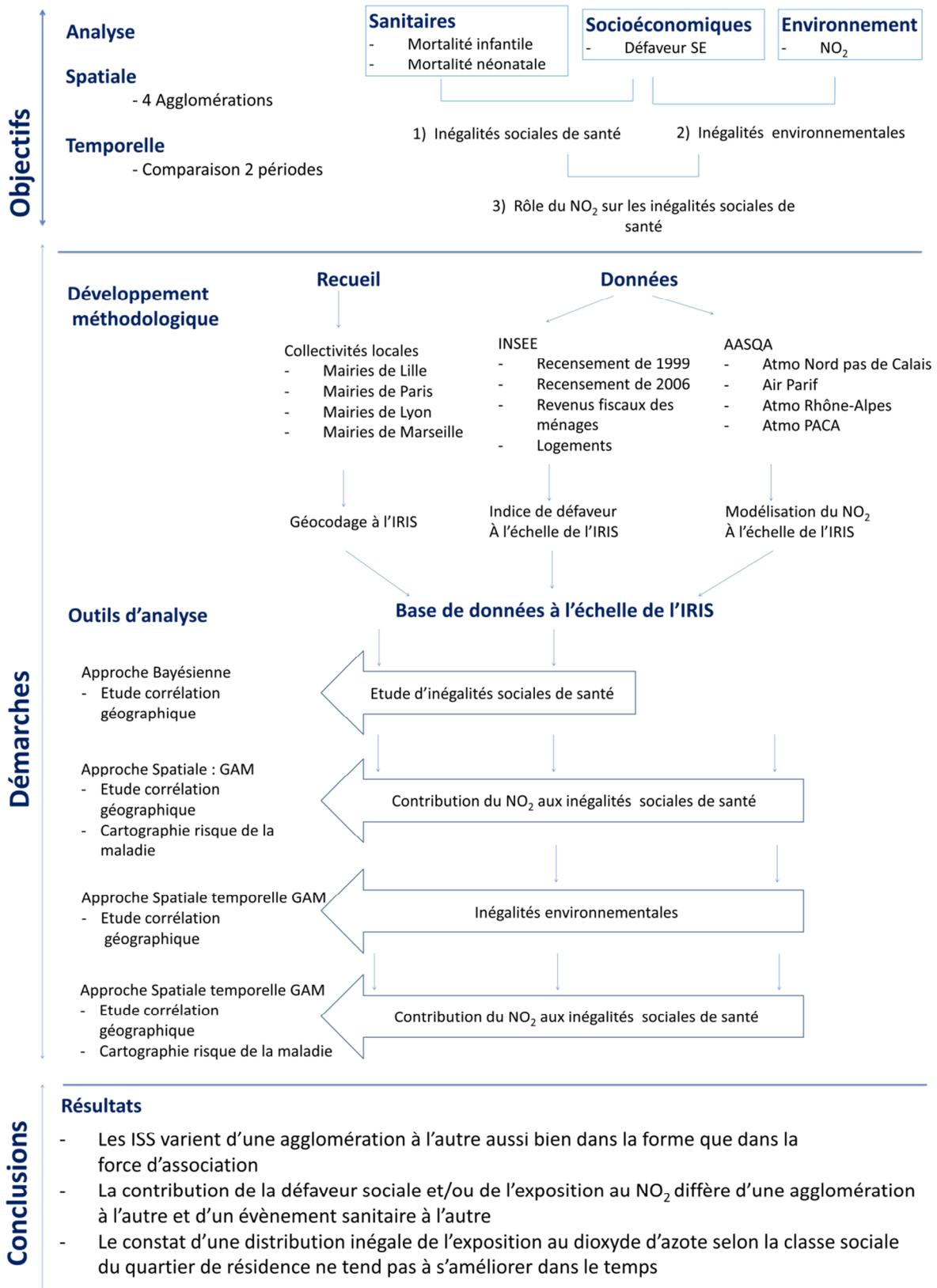
relation entre les évènements de santé et la défaveur du quartier (différentiel de sensibilité). Ensuite, analyser si à défaveur égale, on retrouve une relation entre les évènements de santé et l'exposition au dioxyde d'azote (différentiel de sensibilité)

La première étape se veut méthodologique, présentant et justifiant les méthodes mises en œuvre pour cette thèse. Plus précisément, nous justifierons le choix de la zone d'étude et de l'unité géographique retenue. Nous exposerons la procédure de recueil des données sanitaires. Enfin, les données socio-économiques et environnementales seront définies et le choix des méthodes statistiques adaptées au design et aux objectifs de l'étude seront présentés.

La seconde étape est la présentation des travaux empiriques. Ce chapitre sera présenté sous la forme de trois articles dans le cadre de cette thèse. Chaque partie commence par une présentation synthétique de l'article, suivie du texte intégral et des résultats complémentaires. Le premier article présente les inégalités sociales de santé. Le deuxième article présente une analyse spatiale du rôle du dioxyde d'azote sur les inégalités sociales de santé. Le troisième article présente une analyse spatio-temporelle des inégalités environnementales. Et enfin, une quatrième partie présente une analyse spatio-temporelle du rôle du dioxyde d'azote sur les inégalités sociales de santé.

Finalement, le quatrième et dernier chapitre sera consacré aux contributions aux autres travaux développés au sein du projet Equit'Area, et plus particulièrement, aux études intégrant d'autres expositions environnementales dites « positives » (les espaces verts) ou nuisances environnementales (industries polluantes et bruit) sur les ISS, auxquelles j'ai contribué. Une discussion générale sera engagée ; elle visera à apprécier les forces et les faiblesses de ce projet et à présenter les perspectives et hypothèses de recherche pouvant poursuivre ce travail ainsi que sa contribution à la décision publique en matière de gestion du risque de décès infantile en lien avec le contexte environnemental et social.

Schéma conceptuel de la thèse



Références : Chapitre 1

1. OMS | Principaux concepts relatifs aux déterminants sociaux de la santé [Internet]. WHO. Available from: http://www.who.int/social_determinants/thecommission/finalreport/key_concepts/fr/
2. Moleux, M, Schaetzel, F, Scotton, C. Les inégalités sociales de santé : Déterminants sociaux et modèles d'action. IGAS, Inspection générale des affaires sociales; 2011 mai. Report No.: N°RM2011-061P.
3. Leclerc, A, Kaminski, M, Lang, T. Inégaux face à la santé - Du constat à l'action. INSERM; 2008.
4. Haut conseil de la santé publique. Les inégalités sociales de santé : sortir de la fatalité. HCSP; 2009.
5. Rollet Echalié, C. La politique à l'égard de la petite enfance sous la IIIe République. PUF. 1990. p. 200.
6. 4. OUI à la réduction de la mortalité infantile | HuitFoisOui [Internet]. Available from: <http://www.huitfoisoui.fr/objectif/4-oui-a-la-reduction-de-la-mortalite-infantile/>
7. Smith GC, Fretts RC. Stillbirth. *Lancet*. 2007 Nov 17;370(9600):1715–25.
8. Reidpath DD, Allotey P. Infant mortality rate as an indicator of population health. *J Epidemiol Community Health*. 2003 mai;57(5):344–6.
9. Fantini MP, Stivanello E, Dallolio L, Loghi M, Savoia E. Persistent geographical disparities in infant mortality rates in Italy (1999-2001): comparison with France, England, Germany, and Portugal. *Eur J Public Health*. 2006 Aug;16(4):429–32.
10. Korkodilos M, Thorne A, Earwicker R. The Infant Mortality National Support Team. *Perspectives in Public Health*. 2009 Sep 1;129(5):214–5.
11. Murray CJ, Salomon JA, Mathers C. A critical examination of summary measures of population health. *Bull World Health Organ*. 2000;78(8):981–94.
12. Blaxter M, Social Science Research Council (Great Britain), Great Britain. Dept. of Health and Social Security. The health of children a review of research on the place of health in cycles of disadvantage [Internet]. 1981. Available from: <http://catalog.hathitrust.org/api/volumes/oclc/11260413.html>
13. Barker DJP. The origins of the developmental origins theory. *J Intern Med*. 2007 May;261(5):412–7.
14. De Graaf JP, Steegers EAP, Bonse GJ. Inequalities in perinatal and maternal health. *Curr Opin Obstet Gynecol*. 2013 Apr;25(2):98–108.
15. Johnson TS, Malnory ME, Nowak EW, Kelber S. Using Fetal and Infant Mortality Reviews to improve birth outcomes in an urban community. *J Obstet Gynecol Neonatal Nurs*. 2011 Feb;40(1):86–97.

16. Kinney HC, Thach BT. The sudden infant death syndrome. *N Engl J Med.* 2009 Aug 20;361(8):795–805.
17. Hollowell J, Oakley L, Kurinczuk JJ, Brocklehurst P, Gray R. The effectiveness of antenatal care programmes to reduce infant mortality and preterm birth in socially disadvantaged and vulnerable women in high-income countries: a systematic review. *BMC Pregnancy Childbirth.* 2011;11:13.
18. Hollowell J, Kurinczuk JJ, Brocklehurst P, Gray R. Social and ethnic inequalities in infant mortality: a perspective from the United kingdom. *Semin Perinatol.* 2011 Aug;35(4):240–4.
19. Barbieri, M, Toulemon,L. Les enfants tous égaux devant la mort? Problèmes d'observation et de mesures des différences sociales de la mortalité infantile en France. *Histoires de familles, histoires familiales.* 2005;(156):407–22.
20. Blondel, B, Eb,M, Matet, N, Bréart, G, Jouglà, E. Apport du certificat de décès néonatal à la connaissance de la mortalité en France. *Bulletin Epidémiologique Hebdomadaire.* 2006. p. 29–30.
21. Arntzen A, Mortensen L, Schnor O, Cnattingius S, Gissler M, Andersen A-MN. Neonatal and postneonatal mortality by maternal education--a population-based study of trends in the Nordic countries, 1981-2000. *Eur J Public Health.* 2008 Jun;18(3):245–51.
22. Arntzen A, Samuelsen SO, Bakketeig LS, Stoltenberg C. Socioeconomic status and risk of infant death. A population-based study of trends in Norway, 1967-1998. *Int J Epidemiol.* 2004 Apr;33(2):279–88.
23. Sovio U, Dibden A, Koupil I. Social determinants of infant mortality in a historical Swedish cohort. *Paediatr Perinat Epidemiol.* 2012 Sep;26(5):408–20.
24. Smith LK, Manktelow BN, Draper ES, Springett A, Field DJ. Nature of socioeconomic inequalities in neonatal mortality: population based study. *BMJ.* 2010;341:c6654.
25. Rocheleau CM, Lawson CC, Whelan EA, Rich-Edwards JW. Shift work and adverse pregnancy outcomes: comments on a recent meta-analysis. *BJOG.* 2012 Feb;119(3):378; author reply 379–380.
26. Blondel B, Bréart G. Mortinatalité et mortalité néonatale. *EMC - Pédiatrie.* 2004 Feb;1(1):97–108.
27. Danet, S. Les disparités régionales de santé. A partir des indicateurs associés à la loi relative à la politique de santé publique. *DREES;* 2006 p. 46.
28. Arntzen A, Nybo Andersen AM. Social determinants for infant mortality in the Nordic countries, 1980-2001. *Scand J Public Health.* 2004;32(5):381–9.
29. Mortensen LH, Helweg-Larsen K, Andersen A-MN. Socioeconomic differences in perinatal health and disease. *Scand J Public Health.* 2011 Jul;39(7 Suppl):110–4.
30. Bonzini M, Palmer KT, Coggon D, Carugno M, Cromi A, Ferrario MM. Shift work and pregnancy outcomes: a systematic review with meta-analysis of currently available epidemiological studies. *BJOG.* 2011 Nov;118(12):1429–37.

31. Congdon P, Southall H. Small area variations in infant mortality in England and Wales in the inter-war period and their link with socio-economic factors. *Health Place*. 2004 Dec;10(4):363–82.
32. Blondel, B, Kermarrec, M. Les naissances en 2010 et leur évolution depuis 2003. Enquête nationale périnatale 2010. Paris: INSERM, DGS, DREES; 2011 Mai p. 117.
33. Pike IL. Maternal stress and fetal responses: evolutionary perspectives on preterm delivery. *Am J Hum Biol*. 2005 Feb;17(1):55–65.
34. Rich-Edwards JW, Grizzard TA. Psychosocial stress and neuroendocrine mechanisms in preterm delivery. *Am J Obstet Gynecol*. 2005 May;192(5 Suppl):S30–35.
35. Kramer MR, Hogue CJ, Dunlop AL, Menon R. Preconceptional stress and racial disparities in preterm birth: an overview. *Acta Obstet Gynecol Scand*. 2011 Dec;90(12):1307–16.
36. Homer CJ, James SA, Siegel E. Work-related psychosocial stress and risk of preterm, low birthweight delivery. *Am J Public Health*. 1990 Feb;80(2):173–7.
37. Marmot M, Wilkinson RG. Psychosocial and material pathways in the relation between income and health: a response to Lynch et al. *BMJ*. 2001 May 19;322(7296):1233–6.
38. Neggers Y, Goldenberg R, Cliver S, Hauth J. The relationship between psychosocial profile, health practices, and pregnancy outcomes. *Acta Obstet Gynecol Scand*. 2006;85(3):277–85.
39. Tegethoff M, Greene N, Olsen J, Meyer AH, Meinlschmidt G. Maternal psychosocial adversity during pregnancy is associated with length of gestation and offspring size at birth: evidence from a population-based cohort study. *Psychosom Med*. 2010 May;72(4):419–26.
40. Littleton HL, Bye K, Buck K, Amacker A. Psychosocial stress during pregnancy and perinatal outcomes: a meta-analytic review. *J Psychosom Obstet Gynaecol*. 2010 Dec;31(4):219–28.
41. Mutambudzi M, Meyer JD, Warren N, Reisine S. Effects of psychosocial characteristics of work on pregnancy outcomes: a critical review. *Women Health*. 2011 May;51(3):279–97.
42. Hillemeier MM, Weisman CS, Chase GA, Dyer A-M, Shaffer ML. Women's preconceptional health and use of health services: implications for preconception care. *Health Serv Res*. 2008 Feb;43(1 Pt 1):54–75.
43. Shi L, Macinko J, Starfield B, Xu J, Regan J, Politzer R, et al. Primary care, infant mortality, and low birth weight in the states of the USA. *J Epidemiol Community Health*. 2004 May;58(5):374–80.
44. Macinko J, Starfield B, Shi L. The contribution of primary care systems to health outcomes within Organization for Economic Cooperation and Development (OECD) countries, 1970-1998. *Health Serv Res*. 2003 Jun;38(3):831–65.

45. Macinko JA, Shi L, Starfield B. Wage inequality, the health system, and infant mortality in wealthy industrialized countries, 1970-1996. *Soc Sci Med.* 2004 Jan;58(2):279–92.
46. Macinko J, Starfield B, Shi L. Quantifying the health benefits of primary care physician supply in the United States. *Int J Health Serv.* 2007;37(1):111–26.
47. Pilkington H, Blondel B, Papiernik E, Cuttini M, Charreire H, Maier RF, et al. Distribution of maternity units and spatial access to specialised care for women delivering before 32 weeks of gestation in Europe. *Health Place.* 2010 May;16(3):531–8.
48. Moos MK. Preconceptional wellness as a routine objective for women’s health care: an integrative strategy. *J Obstet Gynecol Neonatal Nurs.* 2003 Aug;32(4):550–6.
49. Moos M-K. Preconceptional health promotion: progress in changing a prevention paradigm. *J Perinat Neonatal Nurs.* 2004 Mar;18(1):2–13.
50. Gray R, Bonellie SR, Chalmers J, Greer I, Jarvis S, Kurinczuk JJ, et al. Contribution of smoking during pregnancy to inequalities in stillbirth and infant death in Scotland 1994-2003: retrospective population based study using hospital maternity records. *BMJ.* 2009 Oct 1;339:b3754.
51. O’Leary CM, Nassar N, Kurinczuk JJ, Bower C. The effect of maternal alcohol consumption on fetal growth and preterm birth. *BJOG.* 2009 Feb;116(3):390–400.
52. Minnes S, Lang A, Singer L. Prenatal tobacco, marijuana, stimulant, and opiate exposure: outcomes and practice implications. *Addict Sci Clin Pract.* 2011 Jul;6(1):57–70.
53. Northstone K, Emmett P, Rogers I. Dietary patterns in pregnancy and associations with socio-demographic and lifestyle factors. *Eur J Clin Nutr.* 2008 Apr;62(4):471–9.
54. Weck RL, Paulose T, Flaws JA. Impact of Environmental Factors and Poverty on Pregnancy Outcomes. *Clinical Obstetrics and Gynecology.* 2008 Jun;51(2):349–59.
55. Greenland S. Ecologic versus individual-level sources of bias in ecologic estimates of contextual health effects. *Int J Epidemiol.* 2001 Dec;30(6):1343–50.
56. Norman P, Gregory I, Dorling D, Baker A. Geographical trends in infant mortality: England and Wales, 1970-2006. *Health Stat Q.* 2008;(40):18–29.
57. Oakley L, Maconochie N, Doyle P, Dattani N, Moser K. Multivariate analysis of infant death in England and Wales in 2005-06, with focus on socio-economic status and deprivation. *Health Stat Q.* 2009;(42):22–39.
58. Padilla C, Lalloué B, Pies C, Lucas E, Zmirou-Navier D, Séverine D. An Ecological Study to Identify Census Blocks Supporting a Higher Burden of Disease: Infant Mortality in the Lille Metropolitan Area, France. *Matern Child Health J.* 2013 Mar 12;
59. Calling S, Li X, Sundquist J, Sundquist K. Socioeconomic inequalities and infant mortality of 46,470 preterm infants born in Sweden between 1992 and 2006. *Paediatr Perinat Epidemiol.* 2011 Jul;25(4):357–65.
60. Padilla CM, Deguen S, Lalloué B, Blanchard O, Beaugard C, Troude F, et al. Cluster analysis of social and environment inequalities of infant mortality. A spatial study in

- small areas revealed by local disease mapping in France. *Sci Total Environ*. 2013 Jun 1;454-455:433–41.
61. Guildea ZE, Fone DL, Dunstan FD, Sibert JR, Cartlidge PH. Social deprivation and the causes of stillbirth and infant mortality. *ArchDisChild*. 2001 Apr;84(4):307–10.
 62. Luo Z-C, Wilkins R, Kramer MS. Effect of neighbourhood income and maternal education on birth outcomes: a population-based study. *CMAJ*. 2006 May 9;174(10):1415–20.
 63. Smith LK, Budd JLS, Field DJ, Draper ES. Socioeconomic inequalities in outcome of pregnancy and neonatal mortality associated with congenital anomalies: population based study. *BMJ*. 2011;343:d4306.
 64. Spencer N. The effect of income inequality and macro-level social policy on infant mortality and low birthweight in developed countries--a preliminary systematic review. *Child Care Health Dev*. 2004 Nov;30(6):699–709.
 65. Regidor E, Pascual C, Martínez D, Calle ME, Ortega P, Astasio P. The role of political and welfare state characteristics in infant mortality: a comparative study in wealthy countries since the late 19th century. *Int J Epidemiol*. 2011 Oct;40(5):1187–95.
 66. Regidor E, Martínez D, Santos JM, Calle ME, Ortega P, Astasio P. New findings do not support the neomaterialist theory of the relation between income inequality and infant mortality. *Social Science & Medicine*. 2012 Aug;75(4):752–3.
 67. Avendano M. Correlation or causation? Income inequality and infant mortality in fixed effects models in the period 1960–2008 in 34 OECD countries. *Social Science & Medicine*. 2012 Aug;75(4):754–60.
 68. Judge K. Inequalities in infant mortality: Patterns, trends, policy responses and emerging issues in Canada, Chile, Sweden and the United Kingdom. *Health Sociology Review*. 2009 Jun;18(1):12–24.
 69. Luo Z-C, Kierans WJ, Wilkins R, Liston RM, Mohamed J, Kramer MS. Disparities in birth outcomes by neighborhood income: temporal trends in rural and urban areas, british columbia. *Epidemiology*. 2004 Nov;15(6):679–86.
 70. Deguen S, Zmirou-Navier D. Social inequalities resulting from health risks related to ambient air quality--A European review. *Eur J Public Health*. 2010 Feb;20(1):27–35.
 71. Evans GW, Kantrowitz E. Socioeconomic status and health: the potential role of environmental risk exposure. *AnnuRevPublic Health*. 2002;23(0163-7525 (Linking)):303–31.
 72. Morello-Frosch R, Pastor M, Porras C, Sadd J. Environmental justice and regional inequality in southern California: implications for future research. *EnvironHealth Perspect*. 2002 Apr;110 Suppl 2(0091-6765 (Linking)):149–54.
 73. Brainard JS, Jones AP, Bateman IJ, Lovett AA, Fallon PJ. Modelling environmental equity: access to air quality in Birmingham, England. *Environment and Planning A*. 2002;34(4):695–716.

74. Mitchell G, Dorling D. An environmental justice analysis of British air quality. *Environment and Planning A*. 2003;35(5):909–29.
75. Mitchell G. Forecasting environmental equity: air quality responses to road user charging in Leeds, UK. *J Environ Manage*. 2005 Nov;77(3):212–26.
76. Chaix B, Gustafsson S, Jerrett M, Kristersson H, Lithman T, Boalt A, et al. Children's exposure to nitrogen dioxide in Sweden: investigating environmental injustice in an egalitarian country. *J Epidemiol Community Health*. 2006 Mar;60(3):234–41.
77. Norton JM, Wing S, Lipscomb HJ, Kaufman JS, Marshall SW, Cravey AJ. Race, wealth, and solid waste facilities in North Carolina. *Environ Health Perspect*. 2007;115(9):1344–50.
78. Rapport OCDE ENV/EPOC/WPNEP. Environment and distributional issues: analysis, evidence and policy implications. 2004.
79. Bowen W. An analytical review of environmental justice research: what do we really know? *Environ Manage*. 2002 Jan;29(1):3–15.
80. Marshall JD. Environmental inequality: Air pollution exposures in California's South Coast Air Basin. *Atmospheric Environment*. 2008 Jul;42(21):5499–503.
81. Yanosky JD, Schwartz J, Suh HH. Associations between measures of socioeconomic position and chronic nitrogen dioxide exposure in Worcester, Massachusetts. *J Toxicol Environ Health Part A*. 2008;71(24):1593–602.
82. Brochu PJ, Yanosky JD, Paciorek CJ, Schwartz J, Chen JT, Herrick RF, et al. Particulate air pollution and socioeconomic position in rural and urban areas of the Northeastern United States. *Am J Public Health*. 2011 Dec;101 Suppl 1:S224–230.
83. Bell ML, Ebisu K. Environmental inequality in exposures to airborne particulate matter components in the United States. *Environ Health Perspect*. 2012 Dec;120(12):1699–704.
84. Finkelstein MM, Jerrett M, Sears MR. Environmental inequality and circulatory disease mortality gradients. *J Epidemiol Community Health*. 2005 Jun;59(6):481–7.
85. Jerrett M, Burnett RT, Kanaroglou P, Eyles J, Finkelstein N, Giovis C, et al. A GIS - environmental justice analysis of particulate air pollution in Hamilton, Canada. *Environment and Planning A*. 2001;33(6):955–73.
86. Buzzelli M, Jerrett M. Geographies of susceptibility and exposure in the city: environmental inequity of traffic-related air pollution in Toronto. *Canadian Journal of Regional Science* [Internet]. 2007 Available from: <http://www.highbeam.com/doc/1G1-176978845.html>
87. Crouse DL, Ross NA, Goldberg MS. Double burden of deprivation and high concentrations of ambient air pollution at the neighbourhood scale in Montreal, Canada. *Soc Sci Med*. 2009 Sep;69(6):971–81.
88. Kruize H, Driessen PPJ, Glasbergen P, van Egmond KND. Environmental equity and the role of public policy: experiences in the Rijnmond region. *Environ Manage*. 2007 Oct;40(4):578–95.

89. Rotko T, Kousa A, Alm S, Jantunen M. Exposures to nitrogen dioxide in EXPOLIS-Helsinki: microenvironment, behavioral and sociodemographic factors. *J Expo Anal Environ Epidemiol*. 2001 Jun;11(3):216–23.
90. Kohlhuber M, Mielck A, Weiland SK, Bolte G. Social inequality in perceived environmental exposures in relation to housing conditions in Germany. *Environmental Research*. 2006 Jun;101(2):246–55.
91. Fairburn J, Butler B, Smith G. Environmental justice in South Yorkshire: locating social deprivation and poor environments using multiple indicators. *Local Environment*. 2009;14(2):139–54.
92. Walker G. Beyond Distribution and Proximity: Exploring the Multiple Spatialities of Environmental Justice. In: Holifield R, Porter M, Executive GWC of EJS, editors. Wiley-Blackwell; 2010 p. 23–46.
93. Forastiere F, Stafoggia M, Tasco C, Picciotto S, Agabiti N, Cesaroni G, et al. Socioeconomic status, particulate air pollution, and daily mortality: differential exposure or differential susceptibility. *Am J Ind Med*. 2007 Mar;50(3):208–16.
94. Laurian L. Environmental Injustice in France. *Journal of Environmental Planning and Management*. 2008;51(1):55–79.
95. Havard S, Deguen S, Zmirou-Navier D, Schillinger C, Bard D. Traffic-related air pollution and socioeconomic status: a spatial autocorrelation study to assess environmental equity on a small-area scale. *Epidemiology*. 2009 Mar;20(2):223–30.
96. Namdeo A, Stringer C. Investigating the relationship between air pollution, health and social deprivation in Leeds, UK. *Environment International*. 2008 Jul;34(5):585–91.
97. Jerrett M, Burnett R, Brook J, Kanaroglou P, Giovis C, Finkelstein N, et al. Do socioeconomic characteristics modify the short term association between air pollution and mortality? Evidence from a zonal time series in Hamilton, Canada. *J Epidemiol Community Health*. 2004 Jan;58(1):31–40.
98. Naess O, Piro FN, Nafstad P, Smith GD, Leyland AH. Air pollution, social deprivation, and mortality: a multilevel cohort study. *Epidemiology*. 2007 Nov;18(6):686–94.
99. Martins M, Fatigati F, Véspoli T, Martins L, Pereira L, Martins M, et al. Influence of socioeconomic conditions on air pollution adverse health effects in elderly people: an analysis of six regions in São Paulo, Brazil. *J Epidemiol Community Health*. 2004 Jan;58(1):41–6.
100. O'Neill M., Jerrett M, Kawachi I, Levy J, Cohen A, Gouveia N, et al. Health, Wealth, and Air Pollution: Advancing Theory and Methods. *Environmental Health Perspectives*. 2003 Sep 2;111(16):1861–70.
101. Havard S, Deguen S, Bodin J, Louis K, Laurent O, Bard D. A small-area index of socioeconomic deprivation to capture health inequalities in France. *Soc Sci Med*. 2008 Dec;67(12):2007–16.

102. Filleul L, Rondeau V, Cantagrel A, Dartigues J-F, Tessier J-F. Do subject characteristics modify the effects of particulate air pollution on daily mortality among the elderly? *J Occup Environ Med.* 2004 Nov;46(11):1115–22.
103. Laurent O, Filleul L, Havard S, Deguen S, Declercq C, Bard D. Asthma attacks and deprivation: gradients in use of mobile emergency medical services. *JEpidemiolCommunity Health.* 2008 Nov;62(11):1014–6.
104. Laurent O, Bard D, Filleul L, Segala C. Effect of socioeconomic status on the relationship between atmospheric pollution and mortality. *JEpidemiolCommunity Health.* 2007;61(8):665–75.

CHAPITRE 2 :

MATÉRIELS ET MÉTHODES

Dans un contexte d'étude écologique, ce chapitre présente tout d'abord la zone d'étude, puis les bases de données utilisées (sanitaires, environnementales et socio-économiques) et enfin le développement méthodologique d'une approche spatiale dans l'étude des inégalités sociales de santé à l'aide de méthodes statistiques adaptées.

I. LE CADRE D'ÉTUDE

Dans cette partie nous précisons les critères qui ont guidé le choix des zones d'études, le type d'étude ainsi que l'unité géographique.

A. Les zones d'étude

Notre étude écologique géographique porte sur quatre agglomérations : Lille, Paris & sa petite couronne, Lyon et Marseille. Le choix des agglomérations s'est basé sur deux critères majeurs.

Le premier est la taille de l'agglomération. En effet, les événements de santé que nous analysons sont rares et il est nécessaire que nous ayons un effectif assez élevé pour avoir une puissance statistique suffisante. Les 4 agglomérations choisies sont les 4 plus importantes agglomérations de France et représentent 26% de la population française (voir section « présentation de agglomérations »).

Le deuxième est une grande diversité en termes de situations sociale, environnementale et de santé au sein même d'une agglomération ainsi qu'entre les agglomérations (1) dans l'objectif d'explorer le caractère stable ou au contraire variable des relations analysées entre les différentes catégories de variables du projet Equit'Area. Les spécificités géographiques

traduisent des compositions sociales, mais également des comportements spécifiques pour faire face à la santé (1), aussi bien en matière d'habitudes alimentaires que de conduites à risques ou de recours aux soins (2). C'est pourquoi elles se manifestent par des contrastes en termes d'espérance de vie ou de mortalité : les taux de décès sont par exemple plus élevés dans le Nord-Pas de Calais et plus faibles en Rhône-Alpes et en Île-de-France.

B. Le type d'étude : étude écologique

Plusieurs approches auraient pu être envisagées pour répondre aux questions et objectifs du projet. Le choix d'une étude écologique s'est basé sur deux critères.

Tout d'abord, le projet Equit'Area a pour perspectives d'étendre l'analyse des inégalités sociales de santé en France sur d'autres événements sanitaires et d'autres zones d'études. C'est pourquoi l'approche écologique permettait d'avoir accès à des données publiques sans mettre en place un recueil de données individuelles avec un suivi de femmes enceintes sur plusieurs années (long et coûteux). Dans les perspectives du projet Equit'Area, l'intégration de données individuelles aux données écologiques est en cours et permettra la mise en place d'analyses multi-niveaux permettant de dégager le rôle des facteurs contextuels de ceux individuels dans la survenue d'un événement sanitaire.

De plus, une littérature récente a établi que l'environnement de vie du quartier de la mère et de l'enfant a une influence sur les issues de grossesse indépendamment des facteurs de risque individuels (3–6). Alors que les études écologiques sur les inégalités sociales de mortalité infantile sont rares, l'approche écologique est fréquemment employée en vue de l'analyse des issues de grossesse (7–9). Cependant, ce type d'étude nécessite une méthodologie rigoureuse afin de minimiser les biais écologiques et tenir compte de la dépendance des unités spatiales. C'est pourquoi nous avons choisi une méthode statistique adaptée (Chapitre 2, partie V), et une unité géographique fine (petite) (voir section « unité géographique »)

Enfin, les études écologiques sont très utiles pour déterminer les zones géographiques à risque défavorables pour la santé. L'intérêt est de cartographier les résultats de manière efficace et

directement utilisable pour la prise de décision et la communication auprès des pouvoirs publics.

C. L'unité géographique : l'IRIS

L'unité géographique utilisée est l'IRIS, découpage infra communal mis en œuvre par l'Institut National de la Statistique et des Études Économiques au cours du recensement de 1999 et de 2006. Ce découpage constitue une partition du territoire de ces communes en " quartiers " dont la population est de l'ordre de 2000 habitants, comparable au Census block Group des USA. Les communes d'au moins 10 000 habitants et la plupart des communes de 5 000 à 10 000 habitants sont découpées en IRIS. La géométrie des IRIS est éditée par l'institut géographique national (IGN) et l'institut national de la statistique et des études économiques (INSEE). Le découpage des quartiers de résidence en IRIS tient compte des obstacles physiques introduisant des ruptures dans le paysage urbain (infrastructures autoroutières, cours d'eau, espaces verts...) (10).

Selon la définition de l'INSEE, ils se déclinent en trois types de zones : les IRIS d'habitat (ils sont homogènes quant au type d'habitat), les IRIS d'activité (deux fois plus d'emplois salariés que de population résidente) et les IRIS divers (parcs de loisirs, zones portuaires, forêts...) (10).

Choix de l'IRIS

Le choix de l'unité géographique grâce à laquelle les différentes données seront recueillies, produites et estimées est une étape préliminaire essentielle à la mise en œuvre d'une étude écologique (11,12). Tout d'abord, l'IRIS a été reconnu par plusieurs auteurs comme étant l'unité géographique adaptée pour explorer la variabilité spatiale d'évènements de santé (13–15). En effet, selon le niveau géographique d'observation retenu (région, département, ville, quartier, etc.), la mise en évidence d'une association entre deux facteurs peut s'avérer plus ou moins concluante (11,12). L'intensité ou le sens de l'association peuvent être fortement modifiés selon la résolution géographique à laquelle les données sont recueillies (12,16).

De plus, en France, l'IRIS constitue la plus petite unité géographique française pour laquelle l'intégralité des données démographiques et socio-économiques du recensement est disponible en France.

Enfin, l'unité optimale doit être assez grande pour fournir des indicateurs de santé stables et assez petite pour être homogène en termes d'exposition, de caractéristiques socio-économiques (12,16). Généralement, plus l'unité géographique est fine (petite), plus ses caractéristiques tendent à être homogènes et à se rapprocher des caractéristiques individuelles des personnes qui la composent (17), ce qui tend à réduire le risque, majeur ici de « biais écologique » (18).

D. Description du cadre de l'étude

Lille

L'agglomération de Lille est composée de 85 communes et de 504 IRIS sur un territoire de 611 km² (1951 hab/km²). Située dans la région du Nord Pas de Calais, 60% de la population totale (1 193 244 hab. en 2008) se concentre surtout dans les villes de Lille, Roubaix, Tourcoing et Villeneuve d'Ascq.

Lyon

L'agglomération de Lyon est la troisième en France avec une population de 1 281 971 habitants en 2008 et une superficie de 515.96km², soit une densité de 2485 hab/km². Composée de 58 communes et de 511 IRIS, elle est la capitale de la région Rhône-Alpes.

Marseille

L'agglomération de Marseille est la deuxième en France avec une population de 1 714 828 hab. en 2008 et une superficie de 3173.5km² soit la plus petite densité 540 hab/km². Composée de 628 IRIS et de 52 communes, elle est la capitale de la région Provence-Alpes Côte d'Azur.

Paris

L'agglomération de Paris et de sa petite couronne est constituée d'un regroupement administratif de 2749 IRIS, 412 communes, 4 départements (Paris, Hauts de Seine, Val de Marne, Seine-Saint Denis) et totalisait 10 354 675 habitants en 2008. Avec une superficie de 105 km², la commune de Paris est une petite commune. Rome est par exemple 12 fois plus étendue (1 285 km²) et Berlin 8,5 fois. En revanche, elle concentre une densité de population de 21 196 hab/km², l'une des plus élevées du monde, et environ 18 % de la population totale de la région Île-de-France, pour à peine 0,9 % de sa superficie.

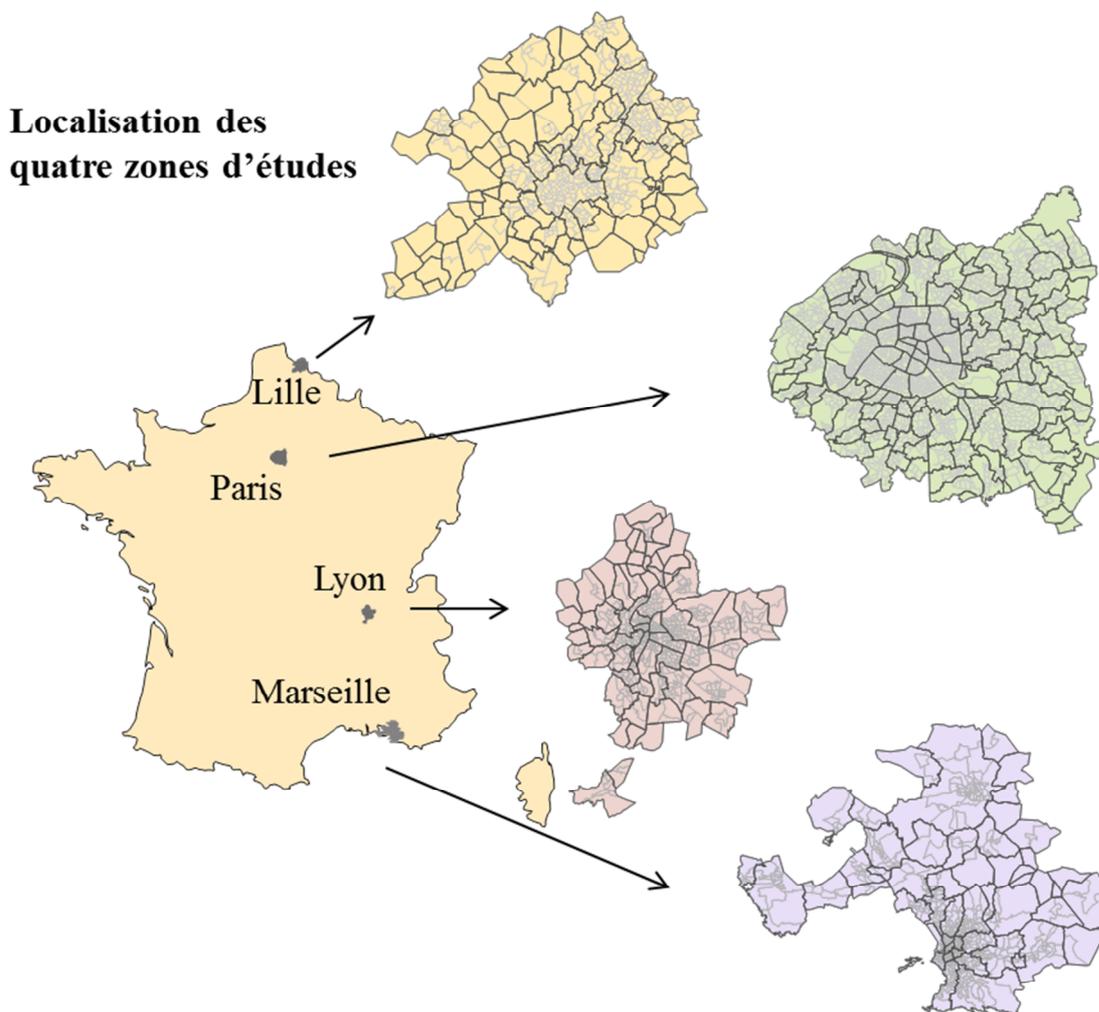


Figure 5: Cartographies des quatre agglomérations du projet Equit' Area.

II. DONNÉES SANITAIRES

Les données sanitaires utilisées dans le cadre de ce travail portent sur la mortalité infantile, incluant la mortalité néonatale dans chacune des 4 agglomérations de l'étude. Dans la présente section, nous présenterons les étapes du processus d'estimation des taux de mortalité infantile et néonatale (recueil des données, géocodage, estimation des taux à l'échelle de l'IRIS), ainsi que les difficultés rencontrées.

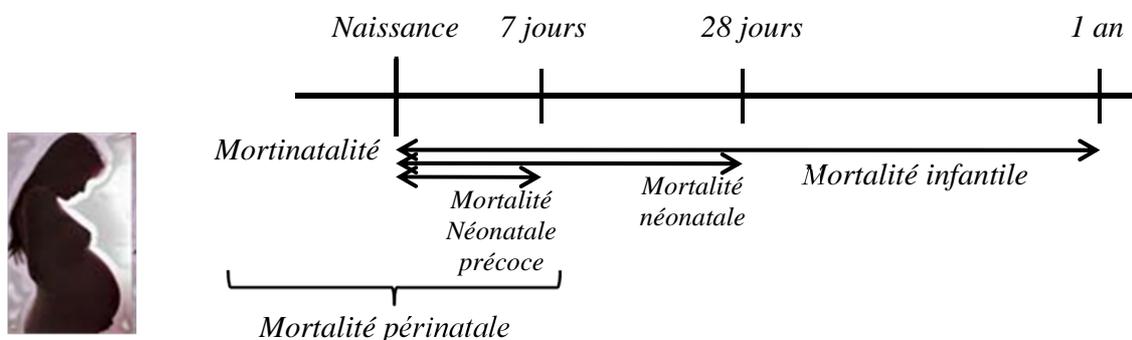
A. Définition

Mortalité infantile

La mortalité infantile est une statistique calculée en rapportant le nombre d'enfants morts avant l'âge d'un an au nombre total d'enfants nés vivants au cours de la même période. Cette statistique est exprimée pour 1 000 naissances vivantes (‰).

Mortalité néonatale

La mortalité néonatale est l'ensemble des enfants nés vivants mais décédés entre la naissance et le 28^e jour de vie. On distingue la « mortalité néonatale précoce », pour les décès durant la première semaine, de la « mortalité néonatale tardive », pour ceux des trois semaines suivantes. La mortalité néonatale est le rapport entre le nombre d'enfants décédés à moins de 28 jours sur le nombre total d'enfants nés vivants au cours de la même période. Elle est aussi exprimée pour 1000 naissances vivantes (‰).



B. Recueil des données sanitaires

Les données sanitaires portent sur le nombre de cas de décès d'enfants de moins de 1an (ou de moins de 1 mois pour la mortalité néonatale) par an et par IRIS. Or ces informations ne sont pas disponibles à une échelle aussi fine. C'est pourquoi, nous avons engagé un processus de recueil de données afin de récupérer l'adresse postale de résidence des parents et la géocoder l à l'échelle de l'IRIS. La procédure sera détaillée dans la partie (1 : Procédure de recueil).

Le choix de recueillir les données dans les mairies s'imposait car ce sont les seules à conserver les adresses postales exactes des défunts (détails partie : acte de décès). Le caractère personnel des informations collectées impliquait une demande d'autorisation auprès des commissions CCTIRS et CNIL (Commission nationale de l'informatique et des libertés). Les autorisations ont été accordées en 2010 et 2012 pour un recueil des données dans le but de mettre en œuvre un traitement de données ayant pour finalité l'agrégation à l'échelle de l'IRIS (Demande d'autorisation n°911149).

L'acte de décès

L'acte de décès est enregistré dans la commune où a eu lieu le décès. L'officier d'état civil remplit un bulletin de décès anonyme pour respecter le secret médical (modèle 7 en annexe) composé de plusieurs volets. Un premier volet adressé à l'attention des mairies comporte les informations suivantes : la date de naissance et de décès, le sexe, la profession du père et de la mère et l'adresse de domicile du défunt.

Certains événements qui se sont déroulés en dehors de la commune ou ayant fait l'objet par ailleurs d'actes authentiques doivent être transcrits sur les registres de la commune. Un bulletin de transcription (modèle 1) est alors établi. Ces transcriptions concernent, par exemple, les décès hors de la commune de personnes domiciliées dans la commune, les jugements d'adoption, etc.

Les autres volets du modèle 7 (modèle en annexe) ne comportent que la commune comme adresse de résidence. Un volet à destination de l'INSERM présente tous les renseignements nécessaires aux statistiques de mortalité. La cause du décès figure sur un certificat

confidentiel de décès rédigé par le médecin qui a constaté le décès (médecin de l'état civil ou médecin traitant faisant fonction) et sera envoyé au CepiDC via l'Agence Régionale de la Santé (ARS).

Procédure de recueil

Avant de commencer le recueil, une liste des communes détaillant les coordonnées précises des mairies composant nos agglomérations a été dressée. Afin d'optimiser le recueil, toutes les communes ont été triées en terme de priorité à visiter selon plusieurs critères :

1. Nombre de décès infantiles ou néonataux par commune (d'après la Base CepiDC)
2. Présence pendant la période d'étude (2000-2009) d'un établissement de santé comportant une unité obstétrique (L'obstétrique est une spécialité médico-chirurgicale qui a pour objet l'étude et la prise en charge de la grossesse et de l'accouchement.)
3. Présence pendant la période d'étude (2000-2009) d'un organisme de santé (hors 2)
4. Autres communes (hors 1,2 &3) triées selon la taille de population et le nombre de naissances.

L'étape suivante est la prise de contact avec les services d'état civil de chaque municipalité. L'objectif est d'avoir le responsable du service d'état service afin de lui présenter le projet et de connaître la forme sous laquelle sont conservées les données (informatique ou non). Si les données ne sont pas informatisées (la majorité des mairies), un déplacement dans la commune est alors nécessaire pour avoir accès aux registres d'état civil et recueillir les données manuellement.

Une fois cette prise de contact téléphonique réalisée, un courrier signé de Denis Zmirou, directeur du Département Santé-Environnement-Travail de l'EHESP est envoyé présentant, les objectifs du projet. Ce courrier est accompagné d'une copie de l'accord de la CNIL pour l'accès aux informations. Que ce soit par courriel ou sur le terrain, une phase de relance téléphonique est nécessaire. Il s'agit de s'assurer que le courrier est bien reçu et que le responsable a donné son accord afin de prendre un rendez-vous pour le recueil.

L'étape suivante consiste à collecter les données proprement dites et les communes prioritaires (celles accueillant un établissement de santé) sont visitées en premier. Nous recueillons donc tous les décès de personnes résidant dans la commune, auxquels s'ajoutent ceux des communes avoisinantes et non résidentes de l'agglomération. Une mise à jour quotidienne permettait de détecter les doublons éventuels (décès collectés dans la commune de décès et de résidence), et de comptabiliser le nombre de cas par commune. Les derniers cas manquants dans les petites communes sont complétés les semaines suivantes afin d'être le plus exhaustif possible.

Exhaustivité des données

Enfin dans le but de déterminer l'exhaustivité des données à notre disposition suite au recueil, le CepiDC de l'INSERM, partenaire de ce projet, nous a fourni le nombre de cas de décès de moins de 1 an et 1 mois par commune et par année.

Géocodage à l'échelle de l'IRIS

Cette étape consiste à géocoder les données sanitaires. A chaque décès a été attribué un numéro d'IRIS à partir de l'adresse postale de résidence des parents à l'aide des bases de données cartographiques Correspondances Adresses – Zonages Urbains – Edition 2004 et Contours IRIS diffusées par l'INSEE. Lorsque l'adresse postale du sujet était incomplète ou erronée, l'utilisation du Système d'Information Géographique (SIG) Arc View version 10.1 (ESRA, redlands, CA) et des plans et index des rues des communes de l'agglomération a permis de corriger l'adresse et ainsi de géocoder. A l'issue de cette démarche, 1 seul décès n'a pas pu être géocodé à l'échelle de l'IRIS pour l'agglomération de Lyon.

Estimation des taux de mortalité

L'estimation du taux de mortalité implique de connaître le nombre de naissances par année et par IRIS. Or l'INSEE possède uniquement une base de données composée du nombre de naissance par IRIS par an de 2005 à 2009 et une base de données du nombre de naissance par commune et par an. Une estimation du nombre de naissance par IRIS a été effectuée selon la méthode suivante :

- Pour les petites communes qui ne comportent qu'1 seul IRIS, le nombre de naissances dans cette commune a été reporté pour l'IRIS correspondant.
- Pour les autres communes, le nombre de naissances par commune disponible par l'INSEE est réparti proportionnellement aux IRIS qui la composent (avec pour proportion de référence les années disponibles).

Une fois ces données complétées pour chaque année et pour chaque IRIS, nous faisons la somme du nombre de cas sur la somme du nombre de naissances par IRIS afin de calculer les taux de mortalité infantile et néonatale sur notre période d'étude de 2000 à 2009 pour 1000 naissances vivantes.

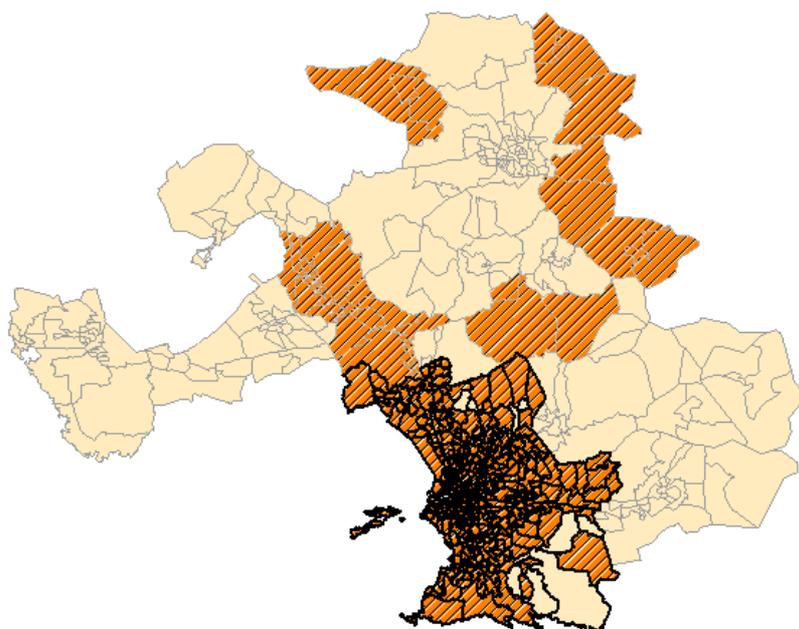
C. Les difficultés rencontrées

Agglomération de Marseille

L'accès aux données a été très difficile pour l'agglomération de Marseille et Aix-en-Provence. Il a été effectué par Hermine Merel étudiante en Master 2 lors de son stage. Des difficultés administratives n'ont pas permis le recueil au sein de l'état civil de la ville d'Aix-en-Provence ; les seuls cas obtenus proviennent des services d'état civil des communes alentours et ne sont pas représentatives des données du CepiDC, à défaut de la ville-centre. Les données manquantes étant trop importantes, Aix-en-Provence est exclue de la zone d'étude de l'Agglomération de Marseille pour le présent travail. Une difficulté supplémentaire a été rencontrée, liée aux retranscriptions d'actes de décès non envoyées dans les communes de domicile. Ainsi, dans les cas où l'enfant n'est pas décédé dans sa ville de résidence mais dans une autre ville de la zone d'étude, les données ne peuvent pas être récupérées et sont alors considérées comme manquantes. La carte suivante présente l'agglomération de Marseille et la zone réduite qui comprend les communes qui ont accepté de collaborer lors du recueil de données.

Légende

-  Marseille centre
-  Zone réduite
-  Agglomération de Aix M arseille



Agglomération de Paris et sa Petite couronne

La faible informatisation des actes de décès de l'état civil a considérablement rallongé l'estimation du temps accordé au recueil de données dans la thèse. De ce fait, j'ai préparé et estimé le recueil pour les quatre agglomérations mais effectué le recueil sur le terrain pour les agglomérations de Lille et de Lyon seulement. La fin du recueil pour la petite couronne de Paris est prévue fin de l'été 2013, ce qui ne permet pas d'effectuer les analyses sur la mortalité infantile et néonatale pour la petite couronne. Néanmoins, un travail collaboratif avec l'équipe INSERM 707, Equipe DS3 (Déterminants sociaux de la santé et du recours aux soins) a permis d'acquérir informatiquement les données issues des actes de décès pour la ville de Paris (Paris intra-muros) pour les années 2004 à 2009.

D. Description des données sanitaires

Le Tableau 1 présente les données des décès recueillies pour les 4 Agglomérations. Sur la période 2000-2009 pour Lille, Lyon, Marseille et la période 2004-2009 pour Paris-ville. A Lille, Lyon et Paris respectivement, 668 cas sur 705, 715 cas sur 741 et 629 sur 684 ont été recueillis, ce qui correspond à un taux d'exhaustivité de 94,7%, 96,5% et 91,9%. A Marseille, si on ne compte que la zone réduite après le recueil de données, 308 cas ont été collectés sur 365, l'exhaustivité est seulement de 84%. De plus, la zone réduite ne représente que 423 IRIS sur les 628 (67%) de l'agglomération de Marseille. Le manque de continuité spatiale de cette zone ne nous permet pas d'effectuer les analyses statistiques. Marseille ne sera donc pas considérée dans les parties suivantes (Chapitre 3 : Partie I,II,IV).

Le taux moyen de mortalité infantile varie de 4,1‰, 3,8‰ à 3,4‰ et les taux de mortalité néonatale de 2,8‰, 2,6‰ à 2,5‰ pour les agglomérations de Lille, Lyon et la ville de Paris pour 1000 naissances vivantes. La mortalité néonatale représente environ 69%,70% et 73% des décès d'enfants de moins de 1 an pour Lille, Lyon et la ville de Paris. La répartition des décès d'enfants pour chaque sexe confirme les résultats de la littérature (19,20), avec une légère surmortalité pour les garçons.

Tableau 1. Mortalité infantile et néonatale dans les agglomérations de Lille, Lyon, Marseille et Paris-ville, 2000-2009.

Période d'étude	Lille		Lyon		Paris-ville	
	2000-2009		2000-2009		2004-2009	
	N	%	N	%	N	%
Mortalité infantile (1000 naissances vivantes)	668	4,1‰	714	3,8‰	629	3,4‰
Mortalité néonatale (1000 naissances vivantes)	461	2,8‰	501	2,6‰	462	2,5‰
Mortalité néonatale / mortalité infantile	461/668	69	502/715	70	462/629	73
Rapport du genre : Garçon/ Total	364/668	54,5	404/715	56,5	358/629	56,9
IRIS sans cas /Total IRIS	204/504	40,5	160/511	31,3	585/992	58,9

Classiquement utilisés pour la représentation cartographique d'évènements rares, les (*Standardised Mortality/Morbidity Ratio*, SMR)⁴ permettent d'analyser les contrastes de la distribution géographique d'une maladie. Nous avons utilisé la standardisation indirecte afin de comparer facilement les cartes de répartition de la mortalité infantile dans les 4 agglomérations (Figures 6 et 7).

⁴Le SMR se calcule, selon la méthode de standardisation indirecte, comme le rapport entre le nombre total d'évènements observés O_i dans l'unité i pendant la période t et le nombre d'évènements attendus E_i sous l'hypothèse de taux d'évènements par la variable de standardisation (généralement l'âge, le sexe) égaux à ceux d'une population de référence.

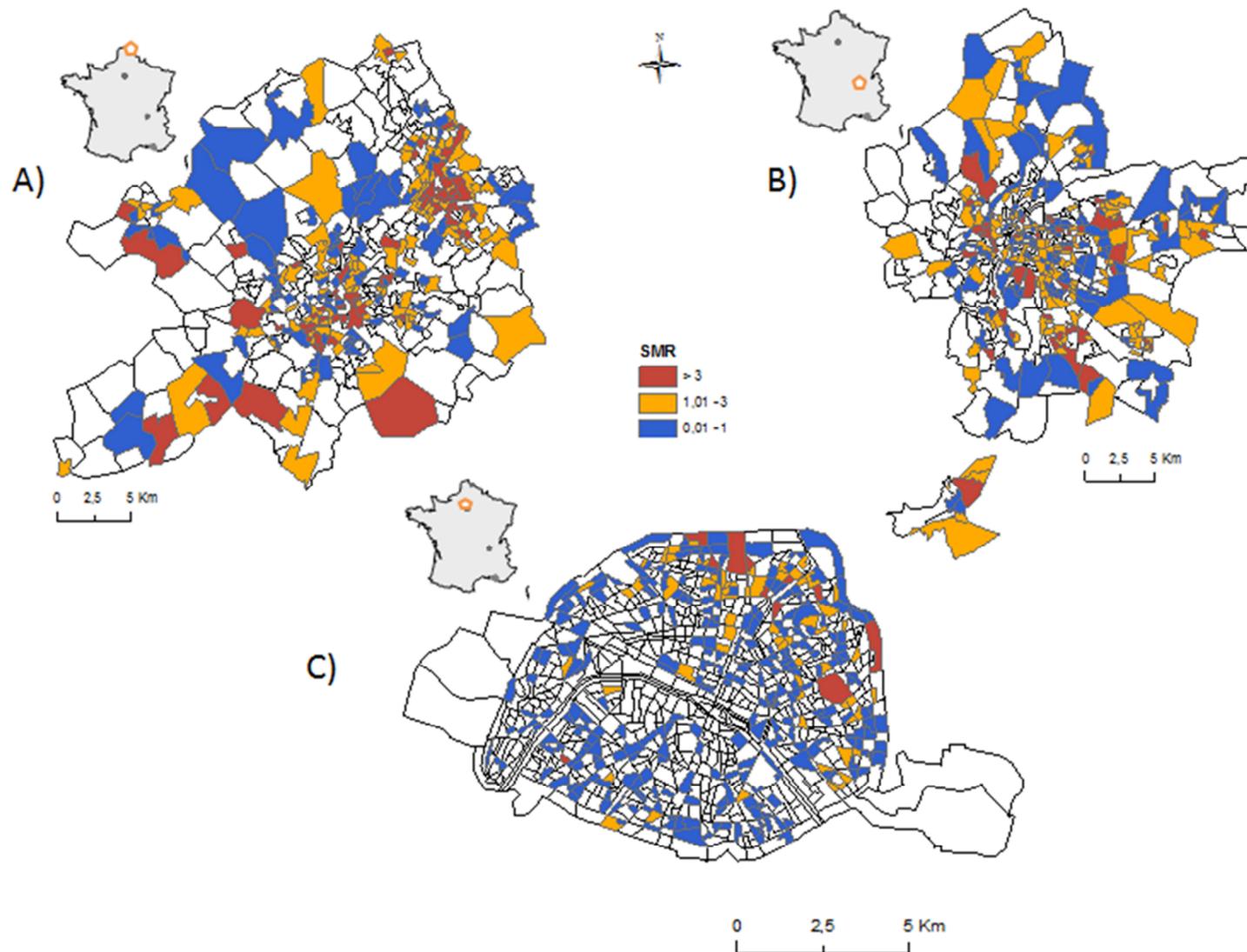


Figure 6: Cartographies des rapports de mortalités standardisés (SMR) de mortalité infantile à A) Lille, B) Lyon, C) Paris sur la période 2000 à 2009.

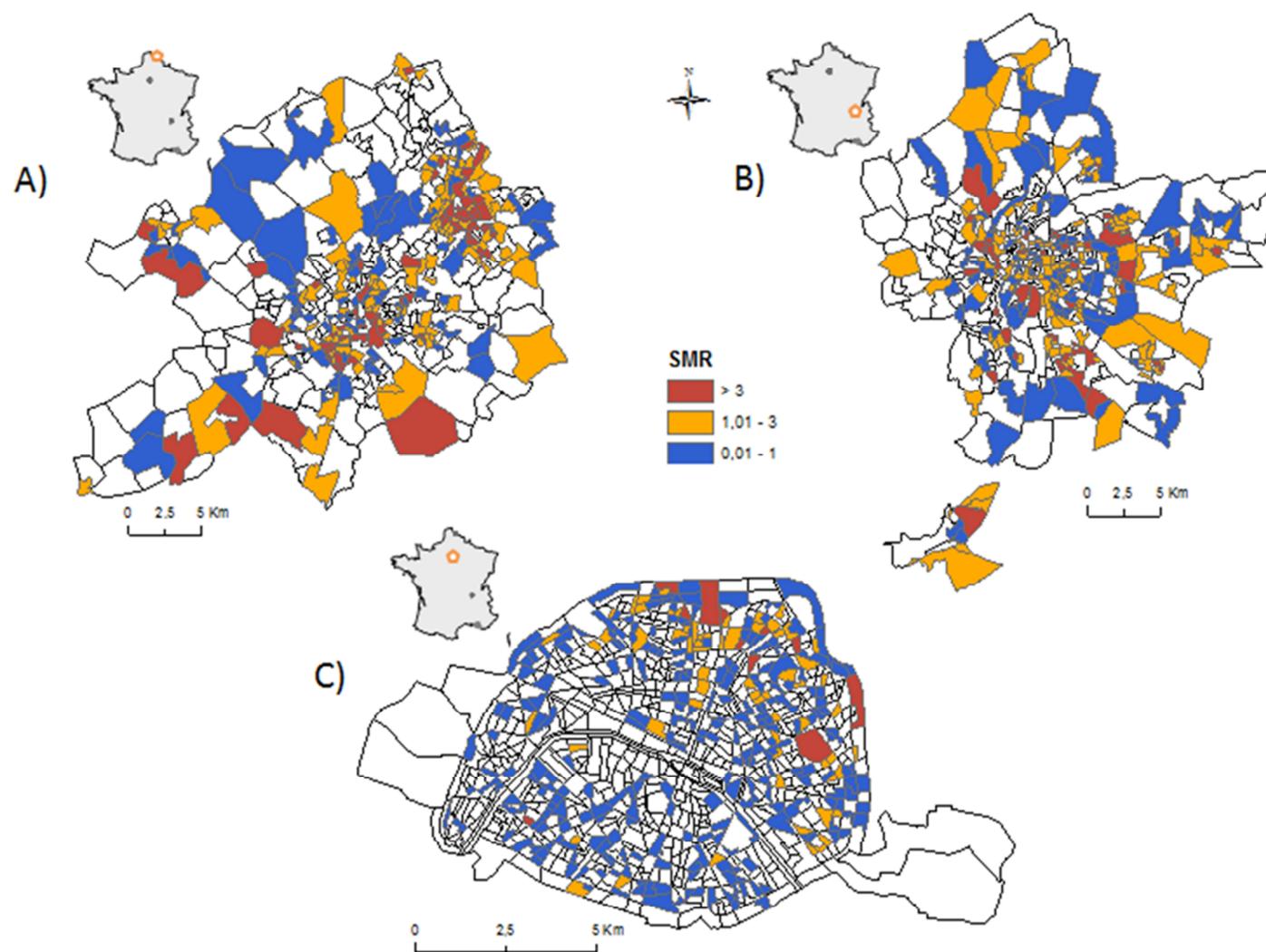


Figure 7 : Cartographies des rapports de mortalités standardisés (SMR) de mortalité néonatale à A) Lille, B) Lyon, C) Paris sur la période 2000 à 2009.

III. DONNÉES SOCIO-ÉCONOMIQUES

A. Choix d'un indice de défaveur

La défaveur socio-économique du quartier se définit comme un concept complexe qui inclut plusieurs dimensions, telles que le chômage, le revenu, l'éducation, le logement, la cohésion sociale (5,21–24). Afin d'analyser cette défaveur socio-économique en relation avec l'exposition à la pollution de l'air et la mortalité infantile ou néonatale nous avons utilisé un indice composite de défaveur socio-économique développé en France à l'échelle de l'IRIS par notre équipe (25).

Notre choix ne s'est pas porté sur des métriques simples utilisées dans la littérature pour caractériser le profil socio-économique du territoire tels que la profession des parents (26), l'éducation des parents (3), le revenu du ménage (27)...). Plusieurs études ont utilisé une seule variable pour représenter le statut socio-économique, néanmoins la limite majeure est de ne pas prendre en compte les multiples dimensions du statut socio-économique du quartier. Une alternative consiste à inclure plusieurs variables socio-économiques dans un même modèle mais au risque d'être confronté au problème de colinéarité (corrélation entre les variables socio-économiques). De plus, il est important de limiter le nombre de variables explicatives dans le modèle afin de garantir une estimation plus précise des paramètres. Bien que souvent utilisés dans plusieurs pays, les indices de Townsend (28) et Carstairs (29,30) sont basés uniquement sur 4 variables (le taux de chômage ; la part des ménages sans véhicule ; la part des logements « surpeuplés (plus de 1,5 habitant par pièce) ; la part des ménages locataires de leur logement (Townsend) ; la part des résidents dont le chef de famille appartient à une classe sociale défavorisée selon une nomenclature définie au Royaume Uni (Carstairs)). Ces indices ont été construits il y a 20 ans et ne peuvent pas prendre en compte la façon dont ces variables ont évolué au cours du temps ni les différentes définitions de celles-ci entre les pays.

C'est pourquoi, notre choix s'est porté sur un indice composite de défaveur incluant plusieurs dimensions, et estimé à l'échelle de l'IRIS. Des précédentes études écologiques ont démontré la capacité de cet indice à capturer les inégalités socio-spatiales en lien avec l'environnement

et la santé (asthme, infarctus du myocarde) pour l'agglomération de Strasbourg en France (31–34)(31,34). Dans notre étude, la méthode a été retravaillée pour améliorer la robustesse de l'indicateur, ce qui a fait l'objet d'un travail spécifique (Lalloué et al (35)). Cet indice intègre un grand nombre de variables et dimensions socio-économiques (la structure familiale, le type de ménage, le statut d'immigration, la mobilité, l'emploi, le revenu, l'éducation et le logement) et peut être appliqué à plusieurs zones géographiques (voir variables Chapitre 3 : partie I). Il peut être plus approprié pour fournir des renseignements utiles aux organismes de santé publique.

La méthode de construction de l'indice composite est présentée dans la section suivante. Pour plus de détails, la méthodologie développée pour construire l'indice a fait l'objet d'un article publié récemment (35).

B. Construction de l'indice de défaveur

Cette section concernera l'ensemble des quatre agglomérations dans leur totalité. La première étape est la création de variables socioéconomiques : 52 variables quantitatives couvrant l'intégralité des thèmes du recensement (revenu, éducation, famille, logement, immigration, emploi, biens matériels) retenues à partir d'une synthèse de la littérature.

Trois bases de données ont été utilisées : la base de données du recensement de 1999, puis celle mise à jour en 2006 issue de l'INSEE, la base de données *Revenus fiscaux des ménages* diffusée par l'INSEE et la Direction générale des Impôts, et la base de données *Logements* diffusée par l'INSEE (36).

La prochaine étape a été réalisée par un autre doctorant de l'équipe Equit'Area. Elle a consisté à créer l'indice de défaveur. La méthode retenue est l'analyse en composantes principales (ACP) ; elle a été conduite sur 52 variables socio-économiques. La stratégie retenue pour restreindre la mesure du niveau socio-économique à un seul indice est celle qui maximise l'inertie de la première composante principale. Des analyses en composantes principales successives sont utilisées pour sélectionner les variables et créer l'indice. Finalement, pour

chaque IRIS une valeur d'indice de défaveur a été calculée. La même méthodologie a été appliquée pour la création de l'indice de 1999 et de 2006 (35) (Article en annexe).

C. Description de l'indice de défaveur

Brièvement, parmi les vingt variables finalement retenues dans l'indice, 15 sont communes aux trois agglomérations de Lille, Lyon et Marseille. L'indice explique au moins 57% de la variance de ces variables dans chacune des 3 agglomérations, avec une contribution de plus de 80% des 15 variables communes. Pour la ville de Paris, l'indice explique 57,8% de la variance de ces variables, avec une contribution de plus de 93% (35). Cet indice, est quantitatif, a été également considéré dans les analyses statistiques en catégories pour permettre de révéler l'existence d'un gradient de défaveur. La figure 8 présente la distribution cartographique de l'indice de défaveur en quintiles issu du recensement de 2006 pour les agglomérations de A) Lille, B) Lyon, C) Marseille et D) Paris. Ces cartes mettent en évidence des contrastes entre les IRIS les plus défavorisés (en rouge : valeurs d'indice les plus élevés) et ceux les plus favorisés (en jaune : valeurs d'indice les plus faibles).

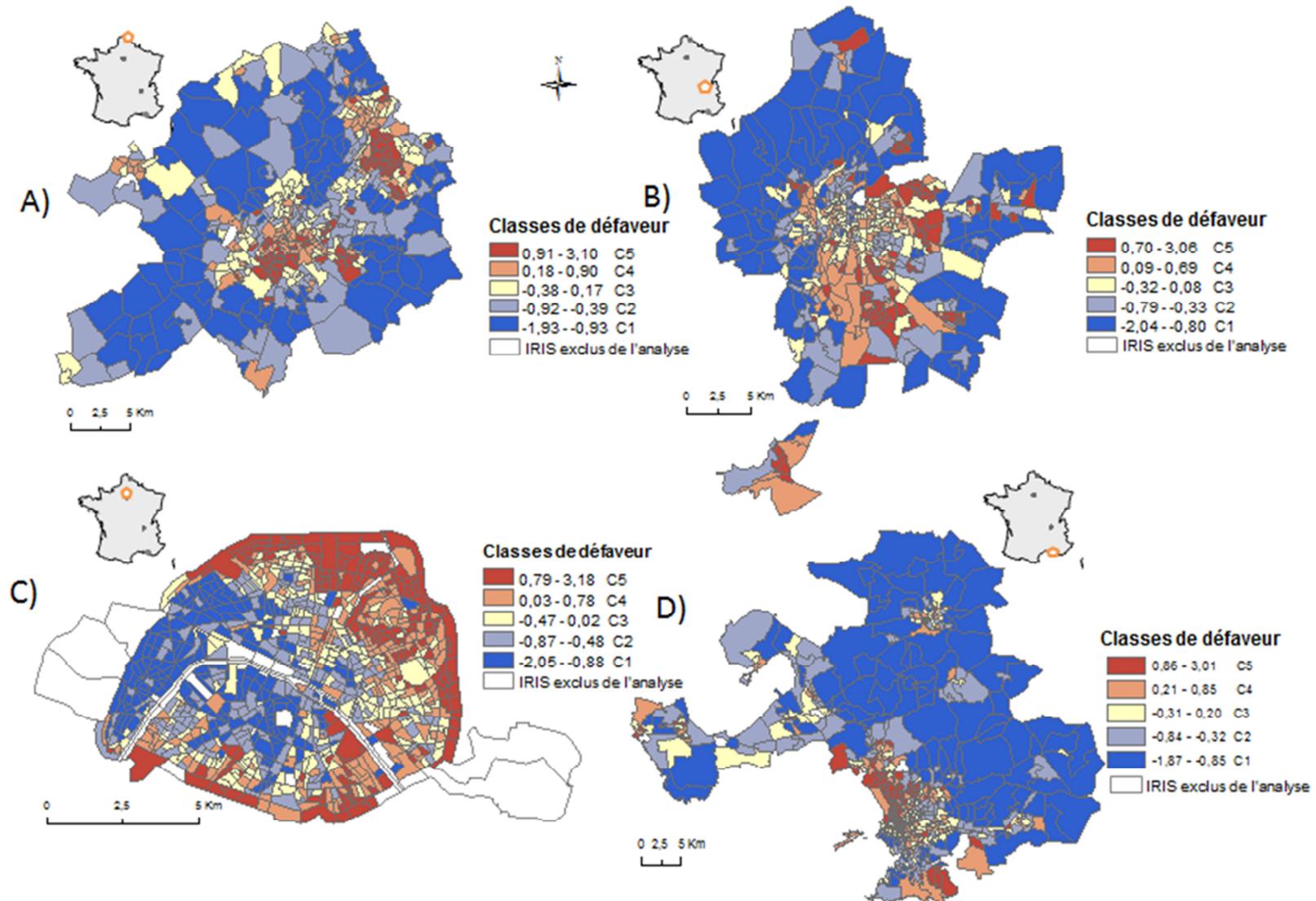


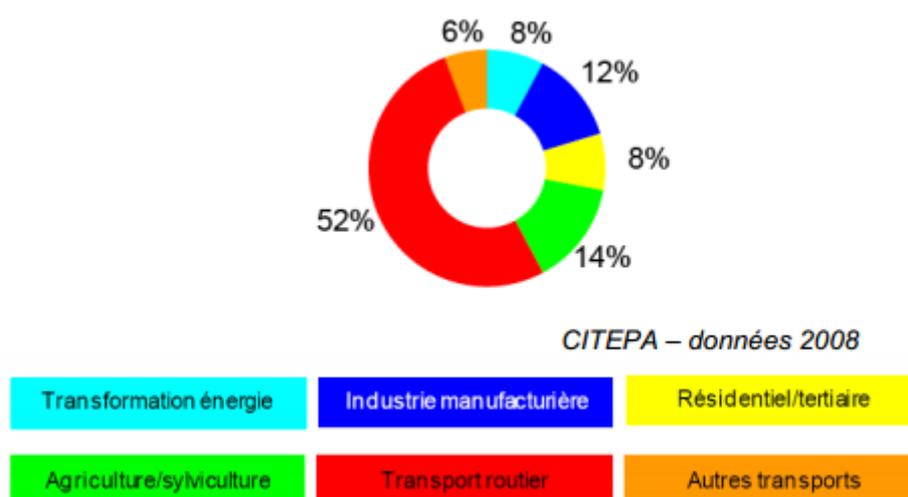
Figure 8 : Cartographies de l'indice de défaveur en quintiles par IRIS à A) Lille, B) Lyon, C) Paris, D) Marseille.

IV. DONNÉES ENVIRONNEMENTALES : LE DIOXYDE D'AZOTE

A. Présentation du dioxyde d'azote

Le dioxyde d'azote (NO_2) est un polluant résultant de tout processus de combustion à haute température. En site urbain, il est principalement d'origine automobile, mais les sources industrielles ou le chauffage résidentiel contribuent aussi aux concentrations mesurées. C'est le NO (monoxyde d'azote) qui est émis à la sortie du pot d'échappement ou des cheminées d'évacuation des fumées de combustion, il est oxydé en quelques minutes en NO_2 . La rapidité de cette réaction fait que le NO_2 est considéré comme un polluant primaire bien qu'il soit en réalité un polluant secondaire. On le retrouve en quantité relativement plus importante à proximité des axes de forte circulation et dans les centres villes. Les oxydes d'azote sont des précurseurs de la pollution photochimique et de dépôts acides. A noter que les niveaux d'oxydes d'azote sont également liés à des processus chimiques dans l'atmosphère et qu'ils dépendent pour une part des conditions météorologiques et de l'ensoleillement (37).

Sources d'émissions (France entière)



Le transport routier est le premier secteur émetteur puisqu'il représente en 2008, 52% des émissions de la France métropolitaine. Depuis 1993, la baisse observée dans le secteur du

transport routier est imputable à l'équipement progressif des véhicules en pots catalytiques, associée aux normes d'émission de plus en plus contraignantes. L'entrée en vigueur de la norme Euro III pour les poids lourds en 2002 et de la norme Euro 4 en 2005 pour les véhicules particuliers, accompagnée d'une stabilité du parc roulant sur la période 2002-2007 (augmentation inférieure à 1%), et même d'une baisse en 2008 (-1,8%), contribue à diminuer les émissions (37).

L'industrie manufacturière représente la deuxième source d'émission avec 12,5% des émissions de la France métropolitaine en 2008. La principale contribution est le secteur de la production d'électricité suite à la hausse de la demande (meilleur confort et développement du matériel électroménager dans les foyers, etc.). Pour ces deux derniers secteurs, la réduction des émissions s'explique, entre autre, par les progrès réalisés par les industriels depuis 1980, en particulier du fait d'une meilleure performance des installations, par la mise en place du programme électronucléaire, par le renouvellement du parc des engins mobiles non routiers de l'industrie et particulièrement dans le sous-secteur du BTP (Bâtiment et Travaux Publics) et par la mise en place dans certaines centrales thermiques de système de traitement du fait de la mise en œuvre des textes transposant la directive sur les grandes installations de combustion (nouvelles valeurs limites d'émission à respecter à partir de 2008) (37).

Les émissions du secteur de l'agriculture/sylviculture (troisième source avec 14,2% des émissions en 2008) sont induites, pour environ 38%, par les sols agricoles suite à l'utilisation de fertilisants azotés et, le reste, par la combustion de produits pétroliers. Les émissions de ce secteur ont baissé de 17,4% depuis 1990 (-38 kt) du fait conjointement du renouvellement du parc des tracteurs et des autres engins et de la réduction de la quantité d'engrais azotés utilisée.

Evolution des concentrations ambiantes de NO₂

En France, la moyenne annuelle des concentrations en NO₂ en fond urbain diminue de façon progressive depuis la fin des années 1990, avec une anomalie en 2003 (année de la canicule), pour se stabiliser ces dernières années (37). Sur les sites localisés en proximité de voies routières (sites trafic), et plus directement affectés par les émissions automobiles, la situation est différente. Les plus grandes agglomérations (Paris, Aix en Provence-Marseille) voient une

nette tendance à la hausse et la valeur limite pour la moyenne annuelle est dépassée ($40 \mu\text{g}/\text{m}^3$ en 2010) en proximité automobile dans plusieurs agglomérations : Aix-en-Provence, Bordeaux, Grenoble, Le Havre, Lille, Lyon, Marseille, Paris, Strasbourg, Toulouse (37).

B. Choix du dioxyde d'azote

Dans la littérature, les résultats suggèrent que les polluants suivants ont une influence sur la mortalité infantile: le dioxyde de soufre (SO_2), l'ozone (O_3), le monoxyde de carbone (CO) ainsi que les particules grosses (PM_{10}), fines ($\text{PM}_{2,5}$) et le dioxyde d'azote (NO_2) (38–41). Le choix du dioxyde d'azote comme variable indicatrice de menace environnementale s'est basé sur plusieurs critères.

Le dioxyde d'azote demeure un problème important de santé publique, avec des niveaux qui sont souvent en dépassement des normes de la directive européenne (2008/50/CE). Ce polluant a été sélectionné car il constitue un bon traceur de la pollution atmosphérique générée par le trafic et présente une hétérogénéité spatiale supérieure à celles des autres polluants (42).

De plus, plusieurs études ont démontré que le dioxyde d'azote est un indicateur de la pollution pour lequel l'exposition varie considérablement entre les groupes socio-économiques (16,43–45) (voir Chapitre 3, III). Le NO_2 est également rapporté comme ayant des effets toxiques en lien direct avec le fœtus (voir Chapitre 3, II). Il affecte le système de défense antioxydant (46), ce qui peut favoriser une peroxydation du liquide du placenta qui perturbera le développement postnatal (47). Son caractère irritant peut entraîner une libération systémique de cytokines qui elles-mêmes déclencheront un accouchement prématuré (48). Ces données restent cependant à confirmer.

Enfin, l'association agréée de surveillance de la qualité de l'air (AASQA) des 4 agglomérations disposait d'une expérience importante en matière de modélisation permettant d'élaborer des cartographies de polluants à une résolution spatiale fine compatible avec les unités de statistiques socio-économiques de l'INSEE (IRIS). L'extension de ce travail aux particules fines $\text{PM}_{2,5}$ est prévue dans les perspectives.

C. Modélisation du dioxyde d'azote

Le but de la modélisation était d'estimer les concentrations annuelles de dioxyde d'azote pour chaque IRIS et par année tout au long de la période d'étude (2002-2009).

Les données ont été obtenues auprès des associations agréées pour la surveillance de la qualité de l'air (AASQA) des territoires concernés : Atmo Nord Pas-de-Calais, AirParif, Atmo Rhône-Alpes et Atmo PACA. Elles ont travaillé conjointement afin d'homogénéiser leur méthode. Dans les paragraphes qui vont suivre nous présenterons brièvement les différentes étapes de modélisation issues du rapport technique des AASQA (36). Les concentrations estimées à l'échelle locale sont obtenues par différents modèles déterministiques: ADMS Urban (Atmospheric Dispersion Modelling System) pour Lille (49,50), SIRANE pour Lyon (51,52) ou encore STREET pour Marseille et Paris (49). Les données d'entrées sont partout de même type : la météorologie (température, vent, humidité, pression), les cadastres d'émissions de polluants (notamment celles du trafic routier) et les mesures de la pollution de fond (stations de mesures) pour les années représentées. Les caractéristiques de ces données d'entrées utilisées pour les différentes agglomérations sont synthétisées dans le tableau 2. L'estimation des concentrations moyennes annuelles de NO₂ de 2002 à 2009 consiste à spatialiser les mesures des stations annuelles en utilisant une méthode d'interpolation spatiale.

Validation des modèles

La performance des modèles a été évaluée en comparant les concentrations de NO₂ mesurées sur les stations aux concentrations annuelles modélisées. Les objectifs de qualité imposés sont en termes d'incertitude : 30% pour le NO₂ en moyenne annuelle selon la directive européenne (Directive 2008/50/CE). En termes de validation, les modèles donnent des résultats très acceptables. Entre les mesures locales (stations) et les concentrations assimilées, les coefficients de détermination oscillent en 0,77 et 0,85 pour Marseille et Aix en Provence, entre 0,63 et 0,74 pour Lille, entre 0,78 et 0,96 pour Lyon.

Tableau 2. Caractéristiques des données d'entrées des modèles

Agglomération	Modèle	Caractéristiques météo	Pollution de fond	Emissions trafic	Emissions surfaciques année	Emissions ponctuelles
Lille	ADMS Urban	Mesure Météo-France à Lesquin pour vitesse et direction du vent, température, précipitations. Modèle MM5 pour les paramètres hauteur de couche limite et flux de chaleur sensible.	Mesure de la station de Salomé (périurbaine)	Modèle de trafic Communauté Urbaine de Lille	2006	31 grandes sources
Lyon	SIRANE	Mesure Météo-France à Bron (St Exupéry pour 2007), horaire de chaque année mesurée (vitesse et direction du vent, température, précipitations, pression).	Mesure de la station de St-Exupéry	Cadastre annuel intégrant des comptages journaliers.	2008	91 grandes sources
Paris	STREET pour la pollution de proximité, ESMERALDA en fond	Modèle MM5 Pour STREET : données de météo moyennes annuelles en Ile-de-France (VV=3m/s et DV = Sud-Ouest)	2002 à 2006 : cartographie ISATIS (géostatistique) à partir des données stations de fond et du cadastre année 2000. 2007 à 2009 : résultats du modèle interrégional ESMERALDA	2002 à 2006 : utilisation d'un réseau trafic régional réactualisé fourni par DRIEA (année 2000), avec calage. 2007 à 2009 : utilisation d'un réseau trafic régional réactualisé fourni par DRIEA (année 2007), avec calage.	2000	200 grandes sources
Marseille	ADMS-Urban	Données tri-horaires du 1er sept 2010 au 31 août 2011 mesurées à la station Météo-France Prado Hippodrome de Marseille	Estimation à partir des mesures réalisées en continu aux stations permanentes d'Air PACA	inventaire des émissions (trafic 2007)	2007 2007	une dizaine de grandes sources ponctuelles

D. Distribution des concentrations du NO₂

Le Tableau 3 résume les statistiques descriptives de la distribution spatiale du NO₂ sur les 4 agglomérations. La valeur limite du NO₂ en moyenne annuelle définie par la Directive de 2008 est de 40 µg/m³ (53); on peut constater que cette limite est dépassée en moyenne sur la période pour 2 des 4 agglomérations, soit à Lyon et Paris. Paris a la moyenne annuelle la plus élevée et Lille la moyenne la plus faible. Marseille et Paris sont caractérisées par une variation interquartile entre les IRIS plus exposés et les IRIS les moins élevés plus importante.

Tableau 3. Statistiques descriptives des concentrations ambiantes du dioxyde d'azote par IRIS dans les 4 agglomérations pour la période 2002-2009.

Agglomération	N†	Moyenne	ET	Min	Q1	Médiane	Q3	Max	ΔIQ
Lille	478	32,1	5,3	21,8	28,8	31,9	35,2	58,2	6,4
Lyon	493	41,3	5,9	29,5	36,6	41,2	45,3	60,1	8,7
Paris	2656	45,8	9,6	23,9	38,7	44,8	52,4	83,3	13,7
Marseille	544	32,5	9,9	8,8	25,3	32,4	39,9	59,4	14,6

N†, le nombre d'IRIS de la zone d'étude. ΔIQ, la variation interquartile (Q3-Q1)

La distribution spatiale des concentrations de NO₂ des 4 agglomérations à l'échelle de l'IRIS sur la période d'étude 2002-2009 est représentée dans la figure 9. Elle met en évidence des concentrations plus élevées dans le centre urbain de Lille, Lyon et Marseille et à proximité des axes de forte circulation des agglomérations. Les concentrations diminuent lorsque l'on s'éloigne de ces villes importantes. Ainsi, les concentrations les plus élevées coïncident avec le trafic automobile dense, particulièrement à Lyon avec le boulevard extérieur et la sortie sud vers l'autoroute A7, à Lille les quartiers qui longent l'A1, A25 et l'axe Lille-Roubaix. A Paris les arrondissements au Nord de la Seine et les quartiers le long du périphérique, ainsi que les quartiers proches du port et de l'aéroport à Marseille ont des concentrations plus élevées.

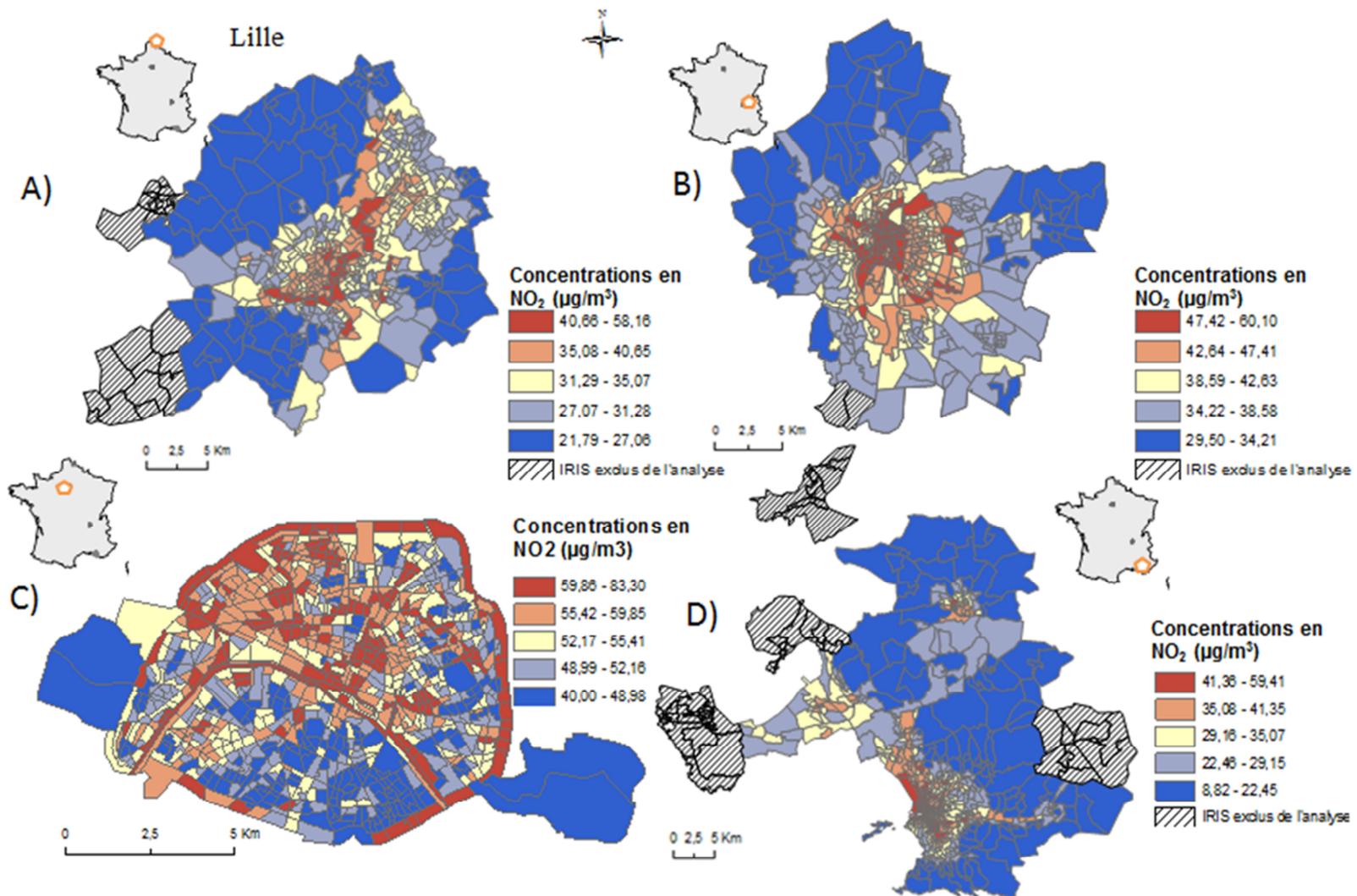


Figure 9 : Cartographies des concentrations moyennes par IRIS du dioxyde d'azote en quintiles pour la période 2002-2009 à A) Lille, B) Lyon, C) Paris, D) Marseille.

V. DES MÉTHODES STATISTIQUES ADAPTÉES : L'ÉPIDÉMIOLOGIE SPATIALE EN SANTÉ PUBLIQUE

Dans le cadre d'une étude écologique, l'investigation des inégalités sociales et environnementales de santé au sein d'un territoire nécessitait de faire appel à des méthodes statistiques adaptées. Le détail des méthodes d'analyse statistique sera décrit au sein des articles consacrés aux différentes étapes de la thèse (Chapitre 3). Cette partie présente les différentes méthodes statistiques employées ainsi que le contexte dans lesquelles elles ont été utilisées. Les trois méthodes utilisées au cours de la thèse sont le SIG, la méthode bayésienne (modèle de Besag, York and Mollié : BYM), et l'approche spatiale à l'aide de modèles *generalized additive models* (GAM).

A. Choix de ces modèles

Compte tenu de la nature des données et du type d'étude, nous devons prendre en compte 1) l'autocorrélation spatiale, 2) un effectif faible des évènements sanitaires. Une approche pour résoudre ces problèmes de représentation consiste à lisser les estimateurs classiques des risques relatifs en ayant recours à une approche spatiale : modèle Bayésien ou GAM, présentées dans les sections suivantes.

L'autocorrélation spatiale

L'autocorrélation spatiale mesure le degré d'interaction et d'interdépendance entre les observations spatialement localisées. En effet, la représentation géographique d'observations n'est généralement pas aléatoire et répond à des schémas de répartition définis. Dans le cas présent, ce qui est observable dans un IRIS particulier dépend de ce qui est observable dans les IRIS voisins. Ainsi, si l'autocorrélation spatiale est positive, les valeurs similaires d'une variable considérée sont regroupées, de sorte que les IRIS voisins se ressemblent davantage que les IRIS éloignés. Si cette autocorrélation n'est pas prise en compte dans les tests statistiques, il peut être conclu à tort à l'existence d'associations statistiquement significatives ou inversement (54).

Le faible effectif des évènements sanitaires.

Lorsque la maladie étudiée est rare et/ou que les unités géographiques sont de petites tailles, les SMR présentent plusieurs limites (55). Tout d'abord, ces estimateurs présentent une variabilité d'autant plus importante que le nombre de cas attendus dans l'unité géographique est faible (56). Ces estimateurs sont donc instables dans les unités peu peuplées et des valeurs extrêmes y sont fréquemment observées. La représentation cartographique des SMR peut alors prendre l'apparence d'une mosaïque de couleurs difficilement interprétable où les colorations les plus fortes – valeurs extrêmes – correspondent le plus souvent aux unités géographiques les moins peuplées. Ce type de représentation donne artificiellement plus de poids aux SMR extrêmes et peut ainsi masquer une éventuelle structure spatiale sous-jacente qui pourrait être le reflet de facteurs environnementaux, par exemple.

B. Approche Bayésienne

Cette méthode statistique a été réalisée en tant qu'étude préliminaire aux analyses spatiales visant à améliorer la compréhension des inégalités socio-spatiales de santé dans nos zones d'études. L'objectif vise à décrire la distribution des cas de mortalité infantile et néonatale selon la défaveur sociale du quartier à l'aide d'une étude de corrélation géographique.

L'approche bayésienne utilisée dans notre étude a fait l'objet d'un travail méthodologique approfondi par Benoit Lalloué, doctorant en mathématiques appliquées. Pour plus de détails, la méthodologie développée a fait l'objet d'un article publié récemment (57).

Lissage des risques relatifs

L'idée sous-jacente du lissage est de réduire la variabilité aléatoire des SMR en partageant l'information qu'apportent les unités géographiques voisines. Le lissage des risques relatifs a été réalisé à l'aide du modèle hiérarchique bayésien proposé par Besag et al en 1991 (connu comme le BYM model) (58,59). L'estimation des risques relatifs s'effectue ainsi vers le SMR moyen de tous les IRIS de la zone d'étude lorsqu'il n'y a pas ou peu de dépendance spatiale (lissage global) et vers le SMR moyen des IRIS voisins lorsque les SMR présentent une structure spatiale (lissage local).

Présentation du modèle BYM

Le modèle Bayésien est un modèle gaussien autorégressif conditionnel structuré selon trois niveaux hiérarchiques. Le premier niveau suppose que le nombre d'événements observés O_i dans l'IRIS i , conditionnellement au risque relatif de survenue de l'événement θ_i , suit une loi de Poisson d'espérance $E_i \cdot \theta_i$, où E_i est le nombre d'événements attendus dans l'IRIS i (Equation 1).

$$O_i \sim \text{Poisson}(E_i \theta_i) \quad (\text{Equation 1})$$

Le deuxième niveau associe au logarithme du risque relatif θ_i une distribution de probabilité définie comme la somme d'un paramètre constant, α , et de deux composantes aléatoires indépendantes u_i et v_i . Ces deux composantes permettent de tenir compte respectivement de la variabilité structurée spatialement (appelée composante de *clustering*) et de la variabilité non structurée (appelée composante d'hétérogénéité) (Equation 2).

$$\text{Log}(\theta_i) = \alpha + u_i + v_i \quad (\text{Equation 2})$$

Ces deux composantes sont considérées comme respectivement distribuées selon une loi normale conditionnelle autorégressive et une loi normale (Equation 3).

$$\begin{aligned} u_i / u_j &\sim N(\bar{u}_i, \sigma_u^2 / m_i) \\ v_i &\sim N(0, \sigma_v^2) \end{aligned} \quad (\text{Equation 3})$$

où u_j correspond aux effets spatiaux pour tous les IRIS autres que l'IRIS i , \bar{u}_i la moyenne de u_i pour tous les IRIS voisins à l'IRIS i , m_i le nombre de voisins de l'IRIS i , et σ_u^2 et σ_v^2 les variances respectives des distributions des composantes u_i et v_i . La définition du voisinage utilisée pour ce travail repose sur une matrice de contiguïté nodale du 1^{er} ordre comme celle présentée dans la note méthodologique du chapitre 2.

Le troisième niveau associe enfin à chacun des paramètres des distributions de probabilité appelées « distributions hyper a priori » : α est ainsi considéré comme distribué selon une loi normale avec une très importante variance, et l'inverse des variances $1/\sigma_u^2$ et $1/\sigma_v^2$

(paramètres également connus sous le nom de précisions) suivent respectivement des distributions Gamma (Equation 4).

$$\begin{aligned}\alpha &\sim N(0, 10000) \\ 1/\sigma_u^2 &\sim Ga(\alpha, \beta) \\ 1/\sigma_v^2 &\sim Ga(\alpha, \beta)\end{aligned}\tag{Equation 4}$$

L'estimation des paramètres de ce modèle fait appel à des intégrales qui ne sont pas calculables par des méthodes analytiques. Il est nécessaire de faire appel à des méthodes de simulation (algorithme de Monte Carlo par chaînes de Markov), mises en œuvre facilement avec le logiciel WinBUGS(60). Les résultats de cette analyse ont été évalués à l'aide du critère de déviance (*Deviance Information Criterion*, DIC). Le DIC met en contraste une mesure de la qualité d'ajustement du modèle avec une mesure de sa complexité. Il s'interprète de la même que le critère d'Akaike, de sorte que plus il est faible plus l'ajustement du modèle est considéré comme bon.

C. Approche Spatiale : GAM

Cette méthode statistique a été réalisée dans deux parties de la thèse (Chapitre3, partie 2 et 4). Dans un premier temps, le but de ces analyses était d'identifier les zones à risque plus élevé de mortalité infantile et néonatale à fine échelle spatiale, des zones d'inégalités sociales, environnementales ou de cumul d'inégalités. Dans un deuxième temps, le but était de connaître l'évolution des inégalités socio-spatiale de santé, et l'évolution du rôle du dioxyde d'azote dans ces inégalités.

Le lissage cartographique des risques

L'objectif du lissage est de fournir de nouveaux estimateurs des risques relatifs à partir des estimations initiales du risque et de la taille de la population de la zone géographique. Le lissage utilisé est le LOESS (locally weighted regression smoothers) qui s'adapte à la taille de la population (61,62). Des précédentes études l'ont utilisées (63–66). Le degré de localité du lissage dépend du pourcentage de zones incluses dans le sous-ensemble, dénommé « span ». Le model GAM permet d'obtenir la taille «optimale» du span en minimisant le critère

d'Akaike (AIC). Si le span est petit, les cartes représenteront des petites surfaces bosselées et si le span est grand, les cartes représenteront des surfaces plus lisses. Dans le but de me former à ces outils et méthodes, j'ai effectué un stage à Boston avec le Professeur Veronica Vieira de la Boston University School of Public Health, qui m'a initiée à ces méthodes innovantes (63–66). Les lecteurs intéressés par ces méthodes pourront néanmoins trouver des descriptions méthodologiques complètes dans les nombreux ouvrages consacrés à la modélisation (GAM) et à leur application, notamment dans le champ de l'épidémiologie spatiale (67–69).

Présentation du modèle GAM

Nous avons utilisé des modèles GAM pour estimer les risques de mortalité infantile à l'échelle de l'IRIS, une forme de régression non-paramétrique ou semi paramétrique avec la capacité d'analyser des données contextuelles avec ajustement sur des covariables. Sans la fonction du lissage « S (X, Y) », le modèle devient une régression de Poisson ordinaire sur les variables.

La formule est la suivante :

$$\text{Log [p (X, Y)]} = \text{S (X, Y)} + \text{offset } \gamma'Z (\text{pop}) + (\text{équation 1})$$

où à gauche, le logarithme du risque de la maladie situé au centroïde de l'IRIS est localisé (X, Y), selon la taille de la population (offset (pop)), et γ est un vecteur de paramètres associé à Z, les covariables.

Les modèles GAM répondent à plusieurs hypothèses dont la deuxième nécessite que la première soit significative. Le schéma d'interprétation des résultats issus de ces modèles se décline en deux étapes :

1. - méthode de détection d'une tendance globale d'agrégation à l'aide du test global de déviance. Cette méthode dite non spécifique vise à tester de manière générale la présence d'agrégation. Hypothèse H_0 : homogénéité du territoire / H_1 : existence de zones de risques plus élevés statistiquement que les autres ou existence de cluster de risque de mortalité infantile ou néonatale.

2. - méthodes de détection spécifique de cluster. Ces méthodes permettent de localiser et tester les agrégats. Un test appelé « Test de permutation » va déterminer les IRIS exactement qui ont statistiquement un risque plus élevé que les autres. Hypothèse H_0 : pas d'association entre le risque de mortalité et la localisation géographique des autres IRIS.

Pour tester l'hypothèse 2, nous avons réaffecté au hasard les coordonnées des IRIS tout en gardant les nombres de cas, la population et les covariables fixes. Nous avons effectué une permutation de la distribution 999 fois en plus du modèle original. Pour chaque permutation, nous avons effectué un modèle GAM en utilisant le span optimal. Nous avons utilisé le logiciel R et le package gam, écrit par Trevor Hastie et Hastie et Tibshirani 1990 (69), pour effectuer les modèles et le logiciel ArcView 10 (ESRI, Inc., Redlands, en Californie) pour cartographier les résultats de nos analyses.

Détection de Clusters de risques

La composante spatiale de la mortalité infantile et néonatale n'a été que très peu étudiée et uniquement abordée par le biais d'étude de corrélation géographique à l'échelle du district (70), du quartier (7,71,72) ou de la région (73). Connaître les particularités spatiales peut permettre d'identifier des causes locales non évidentes d'une agrégation, et de mettre en place des mesures de prévention adaptées.

L'approche spatiale s'intéresse à la description et à la compréhension des variations géographiques du risque de maladie, en particulier à une échelle fine. Ce schéma épidémiologique s'intéresse, dans le présent travail, à l'augmentation localisée de la mortalité infantile ou néonatale.

Le cluster

Un cluster se traduit par des cas plus regroupés dans l'espace que ne le laisse présager le hasard, compte tenu de la distribution de la population. Dans cette situation, le but de l'analyse spatiale consiste à identifier si la distribution spatiale des cas est « normale » (répartition de la maladie au hasard), ou bien s'il existe des agrégats de cas.

Les méthodes de balayage spatial cherchent à détecter l'emplacement des clusters dans la région étudiée. Elles appliquent des fenêtres sur toute la région et dénombrent les cas et les individus à risque à l'intérieur et à l'extérieur de chaque fenêtre. Il existe différentes méthodes de balayage spatial et elles se distinguent entre autres par la construction de la fenêtre qu'elles utilisent.

D. Le système d'information géographique

Un SIG a permis l'édition des résultats des analyses sous forme cartographique. Il a été utilisé pour plusieurs objectifs:

La visualisation de la localisation des données sanitaires, sociales et environnementales : le SIG permet de géoréférencer (attribuer des coordonnées géographiques à un objet afin de le localiser dans l'espace) des données et de mettre en œuvre des données de nature très différentes et qui n'ont souvent pas la même résolution spatiale. Grâce à l'observation de ces données, c'est l'ensemble du contexte environnemental qui peut être décrit et mieux appréhendé.

La construction de clusters (présentation dans la partie approche spatiale): le SIG est utilisé pour combiner les données afin de créer de nouvelles informations pour permettre l'analyse statistique (détection de cluster, régression de Poisson, etc.).

La communication : grâce aux outils du SIG pourra être cartographiée l'information de manière efficace et directement utilisable pour la prise de décision et la communication des résultats d'une étude épidémiologique auprès des pouvoirs publics.

Références : Chapitre 2

1. Rican S, Jouglà E, Salem G. Inégalités sociospatiales de mortalité en France. *Bulletin épidémiologique Hebdomadaire*. 2003;142–5.
2. Pépin, P, ORS Ile de France, ORS Pays de la Loire. Inégalités sociales de santé en France. 2012. Available from: <http://www.mipes.org/-Inegalites-sociales-de-sante-en-.html>
3. Luo Z-C, Wilkins R, Kramer MS. Effect of neighbourhood income and maternal education on birth outcomes: a population-based study. *CMAJ*. 2006 May 9;174(10):1415–20.
4. Marmot M. Social determinants of health inequalities. *Lancet*. 2005 Mar;365(9464):1099–104.
5. Krieger N, Williams DR, Moss NE. Measuring social class in US public health research: concepts, methodologies, and guidelines. *Annu Rev Public Health*. 1997;18:341–78.
6. Meijer M, Röhl J, Bloomfield K, Grittner U. Do neighborhoods affect individual mortality? A systematic review and meta-analysis of multilevel studies. *Soc Sci Med*. 2012 Apr;74(8):1204–12.
7. Dummer TJB, Parker L. Changing socioeconomic inequality in infant mortality in Cumbria. *Arch Dis Child*. 2005 Feb;90(2):157–62.
8. Elo I, Culhane J, Kohler I, O'Campo P, Burke J, Messer L, et al. Neighbourhood deprivation and small-for-gestational-age term births in the United States. *Paediatr Perinat Epidemiol*. 2009 Jan;23(1):87–96.
9. Zeka A, Melly S, Schwartz J. The effects of socioeconomic status and indices of physical environment on reduced birth weight and preterm births in Eastern Massachusetts. *Environ Health*. 2008;7:60.
15. (INSEE). IRIS. Institut National de la statistique et des études économiques. 2013. Available from: <http://www.insee.fr/fr/methodes/default.asp?page=definitions/iris.htm>
11. Haynes KE, Lall SV, Trice MP. Spatial issues in environmental equity. *International Journal of Environmental Technology and Management*. 2001 Jan 1;1(1):17–31.
12. Bowen W. An analytical review of environmental justice research: what do we really know? *Environ Manage*. 2002 Jan;29(1):3–15.
13. Lawson A. *Statistical Methods in Spatial Epidemiology*. New Jersey; 2006.
14. Barcelo MA, Saez M, Saurina C. Spatial variability in mortality inequalities, socioeconomic deprivation, and air pollution in small areas of the Barcelona Metropolitan Region, Spain. *SciTotal Environ*. 2009 Oct 15;407(21):5501–23.
15. Krieger N, Chen J, Waterman P, Soobader M, Subramanian S, Carson R. Choosing area based socioeconomic measures to monitor social inequalities in low birth weight and childhood lead poisoning: The Public Health Disparities Geocoding Project (US). *J Epidemiol Community Health*. 2003 Mar;57(3):186–99.

16. Stroh E, Oudin A, Gustafsson S, Pilesjö P, Harrie L, Strömberg U, et al. Are associations between socio-economic characteristics and exposure to air pollution a question of study area size? An example from Scania, Sweden. *International Journal of Health Geographics*. 2005 Nov 16;4(1):30.
17. Greenland S. Ecologic versus individual-level sources of bias in ecologic estimates of contextual health effects. *Int J Epidemiol*. 2001 Dec;30(6):1343–50.
18. Greenland S, Morgenstern H. Ecological bias, confounding, and effect modification. *Int J Epidemiol*. 1989 Mar;18(1):269–74.
19. Zeitlin J, Saurel-Cubizolles M-J, De Mouzon J, Rivera L, Ancel P-Y, Blondel B, et al. Fetal sex and preterm birth: are males at greater risk? *Hum Reprod*. 2002 Oct;17(10):2762–8.
20. Zeitlin J, Ancel P-Y, Larroque B, Kaminski M, EPIPAGE Study. Fetal sex and indicated very preterm birth: results of the EPIPAGE study. *Am J Obstet Gynecol*. 2004 May;190(5):1322–5.
21. Braveman PA, Cubbin C, Egerter S, Chideya S, Marchi KS, Metzler M, et al. Socioeconomic status in health research: one size does not fit all. *JAMA*. 2005 Dec 14;294(22):2879–88.
22. Gordon D. Census based deprivation indices: their weighting and validation. *J Epidemiol Community Health*. 1995 Dec;49 Suppl 2:S39–44.
23. Galobardes B, Shaw M, Lawlor DA, Lynch JW, Davey Smith G. Indicators of socioeconomic position (part 2). *J Epidemiol Community Health*. 2006 Feb;60(2):95–101.
24. Morris R, Carstairs V. Which deprivation? A comparison of selected deprivation indexes. *J Public Health Med*. 1991 Nov;13(4):318–26.
25. Lalloué B, Monnez J-M, Padilla C, Kihal W, Le Meur N, Zmirou-Navier D, et al. A statistical procedure to create a neighborhood socioeconomic index for health inequalities analysis. *Int J Equity Health*. 2013;12:21.
26. Costa G, D'errico A. Inequalities in health: do occupational risks matter? *Eur J Public Health*. 2006;16(4):340.
27. Branis M, Linhartova M. Association between unemployment, income, education level, population size and air pollution in Czech cities: Evidence for environmental inequality? A pilot national scale analysis. *Health & Place*. 2012 Sep;18(5):1110–4.
28. Townsend, P, Phillimore, P, Beattie, A. *Health and deprivation: inequality and the North*. Routledge. 1988;
29. Carstairs V, Morris R. Deprivation: explaining differences in mortality between Scotland and England and Wales. *BMJ*. 1989 Oct 7;299(6704):886–9.
30. Carstairs V. Deprivation indices: their interpretation and use in relation to health. *J Epidemiol Community Health*. 1995 Dec;49 Suppl 2:S3–8.

31. Havard S, Deguen S, Bodin J, Louis K, Laurent O, Bard D. A small-area index of socioeconomic deprivation to capture health inequalities in France. *Soc Sci Med*. 2008 Dec;67(12):2007–16.
32. Deguen S, Lalloue B, Bard D, Havard S, Arveiler D, Zmirou-Navier D. A small-area ecologic study of myocardial infarction, neighborhood deprivation, and sex: a Bayesian modeling approach. *Epidemiology*. 2010 juillet;21(4):459–66.
33. Laurent O, Pedrono G, Segala C, Filleul L, Havard S, Deguen S, et al. Air pollution, asthma attacks, and socioeconomic deprivation: a small-area case-crossover study. *Am J Epidemiol*. 2008 Jul 1;168(1):58–65.
34. Havard S, Deguen S, Zmirou-Navier D, Schillinger C, Bard D. Traffic-related air pollution and socioeconomic status: a spatial autocorrelation study to assess environmental equity on a small-area scale. *Epidemiology*. 2009 Mar;20(2):223–30.
35. Lalloué B, Monnez J-M, Padilla C, Kihal W, Le Meur N, Zmirou-Navier D, et al. A statistical procedure to create a neighborhood socioeconomic index for health inequalities analysis. *Int J Equity Health*. 2013;12:21.
35. Fiches Techniques. Rapport technique_Equit'Area. Atmo_PACA. 2013. Available from: http://www.atmopaca.org/fiches_thematiques.php
37. Bilan de la qualité de l'air en France en 2009 et des principales tendances observées au cours de la période 2000-2009. MEEDDM (Ministère de l'Ecologie, de l'Energie, du Développement durable et de la Mer); 2010 p. 48.
38. Woodruff TJ, Darrow LA, Parker JD. Air pollution and postneonatal infant mortality in the United States, 1999-2002. *Environ Health Perspect*. 2008 Jan;116(1):110–5.
39. Son J-Y, Cho Y-S, Lee J-T. Effects of air pollution on postneonatal infant mortality among firstborn infants in Seoul, Korea: case-crossover and time-series analyses. *Arch Environ Occup Health*. 2008;63(3):108–13.
40. Scheers H, Mwalili SM, Faes C, Fierens F, Nemery B, Nawrot TS. Does air pollution trigger infant mortality in Western Europe? A case-crossover study. *Environ Health Perspect*. 2011 Jul;119(7):1017–22.
41. Hajat S, Armstrong B, Wilkinson P, Busby A, Dolk H. Outdoor air pollution and infant mortality: analysis of daily time-series data in 10 English cities. *J Epidemiol Community Health*. 2007 Aug;61(8):719–22.
42. Jerrett M, Arain A, Kanaroglou P, Beckerman B, Potoglou D, Sahuvaroglu T, et al. A review and evaluation of intraurban air pollution exposure models. *J Expo Anal Environ Epidemiol*. 2004 online;15(2):185–204.
43. Chaix B, Gustafsson S, Jerrett M, Kristersson H, Lithman T, Boalt A, et al. Children's exposure to nitrogen dioxide in Sweden: investigating environmental injustice in an egalitarian country. *J Epidemiol Community Health*. 2006 Mar;60(3):234–41.
44. Rotko T, Kousa A, Alm S, Jantunen M. Exposures to nitrogen dioxide in EXPOLIS-Helsinki: microenvironment, behavioral and sociodemographic factors. *J Expo Anal Environ Epidemiol*. 2001 Jun;11(3):216–23.

45. Yanosky JD, Schwartz J, Suh HH. Associations between measures of socioeconomic position and chronic nitrogen dioxide exposure in Worcester, Massachusetts. *J Toxicol Environ Health Part A*. 2008;71(24):1593–602.
46. Tabacova S, Balabaeva L, Little RE. Maternal exposure to exogenous nitrogen compounds and complications of pregnancy. *Arch Environ Health*. 1997 Oct;52(5):341–7.
47. Tabacova S, Baird DD, Balabaeva L. Exposure to oxidized nitrogen: lipid peroxidation and neonatal health risk. *Arch Environ Health*. 1998 Jun;53(3):214–21.
48. Walters DM, Breyse PN, Wills-Karp M. Ambient urban Baltimore particulate-induced airway hyperresponsiveness and inflammation in mice. *Am J Respir Crit Care Med*. 2001 Oct 15;164(8 Pt 1):1438–43.
49. Carruthers DJ, Edmunds HA, Lester AE, McHugh CA, Singles RJ. Use and validation of ADMS-Urban in contrasting urban and industrial locations. *International Journal of Environment and Pollution*. 2000;14(1/2/3/4/5/6):364–74.
50. McHugh C, Carruthers D, Edmunds H. ADMS-Urban: an air quality management system for traffic, domestic and industrial pollution. *Int J Environ Pollution*. 1997;8:437–40.
51. Soulhac L, Salizzoni P, Cierco FX, Perkins R. The model SIRANE for atmospheric urban pollutant dispersion; part I, presentation of the model. *Atmospheric Environment*. 2011 Dec;45(39):7379–95.
52. Soulhac L, Salizzoni P, Mejean P, Didier D, Rios I. The model SIRANE for atmospheric urban pollutant dispersion; PART II, validation of the model on a real case study. *Atmospheric Environment*. 2012 Mar;49:320–37.
34. OMS | Santé et qualité de l'air. WHO. Available from: <http://www.who.int/mediacentre/factsheets/fs313/fr/>
54. Waller, LA, Gotway, CA. Global indexes of spatial autocorrelation. *Applied Spatial Statistics for Public Health Data.*; 2004.
55. Richardson S. Modélisation statistique des variations géographiques en épidémiologie. *Rev Epidemiol Sante Publique*. 1992;(40):33–45.
56. Esteve, J, Benhamou, E, Raymond, L. Méthodes statistiques en épidémiologie descriptive. Les éditions INSERM. Paris; 1993.
57. Padilla C, Lalloué B, Pies C, Lucas E, Zmirou-Navier D, Séverine D. An Ecological Study to Identify Census Blocks Supporting a Higher Burden of Disease: Infant Mortality in the Lille Metropolitan Area, France. *Matern Child Health J*. 2013 Mar 12;
58. Besag J, York J, Mollié, A. Bayesian image restoration, with two applications in spatial statistics.
59. Mollie A. Bayesian mapping of Hodgkin's disease in France. Oxford: Oxford University Press; 2000.
60. Lawson A, Browne W, Vidal Rodeiro C. Disease Mapping with WinBUGS and MLWin. Chichester, United Kingdom; 2003.

61. Cleveland WS. LOWESS: A program for smoothing scatterplots by robust locally weighted regression. *American Statistics*. 1981;35–54.
62. Cleveland WS, Grosse E, Shyu MJ. Local regression models. *Statistical Models in S*. New York: Chapman and Hall; 1992.
63. Vieira V, Webster T, Weinberg J, Aschengrau A. Spatial-temporal analysis of breast cancer in upper Cape Cod, Massachusetts. *Int J Health Geogr*. 2008;7:46.
64. Vieira V, Webster T, Weinberg J, Aschengrau A, Ozonoff D. Spatial analysis of lung, colorectal, and breast cancer on Cape Cod: an application of generalized additive models to case-control data. *Environ Health*. 2005;4:11.
65. Webster T, Vieira V, Weinberg J, Aschengrau A. Method for mapping population-based case-control studies: an application using generalized additive models. *Int J Health Geogr*. 2006;5:26.
66. Kelsall, J, Wakefield J. Modeling spatial variation in risk: a geostatistical approach. *American Statistical Association*. 2002;692–701.
67. Kiffer CRV, Camargo ECG, Shimakura SE, Ribeiro PJ Jr, Bailey TC, Pignatari ACC, et al. A spatial approach for the epidemiology of antibiotic use and resistance in community-based studies: the emergence of urban clusters of *Escherichia coli* quinolone resistance in Sao Paulo, Brasil. *Int J Health Geogr*. 2011;10:17.
68. Wood, SN. Stable and efficient multiple smoothing parameter estimation for generalized additive models. *Journal of American Statistical Association*. 2004;673–86.
42. Hastie T, Tibshirani R. *Generalized Additive Models*. London. 1990. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/1964808>
70. Rosicova K, Madarasova Geckova A, van Dijk JP, Kollarova J, Rosic M, Groothoff JW. Regional socioeconomic indicators and ethnicity as predictors of regional infant mortality rate in Slovakia. *Int J Public Health*. 2011 Oct;56(5):523–31.
71. Guildea ZE, Fone DL, Dunstan FD, Sibert JR, Carlidge PH. Social deprivation and the causes of stillbirth and infant mortality. *Arch Dis Child*. 2001 Apr;84(4):307–10.
72. Smith LK, Manktelow BN, Draper ES, Springett A, Field DJ. Nature of socioeconomic inequalities in neonatal mortality: population based study. *BMJ*. 2010;341:c6654.
73. Dallolio L, Di Gregori V, Lenzi J, Franchino G, Calugi S, Domenighetti G, et al. Socio-economic factors associated with infant mortality in Italy: an ecological study. *Int J Equity Health*. 2012;11:45.

CHAPITRE 3 :

EXPOSITION AU DIOXYDE D'AZOTE ET INÉGALITÉS SOCIO-SPATIALES AU REGARD DE LA MORTALITÉ NÉONATALE ET INFANTILE

Dans ce troisième chapitre, nous présenterons les résultats décrivant la relation entre les 3 familles de variables qui sont au cœur de ce projet de recherche.

- i) l'exposition à la pollution atmosphérique, ici le dioxyde d'azote,
- ii) les caractéristiques socio-économique du lieu de résidence et,
- iii) l'évènement sanitaire, ici la mortalité néonatale et infantile.

Ce chapitre sera structuré de la façon suivante :

- Dans la première partie, nous nous sommes intéressés aux inégalités socio-spatiales de santé sur les agglomérations de Lille, Paris (intra-muros) et Lyon (Marseille a été exclu de cette partie au vu des problèmes rencontrés dans la collecte des données sanitaires conduisant à un taux d'exhaustivité trop faible).
- Dans la deuxième partie, nous explorerons spatialement les inégalités sociales de santé décrites dans la partie précédente et l'influence possible de l'exposition au NO₂.
- Forte des résultats des parties précédentes et au vu de l'évolution temporelle décroissante des concentrations moyennes des NO₂ documentée sur la période de l'étude, la 3^{ème} partie s'est plus particulièrement intéressée à l'évolution spatio-temporelle des inégalités environnementales sur les 4 agglomérations françaises.
- Enfin, dans une 4^{ème} et dernière partie, nous explorons l'évolution spatio-temporelle des inégalités sociales de santé et l'influence possible du NO₂. Cette dernière partie, combinée aux résultats des 3 parties vise à répondre aux objectifs de la thèse et explore le rôle de l'exposition à la pollution atmosphérique sur les inégalités sociales de santé en dégagant les spécificités de chaque agglomération et les déterminants communs.

Pour chaque étude, une brève présentation du contexte, les articles publiés ainsi que les résultats complémentaires seront présentés pour les agglomérations de Lille, Lyon et Paris. L'agglomération de Marseille est ajoutée pour les parties n'incluant pas les données sanitaires. La description des données, des méthodes employées, des résultats ainsi que les perspectives sont plus particulièrement détaillées dans les articles publiés.

I. **ETUDE DES INÉGALITÉS SOCIO-SPATIALES DE SANTÉ**

Les inégalités sociales de santé peuvent présenter des visages très différents en France, selon les profils et l'histoire des territoires. C'est pourquoi, nous avons examiné s'il existait des inégalités sociales dans la survenue du décès dans le premier mois ou la première année au sein des 2 agglomérations de Lille et Lyon et de la ville de Paris. Pour cela, nous avons mis en œuvre des modèles statistiques permettant de prendre en compte l'autocorrélation spatiale et la rareté de l'évènement sanitaire.

A. **Contexte**

Un nombre croissant d'articles dans la littérature scientifique ont démontré que la grossesse et les trois premières années de vie d'un enfant sont cruciaux en termes de devenir de la santé physique et émotionnelle à l'âge adulte (1,2). Le développement du cerveau ainsi que les changements biochimiques et hormonaux opérés au cours de cette période, peuvent être sensibles et facilement influencés par les facteurs environnementaux (contextuels) et relationnels (1).

Cette fragilité peut être aggravée par plusieurs facteurs, dont le plus connu est le contexte social. Le contexte social peut influencer la survenue de la grossesse, son déroulement et son suivi, la santé de la femme enceinte, et aussi le devenir du fœtus et la survie du nouveau-né. La survenue d'une grossesse chez les adolescentes est plus élevée chez les familles économiquement défavorisées ou monoparentales (absence du père biologique du nouveau-né notamment) (3). Les caractéristiques socio-économiques des parents définies selon plusieurs indicateurs comme la nationalité, le niveau d'éducation et les ressources du ménage influencent également cette période de fragilité biologique (4,5).

Au-delà du contexte social, la localisation géographique du lieu de résidence est également un facteur important à considérer dans l'étude des inégalités sociales de santé. Les inégalités spatiales et les inégalités sociales interagissent à travers les facteurs suivants : la distance, la configuration spatiale et la division sociale de l'espace (6). La position géographique de

l'adresse de résidence des parents peut permettre un accès direct ou indirect, difficile ou facile aux infrastructures de soins, aux professionnels de santé, faciliter ou contrarier la cohésion ou l'isolement social (6). De nombreux facteurs d'accessibilité (un réseau de transport efficient, la disponibilité d'une voiture, un agencement des personnels de santé) peuvent être associés au niveau socio-économique du territoire. Les quartiers les mieux désirables et bien aménagés sont aussi les plus onéreux (6).

Nous déterminerons dans cette partie si les populations qui habitent dans des quartiers socio-économiquement plus défavorisés ont un risque de mortalité infantile ou néonatale plus élevé que dans les quartiers plus favorisés, et si ce risque suit un gradient selon la défaveur socio-économique du quartier de résidence. Un indicateur composite de défaveur socio-économique a été utilisé afin d'englober les multiples dimensions qui caractérisent la défaveur en milieu urbain. Les données sanitaires ont été recueillies dans les mairies sur la période 2000-2009 pour Lille et Lyon et 2004-2009 pour Paris. Nous avons utilisé des modèles hiérarchiques bayésiens afin de réduire la variabilité issue de l'autocorrélation spatiale dans le calcul de risques relatifs (RR) et d'évaluer les associations entre la mortalité infantile, néonatale et la défaveur socio-économique.

B. Publication et résultats

1. Article publié

Article joint: "An Ecological Study to Identify Census Blocks Supporting a Higher Burden of Disease: Infant Mortality in the Lille Metropolitan Area, France » publié par le journal Maternal and Child Health

An Ecological Study to Identify Census Blocks Supporting a Higher Burden of Disease: Infant Mortality in the Lille Metropolitan Area, France

Cindy Padilla · Benoit Lalloué · Cheri Pies ·
Emminarie Lucas · Denis Zmirou-Navier ·
Deguen Séverine

© Springer Science+Business Media New York 2013

Abstract In France, reducing social health inequalities has become an explicit goal of health policies over the past few years, one of its objectives is specifically the reduction of the perinatal mortality rate. This study investigates the association between infant mortality and social deprivation categories at a small area level in the Lille metropolitan area, in the north of France, to identify census blocks where public authorities should prioritize appropriate preventive actions. We used census data to establish a neighbourhood deprivation index whose multiple dimensions encompass socioeconomic characteristics. Infant mortality data were obtained from the Lille metropolitan area municipalities to estimate a death rate at the census tract level. We used Bayesian hierarchical models in order to reduce the extra variability when computing relative risks (RR) and to assess the associations between infant mortality and deprivation. Between 2000 and 2009, 668 cases of infant

death occurred in the Lille metropolitan area (4.2 per 1,000 live births). The socioeconomic status is associated with infant mortality, with a clear gradient of risk from the most privileged census blocks to the most deprived ones (RR = 2.62, 95 % confidence interval [1.87; 3.70]). The latter have 24.6 % of families who were single parents and 29.9 % of unemployed people in the labor force versus 8.5 % and 7.7 % in the former. Our study reveals socio-spatial disparities in infant mortality in the Lille metropolitan area and highlights the census blocks most affected by the inequalities. Fine spatial analysis may help inform the design of preventive policies tailored to the characteristics of the local communities.

Keywords Infant mortality · Small-area analysis · Social health inequalities · Census tracts · Deprivation index · Bayesian models

C. Padilla (✉) · B. Lalloué · E. Lucas · D. Zmirou-Navier ·
D. Séverine
EHESP School of Public Health (Ecole des Hautes Etudes en
Santé Publique), Rennes, Sorbonne Paris Cité, avenue du
Professeur Léon Bernard, 35043 Rennes Cedex, France
e-mail: cindy.padilla@ehesp.fr

B. Lalloué
e-mail: benoit.lalloue@ehesp.fr

E. Lucas
e-mail: emminarie@gmail.com

D. Zmirou-Navier
e-mail: denis.zmirou@inserm.fr

D. Séverine
e-mail: severine.deguen@ehesp.fr

C. Padilla · B. Lalloué · D. Zmirou-Navier · D. Séverine
INSERM U1085 (IRSET), Rennes, France

C. Padilla
French Environment and Energy Management Agency 20,
avenue du Grésillé-BP 90406, 49004 Angers Cedex, France

B. Lalloué
Institut Elie Cartan UMR 7502, Nancy University, CNRS,
INRIA, Vandoeuvre-les-Nancy, France

C. Pies
School of Public Health, University of California, Berkeley,
CA 94720, USA
e-mail: cpies@berkeley.edu

D. Zmirou-Navier
Lorraine University Medical school, Vandoeuvre-les-Nancy,
France

Introduction

Infant mortality is recognized by the World Health Organization (WHO) as an indicator of the health status of a population and of the effectiveness of the health care system [1]. Large infant mortality disparities between countries were highlighted by the Organization for Economic Cooperation and Development in 2008. At that time, Japan, Finland and Sweden registered the lowest infant mortality rates, with 2.8 deaths per 1,000 live births, while New Zealand, Canada and the United States reported the highest ones (>5 deaths per 1,000 live births) [1]. In France, the infant mortality rate declined steadily from 8.6 to 3.5 deaths per 1,000 live births between 1990 and 2005 but, in recent years, there has been a small increase in this rate ($3.8 \cdot 10^{-3}$ in 2008 [2]). Such disparities are also visible at a finer spatial scale, such as between or within regions or provinces. For example, whereas the infant mortality rate of the Champagne-Ardennes region (north-eastern France) was $5 \cdot 10^{-3}$, the Provence-Alpes-Côte d'Azur region (southern France) recorded a rate of $3 \cdot 10^{-3}$ [2].

Several individual risk factors of infant mortality have been well documented in the literature, including a wide range of health related conditions, such as prematurity, low birth weight [3], gestational age, intrauterine growth restriction or congenital anomalies [4–6], and also maternal age (teenage mothers and mothers aged 40 and over), multiple births, smoking habits [7], and quality of prenatal care [5, 8]. Socioeconomic characteristics have also been reported related to increased risk of infant mortality, such as mother's race or ethnicity, maternal educational level [6] and marital status (infants of unmarried mothers have higher risks than those of married women) [4]. In the US, David et al. suggest that social rather than genetic factors underlie the racial gap. For example, birth weight is similar for neonates of U.S.-born white mothers and Africa-born black mothers, but it is significantly lower for neonates of second-generation African immigrants. [9] In several studies, after controlling for several well-known individual risk factors, significant associations between maternal or parental socioeconomic status and infant mortality persisted [4, 10–13].

Recent literature has established that the neighborhood environment of mother and child has an influence on future birth outcomes independently of individual risk factors [14–19]. A Spanish ecological study conducted in the Barcelona metropolitan area demonstrated a relationship between socioeconomic deprivation and all causes of mortality [20]. In France, Deguen et al. [21] found in the Strasbourg metropolitan area that the risk of myocardial infarction increased with the deprivation index in the community. With regard to infant mortality and socioeconomic status/factors, few such studies have been published. In one example, Guildea [15] demonstrated a clear socio-spatial

gradient in Wales. While ecological studies on infant mortality and socio-economic status are scarce, this study design allowed to show associations between birth outcomes (low birth weight, small for gestational age) and the socioeconomic status (SES) [22–24].

Census blocks have been recognized by several authors as the best geographic unit for exploring the spatial variability in many health indicators [16, 25, 26]. To identify geographic areas with an unfavourable infant mortality risk and help designing health policies, ecological studies are very useful. However, this type of study requires a rigorous methodology. That is why we used a robust methodology: a Bayesian hierarchical model. The Bayesian hierarchical model is commonly used to measure spatially structured variation in risk of rare diseases; it allowed us to examine the influence of contextual factors and to take into account the spatial autocorrelation which expresses the non-independence of geographical observations.

In this context, we designed an ecological study to investigate social health inequalities related to infant mortality in the Lille Metropolitan area (LMA) with a view to identify the deprived census blocks where public authorities should design relevant preventive actions in priority.

Methods

Study Setting and Geographical Level

Our study setting is the LMA, an urban area of 85 municipalities (about 612 km²) located in the Nord-Pas-de-Calais district in northern France. The LMA had 1,106,885 inhabitants in 2007. The small-area level used was the French census block, a sub-municipal division designated by the National Census Bureau; its French acronym is IRIS. The LMA is subdivided into 504 French census blocks with a mean population of 2,165 inhabitants (range 0–4,473 inhabitants). It is the smallest administrative unit in France for which socioeconomic and demographic information is available from the national census. The division of neighborhoods into census blocks takes into account the physical obstacles such as major roadways, bodies of water and green spaces and aims to maximize their homogeneity in term of population size and socioeconomic characteristics. According to Krieger et al. [16] census blocks may be a relevant spatial scale for measuring socioeconomic disparities and assessing socioeconomic gradients in health or environmental exposure.

Health Data

The total number of cases was obtained at the municipality level from the Epidemiological Center for Medical Causes

of Death (CépiDc-Inserm). We extracted data from all the municipalities of LMA between 2000 and 2009. Infant mortality rates were computed as the ratio of total number of cases to the total number of births during that period. In order to derive these rates at the French census block level, data also had to be obtained from the City Halls of the LMA municipalities. In France, only death and birth registries of City Halls have the postal addresses required to geo-code the death events at the census block level. By comparing the CepiDc and City Hall data, we were able to assess the exhaustiveness of the information. Each event was assigned to the subject's French census block of residence after geocoding the parents' address using an INSEE software (Correspondance Adresses-Zones Urbaines 2004) that contains a census block code for each postal address in the LMA. The French National Commission for Digitalized Information and Liberty (CNIL) gave its consent to retrieve geocode and analyze the health data.

Neighborhood Deprivation Index

To characterize the neighborhood deprivation, we used a socioeconomic deprivation index developed in France at the census block level (more details are given elsewhere [27, 28]). To start, 52 socioeconomic variables covered by the INSEE 1999 census and reflecting different dimensions of deprivation (income, educational level, job, housing characteristics, ownership of basic goods, family structure, etc.) were selected. Principal-components analysis was used to synthesize information from these data. To construct a single numeric index over all the blocks, we maximized the inertia of the first component by deleting all the variables weakly correlated with it and the variables with a contribution lower than the average. This method allowed us to identify an axis composed of 19 variables that explained 60.7 percent of the inertia of the initial variables which was used as a deprivation index. In our study, the method has been reworked to improve the robustness of the indicator. We have demonstrated the ability of this index to capture socio-spatial inequalities in health and environmental outcomes [21, 29]. Census blocks were categorized into five deprivation categories corresponding approximately to the quintiles of the index distribution; the first category comprised the most privileged blocks, and the fifth the most deprived [29].

Statistical Analysis

To represent the geographic variation of the relative risk of infant mortality, the classical approach is to calculate the standardized mortality ratio (i.e. the ratio between the observed number of cases and the expected number of cases). Nevertheless, the standardized mortality ratio could

be problematic in case of rare events, which can produce estimations with very large variance and not representative of the reality. One way solve this problem is to estimate relative risks using Bayesian modelling.

We successively applied two models to assess the effect of spatial autocorrelation and heterogeneity on the associations of interest. To accommodate the spatial dependence of the data and the variability of infant mortality rates due to the small number of events per geographic unit, we used a hierarchical Bayesian modeling approach [18] proposed by Besag et al. [30], the intrinsic Gaussian conditional autoregressive model. This model controls both the spatial heterogeneity and the autocorrelation component of our data. It is structured into three hierarchical levels.

The first level is defined following the classical methodological process. Let O_i represent the number of infant deaths in census block i , $i = 1, \dots, n$ (in our framework, i ranges from 1 to 506, the total number of census blocks). We assume that the model is described by the following equation:

$$O_i | \theta_i \sim \text{Poisson}(E_i \theta_i) \tag{1}$$

where E_i denotes the expected cases in the census block i under the assumption of equal risk among all the study area and θ_i is the relative risk of the health event in the census block i (the maximum likelihood estimator of θ_i is the SMR $\frac{O_i}{E_i}$).

The second level assigns to θ_i a log-normal prior distribution defined by $\text{Log}(\theta_i) = \alpha + \sum_k X_{ik} \beta_k + u_i + v_i$ (2) where α is the intercept in the least deprived blocks, X_{ik} are census block-specific covariates with corresponding parameters β_k , and u_i and v_i correspond to two census block-specific random components that model the effects varying in a spatially structured way (i.e., census block clustering) and in an unstructured way (i.e., census block heterogeneity) respectively. We introduced the neighborhood deprivation index into the model in different ways: directly with the quantitative SES index ($X_{ik}=X_i$), as two covariates (X_{ik} with $k \in \{1;2\}$) and as four covariates (X_{ik} with $k \in \{1;2;3;4\}$) corresponding to all the three or five deprivation categories except the first one (the most privileged), used as the reference. The second level also assigns a Gaussian prior distribution and a normal prior distribution to u_i and v_i respectively (Eq. 3).

$$\begin{aligned} u_{-i} | u_{-i} &\sim N(\bar{u}_i, \sigma_u^2 / m_i) \\ v_i &\sim N(0, \sigma_v^2) \end{aligned} \tag{3}$$

where u_{-i} denotes the spatial effects for all blocks other than the block i , \bar{u}_i is the mean of the neighbors' u_i of the block i (adjacency being used as neighborhood criterion), m_i is the number of neighbors of the block i , and σ_u^2 and σ_v^2

are the variance parameters of the u_i and v_i distributions, respectively.

The third level assigns hyper prior distributions to different parameters. Without prior expectations about the covariate effects, a non-informative proper prior distribution is put on the intercept α and the regression coefficients β_k .

$$\alpha \sim N(0, 10000)$$

$$\beta_k \sim N(0, 10000)$$

Classically used for numerical considerations, gamma prior distributions have been associated with the inverse of the variance parameters σ_u^2 and σ_v^2 (Eq. 4).

$$\begin{aligned} 1/\sigma_u^2 &\sim Ga(0.01, 0.01) \\ 1/\sigma_v^2 &\sim Ga(0.01, 0.01) \end{aligned} \quad (4)$$

Parameters for the gamma hyperprior distributions were selected after sensitivity analysis using the Deviance Information Criterion (DIC) developed by Spiegelhalter [31] to compare the different models.

Running Model

All models were fitted via Markov Chain Monte Carlo algorithms with WinBUGS software, version 1.4.1 (MRC Biostatistics Unit, Cambridge, UK). Parameter means and 95 % credible intervals (CIs) were estimated from two independent chains of 30,000 iterations after a burn-in of 30,000 iterations. The convergence of chains was checked with the Gelman and Rubin convergence diagnostic as modified by Brooks and Gelman [32]. Relative risk maps were drawn with ArcGIS version 9.1 (ESRI, Redlands, CA).

Results

Over the 10 years of the study period, 13,045 infants were born among whom 705 died and were registered by the CépiDc in LMA. Of the infants who died, 668 cases could be retrieved from the city hall sources (94.8 %). Missing data are dispersed over all the municipalities. Eight infant deaths (1.2 %) were not included in the analysis because the address information was not available. During the 2000–2009 period, the mean infant mortality rate was 4.2 per 1,000 live births (Standard Deviation = 7.2/1,000) and the median value was 3.8/1,000. Figure 1 illustrates the spatial distribution of infant mortality rates by quintiles and shows large spatial variations across the census blocks of LMA, with a null value in 43 % of the census blocks.

Figure 2 shows the spatial distribution of the deprivation index and also presents strong contrasts across the LMA. These contrasts show the most deprived blocks in the

largest, densely populated cities, particularly Lille, Roubaix, and Tourcoing, and the most privileged blocks (categories 1 and 2) on the outskirts of the metropolitan area. The latter have the best living conditions. The main socioeconomic characteristics presented in Table 1 highlight large variations in our study population, contrasting the most privileged blocks and the most deprived ones: the proportion of single parent families (8.5 vs. 24.6 %), the unemployment rate (7.7 vs. 29.9 %), the proportion of buildings and subsidized housing (7.9 vs. 65.9 %), people

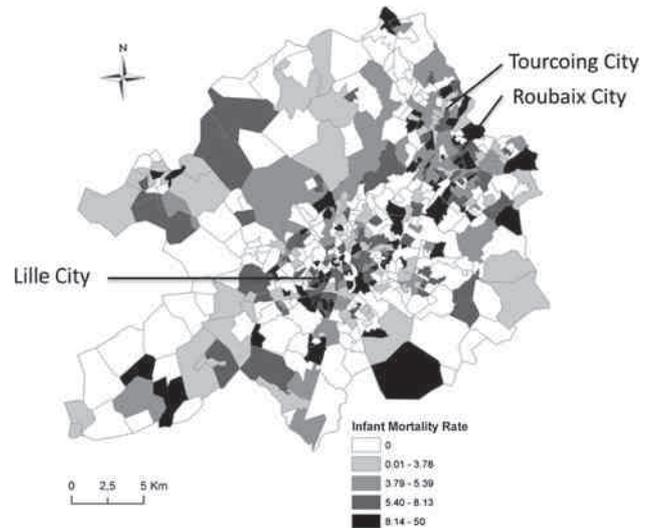


Fig. 1 Map of the infant mortality rates in the Lille metropolitan area at the census block level

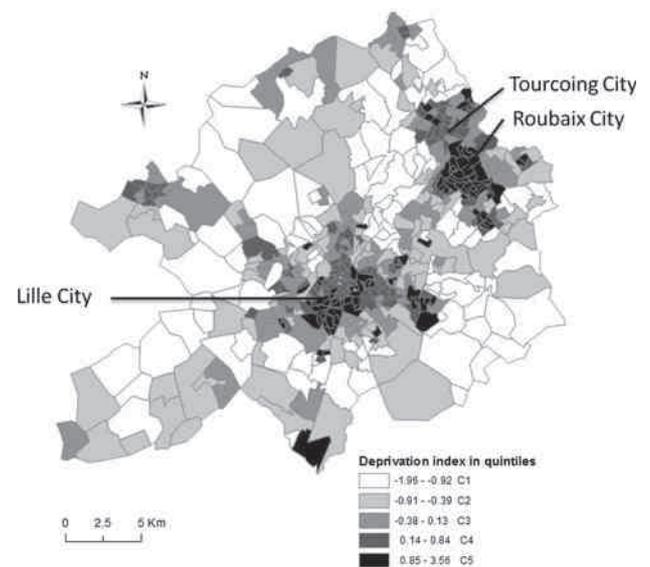


Fig. 2 Map of the deprivation index in the Lille metropolitan area at the census block level from lowest deprivation (white) to highest (black)

Table 1 Mean values of the socioeconomics characteristics, according to deprivation index categories

Variables	Census domain	Deprivation categories				
		C1 ^a	C2	C3	C4	C5
Foreigners in the population (%)	Immigration status	1.63	2.44	4.49	6.62	12.84
Unemployed people in the labor force (%)	Job	7.71	10.15	14.51	18.14	29.96
People unemployed >1 year in the labor force (%)	Job	4.15	5.53	8.69	10.76	18.54
People with insecure jobs in the labor force (%)	Job	7.64	9.38	11.62	14.22	17.18
People with stable jobs in the labor force (%)	Job	73.47	72.69	67.86	61.92	48.48
Blue collar workers in the labor force (%)	Job	14.03	18.39	24.10	23.10	25.76
Single-parent families (%)	Family	8.50	11.39	14.35	17.36	24.58
People aged >15 years with no more than an elementary education (%)	Educational level	12.24	11.37	9.43	9.06	7.55
People aged >15 years with general or vocational maturity certificates (%)	Educational level	11.27	9.75	7.37	7.29	5.01
People aged >15 years with at least a lower postsecondary education (%)	Educational level	13.09	9.52	6.93	9.41	5.23
Primary residences that are houses or farms (%)	Housing	88.96	73.59	66.38	43.92	30.53
Primary residences that are multiple dwelling units (%)	Housing	7.89	23.00	28.40	51.62	65.93
Subsidized housing among all primary residences (%)	Housing	1.69	7.43	15.42	25.48	53.64
Non owner-occupied primary residences (%)	Housing	20.04	34.55	45.07	62.09	79.88
Primary residence with more than 1 person/room (%)	Housing	3.53	5.03	6.52	7.92	13.92
Households without a car (%)	Housing	9.27	16.88	23.84	33.08	43.15
Households with ≥2 cars (%)	Housing	48.29	32.35	23.49	15.74	10.86
Mean no. people/room	Housing	59.88	62.51	65.02	69.62	77.92
Median income/consumption unit (in Euros/year)	Income	28,365	24,321	22,067	19,128	14,719

^a C1 is the least deprived category, C5 is the most

Table 2 Relative risk (RR) and 95 % credible interval (CI) of infant mortality according to census block deprivation index, for several models

Models	With spatial autocorrelation			Without spatial autocorrelation		
	RR (95 % CI)	Deviance	DIC ^a	RR (95 % CI)	Deviance	DIC
Model 1: quantitative deprivation index						
SES	1.60 (1.38–1.84) ^b	1,415.14	1,491.25	1.59 (1.43–1.78) ^b	1,508.81	1,510.85
Model 2: deprivation in five categories ^c						
C1 ^d	1.00	1,419.73	1,493.32	1.00	1,504.63	1,509.64
C2	1.37 (0.96–1.97)			1.36 (0.98–1.91)		
C3	1.67 (1.19–2.36)			1.65 (1.21–2.29)		
C4	2.36 (1.69–3.33)			2.41 (1.79–3.28)		
C5	2.62 (1.87–3.70)			2.72 (2.05–3.64)		
Model 3: without covariates		1,416.90	1,506.95		1,579.19	1,580.18

Table 2 presents relative risks of infant mortality and 95 % CI for three models. Model 1: adjusted on the quantitative deprivation index (continuous), Model 2: adjusted on five deprivation categories and Model 3 without adjustment. These models have been tested with and without spatial autocorrelation. According to the DIC, the models with the best results are those including the deprivation covariates (quantitative or five classes) and with spatial terms. Taking the spatial autocorrelation into account, population located in the most deprived IRIS category (C5) has a 2.62 times greater risk of infant mortality than those located in the less deprived IRIS category (C1)

^a Deviance Information Criterion

^b denotes the relative risk. For an increase of 1 unit of the deprivation index, the infant mortality risk increases by β

^c C1 is the lower deprived category and C5 the most deprived

^d C1 is the reference category

with no more than an elementary education level (7.6 vs. 12.2 %) and a low rate of non-owners of their house or a car (Table 1).

Table 2 shows relative risks of infant mortality and 95 % CI estimated by Bayesian models. Three models were used; the first one according to the deprivation index

on a quantitative scale (Model 1), the second according to the index split into five deprivation categories using quintiles (Model 2) and the last one without adjustment (Model 3). These models have been tested with and without spatial autocorrelation. All results demonstrate a significant influence of deprivation on infant mortality risk, with a clear gradient of risk from the most privileged category (category 1, used as reference) to the most deprived. The latter category presents a 2.62-fold greater risk than the reference category ($RR = 2.62$, $95\% \text{ CI} = [1.87; 3.70]$). Table 2 summarizes the goodness-of-fit of the models, with and without covariates and spatial components. The Deviance Information Criterion is smaller in the models including the deprivation covariates (quantitative or five classes) and in models with spatial terms. Figure 3 shows the geographical location of the estimated relative risks of infant mortality of Model 2 after taking into account spatial autocorrelation at the census block level. The map of smoothed relative risks appears more homogeneous than the infant mortality rate map (Fig. 1). The noise caused by the random instability has been taken into account by the model. However, some IRIS, especially those located in the center and the north east of the city, have still higher risks of infant mortality according to the deprivation categories at the census block level.

Discussion

Using data aggregated at the French census block level and an appropriate statistical method, our study reveals socio-

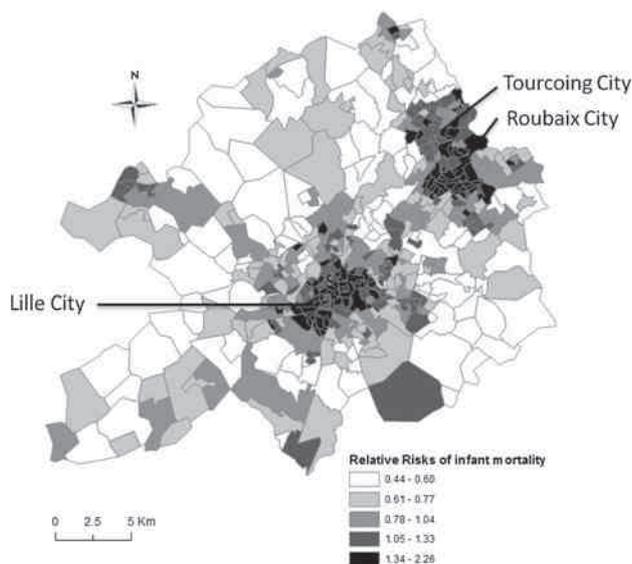


Fig. 3 Map of the estimated relative risks of infant mortality in quintiles from Model 2 with spatial autocorrelation term at the census block level in Lille metropolitan area

spatial disparities in infant mortality in the Lille metropolitan area and highlights the census blocks most affected by the inequalities. The risk is multiplied by 2.62 for the most disadvantaged census blocks compared to the most advantaged. The most advantaged have the best living conditions, with the highest percentages in the following variables: more than two cars in the household, home ownership, stable employment, education level; whereas the most disadvantaged are characterized by the highest percentages in unemployment rate or insecure jobs, building (versus single house) or subsidized housing, few material resources and a low education level (Table 1).

To our knowledge, our study is the first one in France demonstrating the existence of social health inequalities regarding infant mortality at the census block level. Several geographical studies had explored health inequalities for a panel of pregnancy outcomes [15, 16, 23, 33]. Guildea, in Wales, showed that post-neonatal mortality (death occurring between 28 days of life to 1 year) is strongly associated with the Townsend neighbourhood deprivation index; the risk is 1.53 times higher in poor districts compared to privileged districts [15]. An American study conducted over a long period (1969–2001) revealed the persistence of a significant infant mortality gradient according to the neighbourhood deprivation index [34].

A positive trend of increased infant mortality in the disadvantaged neighbourhoods was observed. Also, the percentage of people with a low level of education in the disadvantaged census blocks is three times higher than in the privileged ones (30.2 vs. 9.4 %). Several multilevel studies found an association between the neighborhood SES adjusted on individual maternal education and birth outcomes [18, 34–36]. Singh et al. [34] used a comprehensive area-based deprivation index and an individual measure of maternal education to analyze the extent to which socioeconomic patterns in US population were related to infant, neonatal, and postneonatal mortality. Similarly, in Norway, a population based study showed an inverse association between the level of education of parents and neonatal mortality [37].

The quality and quantity of prenatal care are also related to socioeconomic status. It is well documented that late entry into prenatal care is strongly associated with prematurity [18]. In the United States, early prenatal care varies according to the socioeconomic, physical and social characteristics of the neighborhood, determined at the census block level, and reduces the risk of low birth weight [19]. A study conducted in ten European countries, including France, demonstrates that disparities in access to health care and distance from the mother's residence to the nearest maternity and neonatal intensive care units vary across regions [38]. Urban planning and social characteristics of the urban area partly explain inequalities in access

to health care, which encompasses proximity and diversity of services [39].

Socioeconomic status can also be seen as an indirect determinant of infant mortality. Because it generates stressful situations, social deprivation may shorten the gestation of pregnant women and therefore constitute a risk factor for early mortality after birth. An American study revealed a strong association between maternal stress during pregnancy and low birth weight or prematurity, taking into account socioeconomic variables such as income, race, household crowding and level of education in the neighbourhood [40]. At another level, neighbourhood social cohesion may play a role in modulating the effect of the socioeconomic status of the household on health events related to births [33]. Mothers receiving low or moderate social support and living in low income neighbourhoods had an increased risk of giving birth to premature or low birth weight infants compared to those receiving high social support, in a study conducted in the United States [40]. Also, an extensive epidemiologic literature strongly suggests that genetic factors fail to explain the disparities in birth outcomes according to race in the US, which are better explained by social and environmental determinants [9]. In our study, the most disadvantaged blocks averaged 25.3 % single parent families compared to only 9 % in the most privileged blocks.

Our results are based on relatively complete data, since only eight cases were discarded because they could not be geocoded at the census block level due to lack of information on the residence address. Given the data availability at the time of the study, the deprivation index is composed of data from the 1999 French population census. Due to the ecological design of our study, the results obtained at the block level cannot be extrapolated to the individual level. Only a multilevel analysis that would include both individual and contextual variables, social and demographic factors (e.g. maternal age, marital status, education level and occupational status of parents) would make it possible to refine the results. Hence, we do not have information on behaviours that influence the health of the newborn (maternal smoking during pregnancy, consumption of fruits and vegetables, exercise, etc.) or the regularity and quality of care prenatal received.

However, the deprivation index is robust and validated, consisting of variables covering all areas of deprivation (including education, occupation...) that are commonly used in the literature [21, 27, 41]. The spatial scale, census block, is the smallest area available in France (about 2,000 inhabitants on average) so that the observed associations are likely to converge towards estimates that would have been observed at the individual level, should such data be available. Other factors that may affect the exposure of the mother to environmental stressors during pregnancy or

during the first year of life of the child, such as change of residence during pregnancy or professional mobility, quality of housing, proximity to major roads and associated pollution, occupational exposures or, on the contrary, proximity to green spaces, were not considered in our study.

One strength of this study is that it takes into account the spatial autocorrelation in the relationship between socioeconomic status and infant mortality. The need to account for this phenomenon was reported previously [27, 41], showing how it could influence results and produce spurious associations [21, 27, 41]. In the present case, the estimates remain significant with or without spatial autocorrelation, but the accuracy of the estimates and the fit of the model were improved.

The deprivation index used to measure inequalities at a fine geographical scale is a second strength of the present study. It is based on a rigorous methodology and includes 19 socioeconomic indicators encompassing the various domains of deprivation. This method allowed us to identify specific socioeconomic determinants of disparities in infant mortality across the Lille metropolitan area. It provides useful information to public health bodies in order to devise targets for local or more general preventive actions. In other words, our approach not only allows to flag communities where the risk is highest, but also provides information on the social and economic characteristics of these communities upon which appropriate and focused preventive policies can be devised and implemented, for instance, enhancing the offer of prenatal health and social services in the concerned census blocks, in particular in direction of single mothers and of migrant families.

In conclusion, this study confirms that the most disadvantaged populations are at higher risk of infant mortality in a large European metropolitan area. In a more practical sense, it illustrates how an ecological study conducted on a small geographical scale can become a tool to identify precisely the communities where preventive actions should be undertaken. The nature of the appropriate preventive actions depends on the local population characteristics. To the current evidence, our study adds the possibility to use this type of analysis and mapping to describe some important socioeconomic patterns of the local communities in a way that helps decision makers and community representatives to elaborate on the most relevant policies and actions. Clearly, these policies and actions not only involve health and prenatal care but call for an array of policies regarding, among others, employment, access to subsidized housing and to good quality education.

Acknowledgments The authors thank all scientific researchers of the Equit Area project, personnel of the local council of Lille Metropolitan area and Grégoire Rey of the CeperiDC institute for their

participation. Equit'Area project was supported by the French National Research Agency (ANR), the EHESP School of Public Health, the French Health General Directorate (DGS), the Environment and Energy Management Agency (ADEME) and the Nord-Pas de Calais region.

Conflict of interest None declared.

References

1. Organisation for Economic Co-operation and Development (OECD) (2008). Infant mortality rates in OCDE countries 2008. Available: <http://www.oecd.org/dataoecd/30/54/43136879.xls>; 20-6-2010.
2. The National Institute of Statistics and Economic Studies (INSEE). Infant mortality rate in France metropolitan. Available: http://www.indices.insee.fr/bsweb/servlet/bsweb?action=BS_SERIE&BS_IDBANK=000436398&BS_IDARBO=01000000000000; 2011.
3. Moser, K., Macfarlane, A., Chow, Y. H., Hilder, L., & Dattani, N. (2007). Introducing new data on gestation-specific infant mortality among babies born in 2005 in England and Wales. *Health Statistics Quarterly*, 35, 13–27.
4. Euro-peristat project. (2010). European perinatal health report 2008. Available: <http://www.europeristat.com>; 2010.
5. Mathews, T. J., & MacDorman, M. F. (2008). Infant mortality statistics from the 2005 period linked birth/infant death data set. *National Vital Statistics Reports*, 57(2), 1–32.
6. Oakley, L., Maconochie, N., Doyle, P., Dattani, N., & Moser, K. (2009). Multivariate analysis of infant death in England and Wales in 2005–06, with focus on socio-economic status and deprivation. *Health Statistics Quarterly*, 42, 22–39.
7. Gray, R., Bonellie, S. R., Chalmers, J., Greer, I., Jarvis, S., Kurinczuk, J. J., et al. (2009). Contribution of smoking during pregnancy to inequalities in stillbirth and infant death in Scotland 1994–2003: Retrospective population based study using hospital maternity records. *British Medical Journal*, 1(339), b3754. doi: 10.1136/bmj.b3754.
8. Brunner, E. (1997). Stress and the biology of inequality. *British Medical Journal*, 314(7092), 1472–1476.
9. David, R., & Collins, J. Jr. (2007). Disparities in infant mortality: What's genetics got to do with it? *American Journal of Public Health*, 97(7), 1191–1197.
10. Reading, R. (1997). Poverty and the health of children and adolescents. *Archives of Disease in Childhood*, 76(5), 463–467.
11. Lopez-Azpiazu, I., Sanchez-Villegas, A., Johansson, L., Petkeviciene, J., Prattala, R., & Martinez-Gonzalez, M. A. (2003). Disparities in food habits in Europe: systematic review of educational and occupational differences in the intake of fat. *Journal of Human Nutrition & Dietetics*, 16(5), 349–364.
12. Irala-Estevéz, J. D., Groth, M., Johansson, L., Oltersdorf, U., Prattala, R., & Martínez-González, M. A. (2000). A systematic review of socio-economic differences in food habits in Europe: consumption of fruit and vegetables. *European Journal of Clinical Nutrition*, 54(9), 706–714.
13. Gilman, S. E., Abrams, D. B., & Buka, S. L. (2003). Socioeconomic status over the life course and stages of cigarette use: Initiation, regular use, and cessation. *Journal of Epidemiology and Community Health*, 57(10), 802–808.
14. Diez Roux, A. V. (2007). Neighborhoods and health: where are we and where do we go from here? *Revue d'épidémiologie et de santé publique*, 55(1), 13–21.
15. Guildea, Z. E., Fone, D. L., Dunstan, F. D., Sibert, J. R., & Cartlidge, P. H. (2001). Social deprivation and the causes of stillbirth and infant mortality. *Archives of Disease in Childhood*, 84(4), 307–310.
16. Krieger, N., Chen, J. T., Waterman, P. D., Soobader, M. J., Subramanian, S. V., & Carson, R. (2003). Choosing area based socioeconomic measures to monitor social inequalities in low birth weight and childhood lead poisoning: The public health disparities geocoding project (US). *Journal of Epidemiology and Community Health*, 57(3), 186–199.
17. Marmot, M. (2007). Achieving health equity: from root causes to fair outcomes. *Lancet*, 370(9593), 1153–1163.
18. Luo, Z. C., Wilkins, R., & Kramer, M. S. (2006). Effect of neighbourhood income and maternal education on birth outcomes: A population-based study. *Canadian Medical Association Journal*, 174(10), 1415–1420.
19. O'Campo, P., Xue, X., Wang, M. C., & Caughy, M. (1997). Neighborhood risk factors for low birthweight in Baltimore: A multilevel analysis. *American Journal of Public Health*, 87(7), 1113–1118.
20. Cambra, K., Martínez-Rueda, T.,onso-Fustel, E., Cirarda, F.B., Ibanez, B., Esnaola, S., Calvo, M., Aldasoro, E., Montoya, I. (2010) Mortality in small geographical areas and proximity to air polluting industries in the Basque Country (Spain). *Occup Environ Med*.
21. Deguen, S., Lalloue, B., Bard, D., Havard, S., Arveiler, D., & Zmirou-Navier, D. (2010). A small-area ecologic study of myocardial infarction, neighborhood deprivation, and sex: A Bayesian modeling approach. *Epidemiology*, 21(4), 459–466.
22. Dummer, T. J., Dickinson, H. O., Pearce, M. S., Charlton, M. E., & Parker, L. (2000). Stillbirth risk with social class and deprivation: No evidence for increasing inequality. *Journal of Clinical Epidemiology*, 53(2), 147–155.
23. Elo, I. T., Culhane, J. F., Kohler, I. V., O'Campo, P., Burke, J. G., Messer, L. C., et al. (2009). Neighbourhood deprivation and small-for-gestational-age term births in the United States. *Paediatric and Perinatal Epidemiology*, 23(1), 87–96.
24. Zeka, A., Melly, S. J., & Schwartz, J. (2008). The effects of socioeconomic status and indices of physical environment on reduced birth weight and preterm births in Eastern Massachusetts. *Environmental Health*, 7, 60.
25. Barcelo, M. A., Saez, M., & Saurina, C. (2009). Spatial variability in mortality inequalities, socioeconomic deprivation, and air pollution in small areas of the Barcelona Metropolitan Region, Spain. *The Science of the Total Environment*, 407(21), 5501–5523.
26. Lawson, A. (2006). Statistical methods in spatial epidemiology. (2nd edn). New Jersey.
27. Havard, S., Deguen, S., Bodin, J., Louis, K., Laurent, O., & Bard, D. (2008). A small-area index of socioeconomic deprivation to capture health inequalities in France. *Social Science and Medicine*, 67(12), 2007–2016.
28. Elliott, P., Wakefield, J., & Best, N. (2000). *Spatial epidemiology: Methods and applications*. Oxford: Oxford University Press ed.
29. Laurent, O., Filleul, L., Havard, S., Deguen, S., Declercq, C., & Bard, D. (2008). Asthma attacks and deprivation: Gradients in use of mobile emergency medical services. *Journal of Epidemiology and Community Health*, 62(11), 1014–1016.
30. Besag, J., York, J., Mollié, A. Bayesian image restoration, with two applications in spatial statistics.
31. Spiegelhalter, D., Best, N., & Carlin, B. (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society: Series B*, 64, 583–640.
32. Brooks, S., & Gelman, A. (1998). A alternative methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7, 434–455.

33. Blumenshine, P., Egerter, S., Barclay, C. J., Cubbin, C., & Braveman, P. A. (2010). Socioeconomic disparities in adverse birth outcomes: A systematic review. *American Journal of Preventive Medicine*, *39*(3), 263–272.
34. Singh, G. K., & Kogan, M. D. (2007). Persistent socioeconomic disparities in infant, neonatal, and postneonatal mortality rates in the United States, 1969–2001. *Pediatrics*, *119*(4), e928–e939.
35. O'Campo, P., Xue, X., Wang, M. C., & Caughy, M. (1997). Neighborhood risk factors for low birthweight in Baltimore: A multilevel analysis. *American Journal of Public Health*, *87*(7), 1113–1118.
36. Pearl, M., Braveman, P., & Abrams, B. (2001). The relationship of neighborhood socioeconomic characteristics to birthweight among 5 ethnic groups in California. *American Journal of Public Health*, *91*(11), 1808–1814.
37. Arntzen, A., Samuelsen, S. O., Bakketeig, L. S., & Stoltenberg, C. (2004). Socioeconomic status and risk of infant death. A population-based study of trends in Norway, 1967–1998. *International Journal of Epidemiology*, *33*(2), 279–288.
38. Pilkington, H., Blondel, B., Papiernik, E., Cuttini, M., Charreire, H., Maier, R. F., et al. (2010). Distribution of maternity units and spatial access to specialised care for women delivering before 32 weeks of gestation in Europe. *Health & place*, *16*(3), 531–538.
39. Charreire, H., & Combier, E. (2009). Poor prenatal care in an urban area: A geographic analysis. *Health & place*, *15*(2), 412–419.
40. Nkansah-Amankra, S., Dhawain, A., Hussey, J. R., & Luchok, K. J. (2010). Maternal social support and neighborhood income inequality as predictors of low birth weight and preterm birth outcome disparities: Analysis of South Carolina pregnancy risk assessment and monitoring system survey, 2000–2003. *Maternal and Child Health Journal*, *14*(5), 774–785.
41. Havard, S., Deguen, S., Zmirou-Navier, D., Schillinger, C., & Bard, D. (2009). Traffic-related air pollution and socioeconomic status: A spatial autocorrelation study to assess environmental equity on a small-area scale. *Epidemiology*, *20*(2), 223–230.

2. Compléments

Méthode

L'article publié portant uniquement sur la mortalité infantile dans l'agglomération de Lille, cette partie vise à apporter des compléments sur les autres sites d'étude (l'agglomération de Lyon et la ville Paris) et concerne aussi la mortalité néonatale. Les résultats sont issus des mêmes méthodes statistiques (méthode bayésienne) que l'article publié. L'indice de défaveur a cependant été actualisé à partir des données du recensement de 2006 (celles-ci n'étaient pas encore disponibles au moment de l'écriture de l'article). Dans un souci de comparabilité et afin de commenter les résultats issus des 3 agglomérations, les résultats de Lille (publiés dans l'article) ont été mis à jour avec les données du recensement de 2006.

Description des évènements sanitaires

A Lille, Lyon et Paris, 668, 714 et 629 cas respectivement ont été recueillis auprès des mairies, soit un taux d'exhaustivité de 94,7%, 96,5% et 91,9%, respectivement. Le taux moyen de mortalité infantile est de 4,2 à Lille, 3,4 à Paris et 3,8 pour 1000 naissances vivantes à Lyon. Comme pour la mortalité infantile, le taux de mortalité néonatale est le plus élevé à Lille (2,8‰). A Lyon et Paris, il est comparable avec 2,6 et 2,5 pour 1000 naissances vivantes, respectivement.

Description de la défaveur socio-économique

A Lille et Lyon, les quartiers les plus défavorisés présentent environ 3 fois plus de non propriétaire du logement (p value $<0,001$) et entre 6 et 7 fois plus de logement sociaux proportion d'HLM) ($p <0,001$, Tableau 4) en comparaison des quartiers favorisés. Plus spécifiquement, dans les quartiers défavorisés de Lille, 23,8% de la population active est au chômage et seulement 54,6% occupent des emplois stables versus 7,6% et 74,7% dans les quartiers les plus favorisés ($p <0,001$, Tableau 4). A Lyon, dans les quartiers défavorisés, 22% de la population sont des immigrés étrangers et 20,4% de la population active sont ouvriers contre 7% et 9,1% dans les quartiers favorisés (Tableau 4). A Paris, les quartiers défavorisés ont 38 fois plus de logements sociaux HLM ($p <0,001$), 3 fois plus d'ouvriers ($p <0,001$), et

79,3% de non propriétaire du logement contre 59,1% dans les quartiers favorisés ($p < 0,001$, Tableau 4).

Analyse des inégalités socio-spatiales de santé

Les tableaux 5 et 6 présentent les risques relatifs de mortalité infantile et néonatale estimés par les modèles bayésiens. La défaveur socio-économique a été considérée, d'une part, en tant que variable quantitative afin d'analyser la relation avec l'évènement de santé et en classes afin de détecter d'éventuel gradient socio-économique du risque de mortalité infantile ou néonatale. Les classes de défaveur sont catégorisées en tertiles et quintiles ; cependant seuls les meilleurs modèles minimisant le critère du DIC sont été présentés dans les tableaux (DIC : Deviance Information Criteria, critère statistique de meilleur qualité d'ajustement du modèle).

On peut noter que les DIC sont plus faibles pour les modèles intégrant la matrice d'autocorrélation spatiale confirmant l'importance de prendre en compte ce paramètre dans les analyses statistiques de données agrégées sur une unité géographique. Les DIC sont également plus faibles après ajustement sur la défaveur socio-économique (en quantitatif comme en classe). Une partie de la variabilité observée dans les taux de mortalité est expliquée par le niveau de défaveur socio-économique du quartier de résidence. A Lille, on constate des évidents gradients du risque de mortalité infantile et néonatale des quartiers les plus favorisés vers les plus défavorisés. A Paris, les populations issus des quartiers défavorisés sont également plus à risque de mortalité infantile (RR = 1,73, IC à 95% [1,26-2,36]) et néonatale (RR = 1,61, IC à 95% [1,13-2,32]) que les quartiers favorisés. A Lyon, alors qu'on observe un gradient du risque de mortalité néonatale, seuls les quartiers les plus défavorisés sont à risque pour ce qui concerne la mortalité infantile (RR = 1,61, IC à 95% [1,30-2,00]).

Tableau 4: Valeurs moyennes des caractéristiques socio-économiques de l'IRIS selon l'appartenance à la classe de défaveur

Variables	Domaine	Classes de défaveur								
		Lille			Lyon			Paris		
		C1 ^b	C2	C3	C1 ^b	C2	C3	C1 ^b	C2	C3
Population immigré étrangère (%)	Immigration	3,55	6,25	14,14	7,02	10,76	22,03	17,49	18,53	24,13
Chômeurs au sein de la population active (%)	Emploi	7,56	12,94	23,76	6,90	10,00	17,71	8,38	9,99	14,14
Emploi précaire au sein de la population active (%)	Emploi	23,11	25,79	32,24	24,47	27,46	30,14	22,52	25,35	26,58
Emploi stable au sein de la population active (%)	Emploi	74,73	70,29	54,62	72,26	68,87	61,81	63,84	65,48	65,23
Cadres au sein de la population active (%)	Emploi	16,25	10,27	7,26	19,51	16,11	6,92	38,60	34,78	20,11
Ouvriers au sein de la population active (%)	Emploi	11,48	18,28	19,93	9,10	11,31	20,42	3,06	5,20	10,11
Famille monoparentale (%)	Famille	7,64	10,59	14,05	7,49	7,97	12,17	5,44	6,12	10,77
Pop >15 ans avec un CAP-BEP (%)	Niveau Education	17,41	18,67	16,05	15,92	14,58	18,16	4,08	6,58	10,94
Pop >15 ans avec un Bac à Bac+2 (%)	Niveau Education	26,97	21,71	17,72	13,41	12,00	7,49	9,39	10,80	9,81
Pop >15ans avec au minimum un Bac+2 (%)	Niveau Education	15,25	10,74	9,26	19,64	17,62	8,49	43,77	36,40	21,26
Maison ou ferme comme résidence principale (%)	Logement	85,88	63,22	32,66	47,16	13,36	10,31	1,06	0,95	1,40
HLM comme résidence principale (%)	Logement	7,65	19,65	41,75	5,86	13,09	44,55	1,82	8,26	38,69
Immeuble comme résidence principale (%)	Logement	13,33	35,07	66,14	52,32	84,98	88,14	94,16	95,86	95,49
Non propriétaire dans résidence principale (%)	Logement	22,09	45,76	74,53	32,43	56,03	73,20	59,08	64,80	79,28
Ménages sans voitures (%)	Logement	10,55	21,82	39,62	11,47	27,72	32,78	51,19	60,62	60,46
Ménages avec plus de deux voitures (%)	Logement	45,36	26,48	13,48	42,21	19,89	16,69	8,87	4,85	4,49
Revenu médian (en Euros)	Logement	20977	16146	11794	23035	19324	13482	33929	24659	16981

^b C1 classe la moins défavorisée, C3 classe la plus défavorisée.

Tableau 5 : Risques relatifs (RR) et intervalle de crédibilité à 95% de la mortalité infantile de plusieurs modèles selon l'ajustement à la défaveur socio-économique.

Modèles			Sans autocorrelation spatiale			Avec autocorrelation spatiale		
			RR (95% CI)	Deviance	DIC ^a	RR (95% CI)	Deviance	DIC
Lille 2000-2009	Modèle 1: indice de défaveur	SES	1,30 ^β (1,21-1,39)	1386,09	1388,09	1,29 ^β (1,19-1,40)	1336,19	1384,37
	Modèle 2: indice de défaveur en 3 classes*	C1†	1,00			1,00		
		C2	1,30 (1,03-1,64)	1382,71	1385,71	1,29 (1,01-1,65)	1341,29	1383,54
		C3	2,03 (1,65-2,50)			1,98 (1,58-2,49)		
Modèle 3: sans covariables				1437,02	1438,01		1332,86	1402,09
Lyon 2000-2009	Modèle 1: indice de défaveur	SES	1,23 ^β (1,15-1,32)	1496,60	1498,60	1,21 ^β (1,11-1,31)	1453,71	1495,01
	Modèle 2: indice de défaveur en 3 classes*	C1†	1,00			1,00		
		C2	1,14(0,92-1,42)	1495,73	1498,75	1,16 (0,92-1,46)	1451,02	1494,05
		C3	1,67 (1,37-2,03)			1,61 (1,30-2,00)		
Modèle 3: sans covariables				1527,56	1528,56		1451,54	1507,3
Paris 2004-2009	Modèle 1: indice de défaveur	SES	1,60 ^β (1,38-1,84)	2147,99	2150	1,20 ^β (1,11-1,31)	2014,26	2125,50
	Modèle 2: indice de défaveur en 5 classes*	C1†	1,00			1,00		
		C2	1,37 (0,96-1,97)			1,25 (0,90-1,73)		
		C3	1,67 (1,19-2,36)	2152,84	2157,82	1,13 (0,81-1,57)	2017,86	2132,03
		C4	2,36 (1,69-3,33)			1,37 (0,99-1,88)		
		C5	2,62 (1,87-3,70)			1,73 (1,26-2,36)		
Modèle 3: sans covariables				2176,86	2177,86		2024,13	2138,40

^a Deviance Information Criterion

^β désigne les risques relatifs, Pour une augmentation de 1 unité de la défaveur, le risque de mortalité infantile augmente de β ,

*C1 est la classe la plus favorisée et C5 la plus défavorisée, †C1 est la classe de référence

Tableau 6: Risques relatifs (RR) et intervalle de crédibilité à 95% de la mortalité néonatale de plusieurs modèles selon l'ajustement à la défaveur socio-économique.

Modèles			Sans autocorrelation spatiale			Avec autocorrelation spatiale		
			RR (95% CI)	Deviance	DIC ^a	RR (95% CI)	Deviance	DIC
Lille 2000-2009	Modèle 1: indice de défaveur	SES	1,28 ^β (1,18-1,39)	1194,97	1196,98	1,28 ^β (1,16-1,42)	1135,78	1189,22
	Modèle 2: indice de défaveur en 3 classes*	C1†	1,00			1,00		
		C2	1,41 (1,06-1,87)	1194,38	1197,40	1,40 (1,04-1,91)	1138,39	1190,88
		C3	2,01 (1,56-2,60)			1,97 (1,48-2,63)		
	Modèle 3: sans covariables			1226,08	1227,08		1135,29	1200,75
Lyon 2000-2009	Modèle 1: indice de défaveur	SES	1,20 ^β (1,10-1,30)	1307,35	1309,36	1,21 ^β (1,09-1,34)	1209,71	1287,98
	Modèle 2: indice de défaveur en 3 classes*	C1†	1,00			1,00		
		C2	1,44 (1,11-1,89)	1301,23	1304,25	1,46 (1,09-1,97)	1207,84	1284,39
		C3	1,82 (1,42-2,36)			1,83 (1,38-2,43)		
	Modèle 3: sans covariables			1322,78	1323,78		1212,83	1296,05
Paris 2004-2009	Modèle 1: indice de défaveur	SES	1,18 ^β (1,08-1,28)	1768,07	1770,07	1,17 ^β (1,05-1,29)	1675,53	1758,35
	Modèle 2: indice de défaveur en 5 classes*	C1†	1,00			1,00		
		C2	1,18 (0,82-1,71)			1,18 (0,81-1,74)		
		C3	1,27 (0,90-1,81)	1772,72	1777,73	1,27 (0,88-1,86)	1675,04	1763,81
		C4	1,43 (1,03-2,03)			1,43 (1,00-2,07)		
		C5	1,63 (1,18-2,26)			1,61 (1,13-2,32)		
Modèle 3: sans covariables			1780,30	1781,30		1679,31	1764,79	

^a Deviance Information Criterion

^β désigne les risques relatifs, Pour une augmentation de 1 unité de la défaveur, le risque de mortalité infantile augmente de β ,

*C1 est la classe la plus favorisée et C5 la plus défavorisée, †C1 est la classe de référence,

C. Discussion

Notre étude confirme l'existence d'inégalités socio-spatiales en ce qui concerne la mortalité infantile et néonatale. Cette relation varie d'une agglomération à l'autre aussi bien dans sa forme que dans sa force d'association. Alors qu'à Lille on constate un gradient évident du risque de mortalité infantile et néonatale des quartiers les plus défavorisés aux quartiers les plus favorisés, à Paris ce sont uniquement les quartiers les plus défavorisés qui sont statistiquement plus à risque de mortalité infantile et néonatale. Contrairement aux deux autres agglomérations, Lyon se caractérise par une diversité de la relation selon la date de survenue du décès du nouveau-né. Tandis que seuls les quartiers les plus défavorisés sont les plus à risque de mortalité infantile comme à Paris, on constate un gradient de mortalité néonatale selon la défaveur socio-économique du quartier similairement à Lille. Les conditions socio-économiques de la zone de résidence des parents semblent donc influencer différemment la mortalité infantile et la mortalité néonatale. Le taux de mortalité infantile rapporte la période globale néonatale et post-néonatale où la surveillance sanitaire est moins étroite et où la mère est plus laissée à ses propres expériences. La mortalité post-néonatale est considérée comme liée aux causes exogènes (nutritionnelles, infectieuses accidentelles) (7–10). Les causes de la mortalité néonatale sont endogènes et couvrent une période où la prise en charge collective, d'accès aux soins est maximale, néanmoins cette part de la mortalité infantile est plus sensible aux visites prénatales, congés maternité, Plan Périnatalité, etc (7–10).

II. ETUDE DU RÔLE DU NO₂ DANS LES INÉGALITÉS SOCIO-SPATIALES DE SANTÉ

A. Contexte

Les facteurs environnementaux peuvent affecter la santé des individus et de leur communauté. De plus, certains groupes d'individus de par leurs particularités intrinsèques ont une vulnérabilité plus grande aux facteurs environnementaux. Les enfants, les personnes âgées et les femmes enceintes sont particulièrement sensibles aux effets néfastes de la pollution de l'air (11).

Mécanismes expliquant la relation entre exposition à la pollution atmosphérique et aux issues de grossesses défavorables

La grossesse constitue une période particulièrement sensible (développement des organes, forte prolifération cellulaire et évolution du métabolisme du fœtus lors du développement in utero) aux toxines contenues dans l'air pollué ou aux processus biologiques induits par la réaction au stress inflammatoire (12,13). Plusieurs mécanismes biologiques ont été décrits dans la littérature pour expliquer l'influence de la pollution de l'air sur les issues de grossesse (14). Une exposition à un gaz toxique peut entraîner une inflammation des poumons, et donc une libération systémique de cytokines qui peut conduire au déclenchement prématuré de l'accouchement (15). La production d'une inflammation (du placenta, des poumons, ou systémique) et le stress oxydatif, peuvent augmenter la vulnérabilité de la mère aux infections ; ces infections pouvant entraîner une issue de grossesse défavorable ainsi que des complications à l'accouchement (16–18). L'exposition au dioxyde d'azote, plus particulièrement, affecte le système de défense antioxydant du corps humain (19). Il peut en découler une peroxydation du liquide du placenta de nature à perturber le développement postnatal (20). Enfin, le NO₂ peut aussi avoir des effets toxiques directement sur le fœtus (21). Ainsi, même après ajustement sur l'âge gestationnel, le lien entre l'exposition au dioxyde d'azote et les issues de grossesse tels que le taux de naissance avant terme, la diminution du poids à la naissance, le faible poids pour un âge gestationnel donné est bien documenté dans la littérature (13,14,21,22,22–26). Ces facteurs sont par ailleurs reconnus pour être des facteurs de risque de la mortalité néonatale et infantile (27–30).

Le rôle combiné de la défaveur socio-économique et du NO₂

Des études ont démontré que l'effet de la pollution atmosphérique pouvait amplifier un risque d'issue de grossesse défavorable, déjà élevé chez des personnes résidant dans des quartiers défavorisés (31–34). En effet, cette combinaison d'exposition peut entraîner une plus grande vulnérabilité comme un cumul inégalité que Jerrett et al 2001 appelait le « triple jeopardy » : faible statut socio-économique, faible condition environnementale et faible santé (35).

La suite de ce chapitre présente l'intégralité de l'article accepté par *STOTEN*. L'objectif de l'étude est de détecter les zones à risque plus élevé de mortalité infantile et néonatale à fine échelle spatiale et de chercher à les expliquer par la distribution spatiale de défaveur socio-économique et /ou environnementale.

B. Publication et résultats

1. Article publié

Article joint : "Cluster analysis of social and environment inequalities of infant mortality. A spatial study in small areas revealed by local disease mapping in France." Article publié dans le journal Science of the Total Environment.



Contents lists available at SciVerse ScienceDirect

Science of the Total Environment

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

Cluster analysis of social and environment inequalities of infant mortality. A spatial study in small areas revealed by local disease mapping in France

Cindy M. Padilla ^{a,b,c,*}, Severine Deguen ^{a,b}, Benoit Lalloue ^{a,b,e}, Olivier Blanchard ^{a,b}, Charles Beaugard ^d, Florence Troude ^d, Denis Zmirou Navier ^{a,b,e}, Verónica M. Vieira ^{f,g}

^a EHESP School of Public Health –Sorbonne Paris Cité – Rennes, France

^b INSERM U1085-IRSET – Research Institute of Environmental and Occupational Health, Rennes, France

^c French Environment and Energy Management Agency, Angers, France

^d Official Air Quality Monitoring Associations (AASQA), Atmo Nord Pas-de-Calais, Air Rhône-Alpes, France

^e Lorraine University Medical School–Institut Elie Cartan UMR 7502, Nancy University, CNRS, INRIA, Vandoeuvre-les-Nancy, France

^f Department of Environmental Health, Boston University School of Public Health, Boston, MA 02118, USA

^g Program in Public Health, Chao Family Cancer Center, University of Irvine, CA 92697, USA

HIGHLIGHTS

- Existence of different spatial health inequalities across two metropolitan areas in France
- Evidence of clusters of elevated infant mortality due to the neighborhood SES
- Social and environment exposures contribute to spatial inequalities.
- The spatial analysis method is an important tool to determine areas of interventions.

ARTICLE INFO

Article history:

Received 25 May 2012

Received in revised form 7 January 2013

Accepted 8 March 2013

Available online xxxx

Keywords:

Air pollution

Infant mortality

Deprivation index

Small areas

Generalized additive models

ABSTRACT

Mapping spatial distributions of disease occurrence can serve as a useful tool for identifying exposures of public health concern. Infant mortality is an important indicator of the health status of a population. Recent literature suggests that neighborhood deprivation status can modify the effect of air pollution on preterm delivery, a known risk factor for infant mortality. We investigated the effect of neighborhood social deprivation on the association between exposure to ambient air NO₂ and infant mortality in the Lille and Lyon metropolitan areas, north and center of France, respectively, between 2002 and 2009. We conducted an ecological study using a neighborhood deprivation index estimated at the French census block from the 2006 census data. Infant mortality data were collected from local councils and geocoded using the address of residence. We generated maps using generalized additive models, smoothing on longitude and latitude while adjusting for covariates. We used permutation tests to examine the overall importance of location in the model and identify areas of increased and decreased risk.

The average death rate was 4.2‰ and 4.6‰ live births for the Lille and Lyon metropolitan areas during the period. We found evidence of statistically significant precise clusters of elevated infant mortality for Lille and an east-west gradient of infant mortality risk for Lyon. Exposure to NO₂ did not explain the spatial relationship. The Lille MA, socioeconomic deprivation index explained the spatial variation observed.

These techniques provide evidence of clusters of significantly elevated infant mortality risk in relation with the neighborhood socioeconomic status. This method could be used for public policy management to determine priority areas for interventions. Moreover, taking into account the relationship between social and environmental exposure may help identify areas with cumulative inequalities.

© 2013 Elsevier B.V. All rights reserved.

* Corresponding author at: EHESP School of Public Health –Sorbonne Paris Cité – Rennes, France. Tel.: +33 299022600.

E-mail addresses: cindy.padilla@ehesp.fr (C.M. Padilla), severine.deguen@ehesp.fr (S. Deguen), benoit.lalloue@ehesp.fr (B. Lalloue), olivier.blanchard@ehesp.fr (O. Blanchard), c.beaugard@atmo-npd.fr (C. Beaugard), ftroude@air-rhonealpes.fr (F. Troude), denis.zmirou@inserm.fr (D.Z. Navier), vvieira@uci.edu (V.M. Vieira).

1. Introduction

Infant mortality (death less than one year of age) is recognized as a key indicator of the health status of a population (OECD-Organization for Economic Co-operation and Development, 2010). Several studies have investigated the association between air pollution and infant

mortality in countries with relatively high levels, as well as in countries with lower pollution levels (Tsai et al., 2006; Woodruff et al., 2008; Vrijheid et al., 2012; Romieu et al., 2004; Ritz et al., 2006; Lin et al., 2004; Kaiser et al., 2004; Hajat et al., 2007). The recent literature has established that the neighborhood environment of mother and child has an influence on future birth outcomes independently of individual risk factors (O'Campo et al., 1997; Ponce et al., 2005; Luo et al., 2006; G en ereux et al., 2008; Zeitlin et al., 2011).

The neighborhood socioeconomic status (SES) has been mentioned as an important determinant of birth outcomes, in combination with air pollution (Ponce et al., 2005; Carbajal-Arroyo et al., 2011). Low SES populations may be more susceptible to air pollution than those with higher SES, as several factors more prevalent in disadvantaged populations may modify the pollution–mortality relationship (Yi et al., 2010). Genereux et al. have shown that area–level maternal education and the percent of low income families were associated with the distance between the residence and the nearest highway, which, in turn, were related to differences in exposure to air pollution and the probability of preterm birth (G en ereux et al., 2008). In two studies performed in Mexico (Carbajal-Arroyo et al., 2011; Romieu et al., 2004), the risk of death was significantly higher in infants from low and/or medium-SES areas than in those from high SES areas. Most of these studies are focused in the United States, Canada (Salihu et al., 2012; Ponce et al., 2005; G en ereux et al., 2008; Jerrett et al., 2005b) or countries in economic transition (Carbajal-Arroyo et al., 2011; Romieu et al., 2004; Yi et al., 2010). The number of studies in Europe is very limited (Scheers et al., 2011; Vrijheid et al., 2012).

To identify geographic areas with an unfavorable infant mortality risk and provide relevant data to design local health policies, ecological studies are useful. Studies are particularly useful when a fine resolution scale of such areas allows consideration of the territory in terms of social and environmental characteristics. However, this type of study requires a rigorous methodology in order to minimize ecological biases and to account for the dependency of spatial units. An original statistical method applicable in spatial epidemiologic settings is a generalized additive model (GAM) which can be applied with locally weighted regression smoothers (LOESS) to account for geographic location as a possible predictor of the infant mortality rate (Vieira et al., 2005, 2008; Webster et al., 2006). GAMs provide a spatial representation of health risks, which may be a useful tool to understand the distribution of disease, identifying areas of high disease prevalence, and therefore to set up focused public health interventions (Gatrell and Bailey, 1996; Jerrett et al., 2010).

In this paper, we assess social and environmental inequalities in the spatial distribution of infant mortality in two major metropolitan areas in France. This study has several objectives: i) to detect spatial variations of infant mortality across census blocks, ii) to identify areas of significantly increased and decreased risk adjusted by known risk factors (social characteristics and air pollution, both determined at a neighborhood level), and iii) to illustrate the relevance of spatial epidemiology techniques using generalized additive models, smoothing on longitude and latitude, while adjusting for covariates.

2. Materials and methods

2.1. Study sites and study design

The study is ecological and investigates the spatial distribution of infant mortality in two major metropolitan areas (MAs) in France. The Lille metropolitan area (Nord-Pas-de Calais region, northern France), named Lille M etropole, has an approximate population of 1.1 million inhabitants divided into 85 municipalities and 506 census blocks, for a total area of 611.45 km². The Lyon metropolitan area (Rh one-Alpes region, mid-eastern France), named Grand Lyon, is subdivided into 58 municipalities and 510 census blocks for a total population of approximately 1.2 million inhabitants in an area of 527.15 km².

The statistical unit is the sub-municipal French census block (called IRIS “ lot Regroup e pour l'Information Statistique”) defined by the National Institute of Statistics and Economic Studies (INSEE). It is the smallest administrative unit for which socioeconomic and demographic data are available in France. This geographical unit averages 2000 inhabitants and is constructed to be as homogenous as possible in terms of socio-demographic characteristics and land use. The delineations of the census blocks provided by INSEE also take into account the urban landscape and obstacles that could divide it, such as major traffic roads, green places and water bodies. These two metropolitan areas are of particular interest because they exhibit contrasts in their urban landscape and in some important demographic and socioeconomic characteristics.

2.2. Health outcome

Infant mortality is defined as the number of babies who died during their first year of life per number of births that occurred during this time period. Cases were collected from death certificates in the city hall of each municipality in the MA and the parental addresses were geocoded to the census blocks. A total of 516 and 684 cases of infant deaths in Lille MA and the Lyon MA, respectively, occurred during the period 2002–2009. Fig. 1. A illustrate the spatial distribution of the prevalence of infant mortality by tertiles at the census blocks level of Lille and Lyon MAs.

2.3. Potential cofounders

2.3.1. Deprivation index

For the analysis of socioeconomic disparities, a deprivation index was constructed for all census blocks of the metropolitan areas of Lille and Lyon. The detailed methodological development of this index has been described elsewhere (Havard et al., 2008). In short, the socioeconomic data were obtained from the 2006 national census and provided counts of population, households or residences at the census block level classified by social, economic and demographic characteristics. Using these raw data, we constructed 48 indicators at the census block level according to INSEE's definitions. These variables can be divided into 5 domains: family and household, immigration status and mobility, employment and income, education, housing. Principal components analysis was used to synthesize information from these data. To construct a single numeric index for all of the blocks, we maximized the inertia of the first component by deleting all of the variables only weakly correlated with it and the variables with a contribution lower than the average. This allowed us to identify an axis, composed of 21 variables, which explained 63% of the inertia of the initial variables for the Lille MA and 54 for the Lyon MA.

The socioeconomic variables included in both MAs were Foreigners (%), Immigrant population (%), Single-parent families (%), Unemployed people (%), Employed workers (%), People with stable job (%), Non-owner occupying primary residence (%), Population 15 years and over without diploma (%), Population 15 years and over with post-secondary or secondary diploma (%), Individual house as a primary residence (%), Apartment building as a primary residence (%), Primary residence with a minimum surface area of 100 m² (%), Subsidized housing among all primary residences (%), Primary residence with a garage or other parking space (%), Households without a car (%), Households with 2 or more cars (%), and Median income per consumption unit. Some variables are specific to one MA, as People aged 25 years or younger (%) (Lille), People with insecure job (Lille) (%), Self-employed people (Lyon) (%), Managers workers (Lyon) (%), Blue-collar workers (Lyon) (%). Fig. 1.B shows the spatial distribution of the deprivation index by tertiles on a map of the census blocks of Lille and Lyon MA.

2.3.2. Air pollution concentrations

Annually ambient concentrations of nitrogen dioxide (NO₂) were modeled by the local air quality monitoring network (Atmo Nord Pas-de-Calais, Air Rh one-Alpes) for each block throughout the entire

Lille Metropolitan area

Lyon Metropolitan area

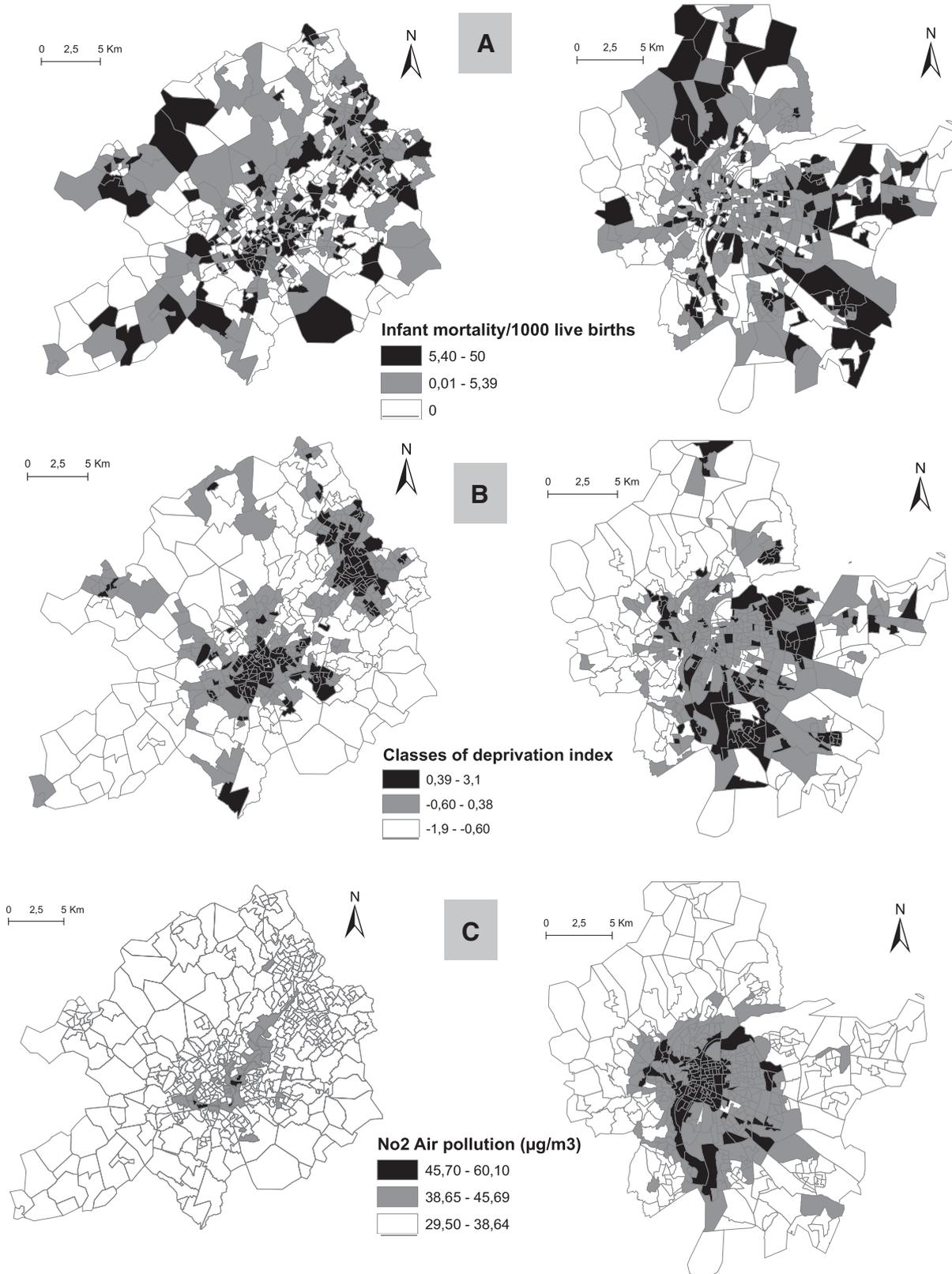


Fig. 1. A. Spatial distribution of infant mortality by census block in the Lyon and Lille MAs during 2002–2009. B. Spatial distribution of the deprivation index by census block in Lyon and Lille MA during 2002–2009. C. Spatial distribution of the NO₂ air pollution by census block in Lyon and Lille MA during 2002–2009.

study period (2002–2009). The two networks developed and tested a methodological approach to describe and characterize disparities in environmental exposures at a local scale for that period. They used different deterministic models: ADMS (Atmospheric Dispersion Modelling System) Urban for the Lille MA (Carruthers et al., 2000; McHugh C et al., 1997) and SIRANE for the Lyon MA (Soulhac et al., 2011, 2012). These models integrate meteorological data: air temperature, wind speed and direction, relative humidity, barometric pressure (supplied by Météo France, the French meteorologic service), emission sources according to their contribution to ambient air pollution, and background pollution measurements as input parameters. Selected emission sources were linear sources (main roads), surface sources (diffuse road sources and residential and tertiary emissions) and important point sources: 31 for the Lille MA and 91 for the Lyon MA (the main polluting industries). The Agglomerative Hierarchical Clustering technique (AHC) was chosen to associate each census block a measuring background permanent station, then assign daily variations. In total, 18 stations were used for the Lille MA and 31 for the Lyon MA. For Lille, in 2009 the mean concentration was modeled using the ADMS model. To reconstruct annual mean concentrations of NO₂ for 2002–2008, the method used consists of a spatial interpolation of data stations that uses spatial concentrations of 2009 as an auxiliary variable. Kriging gives accurate results at the stations for the years 2002–2008. For Lyon, the annual mean concentrations were calculated with the SIRANE model without spatial interpolation, nor measurement assimilation. In a recent review, Jerrett et al. demonstrated the effectiveness and reliability of this type of model for assessing air quality in health effect assessment research (Jerrett et al., 2010). Fig. 1.C shows the spatial distribution of the NO₂ concentrations by tertiles on a map of the census blocks of the Lille and Lyon MAs.

2.4. Statistical methods

2.4.1. Geocoding and descriptive data

Through the efforts to collect cases information, we obtained the parental address of residence of each case, with the authorization of the national CNIL committee on digitalized information and privacy. Residential addresses were matched to the corresponding census blocks using map databases (Correspondence Adresses-Zones Urbaines, 2004) a software issued by INSEE. The following data were available: longitude and latitude of the centroid of each census block, reported number of cases, total births, the modeled NO₂ concentration (µg/m³) and the deprivation index reported in tertiles of the distribution. For east/west comparisons, the Lyon MA was divided using geographic boundaries: the “Rhône” river that flows north-east to south. No geographic boundary divided the Lille MA into eastern and western parts, so we used a vertical line to divide it into two equal surfaces.

Census blocks without any birth (for example, an industrial census block or a park) were excluded from the analysis. We excluded 7 census blocks without births and 2 without socioeconomic information in both Lille and Lyon MAs. We also excluded 24 census blocks (4.5%) in Lille and 13 census blocks in Lyon (2.5%) that had no information on air pollution. The final dataset for Lille included 471 census blocks and 488 census blocks for Lyon during the years 2002 to 2009.

2.4.2. Local disease mapping

We used generalized additive models to estimate census block infant mortality risk, a form of non-parametric or semi parametric regression with the ability to analyze area-based data adjusting for covariates.

We modeled location, a potential proxy measure of unknown exposure or uncontrolled risk factors, using a smooth (S) of longitude (X) and latitude (Y) with a Poisson link function.

$$\text{Log [p(X, Y)]} = S(X, Y) + \text{offset (pop)} + \gamma'Z \quad (1)$$

where the left-hand side is the logarithm of the disease risk at the census block's centroid (X,Y), according to the size of the population

(offset(pop)), and γ is a vector of parameters associated with Z, the vector of covariates.

The model is semi-parametric because it includes both nonparametric and parametric components. Without the smooth function, S(X,Y), the model becomes an ordinary Poisson regression on the covariates. Omitting the covariates produces a crude (unadjusted) map. We used a LOESS smooth which adapts to changes in population density previously used in case control studies (Vieira et al., 2005, 2008; Webster et al., 2006; Kelsall and Wakefield, 2002).

The amount of smoothing depends on the percentage of the data points in the neighborhood, referred to as the span size. GAMs also allow selection of “optimal” span size. We used R software using the GAM package, which is an implementation of the GAM framework of Hastie and Tibshirani 1990, to perform the generalized additive modeling and ArcView 9.3 software (ESRI, Inc., Redlands, California) to map the results of our analyses. We determined the optimal amount of smoothing for each map by minimizing the Akaike's Information Criterion (AIC). Small span sizes produce bumpier surfaces and larger span sizes produce smoother surfaces. As the span size increases, the amount of bias in the fit increases and the variance decreases.

GAMs also provide a framework for testing hypotheses. There are a number of ways to test the global null hypothesis that disease status does not depend on location, i.e., that the map is flat. Similar to analysis of variance in ordinary linear regression, we examined the overall significance of location using the difference in deviance of the complete model (Eq. (1)) and the reduced model omitting the smoothing term. The R software provides an approximate p-value for this statistic assuming a chi square distribution. Because the latter assumption is in general not true for GAMs, we calculated the p-value using a permutation test (Vieira et al., 2002). To test the null hypothesis of no association between infant mortality rate and location, we randomly reassigned the coordinates of the census blocks while keeping the case counts, population, and covariates fixed. We sampled from the null permutation distribution 999 times in addition to the original model. For each permutation, we ran the GAM using the optimal span of the original data and computed the deviance statistic. We divided the rank of the observed value by 1000 to obtain the approximate permutation p-value. If the deviance global statistic indicated that location was significant at the 0.05 level, we then identified areas with significantly increased or decreased risk. We did this by obtaining a distribution of the log risk at every census block using the same set of permutations we used for calculating the global statistics. The areas of significantly decreased risk (“cold spots”) include all census blocks that rank in the lower 2.5% of the census blocks distributions. Areas of significantly elevated risk (“hot spots”) include all census blocks that rank in the upper 2.5% of the census blocks distributions. (Vieira et al., 2005; Young et al., 2010).

We first performed a spatial analysis using the crude model to determine the unadjusted geographic variation in infant mortality. Spatial patterns in the underlying crude analysis could be due to a number of factors with a geographic component. In this study, we were primarily interested in spatial patterns that can be explained by the deprivation index or the NO₂ concentrations. To assess the contribution of these factors to the underlying spatial patterns, we performed adjusted analyses with the deprivation index alone, the NO₂ concentration alone, the deprivation index and NO₂ concentration together, and with their interaction.

3. Results

3.1. Descriptive statistics

The infant mortality rate is equal to 4.2 and 4.6 per 1000 live births for the Lille and Lyon MAs, respectively, between 2002 and 2009. The classes of tertiles of the socioeconomic deprivation and the NO₂ concentration for both MAs are summarized in the Table 1. We note that the classes of tertiles of NO₂ concentrations of the Lyon MA are higher than in the Lille MA (Table 1). Fig. 2 shows the temporal trends of the

NO₂ concentrations during the period 2002 to 2009. In the Lille MA, the meteorology was penalizing in terms of dispersion during years 2003, 2005, which led to higher annual levels of nitrogen dioxide. The pattern is quite stable in both graphs, but the mean NO₂ in the Lyon MA is higher than 40 µg/m³, which is the limit value set by the European community.

Fig. 3 reveals two different relationships between the deprivation index and the NO₂ concentrations. For Lille, the relationship follows a linear trend, the most deprived population living in the most exposed census blocks, whereas in Lyon, the medium class of deprivation is the most exposed. The p-values for an anova-test to compare the mean NO₂ concentrations between the classes of deprivation are significant for both MAs ($p < 0.0001$).

3.2. Spatial analysis

The crude (unadjusted) maps show significant geographic variation based on the global statistics, with $p = 0.005$ for the Lyon and $p < 0.001$ for the Lille MA (Table 2, Figs. 4A and 5A). For Lille, multiple clusters were observed (i.e., bumpier surface) due to a smaller and precise span = 0.30. Two significant urban hotspots are visible that include the two major cities of the metropolitan area, Lille city in the center and Roubaix, a city in the north east of the metropolitan area (Fig. 4A). Unlike, for Lyon, there is a visible significant spatial gradient of risk from northwest to southeast ($p = 0.005$), with a cluster of significant elevated mortality situated in the east of the map. Fig. 5A shows a different pattern of spatial variability as indicated by the large span (span = 0.95). The span denotes the amount of smoothing. The degree of smoothing obtained is a compromise between high sensitivity (high-risk areas correctly identified) and high specificity (non high-risk areas correctly identified). In our model each span was calculated to be optimal by minimizing the AIC in the calculation. In the Appendix, we also note contrasts in some important demographic and socioeconomic characteristics between east and west of the Lyon MA.

After adjustment for NO₂ concentrations alone, Fig. 4B shows little differences compared to the results from the crude analysis in the Lille MA: the same location pattern remains significant ($p = 0.007$), suggesting that exposure to NO₂ does not explain the spatial pattern of infant mortality (Table 2-A). The same holds true in the Lyon MA ($p = 0.001$, Fig. 5B and Table 2-B).

After adjusting for deprivation alone, the results differ depending on the MA. For Lille, Fig. 4C shows no longer any area of statistically significant risk and Table 2-A exhibits a borderline significant p-value ($p = 0.07$). This strongly suggests that socioeconomic status explains a great part of the spatial variability of infant mortality. Populations living in census blocks of higher deprivation have a significant greater risk of infant mortality than populations in the census blocks of the lower deprivation (crude model, $p < 0.001$). Fully adjusting the model with the two variables (SES and NO₂) explained more of the spatial pattern with a larger p-value $p = 0.124$ (Table 2-A) and less pronounced hot and cold spots than in the crude and SES-only adjusted maps.

Adding the interaction term, we found similar p value and maps, suggesting that the interaction term does not explain any of the spatial variation.

For Lyon, after adjusting for deprivation alone Fig. 5C shows that the magnitude of the infant mortality decreases compared to figures derived from the crude analysis (Fig. 5A and Table 2-B); however, there is little change in the geographic pattern, and the global statistic for location remains significant ($p = 0.025$), suggesting that socioeconomic status alone does not entirely explain the spatial variability. The full model adjusting with the two variables (SES and NO₂) explained more of the spatial pattern with a larger but still borderline p-value $p = 0.085$. This p value for Lyon MA means that spatial variation may be due to insufficient adjusting of SES or for other risk factors not taken into account in the model that could explain this remaining heterogeneity in the distribution of infant mortality. When NO₂ is not accounted for (Fig. 5C), compared to the full model (Fig. 5D), infant mortality is actually lower in some geographic areas in the south of the MA. Exposure to air pollution and the socioeconomic status are not correlated. Populations living in census blocks of higher deprivation are not exposed to the higher levels of NO₂ concentration. The most exposed census blocks are in the middle class of deprivation. No interaction is shown between the two factors.

4. Discussion

We used generalized additive models to explore the spatial variation of infant mortality in two major French metropolitan areas. Our results highlight differences in the spatial inequality patterns of infant mortality across the two metropolitan areas.

For the Lille MA, socioeconomic deprivation explained the spatial variations of infant mortality across the different census blocks. More precisely, the two significant clusters of elevated infant mortality detected in the cities of Lille and Roubaix (a city near the Belgium border in the north east of the metropolitan area) are no longer apparent after adjustment for socioeconomic status at the census block level. In the Lyon MA, adjusting for socioeconomic status did not eliminate the clusters of elevated infant mortality. For both MAs, a positive trend of infant mortality from the most deprived to the less deprived census blocks was found. Several studies demonstrated that the socioeconomic status is an important risk factor for a variety of birth outcomes (Guildea et al., 2001; Krieger et al., 2003; Elo et al., 2009; Zeka et al., 2008; Blumenshine et al., 2010; Singh and Kogan, 2007; Luo et al., 2006; Arntzen et al., 2004). During pregnancy, mothers are likely to face multiple stressful life events, including lone-mother, unemployment, and little resources to deal with these conditions (Miranda et al., 2009; O'Neill et al., 2003; Larson, 2007; Nkansah-Amankra et al., 2010a,b; Lu and Chen, 2004). To improve the methodology employed in ecological studies of health inequalities, it would be necessary to carry out studies with individual data in which risk factors related to specific causes of infant mortality are analyzed. These parental factors include poor health status (for example, diabetes, obesity and chronic obstructive lung diseases), toxicants such as nicotine, caffeine, cocaine or alcohol (Patra et al., 2011; O'Leary et al., 2009; Crane et al., 2011)

Table 1

Descriptive statistics of the main confounders, classes of deprivation and classes of NO₂ air pollution for the Lille and Lyon MAs during the period 2002–2009.

		Lille MA (N = 479 census blocks)		Lyon MA (N = 491 census blocks)	
		mean ± SD	[min; max]	mean ± SD	[min; max]
Classes of deprivation (tertiles)	Global	0.02 ± 1.01	[−1.93; 3.10]	−0.02 ± 0.99	[−2.04; 3.06]
	1st	−1.04 ± 0.28	[−1.93; −0.60]	−1.00 ± 0.42	[−2.04; −0.47]
	2nd	−0.12 ± 0.27	[−0.59; 0.38]	−0.13 ± 0.19	[−0.48; 0.20]
	3rd	1.19 ± 0.64	[0.39; 3.10]	1.06 ± 0.74	[0.20; 3.06]
Classes of NO ₂ air pollution (tertiles)	Global	32.06 ± 5.31	[21.79; 58.16]	47.78 ± 5.81	[29.49; 60.10]
	1st	26.53 ± 2.39	[21.79; 30.00]	34.88 ± 2.19	[29.49; 37.97]
	2nd	31.91 ± 1.05	[30.05; 33.82]	41.06 ± 1.76	[38.04; 43.73]
	3rd	37.75 ± 3.83	[33.84; 58.16]	47.78 ± 3.18	[43.73; 60.10]

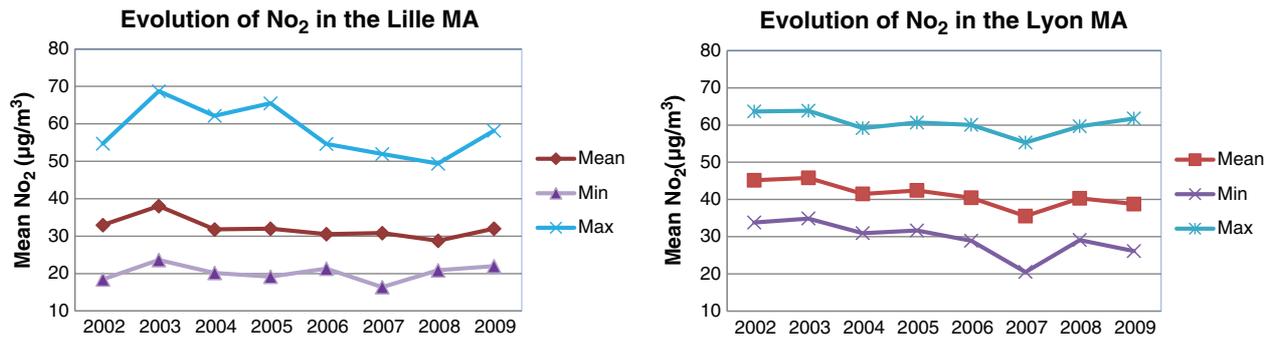


Fig. 2. Yearly averages of the NO_2 ambient air concentrations during the period 2002–2009 for the Lille and Lyon MAs. Fig. 2 shows the temporal trends of the NO_2 concentrations during the period 2002 to 2009. In the Lille MA, the meteorology was penalizing in terms of dispersion during years 2003, 2005, which led to higher annual levels of nitrogen dioxide. The pattern is quite stable in both graphs, but the mean NO_2 in the Lyon MA is higher than $40 \mu\text{g}/\text{m}^3$, which is the limit value set by the European community.

and multiple exposures to pollution (passive smoking, occupational exposure) that could act in addition to or in synergy with access to healthcare. In the conclusion of their extensive analysis of the epidemiologic literature, David and Collins (2007) state that genetic factors fail to explain the strong disparities in birth outcomes according to race in the US, which are better explained by social determinants.

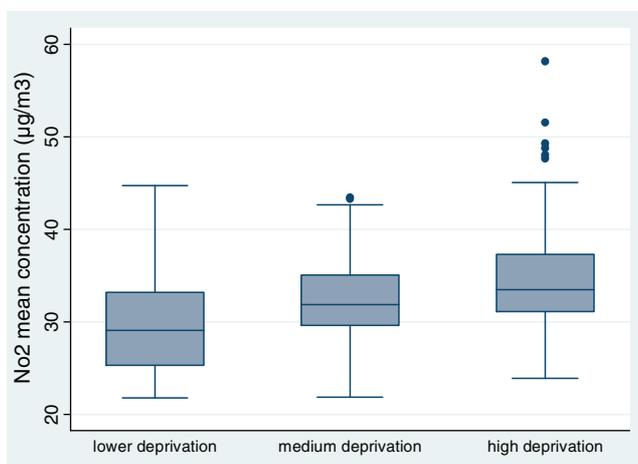
In the two MAs, exposure to ambient air NO_2 did not explain the spatial distribution of infant mortality observed in the crude analysis, though the size of the areas of statistically elevated risk did change. Some authors hypothesize that air pollution contributes to creating or accentuating the socioeconomic inequalities that exist for various illnesses, including cancer, asthma and cardiovascular diseases (Forastiere et al., 2007; Barceló et al., 2009; Jerrett et al., 2005b). It is now highly suspected to contribute to preterm births, intra-uterine growth and perinatal mortality (Bell et al., 2007; Miranda et al., 2009; Slama et al., 2008, 2009; Weck et al., 2008; Woodruff et al., 2008). In the city of Sao Paulo, Brazil, logistic regression revealed a gradient of increasing risk of an early neonatal death with higher exposure to traffic-related air pollution (De Medeiros et al., 2009). In Europe, Scheers et al. (2011) found that risk levels infant mortality increased by 4% for a $10 \mu\text{g}/\text{m}^3$ increase in daily mean PM_{10} among European children under the age of one. In our study setting, NO_2 concentrations stand as a global indicator of air pollution associated with traffic and industrial emissions (Chaix et al., 2006; Vrijheid et al., 2012).

After the final multivariate models were constructed, we tested full models and two-way interactions between pollution and the socioeconomic status to control for confounding. Spatial variation in the risk of

infant mortality no longer exists in Lille after adjusting for deprivation, and adding the NO_2 variable confirms this pattern. In Lyon, adjusting for confounders did change the magnitude of the risk but some spatial variation still exists in the south of the MA, possibly due to other risk factors not taken into account in our analysis. No interaction is shown between the two factors in the Lille or in the Lyon MAs. Some studies suggests that variables such as the socioeconomic status may modify the air pollution–mortality association (Barceló et al., 2009; Jerrett et al., 2004; Martins et al., 2004). Jerrett et al in 2005, in a study of Hamilton (Canada), found that high mortality was associated with exposure to air pollution (especially SO_2) in a citywide model and in intra-urban zones with lower socioeconomic characteristics. In our case, we found no significant results. This finding must be tempered with recognition that the interaction terms themselves tended to be collinear with one or more of the independent variables.

The interpretation of our findings must also consider some weaknesses, notably in exposure assessment. First, the use of dispersion models such as the ADMS-Urban or SIRANE may be limited by the extensive amount of input data that are required. Uncertainty may arise from data sources, estimation methods or measurement tools. However, the models give acceptable results. For the Lille MA, five monitoring sites were available to evaluate the model's results and the correlations varied between 0.63 and 0.74. For the Lyon MA, four monitoring sites were available for this comparison and showed high correlations between the model's predictions and the measured NO_2 values (0.78–0.96). NO_2 concentrations in the Lille MA were modeled only for 2009, and kriging was used for previous years, which yields some uncertainty.

A. Lille MA



B. Lyon MA

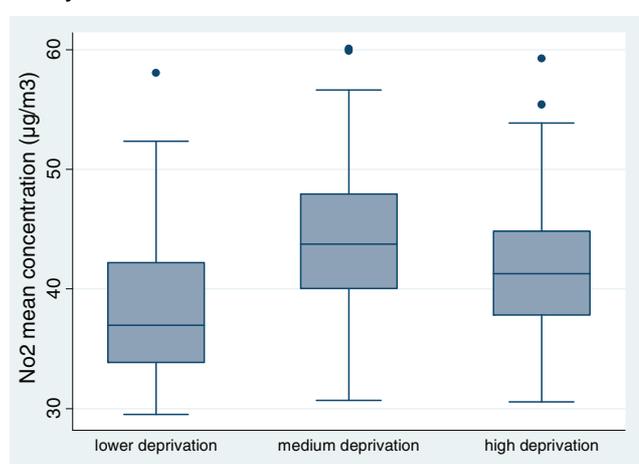


Fig. 3. Comparison of the NO_2 average concentrations according to the classes of deprivation in the Lille and Lyon MAs, period 2002–2009.

Table 2

A and B: Summary of the infant mortality models in the Lille and Lyon MAs. The global p-value denotes whether the smooth term for location is significant in the model.

Lille MA	Span	Deviance p value Global
A.		
Crude model	0.30	<0.001
Adjusted by NO ₂	0.30	0.007
Adjusted by SES	0.95	0.070
Full adjusted	0.95	0.124
Interaction model	0.95	0.125
B.		
Crude model	0.95	0.005
Adjusted by NO ₂	0.95	0.001
Adjusted by SES	0.95	0.025
Full adjusted	0.95	0.085
Interaction model	0.95	0.047

Models such as land-use regression, which are less complex to implement and can also provide reliable estimates of traffic-related air pollution, could have been a relevant alternative (Jerrett et al., 2010).

This work has some notable strengths. Small area analysis allows a deeper understanding of the geographic patterns of health inequalities and is essential for revealing local-level inequalities that are often masked when health estimates are produced at large area scales (cities, counties, and states). Laurent 2007 showed that several studies that had used socioeconomic characteristics measured at an aggregate level (municipality, county or region) did not find the effect of pollution to vary across different areas, whereas many studies that measured socioeconomic conditions at a more disaggregate level (district, neighborhood or census blocks) did reveal joint effects of the two families of factors.

One other strength of our approach is that we could draw in the maps areas of significantly elevated risk (hot spots) by delineating areas that were above the 97.5% confidence interval. This was done by permutation tests, which are flexible hypothesis tests not making prior assumptions on parameters and outcome distribution. Simulation studies showed that the permutation tests behave considerably better than the corresponding classical tests if measured by the critical values attained. We modeled health risks by GAMs. Differently from classical cluster detection methods, GAMs make it possible to include a non-parametric term to account for spatial variation in the health risk, as well as to adjust for potential confounders, and consider multiple

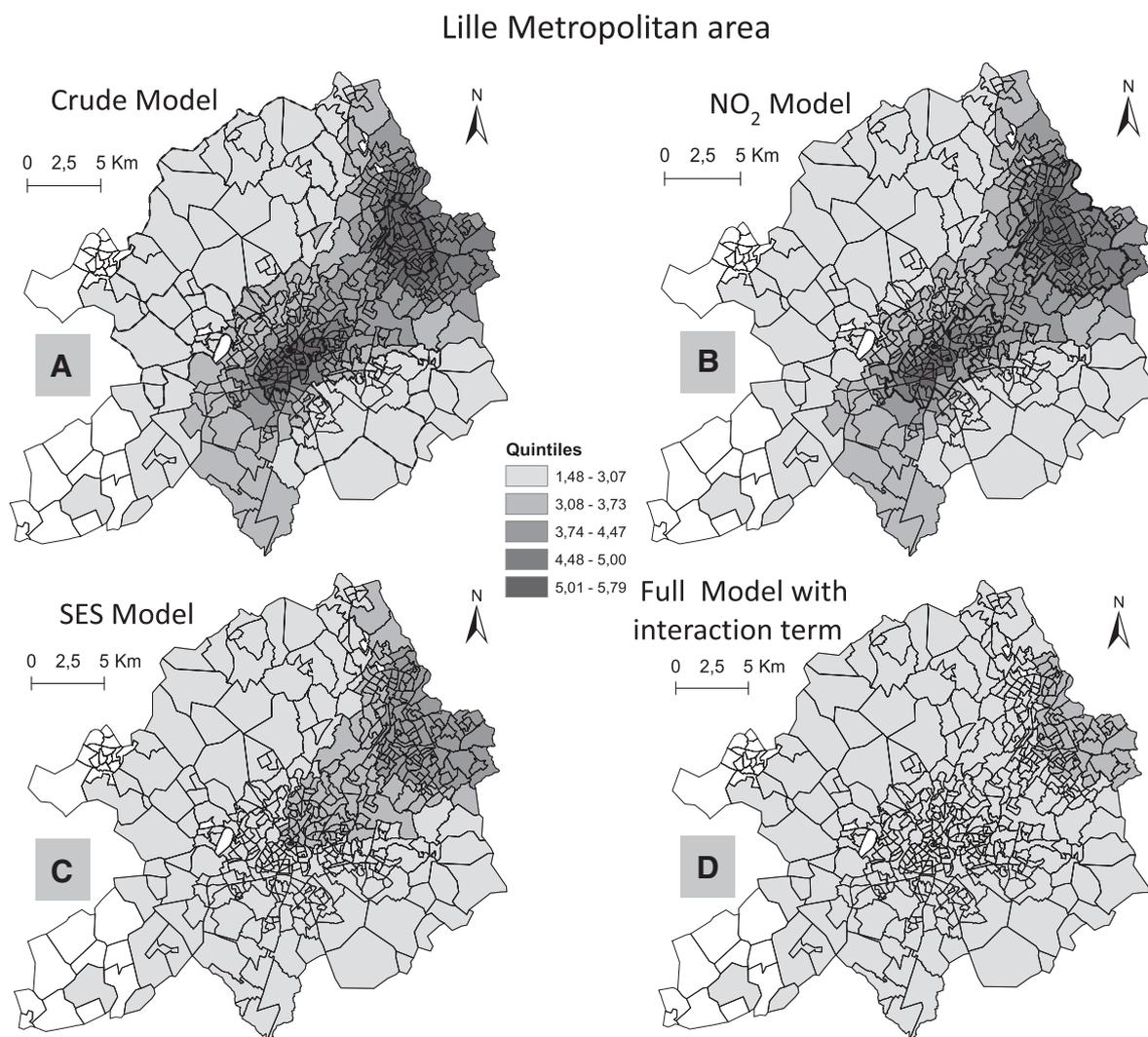


Fig. 4. Prevalence of infant mortality estimated by the GAMs in the Lille Metropolitan Area for crude model (A), according to (B) NO₂ air pollution exposure, (C) deprivation index and (D) with the interaction term. Light grey to dark grey shading indicates lower to higher prevalence. Solid lines identify areas with significantly increased rates (hotspots) and dashed lines identify areas with significantly decreased rates (coldspots).

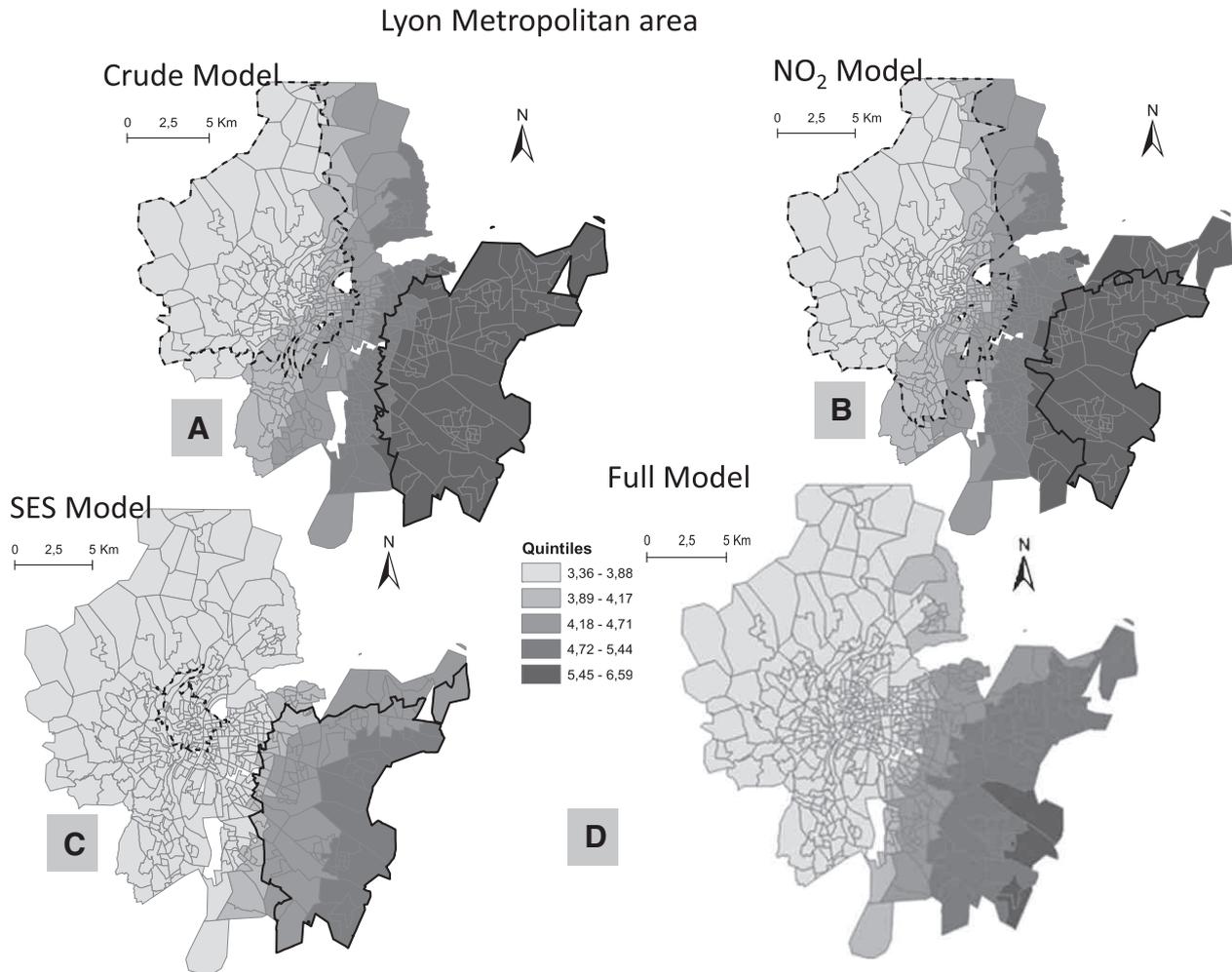


Fig. 5. (A) Prevalence of infant mortality estimated by the GAMs in the Lyon Metropolitan Area for crude model (A), according to (B) NO₂ air pollution exposure, (C) deprivation index and (D) full model. Light grey to dark grey shading indicates lower to higher prevalence. Solid lines identify areas with significantly increased rates (hotspots) and dashed lines identify areas with significantly decreased rates (coldspots).

tests: a global one and a test to determine areas of significantly elevated risks. Compared to other famous cluster methods, the GAM method isn't limited by clusters of predefined shape (i.e., circle or ellipse).

Finally, the procedure to model air pollution concentrations at surface areas of the size of census blocks provides unbiased estimates of exposure to ambient air pollution. Jerrett et al. in 2010 demonstrated the effectiveness and reliability of this type of model for assessing air quality in health effects assessment research. Some studies on reproductive effects of air pollution used surrogate air pollution measures (e.g. average pollutant concentrations at fixed ambient monitoring stations, distance to monitoring sites, vehicular traffic emissions, proximity to highway, and distance to main roads) to estimate maternal but failure to consider spatial variations can lead to exposure misclassification and subsequent bias (Jerrett et al., 2005a).

5. Conclusion

The use of GIS and spatial analysis techniques has been shown to be useful tools to inform public policy and determine areas that warrant specific intervention. In the present setting, these techniques identified clusters of elevated infant mortality in relation with the socioeconomic status, and marginally to air pollution. Moreover, this paper illustrates an approach to take into account the combined effect, and possibly, the

interaction between socioeconomic characteristics and environment exposures and to identify areas which cumulate conditions.

Conflict of interest statement

The authors declare that they have no competing interests.

Acknowledgements

The authors thank all scientific researchers of the Equit'Area project, the personnel of the local councils of the Lille and Lyon Metropolitan Area and Grégoire Rey of the CeperiDC institute for their participation. The Equit'Area project is supported by the French National Research Agency (ANR, contract-2010-PRSP-002-01), the EHESP School of Public Health, the French General Directorate of Health (DGS), the research call launched by the French Institute of Public Health Research (IReSP) in 2010 and the Nord-Pas de Calais region. Cindy Padilla is a recipient of a doctoral grant from the Environment and Energy Management Agency (ADEME). Verónica Vieira was supported by a grant from the National Institute of Environmental Health Sciences (NIEHS; 2P42 ES007381-16A1), National Institutes of Health (NIH). Its contents are solely the responsibility of the authors and do not necessarily represent the views of NIEHS, NIH.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.scitotenv.2013.03.027>.

References

- Arntzen A, Samuelsen SO, Bakketeig LS, Stoltenberg C. Socioeconomic status and risk of infant death. A population-based study of trends in Norway, 1967–1998. *Int J Epidemiol* 2004;33:279–88.
- Barceló MA, Saez M, Saurina C. Spatial variability in mortality inequalities, socioeconomic deprivation, and air pollution in small areas of the Barcelona Metropolitan Region, Spain. *Sci Total Environ* 2009;407:5501–23.
- Bell ML, Ebisu K, Belanger K. Ambient air pollution and low birth weight in Connecticut and Massachusetts. *Environ Health Perspect* 2007;115:1118–24.
- Blumenshine P, Egarter S, Barclay CJ, Cubbin C, Braveman PA. Socioeconomic disparities in adverse birth outcomes: a systematic review. *Am J Prev Med* 2010;39:263–72.
- Carbajal-Arroyo L, Miranda-Soberanis V, Medina-Ramón M, Rojas-Bracho L, Tzintzun G, Solís-Gutiérrez P, et al. Effect of PM10 and O3 on infant mortality among residents in the Mexico City Metropolitan Area: a case-crossover analysis, 1997–2005. *J Epidemiol Community Health* 2011;65:715–21.
- Chaix B, Gustafsson S, Jerrett M, Kristersson H, Lithman T, Boalt A, et al. Children's exposure to nitrogen dioxide in Sweden: investigating environmental injustice in an egalitarian country. *J Epidemiol Community Health* 2006;60:234–41.
- Carruthers DJ, Edmunds HA, Lester AE, McHugh CA, Singles RJ. Use and validation of ADMS-Urban in contrasting urban and industrial locations. *International Journal of Environment and Pollution* 2000;14:364–74.
- Crane JMG, Keough M, Murphy P, Burrage L, Hutchens D. Effects of environmental tobacco smoke on perinatal outcomes: a retrospective cohort study. *BJOG* 2011;118:865–71.
- David R, Collins Jr J. Disparities in infant mortality: what's genetics got to do with it? *Am J Public Health* 2007;97:1191–7.
- De Medeiros APP, Nelson Gouveia, Machado RPP, De Souza MR, Alencar GP, Novaes HMD, et al. Traffic-related air pollution and perinatal mortality: a case-control study. *Environ Health Perspect* 2009;117:127–32.
- Elo I, Culhane J, Kohler I, O'Campo P, Burke J, Messer L, et al. Neighbourhood deprivation and small-for-gestational-age term births in the United States. *Paediatr Perinat Epidemiol* 2009;23:87–96.
- Forastiere F, Stafoggia M, Tasco C, Picciotto S, Agabiti N, Cesaroni G, et al. Socioeconomic status, particulate air pollution, and daily mortality: differential exposure or differential susceptibility. *Am J Ind Med* 2007;50:208–16.
- Gatrell A, Bailey T. Interactive spatial data analysis in medical geography. *Soc Sci Med* 1996;42:843–55.
- Généreux M, Auger N, Goneau M, Daniel M. Neighbourhood socioeconomic status, maternal education and adverse birth outcomes among mothers living near highways. *J Epidemiol Community Health* 2008;62:695–700.
- Guilidea ZE, Fone DL, Dunstan FD, Sibert JR, Cartledge PH. Social deprivation and the causes of stillbirth and infant mortality. *Arch Dis Child* 2001;84:307–10.
- Hajat S, Armstrong B, Wilkinson P, Busby A, Dolk H. Outdoor air pollution and infant mortality: analysis of daily time-series data in 10 English cities. *J Epidemiol Community Health* 2007;61:719–22.
- Havard S, Deguen S, Bodin J, Louis K, Laurent O, Bard D. A small-area index of socioeconomic deprivation to capture health inequalities in France. *Soc Sci Med* 2008;67:2007–16.
- Jerrett M, Burnett R, Brook J, Kanaroglou P, Giovis C, Finkelstein N, et al. Do socioeconomic characteristics modify the short term association between air pollution and mortality? Evidence from a zonal time series in Hamilton, Canada. *J Epidemiol Community Health* 2004;58:31–40.
- Jerrett M, Burnett R, Ma R, Pope III CA, Krewski D, Newbold K, et al. Spatial analysis of air pollution and mortality in Los Angeles. *Epidemiology* 2005a;16:727–36.
- Jerrett M, Buzzelli M, Burnett R, DeLuca P. Particulate air pollution, social confounders, and mortality in small areas of an industrial city. *Soc Sci Med* 2005b;60:2845–63.
- Jerrett M, Gale S, Kontgis C. Spatial modeling in environmental and public health research. *Int J Environ Res Public Health* 2010;7:1302–29.
- Kaiser R, Romieu I, Medina S, Schwartz J, Krzyzanowski M, Künzli N. Air pollution attributable postneonatal infant mortality in U.S. metropolitan areas: a risk assessment study. *Environ Health* 2004;3:4.
- Kelsall J, Wakefield J. Modeling spatial variation in risk: a geostatistical approach. *American Statistical Association*; 2002 [sec. 97].
- Krieger N, Chen J, Waterman P, Soobader M, Subramanian S, Carson R. Choosing area based socioeconomic measures to monitor social inequalities in low birth weight and childhood lead poisoning: the Public Health Disparities Geocoding Project (US). *J Epidemiol Community Health* 2003;57:186–99.
- Larson CP. Poverty during pregnancy: its effects on child health outcomes. *Paediatr Child Health* 2007;12:673–7.
- Lin CA, Pereira LAA, Nishioka DC, Conceição GMS, Braga ALF, Saldiva PHN. Air pollution and neonatal deaths in São Paulo, Brazil. *Braz J Med Biol Res* 2004;37:765–70.
- Lu MC, Chen B. Racial and ethnic disparities in preterm birth: the role of stressful life events. *Am J Obstet Gynecol* 2004;191:691–9.
- Luo Z-C, Wilkins R, Kramer MS. Effect of neighbourhood income and maternal education on birth outcomes: a population-based study. *CMAJ* 2006;174:1415–20.
- Martins MCH, Fatigati F, Véspoli T, Martins L, Pereira L, Martins MA, et al. Influence of socioeconomic conditions on air pollution adverse health effects in elderly people: an analysis of six regions in São Paulo, Brazil. *J Epidemiol Community Health* 2004;58:41–6.
- McHugh C, Carruthers D, Edmunds H. ADMS-Urban: an air quality management system for traffic, domestic and industrial pollution. *Int J Environ Pollution* 1997;8:437–40.
- Miranda ML, Maxson P, Edwards S. Environmental contributions to disparities in pregnancy outcomes. *Epidemiol Rev* 2009;31:67–83.
- Nkansah-Amankra S, Dhawain A, Hussey JR, Luchok KJ. Maternal social support and neighborhood income inequality as predictors of low birth weight and preterm birth outcome disparities: analysis of South Carolina Pregnancy Risk Assessment and Monitoring System survey, 2000–2003. *Matern Child Health J* 2010a;14:774–85.
- Nkansah-Amankra S, Luchok KJ, Hussey JR, Watkins K, Liu X. Effects of maternal stress on low birth weight and preterm birth outcomes across neighborhoods of South Carolina, 2000–2003. *Matern Child Health J* 2010b;14:215–26.
- O'Campo P, Xue X, Wang MC, Caughy M. Neighborhood risk factors for low birthweight in Baltimore: a multilevel analysis. *Am J Public Health* 1997;87:1113–8.
- O'Leary CM, Nassar N, Kurinczuk JJ, Bower C. The effect of maternal alcohol consumption on fetal growth and preterm birth. *BJOG* 2009;116:390–400.
- O'Neill M, Jerrett M, Kawachi I, Levy J, Cohen A, Gouveia N, et al. Health, wealth, and air pollution: advancing theory and methods. *Environ Health Perspect* 2003;111:1861–70.
- OECD—Organization for Economic Co-operation and Development. Infant mortality rates in OCDE countries 2008. Available at <http://www.oecd.org/dataoecd/30/54/43136879.xls> 2010. [June 20].
- Patra J, Bakker R, Irving H, Jaddoe VVW, Malini S, Rehm J. Dose-response relationship between alcohol consumption before and during pregnancy and the risks of low birthweight, preterm birth and small for gestational age (SGA)—a systematic review and meta-analyses. *BJOG* 2011;118:1411–21.
- Ponce NA, Hoggatt KJ, Wilhelm M, Ritz B. Preterm birth: the interaction of traffic-related air pollution with economic hardship in Los Angeles neighborhoods. *Am J Epidemiol* 2005;162:140–8.
- Ritz B, Wilhelm M, Zhao Y. Air pollution and infant death in southern California, 1989–2000. *Pediatrics* 2006;118:493–502.
- Romieu I, Ramírez-Aguilar M, Moreno-Macias H, Barraza-Villarreal A, Miller P, Hernández-Cadena L, et al. Infant mortality and air pollution: modifying effect by social class. *J Occup Environ Med* 2004;46:1210–6.
- Salihi HM, Ghaji N, Mbah AK, Alio AP, August EM, Boubakari I. Particulate pollutants and racial/ethnic disparity in fetoneonatal morbidity outcomes. *Matern Child Health J* 2012;16:1679–87.
- Scheers H, Mwalili SM, Faes C, Fierens F, Nemery B, Nawrot TS. Does air pollution trigger infant mortality in Western Europe? A case-crossover study. *Environ Health Perspect* 2011;119:1017–22.
- Singh GK, Kogan MD. Persistent socioeconomic disparities in infant, neonatal, and postneonatal mortality rates in the United States, 1969–2001. *Pediatrics* 2007;119:e928–39.
- Slama R, Darrow L, Parker J, Woodruff Tracey J, Strickland M, Nieuwenhuijsen M, et al. Meeting report: atmospheric pollution and human reproduction. *Environ Health Perspect* 2008;116:791–8.
- Slama R, Thiebaugeorges O, Goua V, Aussel L, Sacco P, Bohet A, et al. Maternal personal exposure to airborne benzene and intrauterine growth. *Environ Health Perspect* 2009;117:1313–21.
- Soulhac Lionel, Pietro Salizzoni, Cierco F-X, Perkins R. The model SIRANE for atmospheric urban pollutant dispersion; part I, presentation of the model. *Atmos Environ* 2011;45:7379–95.
- Soulhac L, Salizzoni P, Mejean P, Didier D, Rios I. The model SIRANE for atmospheric urban pollutant dispersion; part II, validation of the model on a real case study. *Atmos Environ* 2012;49:320–37.
- Tsai S-S, Chen C-C, Hsieh H-J, Chang C-C, Yang C-Y. Air pollution and postneonatal mortality in a tropical city: Kaohsiung, Taiwan. *Inhal Toxicol* 2006;18:185–9.
- Vieira V, Webster T, Aschengrau A, Ozonoff D. A method for spatial analysis of risk in a population-based case-control study. *Int J Hyg Environ Health* 2002;205:115–20.
- Vieira V, Webster T, Weinberg J, Aschengrau A, Ozonoff D. Spatial analysis of lung, colorectal, and breast cancer on Cape Cod: an application of generalized additive models to case-control data. *Environ Health* 2005;4:11.
- Vieira VM, Webster T, Weinberg J, Aschengrau A. Spatial-temporal analysis of breast cancer in upper Cape Cod, Massachusetts. *Int J Health Geogr* 2008;7:46.
- Vrijheid M, Martinez D, Aguilera I, Ballester F, Basterrechea M, Esplugues A, et al. Socioeconomic status and exposure to multiple environmental pollutants during pregnancy: evidence for environmental inequity? *J Epidemiol Community Health* 2012;66:106–13.
- Webster T, Vieira V, Weinberg J, Aschengrau A. Method for mapping population-based case-control studies: an application using generalized additive models. *Int J Health Geogr* 2006;5:26.
- Weck RL, Paulose T, Flaws JA. Impact of environmental factors and poverty on pregnancy outcomes. *Clin Obstet Gynecol* 2008;51:349–59.
- Woodruff TJ, Darrow LA, Parker JD. Air pollution and postneonatal infant mortality in the United States, 1999–2002. *Environ Health Perspect* 2008;116:110–5.
- Yi O, Kim H, Ha E. Does area level socioeconomic status modify the effects of PM(10) on preterm delivery? *Environ Res* 2010;110:55–61.
- Young RL, Weinberg Janice, Vieira Verónica, Ozonoff A, Webster TF. A power comparison of generalized additive models and the spatial scan statistic in a case-control setting. *International Journal of Health Geographics* 2010;9:37.
- Zeitlin J, Combier E, Levaillant M, Lasbuer L, Pilkington H, Charreire H, et al. Neighbourhood socio-economic characteristics and the risk of preterm birth for migrant and non-migrant women: a study in a French district. *Paediatr Perinat Epidemiol* 2011;25:347–56.
- Zeka A, Melly S, Schwartz J. The effects of socioeconomic status and indices of physical environment on reduced birth weight and preterm births in Eastern Massachusetts. *Environ Health* 2008;7:60.

2. Compléments

Méthode

L'article publié ne concernant que la mortalité infantile sur les agglomérations de Lille et Lyon, cette partie apportera les résultats complémentaires du rôle du dioxyde d'azote sur les inégalités socio-spatiales de mortalité infantile sur la ville de Paris et sur la mortalité néonatale dans ces mêmes zones urbains. Les résultats sont issus des mêmes méthodes statistiques que l'article publié, c'est-à-dire des modèles statistiques additifs généralisés ont permis de prendre en compte l'autocorrélation spatiale et de générer des cartes à l'aide de lissage sur la longitude et la latitude tout en ajustant sur les variables d'intérêt.

Identification des zones à risque de mortalité

Les cartes brutes (c'est-à-dire sans ajustement) mettent en évidence des disparités géographiques significatives pour la mortalité infantile à Paris ($p < 0,001$) et pour la mortalité néonatale à Lille ($p = 0,024$) et Lyon ($p = 0,024$) (Tableaux 7 AB). Ces résultats se basent sur le test global d'homogénéité des risques entre les unités géographiques, ici les IRIS. A Paris, on constate un gradient du risque de mortalité infantile visible du Sud-Ouest au Nord-Est avec un cluster de risque significativement plus élevé au Nord-Est comprenant les 18 et 19^{èmes} arrondissements (Figure 10). A Lille, plusieurs clusters de risque plus élevé de mortalité néonatale sont observés : ils incluent les villes de Lille pour le premier cluster situé au centre de l'agglomération et les villes de Roubaix, Croix, Wasquehal pour le deuxième cluster situé au Nord-Est (Figure 11). A Lyon, un gradient du risque de mortalité néonatale est également observé, allant de l'Ouest vers l'Est avec un cluster de risque plus élevé à l'Est comprenant les villes de Vaux en Velin, Décines Charpieu, Jonage, Meyzieu et Chassieu (Figure 12).

Ajustement sur le NO₂

Après ajustement sur l'exposition au dioxyde d'azote, les clusters de risque de mortalité néonatale de Lyon et de mortalité infantile de Paris ne disparaissent pas et restent significatifs ($p = 0,03$) : l'exposition à la pollution n'explique pas les disparités spatiales de mortalité. A l'inverse à Lille, le cluster de mortalité néonatale disparaît : l'exposition à la pollution explique la variabilité géographique du risque ($p = 0,27$).

Ajustement sur la défaveur socio-économique

Après ajustement sur la défaveur socio-économique, les clusters de risque de chaque agglomération ont disparu. La défaveur socio-économique explique le sur-risque de mortalité néonatale à Lille et Lyon et infantile à Paris.

Combinaison de l'exposition au NO₂ et de la défaveur socio-économique

Après ajustement sur la défaveur socio-économique ainsi que l'exposition au dioxyde d'azote, les modèles expliquent mieux les disparités géographiques du risque de mortalité néonatale sur Lille, Lyon et les disparités géographiques du risque de mortalité infantile sur Paris. Les modèles avec interaction n'apportent pas plus d'information que les modèles complets.

Tableaux 7 AB: Caractéristiques de chaque modèle spatial d'estimation du risque, A) de mortalité infantile pour la ville de Paris, B) de mortalité néonatale de l'agglomération de Lille, Lyon et Paris-ville. *P value = valeur de la déviance du test global dont hypothèse Ho est l'homogénéité spatiale du risque de mortalité infantile ou néonatale.*

A)	Span	Deviance p value Global
Modèle brut	0,90	<0,001
Ajustement sur le SES	0,90	0,074
Ajustement sur le NO2	0,90	0,001
Ajustement complet	0,95	0,132
Modèle avec interaction	0,95	0,110

B)	Lille MA		Lyon MA		Paris MA	
	Span	Deviance p value Global	Span	Deviance p value Global	Span	Deviance p value Global
Modèle brut	0,50	0,024	0,95	0,024	0,95	0,088
Ajustement sur le SES	0,95	0,531	0,95	0,143	0,95	0,674
Ajustement sur le NO2	0,50	0,266	0,95	0,033	0,95	0,129
Ajustement complet	0,95	0,774	0,95	0,525	0,95	0,781
Modèle avec interaction	0,95	0,710	0,95	0,520	0,95	0,749

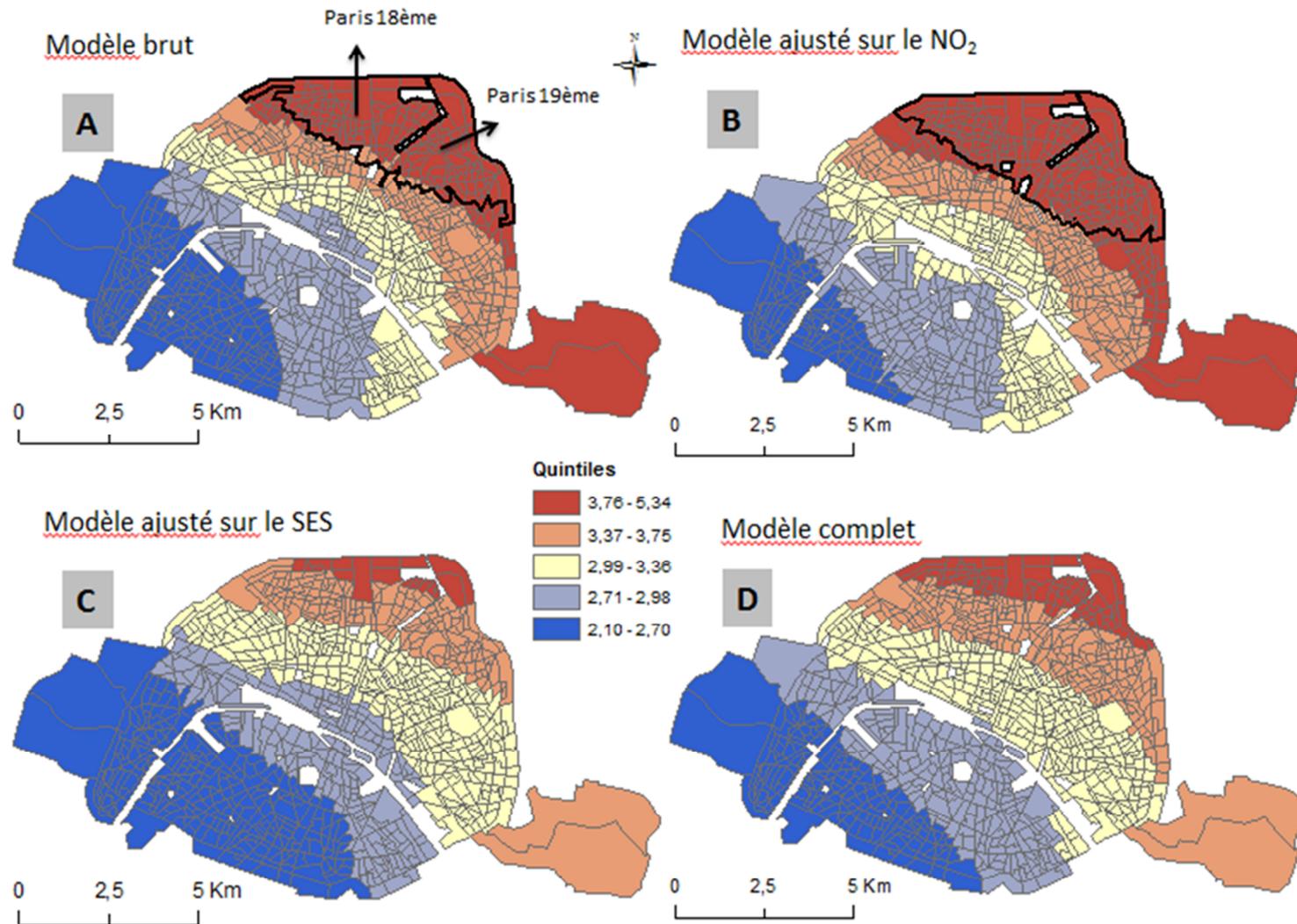


Figure 10: Prévalence de la mortalité infantile estimée par le modèle GAM pour l'agglomération de Paris, issu du modèle Brut A), ajusté sur l'exposition au NO₂ B), ajusté sur la défaveur sociale C) et sur le modèle complet D). La ligne noire et solide identifie les IRIS à risque significativement plus élevés de mortalité infantile.

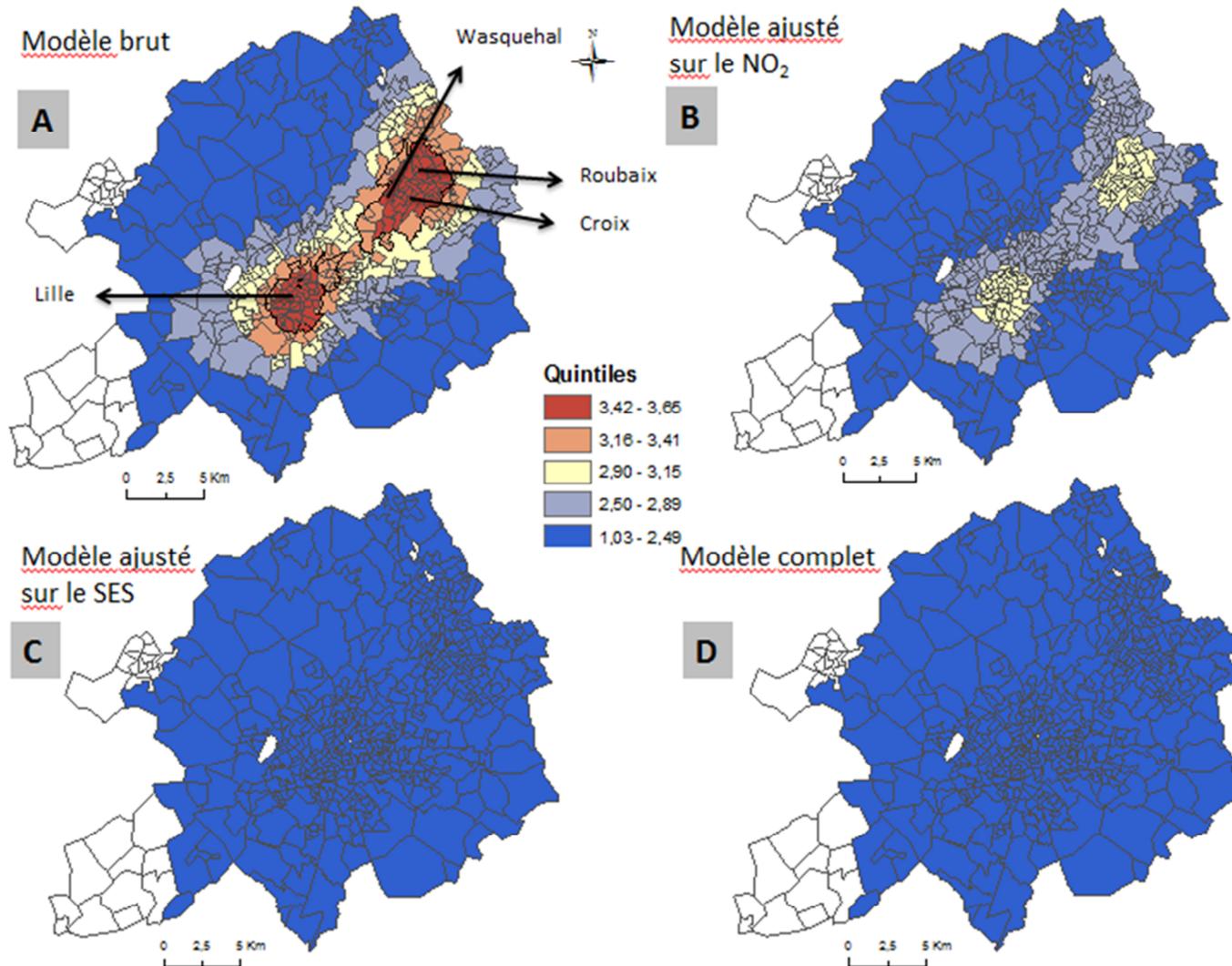


Figure 11 : Prévalence de la mortalité néonatale estimée par le modèle GAM pour l'agglomération de Lille, issu du modèle Brut A), ajusté sur l'exposition au NO₂ B), ajusté sur la défaveur sociale C) modèle complet D). La ligne noire et solide identifie les IRIS à risque significativement plus élevés de mortalité néonatale.

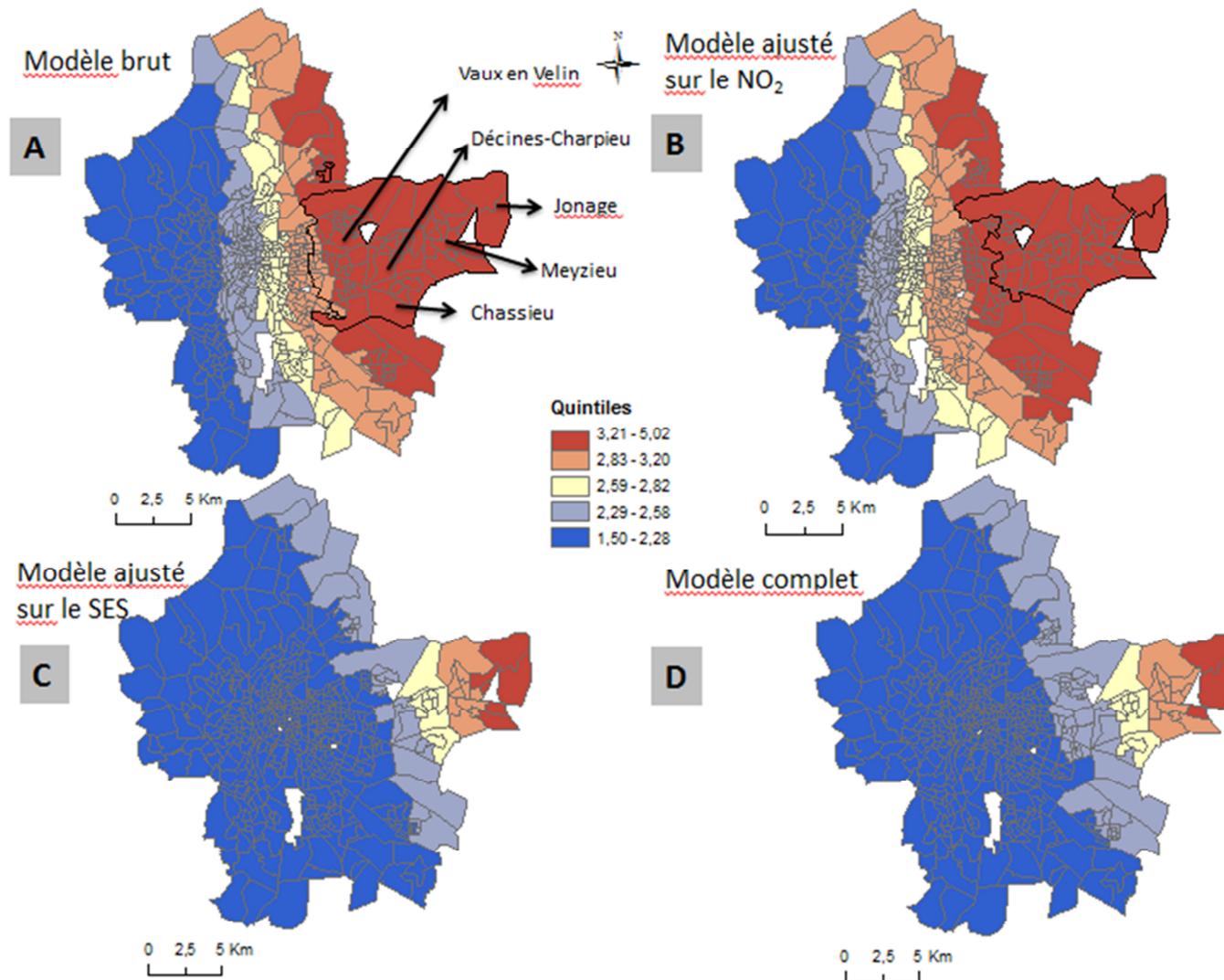


Figure 12 : Prévalence de la mortalité néonatale estimée par le modèle GAM pour l'agglomération de Lyon, issu du modèle Brut A), ajusté sur l'exposition au NO₂ B), ajusté sur la défaveur sociale C) modèle complet D). La ligne noire et solide identifie les IRIS à risque significativement plus élevés de mortalité néonatale.

C. Discussion

Notre étude a mis en évidence que la force ainsi que le sens des associations varient entre les agglomérations et la période de survenue du décès (néonatal ou infantile). Nous avons démontré des zones à risque de mortalité infantile significativement plus élevée dans les 3 agglomérations et des zones à risque significativement plus élevées de mortalité néonatale seulement à Lille et Lyon. Les clusters de sur-risque ne sont pas superposables géographiquement. A Lille, le cluster de mortalité néonatale inclut plus de villes que celui identifié pour la mortalité infantile. A Lyon, au contraire, le cluster de mortalité néonatale est plus petit que celui observé pour la mortalité infantile. Quant à Paris, aucun cluster de mortalité néonatale n'apparaît alors qu'un cluster de mortalité infantile est détecté. La contribution aux disparités spatiales de mortalité de la défaveur socio-économique et/ou de l'exposition au NO₂ diffère d'une agglomération à l'autre. Alors que la défaveur socio-économique semble être le facteur déterminant pour expliquer la variabilité spatiale de la mortalité infantile à Paris, c'est la combinaison de la défaveur socio-économique et de l'exposition au dioxyde d'azote (dans une moindre mesure) qui contribue conjointement à expliquer le sur-risque de mortalité néonatale observé à Lille. A Lyon, en revanche, alors que la combinaison des deux facteurs explique une partie de la variabilité spatiale de la mortalité infantile, la défaveur socio-économique apparaît être le déterminant majeur de la mortalité néonatale.

Mécanismes explicatifs des risques de mortalité : des hypothèses

A l'issue de nos précédents travaux, les résultats de cette étude semblent démontrer que les mécanismes par lesquelles les expositions environnementales contribuent à générer des inégalités socio-spatiales de santé sont différents selon les zones d'étude. A Lille, c'est la combinaison du différentiel d'exposition, et de la vulnérabilité des personnes qui pourrait expliquer nos résultats ; en effet, les populations des quartiers défavorisés apparaissent plus exposées au dioxyde d'azote (Figure 13) et plus à risque de mortalité que les populations des quartiers favorisés (à niveau de pollution égal). L'absence d'interaction significative indique que ces 2 facteurs agissent indépendamment l'un de l'autre, chacun contribuant, par des mécanismes propres décrits dans la première partie de cette thèse, à affecter l'issue de grossesse (mortalité néonatale) ou le très jeune âge (mortalité infantile). Néanmoins, A

l'inverse, la ville de Paris ne présente pas de différentiel sur l'ensemble de la période. A Lyon, on constate un différentiel de vulnérabilité mais pas d'exposition (même si nous ne pouvons pas exclure totalement la possibilité d'un différentiel d'exposition puisque la mesure du NO₂ n'est faite qu'au lieu de résidence). Les quartiers défavorisés sont moins affectés par la pollution (niveaux de NO₂ moins élevés) et pourtant les conséquences sanitaires de cette exposition sont plus marquées que dans les milieux plus aisés. Les explications résultent de divers processus qui conjuguent leur effet.

Tout d'abord, les mères défavorisées peuvent être plus vulnérables aux effets sanitaires de par leur état de santé précaire, leurs comportements face à la santé ainsi que le moindre recours aux soins. En effet, les populations résidant dans les quartiers défavorisés cumulent les facteurs socio-économiques déterminants de l'état de santé, comme un niveau d'éducation et un niveau de ressource économique et matériel faible (36). De plus, ces facteurs de risques se traduisent par des comportements dangereux face à la santé, comme les comportements non hygiéniques pendant la grossesse (les habitudes tabagiques (37), alcool (38), la drogue (39), ne pas manger équilibrer (40) et la sédentarité) et les comportements liés à l'accessibilité économique et physique aux soins (l'assiduité à se faire soigner, l'adhésion aux mesures préventives et le suivi rigoureux pendant la grossesse (5,41,42)). Ensuite, les mères défavorisées peuvent être exposées à d'autres expositions environnementales ; au domicile (un habitat dégradé avec des moisissures, faire le ménage avec des produits d'entretien nocifs (43), moins bien isolé des polluants extérieurs (18), au travail (plus de stress, contact avec des produits nocifs (43), dans les transports.

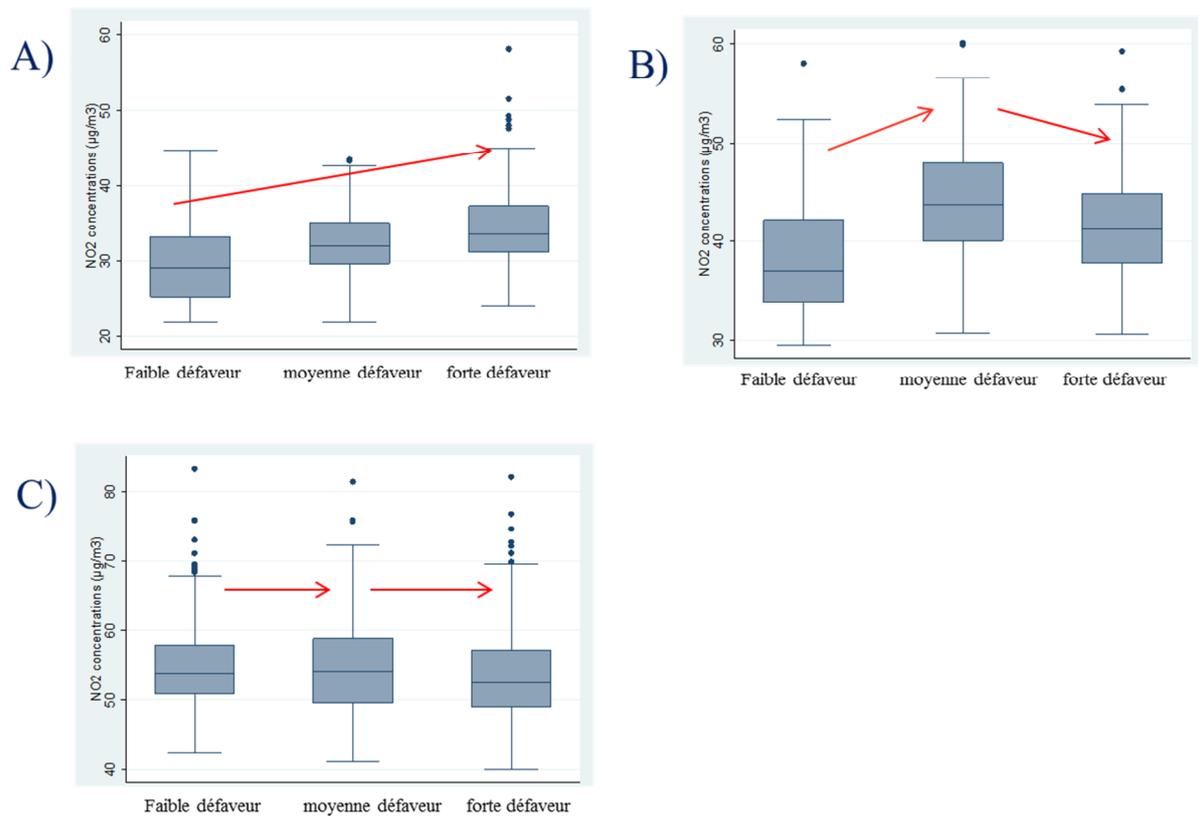


Figure 13: Relation entre la défaveur socio-économique et la concentration en dioxyde d'azote pour les agglomérations de A) Lille, B) Lyon et C) la ville de Paris sur la période 2002-2009.

III. ETUDE SPATIO-TEMPORELLE DES INÉGALITÉS ENVIRONNEMENTALES

Les résultats décrits dans les parties I et II mettent en évidence l'existence manifeste d'inégalités socio-spatiales de santé. Leur réduction constitue une des priorités de la politique de santé publique conduite en France. Le présent projet vise à améliorer la compréhension des inégalités environnementales comme possible contribution à cette situation.

A. Contexte

En France, le premier plan national santé environnement (PNSE) établi sur la période 2004-2008 a permis l'émergence politique de la thématique santé environnement et sa prise en compte par l'ensemble de la société. La prise de conscience de ces inégalités a amené le PNSE2 (44) en 2009 à considérer les inégalités environnementales comme un axe structurant du Plan. Dans ce second plan, les questions d'inégalités liées au contexte socio-économique, et d'inégalités géographiques liées à l'inhomogénéité des altérations de l'environnement sur le territoire sont abordées. Par ailleurs, la Charte de l'environnement promulguée en mars 2005 donne une place de premier ordre à l'égalité environnementale, son article 1er déclarant que « chacun a le droit de vivre dans un environnement équilibré et respectueux de la santé » (45).

En 2004, le rapport d'activités de l'observatoire national des zones urbaines sensibles (ZUS) démontre que parmi les 500 quartiers prioritaires de la politique de la ville (environ 3 millions de personnes concernées), «13% sont traversés et 32% longés par une autoroute, 56% traversés et 83% longés par des voies rapides ou des grandes nationales (15) » c'est-à-dire exposés à la pollution atmosphérique et plus particulièrement au dioxyde d'azote (46). De plus, près de 42% des communes ayant une ZUS sont exposées au risque industriel contre 21% pour les autres (46). En outre, c'est au sein des grandes agglomérations de province et de l'agglomération parisienne que le clivage entre les quartiers aisés et les quartiers modestes est le plus net. La taille de l'agglomération joue sur l'importance des disparités sociales et environnementales constatées : «dans l'agglomération parisienne, un tiers des ménages des quartiers modestes subissent des nuisances fréquentes (47).

En France, plusieurs études ont donc suggéré que les populations les plus défavorisées ont tendance à vivre dans des quartiers caractérisés par un niveau plus élevé de pollution environnementale (48,49). Néanmoins, la conjonction de l'évolution temporelle et de la répartition géographique de ces inégalités environnementales est peu analysée. C'est pourquoi nous avons étudié la relation spatiale et temporelle entre les niveaux de dioxyde d'azote ambiants et les caractéristiques socio-économiques des quartiers dans les agglomérations de Lille, Lyon, Paris & sa petite couronne et Marseille.

B. Publication

Article soumis

Article joint: "Environmental inequalities in France - A spatio-temporal analysis conducted at a small geographical level in four French metropolitan areas." Article soumis au journal Environmental Research.

1 Original Article

2 Air quality and social deprivation in four French metropolitan areas– A spatio-temporal
3 environmental inequality analysis conducted at a small geographical level.

4
5
6
7
8
9
10
11
12
13 **Cindy M PADILLA^{1,2,3}, Wahida KIHAL¹, Verónica M VIEIRA⁴, Philippe ROSSELO⁵, Geraldine LeNIR⁵, Denis**
14 **ZMIROU-NAVIER^{1,2,6}, Severine DEGUEN^{1,2}**

15
16
17
18
19 ¹ EHESP School of Public Health –Sorbonne Paris Cité – Rennes-France

20 ² INSERM U1085-IRSET – Research institute of environmental and occupational health. Rennes, France

21 ³ French Environment and Energy Management Agency - Angers – France

22 ⁴ Program in Public Health, Chao Family Cancer Center, University of Irvine, CA 92697, USA

23 ⁵ Official Air Quality Monitoring Associations (AASQA), Atmo Nord Pas-de-Calais (Lille), Air Rhône-Alpes
24 (Lyon)- France

25 ⁶ Lorraine University medical school, Vandoeuvre-les-Nancy-France

26
27
28
29
30
31
32
33
34 Authors

35
36 W.Kihal wahida.kihal@ehesp.fr

37
38 V.M. Vieira vvieira@uci.edu

39
40 G Le Nir Geraldine.LeNir@air-parif.asso.fr

41
42 P.Rosselo philippe.rosselo@atmo-paca.fr

43
44 D Zmirou-Navier denis.zmirou@inserm.fr

45
46 S Deguen severine.deguen@ehesp.fr

47
48
49
50
51 Corresponding author

52
53 C.M PADILLA. EHESP- Avenue du Professeur Leon Bernard CS 74312, 35043 Rennes, + 33 2 99 02
54 26 00. cindy.padilla@ehesp.fr

30 ABSTRACT

31 Several studies documented that more deprived populations tend to live in areas characterized by
32 higher levels of environmental pollution. Yet, time trends and geographic patterns of this
33 disproportionate distribution of environmental burden remain poorly assessed, especially in Europe.
34 We investigated the spatial and temporal relationship between ambient air NO₂ concentrations and
35 socioeconomic and demographic data in four French Metropolitan Areas (Lille in north of France,
36 Lyon in the center, Marseille in the south and Paris) during two different time periods. The
37 geographical unit used was the census block. The response variable was the NO₂ annual average
38 concentration ($\mu\text{g}/\text{m}^3$) per census block and the explanatory variables were a neighborhood deprivation
39 index and socioeconomic and demographic data derived from the national census. Generalized
40 additive models allowed us to take into account spatial autocorrelation. We found that the strength and
41 direction of the association between deprivation and NO₂ estimates varied between cities. In Paris, the
42 higher social category is exposed to a higher concentration of NO₂, on average, whereas in Lyon and
43 in Lille, the census blocks more likely to have higher concentration were those predominantly of the
44 middle and the lower social categories. Socioeconomic profiles associated with air pollution vary
45 according to the cities and the time period. For Lyon, Marseille and Paris, the proportion of
46 immigrants is an important factor for the first period but not for the second. Conversely, in Marseille
47 and Lille, the proportion of the population with insecure jobs was associated with NO₂ concentrations
48 in the second period. In Paris, relationships with the proportions of no property owners and low
49 income were observed during both time periods. There is clear evidence of a variety of spatial and
50 temporal patterns of environmental inequalities in French metropolitan areas, dependent on their
51 historical social and cultural characteristics. General statements about environmental inequalities are
52 inappropriate.

53
54
55 54 Keywords : air pollution, environmental inequalities, generalized additive models, social determinants,
56
57 55 spatial-temporal analysis
58
59
60 56

1. Introduction

59

60 Environmental justice first emerged in the United States and Canada where it is now an important part
61 of environment and public health policy assessment (Jerrett et al., 2001; Bowen, 2002; Fairburn et al.,
62 2009; Laurent, 2011). The concept draws attention to the questions of whether certain socioeconomic
63 groups, including the economically and politically disadvantaged, bear a disproportionate burden of
64 environmental externalities, and whether policies and practices that relate to sources of nuisances and
65 pollution or, conversely, to wholesome environments (e.g., green spaces) are equitable and fair
66 (Bowen, 2002).

67 A number of ecological studies dealing with environmental equity (or justice) have investigated this
68 topic and assessed population exposure to environmental pollution and socioeconomic characteristics
69 using data collected at different geographic scales. As expected, many studies conclude that groups
70 with a low socioeconomic status (SES) tend to be more highly exposed to air pollutants and toxicants,
71 due especially to the proximity of their homes to pollution sources (e.g. high-traffic roads, industrial
72 facilities and waste disposal sites) (Finkelstein et al., 2005; Chaix et al., 2006; Marshall, 2008; Briggs
73 et al., 2008; Yanosky et al., 2008; Bell and Ebisu, 2012; Brochu et al., 2011; Diekmann and Meyer,
74 2010; Viel et al., 2010).

75 More recently, the issue of uneven distribution of environmental pollution across populations with
76 different socioeconomic status entered into discussions in Europe (Netherlands (Kruize et al., 2007),
77 Finland (Rotko et al., 2001), Sweden (Chaix et al., 2006), Germany (Kohlhuber et al., 2006), UK
78 (Namdeo and Stringer, 2008; Mitchell and Dorling, 2003; Fairburn et al., 2009; Walker, 2010;
79 Jephcote and Chen, 2012), Italy (Forastiere et al., 2007), France (Laurian, 2008; Havard et al., 2009).
80 In contrast with American studies, inconsistent results were obtained from studies performed in
81 Europe (Bowen, 2002). For instance, while some report that populations with low socioeconomic
82 status are more exposed to air pollutants (Kruize et al., 2007; Namdeo and Stringer, 2008), others find
83 that populations with middle socioeconomic status experience higher levels of air pollution (Havard et

84 al., 2009), or report an inverse relationship (Forastiere et al., 2007) . The methodological diversity of
1
2 85 these studies and the variety of their settings may partly explain the heterogeneity of their results. This
3
4 86 heterogeneity might also express the diversity of the urban make-up both across and within European
5
6 87 countries (Deguen and Zmirou-Navier, 2010). According to an Organisation for Economic Co-
7
8 88 operation and Development (*OECD*) report (OCDE Report, 2004), more studies are needed in Europe
9
10 89 to improve our understanding of the underlying mechanisms of environmental inequality.

11
12
13
14 90 A decreasing trend of urban air pollution has been observed in most European countries, an effect of
15
16 91 the important evolution of the general structure of the European economy during the last two decades,
17
18 92 but also of national regulations, in compliance with the European directives (1999/30/EC,
19
20 93 2008/50/EC). In spite of air quality improvements, air pollution remains a major public health research
21
22 94 field, particularly in consideration of social justice.

23
24
25
26 95 In this context, our work concerns two issues. Firstly, we will identify whether urban neighborhoods
27
28 96 are characterized by an uneven distribution of ambient air concentrations of nitrogen dioxide (NO₂)
29
30 97 according to the level of deprivation in four large French metropolitan areas (MAs). Secondly, we will
31
32 98 investigate the time trends of environmental inequalities by comparing two time periods during the last
33
34 99 decade (2002-2005 v.s. 2006-2009) during which a general pattern of air pollution reduction was
35
36 100 observed. The underlying questions which we shall try to answer are threefold: (1) are environmental
37
38 101 inequalities comparable across the four French cities, with regards to air pollution? (2) how do
39
40 102 environmental inequalities change over the time? (3) do the socioeconomic markers of environmental
41
42 103 inequalities differ between the two study periods?

43
44
45
46 104

47
48 105 Nitrogen dioxide was selected because it is known to be a good tracer of urban air pollution generated
49
50 106 by traffic and because its spatial heterogeneity is recognized to be greater than for other air pollutants
51
52 107 (Jerrett et al., 2004). It is also a pollution indicator for which exposure varies substantially among
53
54 108 socioeconomic groups (Rotko et al., 2001; Stroh et al., 2005; Chaix et al., 2006; Yanosky et al., 2008;
55
56 109 Crouse et al., 2009). A principal originality of the study stems from the small-spatial scale of the
57
58 110 analysis which reduces the residual autocorrelation between spatial units, the comprehensive modeling
59
60
61
62
63
64
65

111 of the urban distribution of NO₂ concentrations, and the use of the same statistical methodology for the
112 4 major metropolitan areas.

113

114 **2. Materials and Methods**

115

116 *2.1. Study area and setting*

117 Our study is an ecological study using the smallest geographical level unit with available
118 socioeconomic data in France. The statistical unit is the sub-municipal French census block (called
119 IRIS “Îlot Regroupé pour l’Information Statistique”) defined by the National Institute of Statistics and
120 Economic Studies (INSEE). This geographical unit averages 2000 inhabitants and is constructed to be
121 as homogenous as possible in terms of socio-demographic characteristics and land use. The spatial
122 analysis of environmental inequalities were stratified by two periods of 4 years (2002-2005, 2006-
123 2009) to assess trend over time for the four Metropolitan Areas.

124 The four French metropolitan areas in this study are: Lille (North of France), Marseille (South of
125 France), Lyon (Center of France) and Paris (see localization on Figure 1). The Paris MA includes Paris
126 city and three ‘départements’ surrounding the capital city (named “petite couronne”). These four MAs
127 have been chosen because they exhibit significant differences regarding socioeconomic and
128 demographic characteristics (Table 1).

129

130 *2.2. Neighborhood socioeconomic variables*

131 The census of 1999 for the first period (2002-2005) and the census of 2006 for the second period
132 (2006-2009) were used to estimate socioeconomic variables. We included in the analysis variables
133 which have already been significantly associated with environmental quality and health events in
134 previous studies; this includes the prevalence or proportion of unemployment, insecure jobs, white
135 collar jobs, high level of education , single parents, subsidized housing, immigrants and mobility, non-

136 homeownership and median income (Jerrett et al., 2001; Brainard et al., 2002; Yanosky et al., 2008;
137 Brochu et al., 2011; Branis and Linhartova, 2012).

138 A deprivation index was estimated at the census block level for the four MAs and by period. The
139 methodology developed to construct this index is described in a recently published article (Lalloué et
140 al., 2013). Briefly, this index was constructed by a principal component analysis from a selection of 21
141 socioeconomic and demographic variables from national census. These variables reflected multiple
142 dimensions of deprivation: income, education, employment, family, housing and immigration status.

143

144 *2.3. Air pollution: spatial and temporal patterns*

145 Annual ambient air concentrations of NO₂ were modeled by the local air quality monitoring networks
146 (Atmo Nord Pas-de-Calais for Lille MA, Air Rhône-Alpes for Lyon MA, Air PACA for Marseille
147 MA, Airparif for Paris-petite couronne) for each census block throughout the entire study period
148 (2002–2009). The four air quality monitoring networks developed and tested a methodological
149 approach to describe and characterize disparities in environmental exposures at a local scale for that
150 period. They used different deterministic models: ADMS (Atmospheric Dispersion Modeling System)
151 Urban for the Lille MA (McHugh et al., 1997; Carruthers et al., 2000) , SIRANE for the Lyon MA
152 (Soulhac et al., 2012, 2011), ESMERALDA for the Paris MA (Carruthers et al., 2000) and STREET
153 for the Marseille MA (Carruthers et al., 2000). In 2010, Jerrett et al demonstrated the effectiveness and
154 reliability of these types of models for assessing air quality in health assessment research (Jerrett et al.,
155 2010). More precisely, these models integrate meteorological data (air temperature, wind speed and
156 direction, relative humidity, barometric pressure, supplied by Météo France, the French meteorologic
157 service), emission sources according to their contribution to ambient air pollution, and background
158 pollution measurements as input parameters. Selected emission sources were linear sources (main
159 roads), surface sources (diffuse road sources and residential and tertiary emissions) and important
160 point sources.

161

162 2.4. *Statistical analysis*

1
2
3 163 We first described the association between NO₂ concentrations and deprivation by calculating the
4
5 164 difference, in term of NO₂ concentrations, between opposite areas of deprivation (Table 2-AB). The
6
7 165 most deprived areas correspond to the 80th percentile of the deprivation index distribution (noted P80),
8
9 166 and the less deprived correspond to 20th percentile (noted P20). In addition, in order to compare this
10
11 167 association between the two periods (period 1: 2002-2005 and period 2: 2006-2009), we stratified the
12
13 168 analysis by period for each MA. Then, simple and multivariate regression analyses including several
14
15 169 socioeconomic variables were carried out, adjusting for spatial autocorrelation. Three different
16
17
18 170 findings could be observed:

19
20
21 171 (1) Should the socioeconomic variables which were significantly associated with NO₂ concentrations
22
23 172 during the first study period not remain associated with air quality during the second period, the
24
25 173 conclusion of this observation may be that environmental disparities tend not to level off over time.

26
27
28
29 174 (2) An inverse situation could be observed (significant associations in the second period which were
30
31 175 not present in the first one), suggesting that environmental inequalities appeared recently.

32
33
34 176 (3) The same socioeconomic variables were significantly associated with the NO₂ concentrations in
35
36 177 both periods, meaning that environmental inequalities persisted in the second period. In this last case,
37
38 178 three different situations could be observed: the inequalities increase, decrease or stay stable between
39
40 179 the two periods.

41
42
43
44 180

45
46
47 2.5. *Regression models*

48
49
50 182 We used generalized additive models (GAMs), a form of non-parametric regression with the ability to
51
52 183 analyze area-based data adjusting for covariates and taking into account spatial autocorrelation (Hastie
53
54 184 and Tibshirani, 1990; Wood, 2006, 2004; Kiffer et al., 2011).

55
56
57
58 185 We modeled location, a potential proxy measure of unknown exposure or uncontrolled risk factors,
59
60 186 using a smooth (S) of longitude (X) and latitude (Y) with a gaussian link function.

187 $\text{Log } [p(X,Y)] = S(X,Y) + \gamma'Z$ (equation 1)

188 where the left-hand side is the logarithm of the NO₂ concentrations estimation at the census block's
189 centroid (X,Y), and γ is a vector of parameters associated with Z, the vector of covariates.

190 In our study, Z represents the deprivation index or the socioeconomic variables and γ the coefficient
191 associated with the covariates. The model is semi-parametric because it includes both nonparametric
192 and parametric components. Without the smooth function, S(X,Y), the model becomes an ordinary
193 linear regression on the covariates. The amount of smoothing depends on the percentage of the data
194 points in the neighborhood, referred to as the span size. GAMs also allow selection of "optimal" span
195 size. We determined the optimal amount of smoothing by minimizing the Akaike's Information
196 Criterion (AIC).

197 We first tested all variables in separate models and then all variables were included in the full model.
198 Variables that have no significant influence in the full model were removed. The best model was the
199 model with the smallest Akaike's Information Criterion. Because descriptive analysis had preliminarily
200 revealed that the shape of the relationship between NO₂ concentrations and the deprivation index in
201 Lyon MA was not linear, the deprivation index and the socioeconomic variables were classified into
202 tertiles and used in models of regression as categorical data (Figure 2).

203
204 Census blocks without socioeconomic information (for example, an industrial census block or a park)
205 were excluded from the analysis; i.e., 2 (0.4%) in Lille, 13 (2.5%) in Lyon, 11 (1.7%) in Marseille and
206 54 (1.9%) in Paris. The final dataset included 476, 485, 545 and 2692 census blocks for Lille, Lyon,
207 Marseille and Paris respectively. We also excluded census blocks that had no information on air
208 pollution: 24 (4.7%) in Lille, 13 (2.5%) in Lyon, 72 (11.4%) in Marseille and 3 (0.1%) for Paris.

209 We used the R software using the gam package, written by Trevor Hastie to implement the GAM
210 framework of Hastie and Tibshirani (1990,), and ArcView 9.3 software (ESRI, Inc., Redlands,
211 California) to map the results of our analyses.

212 3. Results

1 213

2 214 3.1. Descriptive statistics

3 214
4
5
6 215 Table 1 summarizes the annual averages of the NO₂ concentrations for the period 2002-2005 and
7
8 216 2006-2009 for the four MAs (Lille, Paris, Lyon and Marseille). The Lyon MA had the largest decrease
9
10 217 of NO₂ concentrations between the two periods from 43.6 µg/m³ (SD=5.4) to 38.6 µg/m³ (SD=6.3) for
11
12 218 the first and the second periods, respectively (-11.5%). The Marseille MA had the smallest difference
13
14 219 (-3.3%) with 33µg/m³ (SD=10.2) for the first period and 31.9µg/m³ (SD=9.8) for the second period.
15
16 220 The Paris MA had the highest level of NO₂ concentrations with 47.4µg/m³ (SD=9.3) for the first and
17
18 221 44.2µg/m³ (SD=9.9) for the second period (-6.8 %). In the Lille, Lyon, Marseille and Paris MAs, the
19
20 222 neighborhoods with the highest concentration are in the central areas of the cities and along major
21
22 223 roads that link these cities (Figures 3).

23
24
25
26
27 224 Neighborhood socioeconomic variables present some similar patterns between the two periods (Table
28
29 225 1). All MAs are characterized by an increased proportion of unsecure jobs and of immigrants and a
30
31 226 decreased proportion of single parent families and median income (expressed in euros). They also
32
33 227 displayed specific patterns depending on the MA. Marseille has the highest proportion of
34
35 228 unemployment and the lowest proportion of subsidized housing whereas Paris has the highest
36
37 229 proportion of white collar jobs and non-homeownership. Figure 4 shows the spatial distribution of the
38
39 230 deprivation index in tertiles and also presents strong contrasts within each MA. In the Lille and
40
41 231 Marseille MAs, the most deprived areas are located in the urban center and in the major cities (Lille
42
43 232 and Roubaix for the Lille MA, and Marseille and Aix-en-Provence for the Marseille MA) (Figure 4).
44
45 233 In Lyon, the suburb seems to be most deprived, whereas in Paris the most deprived areas are
46
47 234 concentrated in the north and the south-eastern part of the MA (Figures 4).

48 235 3.2. Levels of air pollution across social groups

49
50
51
52 236 Figure 2 reveals three different patterns of relationship between the deprivation index and the NO₂
53
54 237 concentrations. For Lille and Marseille MAs, the relationship follows a linear trend, with the most
55
56 238 deprived population living in the most exposed census blocks. While in the Lyon MA, the medium

239 class of deprivation is the most exposed, in the Paris MA, the population living in the census blocks
240 with a higher social level was exposed to higher concentrations of NO₂.

241 Table 2-A reveals environmental inequalities related to NO₂ concentrations. In Lille, Lyon and
242 Marseille MAs, the ratio between the average NO₂ concentrations in the most disadvantaged census
243 blocks (percentile 80 of the distribution of the deprivation index, P80) and the most favored ones
244 (percentile 20, P20) vary between 1.10 and 1.67. Inversely, in Paris MA, the NO₂ concentrations are
245 higher in the most favored compare to the most disadvantaged census blocks (P80/P20 ratio=0.83).

246
247 Table 2-B highlights a decrease of NO₂ concentrations during the study period. However, this
248 reduction differs between the most deprived and the most favored census blocks. In the Lille and Lyon
249 MAs, the reduction between the two periods is lower in the most disadvantaged census blocks than in
250 the most favored one (-9.3% versus -10.3%, and -13.7% versus -16.8%, respectively). By contrast, in
251 Paris and Marseille MAs, the reduction is greater among the most deprived census blocks.

252 253 *3.3. Socioeconomic characteristics and air pollution*

254 Table 3-AB presents the results of the simple regression coefficients adjusting for location, and the
255 multivariate regression coefficient adjusting for location using GAM models of Lyon, Lille, Paris and
256 Marseille MAs for both time periods. Lyon MA was not linear, the deprivation index and the
257 socioeconomic variables were classified into tertiles and used in models of regression as categorical
258 data (C1 is the reference category). Multivariate GAM models represent the best model with the
259 smallest Akaike's Information Criterion. They showed significant associations between NO₂
260 concentrations and the socioeconomic characteristics across all the metropolitan areas. The strength of
261 the association between socioeconomic characteristics and NO₂ concentrations varies according to the
262 MA and the time period of interest. For the Lyon, Marseille and Paris MAs, the proportion of
263 immigrants was linked to air pollution during the first period but not in the second. Conversely, in the
264 Marseille and Lille MAs, the proportion of insecure jobs and, in Lyon MA, the proportion of non-

265 homeownership, became significantly associated with NO₂ concentrations in the second time period.

266 In the Paris MA, the proportion of non-homeownership and the low to medium income were similarly

267 linked to air pollution over the two periods. According to the homogeneity test, the period does not

268 modify the association of NO₂ concentrations with the proportion of non-homeownership (p-

269 interaction=0.78) or with the median income (p-interaction=0.13).

270

271 **4. Discussion**

273 Our study demonstrates the existence of social inequalities at the census blocks level in France,

274 regarding exposure to air pollution at the place of residence. We found an uneven distribution of NO₂

275 concentrations between census blocks where different social groups live in four large French MAs.

276 Our study reveals also differences in the strength and direction of the association between deprivation

277 (using the deprivation index or the socioeconomic variables) and NO₂ concentrations between two

278 time periods. An additional finding concerns the change of socioeconomic markers of environmental

279 inequalities between the two periods.

280

281 Our results confirms that levels of ambient air pollution in a MA can be variable at the local scale and

282 can create disparities between neighborhoods as already shown by previous studies (Jerrett et al.,

283 2005; Briggs et al., 2008; Crouse et al., 2009). Our results show that, in Paris, the most advantaged

284 census blocks were exposed to higher levels of NO₂, on average, whereas in Lyon and in Lille, the

285 most disadvantaged categories experienced greater exposure. This contrasted finding is coherent with

286 the diversity of the results reported in European studies dealing with social inequalities in association

287 with air pollution (Deguen and Zmirou-Navier, 2010). Stroh et al in 2005 also found that exposure to

288 air pollutants of different socioeconomic groups varies considerably between areas (Stroh et al., 2005).

289 Our results for the Lyon MA mimics previous observations from Havard et al in the Strasbourg MA

290 (North-East France); the relation is not linear, with the midlevel deprivation areas being the most

291 exposed to traffic-related air pollution (Havard et al., 2008). Several studies (Gouveia et al., 2003) in

292 Brazil, Branis (2012) in Czech cities, Forestiere (2007) in Rome have found, as in Paris, that the most

1 293 privileged census blocks were located in the most polluted areas. Finally, our finding in the Lille MA
2 294 of higher NO₂ concentrations in deprived areas is consistent with other results reported in other cities.
3
4 295 Yanosky et al (2008), in Massachusetts, found that, after controlling for spatial autocorrelation,
5
6 296 deprived areas were significantly associated with traffic related air pollutants, including nitrogen
7
8 297 dioxide. In Rijnmond (The Netherlands), according to Kruize et al. (2007), and in Leeds (UK),
9
10 298 according to Mitchell et Dorling et al. (2003) and Namdeo et al. (2008), lower income groups live in
11
12 299 places with higher levels of NO₂ than higher income groups.
13

14 300

15
16
17 301 Strong social disparities exist within large metropolitan areas in France. These disparities can be
18
19 302 explained by their history and by the long term local urban policies which contribute through urban
20
21 303 development to different social and economic make- up between cities and neighborhoods. Typically,
22
23 304 neighborhoods close to industrial sources of pollution are where deprived populations live. Past
24
25 305 working conditions, in particular the social and cultural organization of industrial areas inherited from
26
27 306 the first industrial revolution during the 19th century may explain the spatial distribution of the
28
29 307 deprived population in France. For instance, the Lille MA was characterized by major areas of mines
30
31 308 and blue workers lived close to their workplace to increase productivity and reduce time losses in
32
33 309 transport. In Lyon, people in the privileged neighborhoods in part live in the west green countryside at
34
35 310 a reasonable distance from the city center and from pollution sources, whereas the middle class stay in
36
37 311 the city center (where white collar jobs are concentrated) and in the eastern area of Lyon where real
38
39 312 estate price is not as high. The transformation of Paris during the Second Empire (1852-1870) and the
40
41 313 Baron Haussmann works achieved a comprehensive modernization of the capital, both with a
42
43 314 hygienist objective and for security reasons (large open roads favor the rapid move of armed troops)
44
45 315 and the central neighborhoods of the city of Paris became favored.
46
47
48
49

50
51 316
52 317 The housing market is a key factor that biases land use decisions and might explain why some groups
53
54 318 of people suffer from both a low socioeconomic status and poorer air quality at their place of
55
56 319 residence. In Lille, Lyon, Marseille and Paris, the land price in deprived areas is on average 1935,
57
58 320 1863, 1325 and 2271 euros per square meter, and in the favored areas 2479, 2711, 2172 and 5457
59
60
61
62
63
64
65

1 321 euros/m² respectively (chamber of notary, personal communication). One reason is that the presence of
2 322 pollution sources depresses the housing market and provides an opportunity for local authorities to
3
4 323 construct council housing at low cost (Deguen and Zmirou-Navier, 2010). Consistent with previous
5
6 324 studies in Europe (Brainard et al., 2002; Mitchell and Dorling, 2003; Stroh et al., 2005; Chaix et al.,
7
8 325 2006), we found that social and economic resources (income, insecure job, non-homeownership,
9
10 326 single parent and occupation) explain mainly the environmental inequalities (Deguen and Zmirou-
11
12 327 Navier, 2010; Havard et al., 2009).

13
14
15 328

16
17 329 To our knowledge, few studies have investigated the patterns of environmental inequalities between
18
19 330 deprived and favored areas; this is the first study comparing these patterns between several
20
21 331 metropolitan areas (Mitchell, 2005; Tonne et al., 2008). Despite a reduction in NO₂ concentrations
22
23 332 over the study period in the four MAs, we found contrasted results in the temporal differences of
24
25 333 environmental inequalities. In Paris and Marseille, we show that the improvement was more beneficial
26
27 334 to the deprived census blocks; noteworthy is that the difference in air quality between deprived and
28
29 335 more well off areas in Marseille remains large. These finding are consistent with results of Tonne et al
30
31 336 in 2008 who in London, showed a greater reduction in air pollution in deprived areas than in the most
32
33 337 affluent ones after the introduction of the Congestion Charging Zone (Tonne et al., 2008). Comparing
34
35 338 the trend of NO₂ levels between 1993 and 2005 in Leeds Mitchell et al demonstrated that the average
36
37 339 difference between deprived and affluent communities declined from 10.6 mg/m³ in 1993 to 3.7
38
39 340 mg/m³ in 2005 as a result of city-wide improvements in air quality driven by fleet renewal (Mitchell,
40
41 341 2005). In our data, the results in Lille and Lyon show an inverse situation, with an increase in
42
43 342 environmental inequalities.

44
45
46 343

47
48
49 344 The strength of the association between air quality and social characteristics from our results is weak
50
51 345 whatever the MA. This could be due to the consideration of spatial autocorrelation in models which
52
53 346 reduces the strength of the association (Yanosky et al., 2008; Havard et al., 2009). Moreover, in a
54
55 347 study in Switzerland, Diekmann and Meyer (2010) also found a weak association and they explain that
56
57 348 environmental injustice may be less accentuated in the Swiss population than in countries with a
58
59
60
61
62
63
64
65

349 higher degree of socioeconomic segregation like in the United States. Another explanation suggests
1
2 350 that the size of areas considered in the different studies may also influence the strength of the
3
4 351 association. Stroh et al (Stroh et al., 2005) and Briggs et al in 2008 conclude that small area
5
6 352 associations between socioeconomic status and environmental pollution become stronger with an
7
8 353 increased level of spatial aggregation.

10
11
12 354 However, this study does have several strengths. According to Jerrett (2004), NO₂ modeling provides
13
14 355 a powerful method to estimate pollutant concentrations over a fine spatial scale. It gives acceptable
15
16 356 results and shows high correlations between the model's predictions and the measured NO₂ values
17
18 357 obtained from the monitoring stations. The same methodology was used to assess intra urban exposure
19
20
21 358 in each MA and to compare the environmental inequalities between them. In order to minimize the
22
23 359 ecological biases and take into account the dependency of spatial units we used validated models.
24
25 360 Bowen et al (2002), reviewed the empirical research on environmental justice and concluded that
26
27 361 published studies fail to consider important methodological issues about choice of geographic units,
28
29 362 analytical methods and exposure estimates and taking into account spatial autocorrelation among
30
31 363 observations. Lastly, the heterogeneity of the methods used may in part explain discrepancies in the
32
33
34 364 results.

35
36
37 365

40 366 **5. Conclusion**

41 367
42
43
44 368 Our findings provide evidence of environmental inequalities, whose spatial and temporal patterns vary
45
46 369 according to the historical socioeconomic and cultural characteristics of the large French metropolitan
47
48 370 areas that were studied. Traffic related air pollution is an important contributor to environmental
49
50
51 371 inequalities and associated risks and calls for further examination of the role of socioeconomic
52
53 372 characteristics in air pollution epidemiology and risk assessment. Such evidence may inform local or
54
55 373 country-wide policies that would aim to cope with environmental health inequalities.

56
57 374

59 375 **Competing interests**

376 The authors declare that they have no competing interests.

377

378 **Acknowledgements**

379 The authors thank all scientific researchers of the Equit'Area project, the personnel of the local
380 association in air quality AASQA Lille, Lyon, Marseille and Paris Metropolitan Area who work on the
381 harmonization of modeled NO₂. The Equit'Area project is supported by the French National Research
382 Agency (ANR, contract-2010-PRSP-002-01), the EHESP School of Public Health, the French General
383 Directorate of Health (DGS) and the Nord-Pas de Calais region. Cindy Padilla is recipient of a
384 doctoral grant from the Environment and Energy Management Agency (ADEME). Verónica Vieira
385 was supported by a grant from the National Institute of Environmental Health Sciences (NIEHS; 2P42
386 ES007381-16A1), National Institutes of Health (NIH). Its contents are solely the responsibility of the
387 authors and do not necessarily represent the views of NIEHS, NIH.

388

389 **References**

390

391 Bell, M.L., Ebisu, K., 2012. Environmental inequality in exposures to airborne particulate matter
392 components in the United States. *Environ. Health Perspect.* 120, 1699–1704.

393 Bowen, W., 2002. An analytical review of environmental justice research: what do we really know?
394 *Environ Manage* 29, 3–15.

395 Brainard, J.S., Jones, A.P., Bateman, I.J., Lovett, A.A., Fallon, P.J., 2002. Modelling environmental
396 equity: access to air quality in Birmingham, England. *Environment and Planning A* 34, 695 –
397 716.

398 Branis, M., Linhartova, M., 2012. Association between unemployment, income, education level,
399 population size and air pollution in Czech cities: Evidence for environmental inequality? A
400 pilot national scale analysis. *Health & Place* 18, 1110–1114.

401 Briggs, D., Abellan, J.J., Fecht, D., 2008. Environmental inequity in England: small area associations
402 between socio-economic status and environmental pollution. *Soc Sci Med* 67, 1612–1629.

403 Brochu, P.J., Yanosky, J.D., Paciorek, C.J., Schwartz, J., Chen, J.T., Herrick, R.F., Suh, H.H., 2011a.
404 Particulate air pollution and socioeconomic position in rural and urban areas of the
405 Northeastern United States. *Am J Public Health* 101 Suppl 1, S224–230.

406 Carruthers, D.J., Edmunds, H.A., Lester, A.E., McHugh, C.A., Singles, R.J., 2000. Use and validation of
407 ADMS-Urban in contrasting urban and industrial locations. *International Journal of*
408 *Environment and Pollution* 14, 364–374.

409 Chaix, B., Gustafsson, S., Jerrett, M., Kristersson, H., Lithman, T., Boalt, A., Merlo, J., 2006. Children’s
410 exposure to nitrogen dioxide in Sweden: investigating environmental injustice in an
411 egalitarian country. *J Epidemiol Community Health* 60, 234–241.

412 Crouse, D.L., Ross, N.A., Goldberg, M.S., 2009. Double burden of deprivation and high concentrations
413 of ambient air pollution at the neighbourhood scale in Montreal, Canada. *Soc Sci Med* 69,
414 971–981.

415 Deguen, S., Zmirou-Navier, D., 2010. Social inequalities resulting from health risks related to ambient
416 air quality--A European review. *Eur J Public Health* 20, 27–35.

417 Diekmann, A., Meyer, R., 2010. Democratic smog? An empirical study on the correlation between
418 social class and environmental pollution. *Kolner Zeitschrift fur Soziologie und*
419 *Sozialpsychologie* 62, 437.

420 Fairburn, J., Butler, B., Smith, G., 2009. Environmental justice in South Yorkshire: locating social
421 deprivation and poor environments using multiple indicators. *Local Environment* 14, 139–
422 154.

423 Finkelstein, M.M., Jerrett, M., Sears, M.R., 2005. Environmental inequality and circulatory disease
424 mortality gradients. *J Epidemiol Community Health* 59, 481–487.

425 Forastiere, F., Stafoggia, M., Tasco, C., Picciotto, S., Agabiti, N., Cesaroni, G., Perucci, C.A., 2007.
426 Socioeconomic status, particulate air pollution, and daily mortality: differential exposure or
427 differential susceptibility. *Am. J. Ind. Med.* 50, 208–216.

428 Gouveia, N., Hajat, S., Armstrong, B., 2003. Socioeconomic differentials in the temperature-mortality
429 relationship in São Paulo, Brazil. *Int J Epidemiol* 32, 390–397.

430 Hastie, T., Tibshirani, R., 1990. *Generalized Additive Models*. London, Chapman and Hall.

431 Havard, S., Deguen, S., Bodin, J., Louis, K., Laurent, O., Bard, D., 2008. A small-area index of
432 socioeconomic deprivation to capture health inequalities in France. *Soc Sci Med* 67, 2007–
433 2016.

434 Havard, S., Deguen, S., Zmirou-Navier, D., Schillinger, C., Bard, D., 2009. Traffic-related air pollution
435 and socioeconomic status: a spatial autocorrelation study to assess environmental equity on
436 a small-area scale. *Epidemiology.* 20, 223–230.

437 Jephcote, C., Chen, H., 2012. Environmental injustices of children's exposure to air pollution from
1 438 road-transport within the model British multicultural city of Leicester: 2000-09. *Sci. Total*
2 439 *Environ.* 414, 140–151.

3 440 Jerrett, M., Arain, A., Kanaroglou, P., Beckerman, B., Potoglou, D., Sahuvaroglu, T., Morrison, J.,
4 441 Giovis, C., 2005. A review and evaluation of intraurban air pollution exposure models. *J Expo*
5 442 *Anal Environ Epidemiol* 15, 185–204.

6 443 Jerrett, M., Burnett, R., Brook, J., Kanaroglou, P., Giovis, C., Finkelstein, N., Hutchison, B., 2004. Do
7 444 socioeconomic characteristics modify the short term association between air pollution and
8 445 mortality? Evidence from a zonal time series in Hamilton, Canada. *J Epidemiol Community*
9 446 *Health* 58, 31–40.

10 447 Jerrett, M., Burnett, R.T., Kanaroglou, P., Eyles, J., Finkelstein, N., Giovis, C., Brook, J.R., 2001. A GIS -
11 448 environmental justice analysis of particulate air pollution in Hamilton, Canada. *Environment*
12 449 *and Planning A* 33, 955 – 973.

13 450 Jerrett, M., Gale, S., Kontgis, C., 2010. Spatial modeling in environmental and public health research.
14 451 *Int J Environ Res Public Health* 7, 1302–1329.

15 452 Kiffer, C.R.V., Camargo, E.C.G., Shimakura, S.E., Ribeiro, P.J., Jr, Bailey, T.C., Pignatari, A.C.C.,
16 453 Monteiro, A.M.V., 2011. A spatial approach for the epidemiology of antibiotic use and
17 454 resistance in community-based studies: the emergence of urban clusters of *Escherichia coli*
18 455 quinolone resistance in Sao Paulo, Brasil. *Int J Health Geogr* 10, 17.

19 456 Kohlhuber, M., Mielck, A., Weiland, S.K., Bolte, G., 2006. Social inequality in perceived environmental
20 457 exposures in relation to housing conditions in Germany. *Environmental Research* 101, 246–
21 458 255.

22 459 Kruize, H., Driessen, P.P.J., Glasbergen, P., van Egmond, K.N.D., 2007. Environmental equity and the
23 460 role of public policy: experiences in the Rijnmond region. *Environ Manage* 40, 578–595.

24 461 Lalloué, B., Monnez, J.-M., Padilla, C., Kihal, W., Le Meur, N., Zmirou-Navier, D., Deguen, S., 2013. A
25 462 statistical procedure to create a neighborhood socioeconomic index for health inequalities
26 463 analysis. *Int J Equity Health* 12, 21.

27 464 Laurent, É., 2011. Issues in environmental justice within the European Union. *Ecological Economics*
28 465 70, 1846–1853.

29 466 Laurian, L., 2008. Environmental Injustice in France. *Journal of Environmental Planning and*
30 467 *Management* 51, 55–79.

31 468 Marshall, J.D., 2008. Environmental inequality: Air pollution exposures in California's South Coast Air
32 469 Basin. *Atmospheric Environment* 42, 5499–5503.

33 470 McHugh C, Carruthers D, Edmunds H, 1997. ADMS-Urban: an air quality management system for
34 471 traffic, domestic and industrial pollution. *Int J Environ Pollution* 8, 437–440.

35 472 Mitchell, G., 2005. Forecasting environmental equity: air quality responses to road user charging in
36 473 Leeds, UK. *J. Environ. Manage.* 77, 212–226.

37 474 Mitchell, G., Dorling, D., 2003. An environmental justice analysis of British air quality. *Environment*
38 475 *and Planning A* 35, 909 – 929.

39 476 Namdeo, A., Stringer, C., 2008. Investigating the relationship between air pollution, health and social
40 477 deprivation in Leeds, UK. *Environment International* 34, 585–591.

41 478 OCDE Report ENV/EPOC/WPNEP, 2004. Environment and distributional issues: analysis, evidence and
42 479 policy implications.

43 480 Rotko, T., Kousa, A., Alm, S., Jantunen, M., 2001. Exposures to nitrogen dioxide in EXPOLIS-Helsinki:
44 481 microenvironment, behavioral and sociodemographic factors. *J Expo Anal Environ Epidemiol*
45 482 11, 216–223.

46 483 Soulhac, L., Salizzoni, P., Cierco, F.-X., Perkins, R., 2011. The model SIRANE for atmospheric urban
47 484 pollutant dispersion; part I, presentation of the model. *Atmospheric Environment* 45, 7379–
48 485 7395.

49 486 Soulhac, L., Salizzoni, P., Mejean, P., Didier, D., Rios, I., 2012. The model SIRANE for atmospheric
50 487 urban pollutant dispersion; PART II, validation of the model on a real case study. *Atmospheric*
51 488 *Environment* 49, 320–337.

489 Stroh, E., Oudin, A., Gustafsson, S., Pilesjö, P., Harrie, L., Strömberg, U., Jakobsson, K., 2005. Are
1 490 associations between socio-economic characteristics and exposure to air pollution a question
2 491 of study area size? An example from Scania, Sweden. *International Journal of Health*
3 492 *Geographics* 4, 30.

4 493 Tonne, C., Beevers, S., Armstrong, B., Kelly, F., Wilkinson, P., 2008. Air pollution and mortality
5 494 benefits of the London Congestion Charge: spatial and socioeconomic inequalities. *Occup*
6 495 *Environ Med* 65, 620–627.

8 496 Viel, J.F., Hagi, M., Upegui, E., Laurian, L., 2010. Environmental justice in a French industrial region:
9 497 Are polluting industrial facilities equally distributed? *Health Place*.

10 498 Walker, G., 2010. Beyond Distribution and Proximity: Exploring the Multiple Spatialities of
11 499 Environmental Justice, in: Holifield, R., Porter, M., Executive, G.W.C. of E.J.S. (Eds.), *Spaces of*
12 500 *Environmental Justice*. Wiley-Blackwell, pp. 23–46.

14 501 Wheeler, B.W., Ben-Shlomo, Y., 2005. Environmental equity, air quality, socioeconomic status, and
15 502 respiratory health: a linkage analysis of routine data from the Health Survey for England. *J*
16 503 *Epidemiol Community Health* 59, 948–954.

17 504 Wood, SN, 2004. Stable and efficient multiple smoothing parameter estimation for generalized
18 505 additive models. *Journal of American Statistical Association* 673–86.

20 506 Wood, SN, 2006. *Generalized additive models: An introduction with R*. Chapman and Hall/CRC press.

21 507 Yanosky, J.D., Schwartz, J., Suh, H.H., 2008. Associations between measures of socioeconomic
22 508 position and chronic nitrogen dioxide exposure in Worcester, Massachusetts. *J. Toxicol.*
23 509 *Environ. Health Part A* 71, 1593–1602.

24 510

Table 1: Descriptive statistics of the distribution of the socioeconomic variables and NO₂ concentrations according to periods and MAs. Mean and standard deviation for all MA

	LILLE (504 census blocks)			LYON (511 census blocks)			MARSEILLE (628 census blocks)			PARIS (2749 census blocks)		
	2002-2005	2006-2009	p†	2002-2005	2006-2009	p	2002-2005	2006-2009	p	2002-2005	2006-2009	p
NO ₂ concentrations (µg/m ³)	33.7±6.04	30.5±4.6	***	43.6±5.4	38.6±6.3	***	33.0±10.2	31.9±9.8	NS	47.4±9.3	44.2±9.9	***
% unemployment	16.2±9.6	14.7±9.1	***	12.7±6.6	11.6±6.4	***	21.6±11.7	16.2±9.8	***	12.7±5.8	11.9±5.5	***
% single parent	15.3±7.1	10.8±5.5	***	14.2±6	9.2±4.5	***	18.7±8.4	12.1±5.8	***	16.2±6	9.9±5	***
% insecure job	12.1±4.2	27.2±7.6	***	11.6±3.6	27.4±7.9	***	10.6±3.2	25.2±6.9	***	10.7±3.2	22.7±5.5	***
% white collar	13.2±11.1	11.4±8.4	**	15.4±10.7	14±9	**	12.6±10.6	10.6±7.4	***	22.7±14.8	21.8±13.3	***
% higher education	8.1±3.5	9.9±3.7	**	9.3±3.6	11±3.9	**	7.8±3.5	9.5±4.3	**	9.4±2.9	10±3.4	*
% non-homeownership	48.7±25.9	47.9±26.4	*	56.4±21.8	53.7±22.8	***	53.3±22.2	52.7±23.8	NS	64±20.5	61±22	***
% subsidized housing	21.5±26.8	23.4±24.5	***	24.4±29.8	21.2±25	***	19.3±28.1	17.5±25.3	*	29.2±32.5	24.8±28.3	***
% immigrants	4.8±4.5	8.2±6.3	***	7.7±6.6	13.3±8.9	***	5.4±5.7	11.0±8.7	***	12.3±6.8	20.3±9.1	***
Median income (euro)	21952±7130	16339±5411	***	23751±6558	18691±5194	***	21534±7383	16302±5672	***	27636±7823	21638±8343	***

† p values for Student Paired test between periods for each MA.

*p<0.05. **p<0.01. ***p<0.001. NS= not significant

Table 2-A : NO₂ concentrations (µg/m³) in contrasted areas according to socioeconomic deprivation in the 4 metropolitan areas, France, 2002 to 2009.

Metropolitan areas	Number of †(P80) and ‡(P20) census blocks	NO ₂ concentrations in the P80 census blocks	NO ₂ concentrations in the P20 census blocks	P80/P20
Lille	96/478	34.6±5.4	28.7 ± 4.8	1.21
Lyon	98/493	40.2±6.0	36.5 ± 5.1	1.10
Marseille	109/544	37.4 ± 8.7	22.3 ± 6.7	1.67
Paris	532/2656	42.1± 8	50.5 ± 8.9	0.83

†(P80): The most deprived census blocks, corresponding to the 80th percentile of the deprivation index distribution

‡(P20): The less deprived census blocks, corresponding to 20th percentile of the deprivation index distribution

Table 2-B : Evolution of the distribution of NO₂ concentrations (µg/m³) according to extreme areas of deprivation between two periods in the 4 metropolitan areas, France.

Metropolitan areas	NO ₂ concentrations in P80 areas at		NO ₂ concentrations in P20 areas at		Evolution** (%) P80	Evolution (%) P20
	P1*	P2*	P1*	P2*		
Lille	36.3± 6.3	32.9± 4.7	30.8± 5.8	27.6± 4.3	-9.3%	-10.3%
Lyon	43.4± 4.9	37.5± 5.5	40.1± 4.9	33.4± 5.1	-13.7%	-16.8%
Marseille	39.8± 8.3	36.7± 8.6	22.5± 6.6	22.1± 6.6	-7.5%	-1.8%
Paris	44.4± 8.4	40.6± 8.1	50.8± 9.4	49.1± 9.6	-8.6%	-3.3%

† (P80): The most deprived census blocks, corresponding to the 80th percentile of the deprivation index distribution

‡ (P20): The less deprived census blocks, corresponding to 20th percentile of the deprivation index distribution

* P1: first period 2002-2005, P2: second period 2006-2009

** ((NO₂ concentrations at P2 - NO₂ concentrations at P1) / NO₂ concentrations of P1) *100. Example: -9.3% = ((32.9-36.3)/36.3)*100

Table 3: Regression coefficient of the estimated NO₂ concentrations by the neighborhood socioeconomic characteristics for two periods in Lyon, Lille, Marseille and Paris MA for A) 2002-2005 period, B) 2006-2009 period. Only Lyon MA was not linear, the socioeconomic variables were classified into tertiles and used in models of regression as categorical data (C1 is the reference category).

	Lyon			Lille			Marseille			Paris		
	Categories	β Crd	β Adj	β^a Crd	β^b Adj	β Crd	β Adj	β Crd	β Adj	β Crd	β Adj	
% unemployment [†]	C2	0.005	--	0.072 *	0.010	0.027	--	-0.009	-0.093 *			
	C3	0.005	--									
% single parent	C2	0.008 *	0.006 *	0.070	-0.007	0.029	--	0.031 *	--			
	C3	0.001	-0.001									
% insecure job	C2	0.002	--	0.018	-0.087	0.119	--	0.066 *	--			
	C3	0.003	--									
% white collar jobs	C2	-0.003	--	-0.020	-0.004	-0.063 *	-0.056 *	-0.016 *	--			
	C3	-0.008	--									
% higher education		-0.005	--	-0.032	0.081	-0.135	--	-0.077	--			
	C3	-0.005	--									
% non-homeownership	C2	0.007	--	0.020	0.007	0.009	--	0.011 *	0.012 *			
	C3	0.003	--									
% subsidized housing	C2	0.003	--	0.021 *	0.012	0.002	-0.009	0.003	--			
	C3	-0.0002	--									
% Immigrants	C2	0.007 *	0.064 *	0.15 *	0.141 *	-0.002	-0.069 *	0.023	0.045 *			
	C3	0.006 *	0.067 *									
Median Income ‡	C2	38.4	--	-0.056	-0.005	-0.123 *	-0.130 *	-0.039 *	-0.039 *			
(euro)	C3	-27.8	--									

β^a denotes the simple regression coefficients adjusting for location, β^b denotes the multivariate regression coefficient adjusting for location

† For an increase of 1 unit of the Socioeconomic and demographic variables, the NO₂ increases by β .

‡ For an increase of 1000 euros of median income, the NO₂ increases by β .

* Anova test is significant $p < 0.05$

Table 3. B) 2006-2009 period

Categories	Lyon			Lille			Marseille			Paris		
	β Crd	β Adj		β^a Crd	β^b Adj		β Crd	β Adj		β Crd	β Adj	
% unemployment [†]	0.003	--	continuous	0.038	--	continuous	-0.005	-0.051	0.037	--	--	--
% single parent	0.006	--	continuous	0.029	--	continuous	-0.035	-0.094 *	-0.005	-0.093 *	--	--
% insecure job	0.006 *	0.005 *	continuous	0.029	--	continuous	-0.035	-0.094 *	-0.005	-0.093 *	--	--
	-0.0007	-0.005	continuous	-0.036	-0.049 *	continuous	-0.042 *	-0.061 *	-0.007	--	--	--
% white collar jobs	0.008	--	continuous	0.007	--	continuous	-0.104 *	-0.090 *	-0.023 *	--	--	--
% higher education	-0.002	--	continuous	-0.048	--	continuous	-0.083	--	0.071	--	--	--
	-0.004	--	continuous			continuous						
	-0.003	--	continuous			continuous						
	-0.002	--	continuous			continuous						
% non-homeownership	0.009 *	0.011 *	continuous	0.014	--	continuous	0.007	--	0.012 *	0.013 *	--	--
	0.005 *	0.009 *	continuous			continuous						
% subsidized housing	0.006 *	--	continuous	0.014	--	continuous	0.003	--	0.005	--	--	--
	0.004	--	continuous			continuous						
% Immigrants	0.003	--	continuous	0.10 *	0.116 *	continuous	0.012	--	0.034 *	--	--	--
	0.005	--	continuous			continuous						
Median Income ‡	-51.9	--	continuous	-0.07	--	continuous	-0.125	-0.236 *	-0.065 *	-0.076 *	--	--
(euro)	-38.5	--	continuous			continuous						

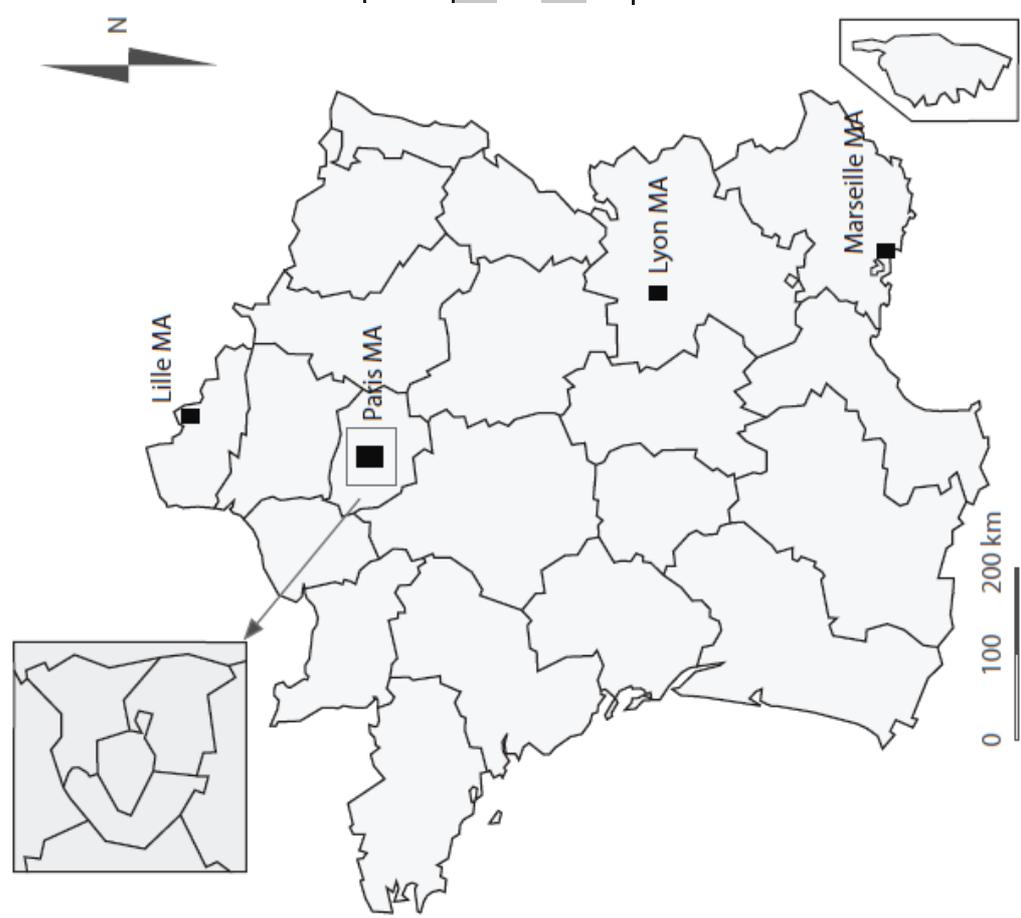
β^a denotes the simple regression coefficients adjusting for location, β^b denotes the multivariate regression coefficient adjusting for location

† For an increase of 1 unit of the Socioeconomic and demographic variables, the No_2 increases by β .

‡ For an increase of 1000 euros of median income, the No_2 increases by β .

* Anova test is significant $p < 0.05$

Figure 1



	Number of Census Blocks	Number of municipalities	Population in 2008 (inhabitants)
Lille	504	85	1.193.244
Paris	2749	412	10.354.675
Lyon	511	58	1.281.971
Marseille	628	52	1.715.096

Figure 1 : Location of the four metropolitan areas in France.

Figure 2

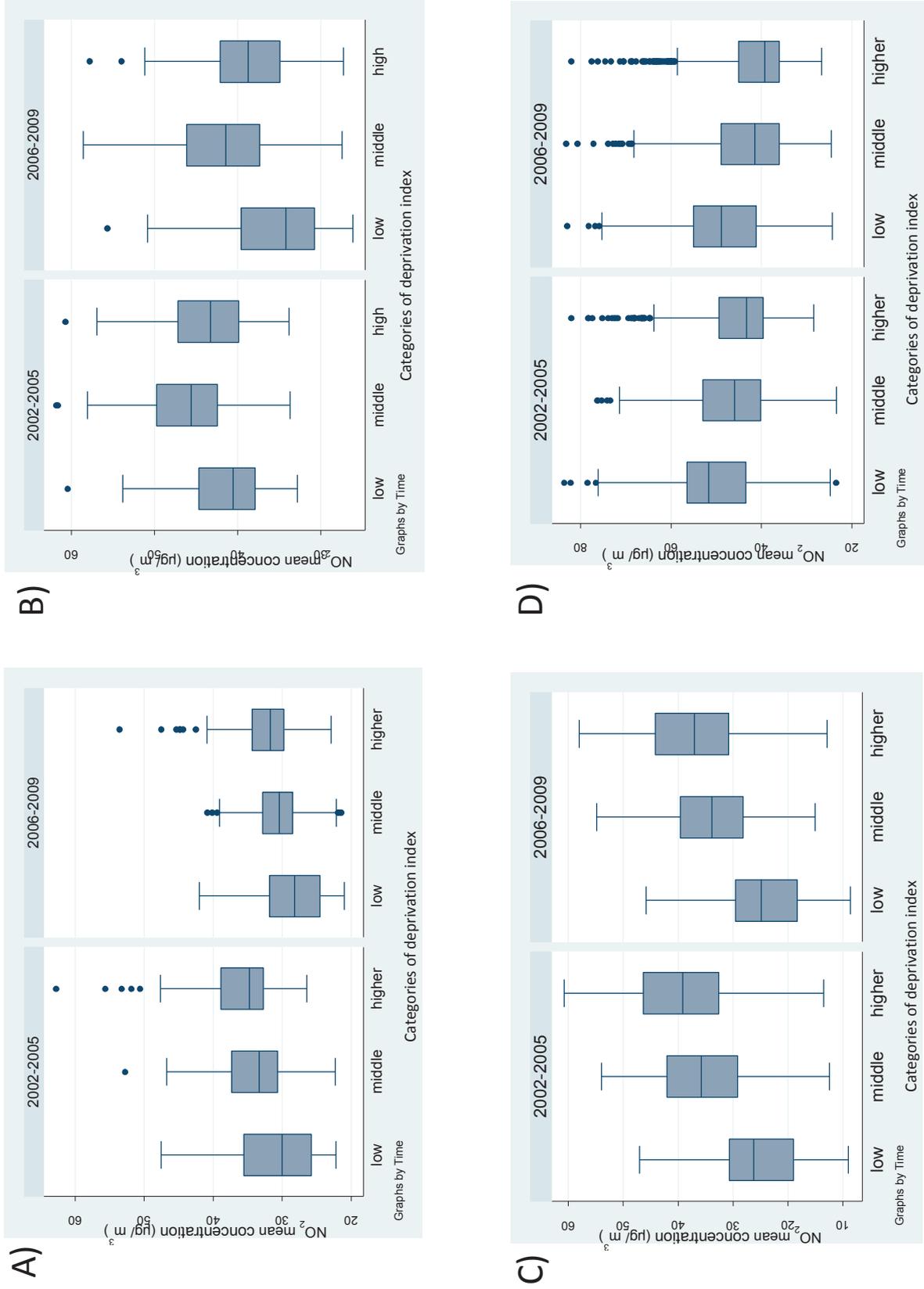


Figure 2 : Evolution of the relationship between the deprivation index and NO₂ mean concentration during the two periods for A) Lille, B) Lyon, C) Marseille, D) Paris MA.

Figure 3

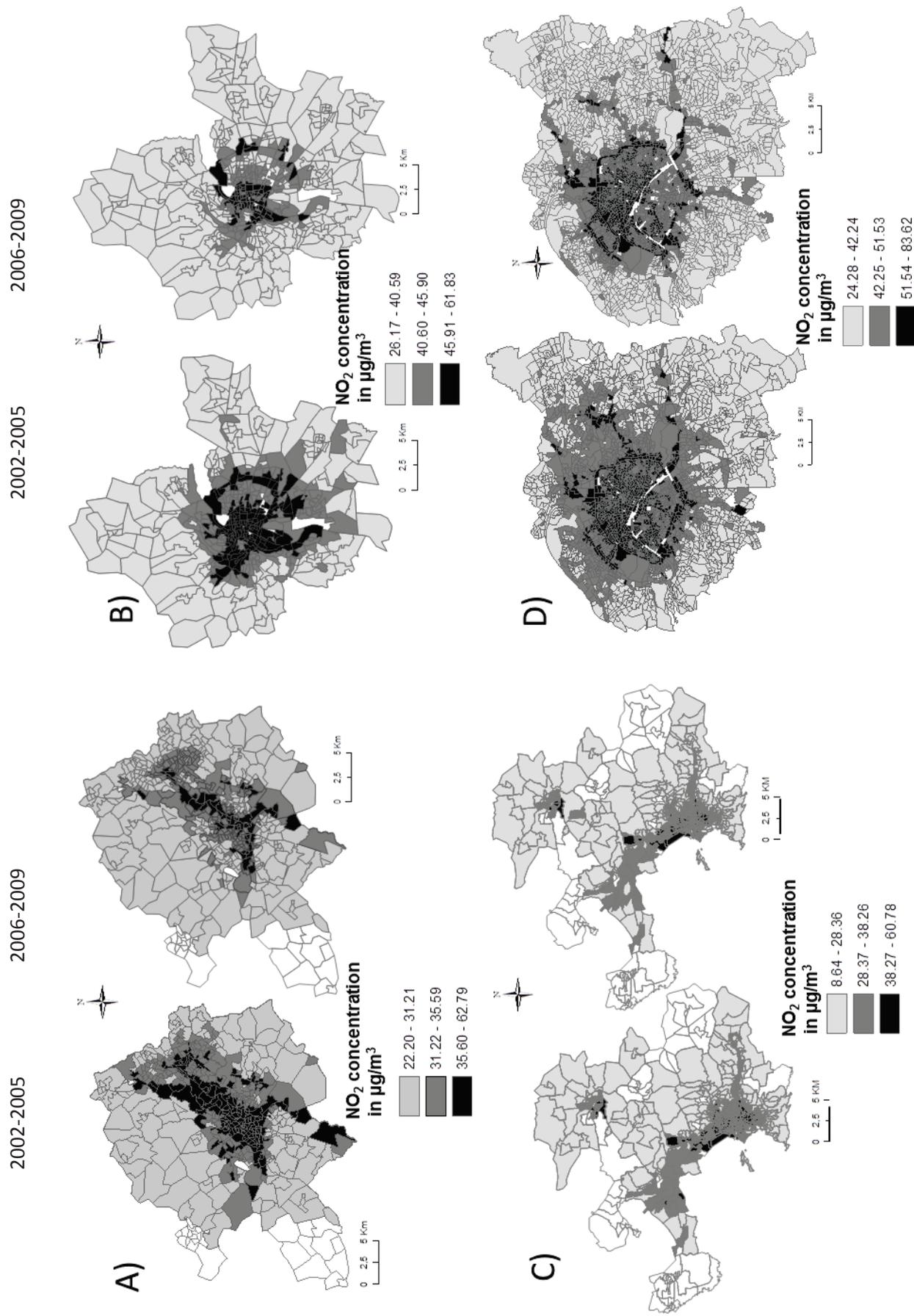


Figure 3. Maps of the spatial distribution of the average of NO₂ concentrations in µg/m³ in tertiles for 2002-2005 and 2006-2009 for A) Lille, B) Lyon, C) Marseille, D) Paris Mias.

Figure 4

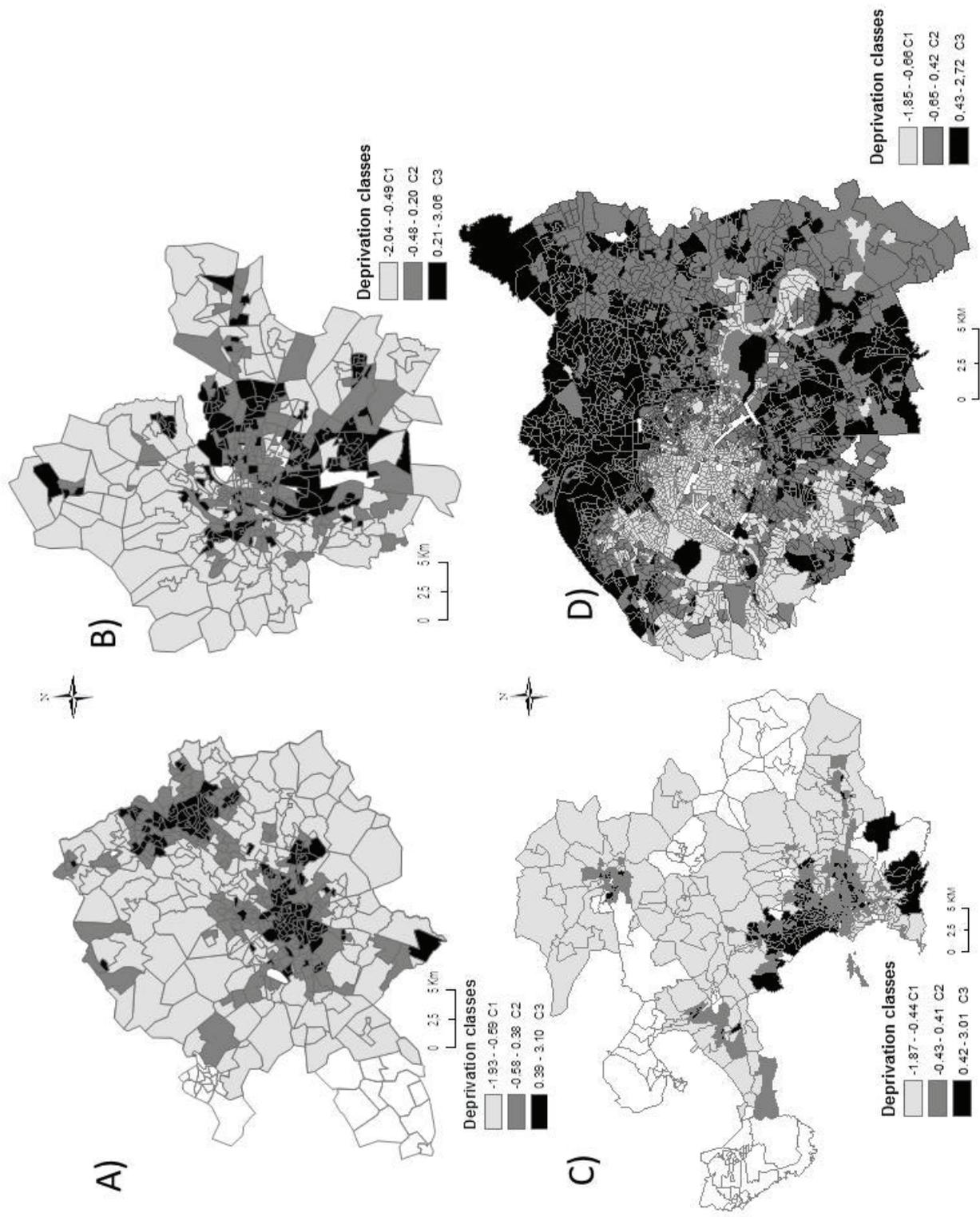


Figure 4. Spatial distribution of the deprivation index of 2006 in tertiles: A) Lille, B) Lyon, C) Marseille and D) Paris MAS

C. Discussion

Le concept d'inégalités environnementales est relativement récent en Europe et retient aujourd'hui l'attention des pouvoirs publics. Il paraît important de prendre en compte l'évolution temporelle des expositions environnementales pour investiguer cette problématique, particulièrement lorsqu'on étudie l'exposition des populations à la pollution atmosphérique. Nos analyses révèlent une distribution inégale de l'exposition au dioxyde d'azote selon le profil social du quartier de résidence. De plus, nous avons constaté que la force et la direction de l'association entre l'exposition et la défaveur du quartier de résidence variaient entre les agglomérations, mettant ainsi en évidence des spécificités locales et soulignant la difficulté de généraliser les observations issues de ce travail sur les inégalités environnementales dans tout le territoire français. Cette étude ajoute un élément nouveau concernant l'évolution temporelle des inégalités environnementales montrant l'intérêt d'une mise en place d'un suivi dans le temps de ces inégalités comme outil d'évaluation des politiques sociales et/ou d'aménagements urbains. Elle a également permis d'approfondir les déterminants sociaux qui conduisent aux renforcements des inégalités environnementales sur lesquelles il faudrait agir en priorité.

IV. **ETUDE DE L'ÉVOLUTION TEMPORELLE DU RÔLE DU NO₂ DANS LES INÉGALITÉS SOCIO-SPATIALES DE SANTÉ**

Nous savons qu'en France entre 2002 et 2009, la mortalité infantile et la mortalité néonatale ont connu deux évolutions différentes successives : une réduction entre 2002-2005 et une stagnation entre 2006-2009 (voir Chapitre 1, Partie I). Suite aux constats observés dans la précédente partie (Partie III : évolution des inégalités environnementales), nous nous sommes interrogés sur l'évolution spatio-temporelle des inégalités sociales de santé.

A. Contexte

En Europe, la littérature a démontré une évolution dans le temps des inégalités sociales de santé liées à la mortalité infantile et néonatale (50-55). Avant les années 2000, plusieurs études démontraient que les inégalités sociales en termes de mortalité infantile et post-néonatale persistaient voir augmentaient depuis plusieurs décennies (50,52,54). Les résultats sont cependant plus controversés selon les pays pour la mortalité néonatale. Dans les pays nordiques, la réduction de la mortalité néonatale semble moins sensible au groupe de défaveur (50,55). En Angleterre, une étude récente montre que le taux de mortalité néonatale était plus de deux fois plus élevé dans les régions les plus défavorisées que dans les moins défavorisées ; cette étude ajoute de plus que l'écart entre les plus défavorisés et les plus favorisés s'est creusé sur la période 1997 à 2007(52). Des études ont démontré également que la baisse des taux de mortalité infantile et post-néonatale était plus importante et plus rapide pour les groupes les plus favorisés que pour les populations les plus défavorisées (51,55). Cette persistance des inégalités peut refléter une polarisation croissante entre les groupes socio-économique en termes de conditions matérielles, sociales mais aussi attitudes et accès aux soins (55).

Si l'existence d'inégalités sociales de mortalité infantile est maintenant bien établie, il en va autrement de leur évolution temporelle. Deux études seulement ont ouvert la voie à cette préoccupation (7,56). Les deux retracent l'évolution différente de la mortalité infantile, néonatale et post-néonatale selon les catégories socio-professionnelles individuelles des deux parents, puis selon les départements pour la première (56) et les régions de France pour la

seconde (7). Ces études ont conclu que les inégalités sociales du risque de mortalité infantile et de néonatale se sont réduites avec une baisse plus importante observée pour la mortalité néonatale.

Dans ce contexte, nous nous proposons d'explorer les évolutions en termes d'inégalités sociales en France pendant la période de 2002 et 2009. Cette dernière partie répondra aux objectifs spécifiques suivants :

- i. Décrire le nombre de cas ainsi que les taux de mortalité par période et par agglomération (Tableau 8) et décrire la différence, en termes de taux de mortalité, entre des zones opposées de défaveur entre les deux périodes (Tableau 9).
- ii. Analyser la distribution spatiale de la mortalité infantile et néonatale selon la défaveur socio-économique du quartier de résidence et l'exposition au NO₂ et investiguer comment cette répartition a pu évoluer entre les deux périodes d'études (Tableau 10 et Figures 14-19).

L'étude a porté sur les agglomérations de Lille, Lyon et la ville de Paris. Deux périodes distinctes ont été définies. Le choix de ces deux périodes découle de la double évolution des taux de mortalité (réduction sur la période 2002-2005 et stagnation sur la période 2006-2009) évoquée dans le contexte de cette partie. La première période et la deuxième période couvrent respectivement les années 2002-2005 et 2006-2009. A Paris, les données disponibles, nous obligent à partager l'étude de la façon suivante : 2004-2006 et 2007-2009. L'indice de défaveur socio-économique a été estimé à partir des données du recensement de 1999 (pour la première période) et de 2006 (pour la seconde période). La moyenne des concentrations en dioxyde d'azote a été estimée pour chacune des deux périodes. Comme pour la partie II, nous avons utilisé les modèles spatiaux de type GAM afin d'estimer les risques de mortalité infantile ou néonatale à l'échelle de l'IRIS en ajustant sur la défaveur socio-économique et l'exposition au NO₂. Ces modèles prennent en compte l'autocorrélation spatiale en incluant un paramètre de localisation précise de l'unité géographique d'étude (plus précisément, ici, il s'agit des coordonnées X, Y du centroïde du quartier) et le faible nombre d'évènements sanitaires (57).

Le plan d'analyse statistique suivant a été déroulé : nous avons tout d'abord effectué des modèles sans ajustement afin de déterminer les IRIS à risque de mortalité plus élevé. Ensuite nous avons ajusté nos modèles sur la défaveur socio-économique et l'exposition au NO₂ successivement afin de déterminer si ces facteurs pouvaient expliquer ces inégales répartitions géographiques du risque de mortalité. Le schéma d'interprétation des résultats issus de ces modèles se décline en deux étapes :

1. Un premier test appelé « Test global de la déviance » va permettre de savoir s'il existe des disparités géographiques du risque de mortalité infantile / néonatale par quartier. Hypothèse Ho : homogénéité du territoire / H1 : existence de zones de risques plus élevé statistiquement que les autres ou existence de cluster de risque de mortalité infantile ou néonatale.
2. Si la p-valeur est significative, un test appelé « Test de permutation » va déterminer précisément les IRIS qui ont un risque statistiquement plus élevé que les autres. Hypothèse Ho : pas d'association entre le risque de mortalité et la localisation géographique des autres IRIS.

B. Résultats

Evolutions des inégalités sociales de santé

Le tableau 8 présente le nombre de cas ainsi que les taux de mortalité infantile et néonatale sur les deux périodes d'étude. Alors que l'agglomération de Lille et la ville de Paris voient leur nombre de décès infantiles diminuer (-18% -11,1% pour Lille et Paris, respectivement) et néonatale (-20%,-8,3%), Lyon a réduit faiblement le nombre de décès néonataux (-2,4%) et a de même augmenté faiblement le nombre de décès infantiles (+1,6%).

Le tableau 9 présente l'évolution de l'inégalité sociale liée à la mortalité infantile et néonatale dans les 3 agglomérations. Les résultats mettent en évidence que cette réduction des taux de mortalité a profité davantage aux quartiers défavorisés à Lille et Paris. En effet, le taux de mortalité néonatale par exemple diminue de 15,9% et 9,8% entre les deux périodes dans les quartiers plus défavorisés de Lille et Paris, respectivement alors qu'il a fortement augmenté dans les quartiers favorisés (37% et 36,4% à Lille et Paris, respectivement), partant de

niveaux plus bas. A Lyon, les inégalités sociales pour la mortalité infantile et néonatale se sont amplifiées entre les deux périodes. on observe une augmentation plus importante des taux de mortalité infantile et néonatale dans les quartiers défavorisés (+19,5%, +23,6%), pour la mortalité infantile et néonatale, respectivement), par rapport aux quartiers favorisés (+16,7%, +12,9%)

Evolutions des disparités géographiques

A Lille, les cartes brutes révèlent deux clusters de mortalité infantile et un petit cluster de mortalité néonatale à la limite de la significativité en première période (test global $p < 0,001$ et $p = 0,04$, Tableau 10). Alors que le cluster de mortalité néonatale localisé dans la commune de Lille va disparaître en deuxième période (test global $p = 0,36$ Figure 14), seul le cluster de mortalité infantile situé à Roubaix se maintient entre les deux (test global $p = 0,003$, Figure 15). Après ajustement sur la défaveur socio-économique, le cluster de mortalité infantile disparaît, ce qui signifie que la défaveur socio-économique des quartiers de la ville de Roubaix explique significativement les sur-risques de mortalité infantile.

Autre schéma, à Paris, on constate un gradient du risque de mortalité infantile visible du Sud-Ouest au Nord-Est en première période et du Sud au Nord en deuxième période. De plus, le cluster de risque de mortalité infantile significativement plus élevé au Nord Est de Paris en première période (19^{ème} et 20^{ème} arrondissements), va évoluer plus au Nord en deuxième période intégrant les 17^{ème}, 18^{ème} et 19^{ème} arrondissements (Figure 16). La défaveur socio-économique caractérise ces arrondissements dont les clusters disparaîtront après ajustement sur la défaveur socio-économique. On ne constate aucun cluster de risque néonatal pour les deux périodes.

A contrario, à Lyon, les clusters de risque de mortalité infantile et néonatal apparaissent seulement en deuxième période avec un gradient du risque de l'Est à l'Ouest. Ils sont situés à l'Est et comprennent les villes de Vaux en Velin, Decines-Charpieu, Chassieu (Figures 17 et 18). Après ajustement sur la défaveur socio-économique, et l'exposition au NO₂ ensemble ou combiné (modèle avec interaction), la répartition du risque de mortalité infantile devient statistiquement homogène sur le territoire (test global $p = 0,11$, Tableau 10). A contrario, la répartition du risque de mortalité néonatale est expliquée par la défaveur socio-économique ($p = 0,126$) et en partie par l'exposition au NO₂ ($p = 0,06$).

Tableau 8 : Evolution du taux de mortalité infantile et néonatale entre les périodes 2002-2005 et 2006-2009 pour les agglomérations de Lille et Lyon et les périodes 2004-2006 et 2007- 2009 pour la ville de Paris.

Agglomérations	Mortalité infantile			Mortalité Néonatale		
	2002-2005	2006-2009	Evolution	2002-2005	2006-2009	Evolution
	Nombre de cas, (Taux de mortalité)	Nombre de cas, (Taux de mortalité)	Mortalité infantile	Nombre de cas, (Taux de mortalité)	Nombre de cas, (Taux de mortalité)	Mortalité Néonatale
Lille	294 (4,55‰)	241 (3,69‰)	-18,03%	209 (3,23‰)	167 (2,55‰)	-20,09%
Lyon	297 (3,95‰)	302 (3,97‰)	1,62%	212 (2,82‰)	207 (2,72‰)	-2,36%
Paris	333 (3,52‰)	296 (3,23‰)	-11,11%	241 (2,54‰)	221 (2,41‰)	-8,29%

*((Nbre de Cas en P2 – Nbre de Cas en P1)/ Nbre de Cas en P1) *100,

Tableau 9 : Evolution du taux de mortalité infantile et néonatale entre les périodes 2002-2005 et 2006-2009 selon la défaveur socio-économique du quartier de résidence pour les 3 agglomérations,

	Mortalité infantile						Mortalité Néonatale					
	Iris défavorisés †		IRIS favorisés‡		Evolution		Iris défavorisés †		IRIS favorisés‡		Evolution	
	2002-2005	2006-2009	2002-2005	2006-2009	2002-2005	2006-2009	2002-2005	2006-2009	2002-2005	2006-2009	2002-2005	2006-2009
	Nombre de cas, (Taux de mortalité)	Evolution défavorisés	Nombre de cas, (Taux de mortalité)	Evolution favorisés	Nombre de cas, (Taux de mortalité)	Evolution défavorisés	Nombre de cas, (Taux de mortalité)	Evolution défavorisés	Nombre de cas, (Taux de mortalité)	Evolution favorisés	Nombre de cas, (Taux de mortalité)	Evolution favorisés
Lille	149 (5,85‰)	135 (5,26‰)	-9,39%	41 (3,04‰)	47 (2,93‰)	14,63%	107 (4,20‰)	90 (3,50‰)	-15,88%	27 (1,74‰)	37 (2,30‰)	37,04%
Lyon	133 (4,53‰)	159 (5,38‰)	19,55%	48 (2,61‰)	56 (2,98‰)	16,67%	89 (3,04‰)	110 (3,72‰)	23,59%	31 (1,68‰)	35 (1,86‰)	12,90%
Paris	156 (4,12‰)	146 (3,99‰)	-6,41%	61 (2,46‰)	75 (3,13‰)	22,95%	112 (2,96‰)	101 (2,76‰)	-9,82%	44 (1,77‰)	60 (2,50‰)	36,36%

† Populations résidants dans les quartiers issus du troisième tertile de défaveur,

‡ Populations résidants dans les quartiers issus du premier tertile de défaveur,

*((Nbre de Cas en P2 – Nbre de Cas en P1)/ Nbre de Cas en P1) *100

Tableau 10 : Caractéristiques de chaque modèle spatial d'estimation du risque A) de mortalité infantile pour les agglomérations de Lille, Lyon et Paris-ville B) de mortalité néonatale pour les agglomérations de Lille, Lyon et Paris-ville. P valeur de la déviance du test global dont hypothèse Ho est l'homogénéité spatiale du risque de mortalité infantile ou néonatale.

A)	Lille MA		Lyon MA		Paris MA	
	Span	Deviance p value	Span	Deviance p value	Span	Deviance p value
		Global		Global		Global
Première période : 2002-2004			2004-2006			
Modèle brut	0,45	0,006	0,65	0,155	0,95	0,018
Ajustement sur le SES	0,95	0,519	0,95	0,368	0,95	0,394
Ajustement sur le NO2	0,45	0,001	0,55	0,128	0,95	0,027
Deuxième période : 2006-2009			2007-2009			
Modèle brut	0,75	0,003	0,65	<0,001	0,90	0,019
Ajustement sur le SES	0,85	0,173	0,95	0,024	0,95	0,163
Ajustement sur le NO2	0,80	0,036	0,90	0,004	0,95	0,004

B)	Lille MA		Lyon MA		Paris MA	
	Span	Deviance p value	Span	Deviance p value	Span	Deviance p value
		Global		Global		Global
Première période : 2002-2004			2004-2006			
Modèle brut	0,40	0,042	0,95	0,357	0,95	0,062
Ajustement sur le SES	0,95	0,856	0,95	0,494	0,95	0,466
Ajustement sur le NO2	0,40	0,136	0,95	0,348	0,95	0,077
Deuxième période : 2006-2009			2007-2009			
Modèle brut	0,95	0,357	0,95	0,028	0,95	0,747
Ajustement sur le SES	0,95	0,679	0,95	0,126	0,95	0,953
Ajustement sur le NO2	0,95	0,738	0,95	0,060	0,95	0,777

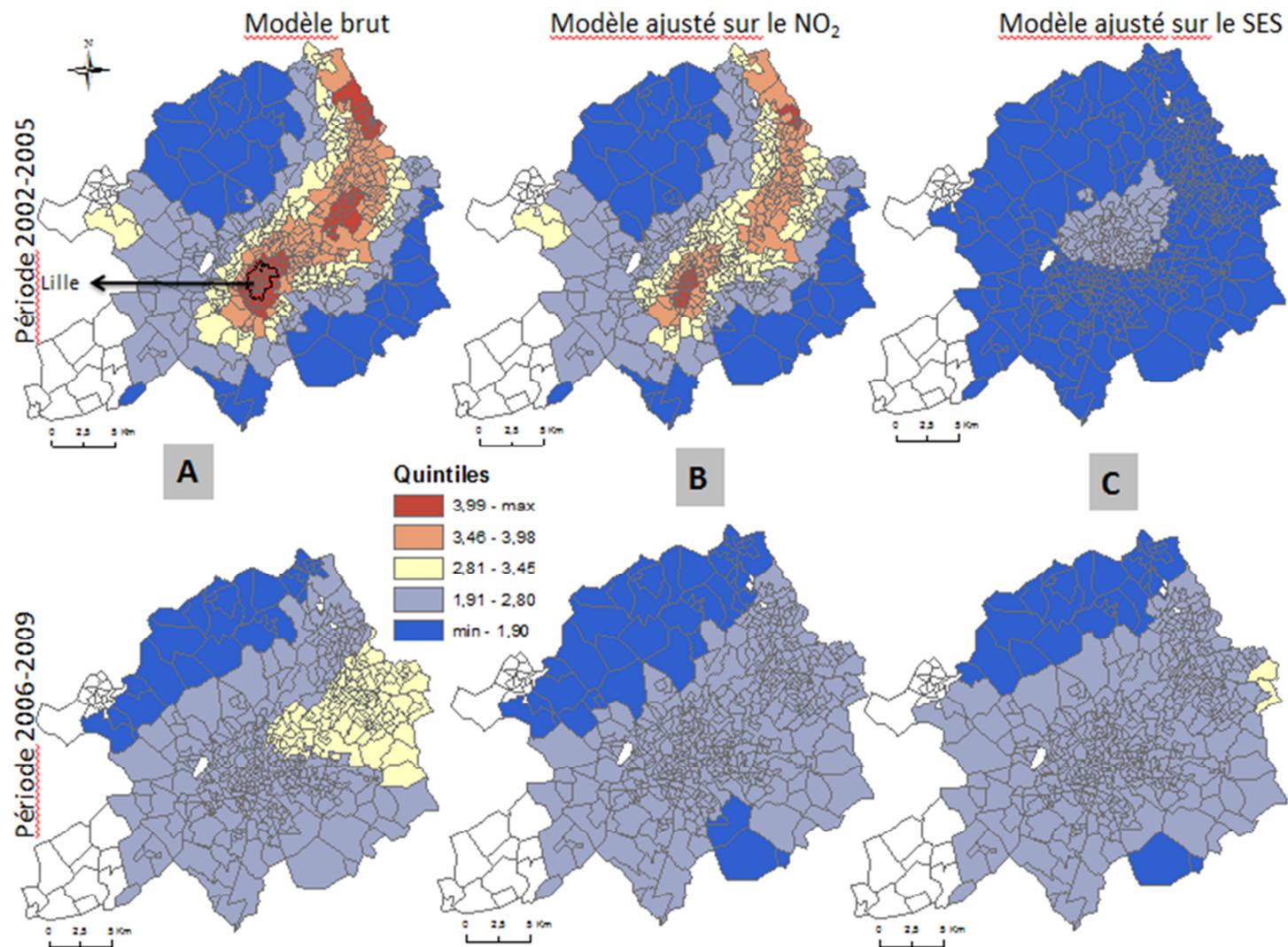


Figure 14 : Prévalence de la mortalité néonatale estimée par le modèle GAM pour l’agglomération de Lille, issu du modèle Brut A), ajusté sur l’exposition au NO₂ B), ajusté sur la défaveur sociale C) sur les deux périodes d’études. La ligne noire et solide identifie les IRIS à risque significativement plus élevés de mortalité néonatale.

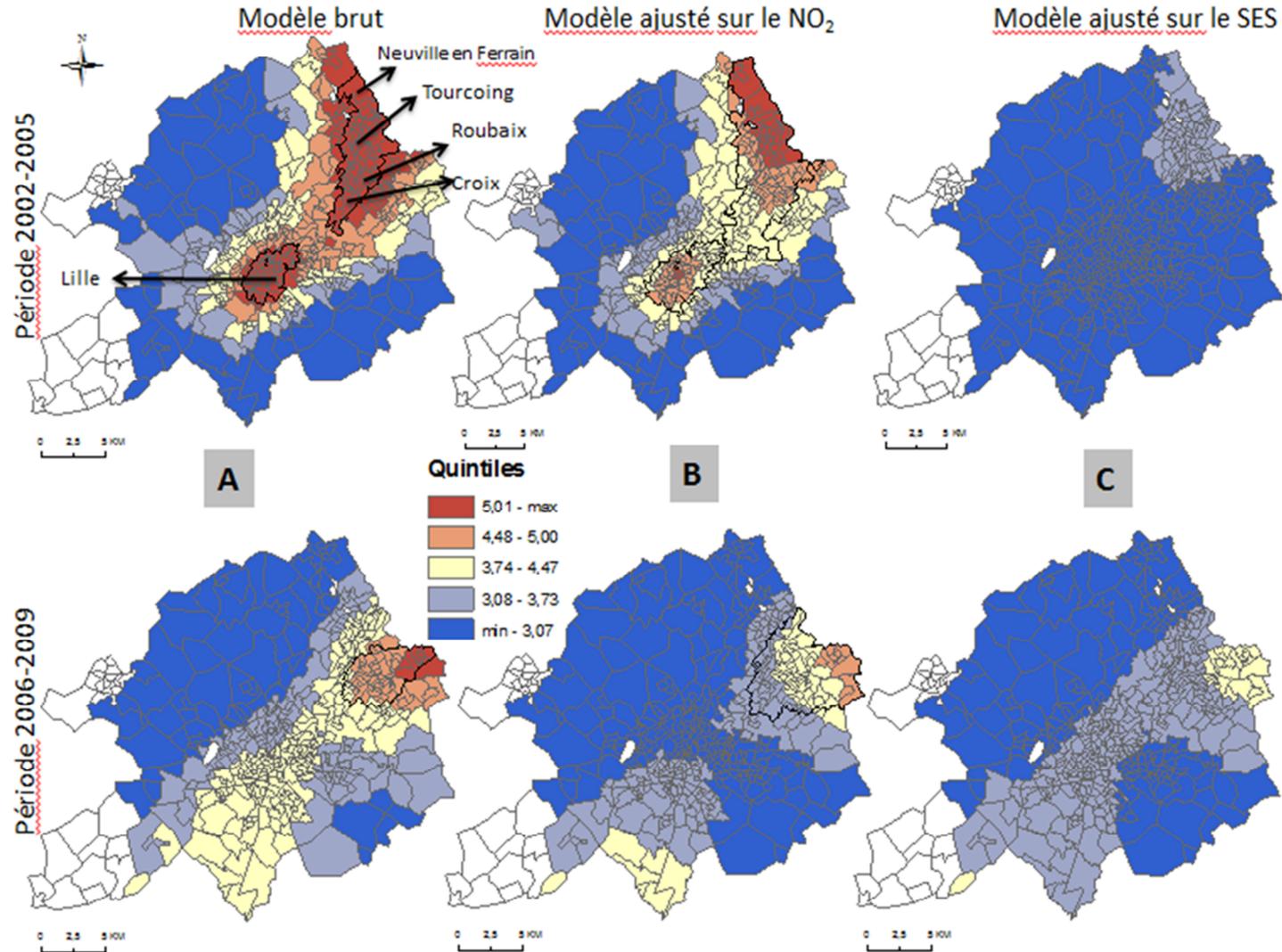


Figure 15 : Prévalence de la mortalité infantile estimée par le modèle GAM pour l'agglomération de Lille, issu du modèle Brut A), ajusté sur l'exposition au NO₂ B), ajusté sur la défaveur sociale C) sur les deux périodes d'études. La ligne noire et solide identifie les IRIS à risque significativement plus élevés de mortalité infantile.

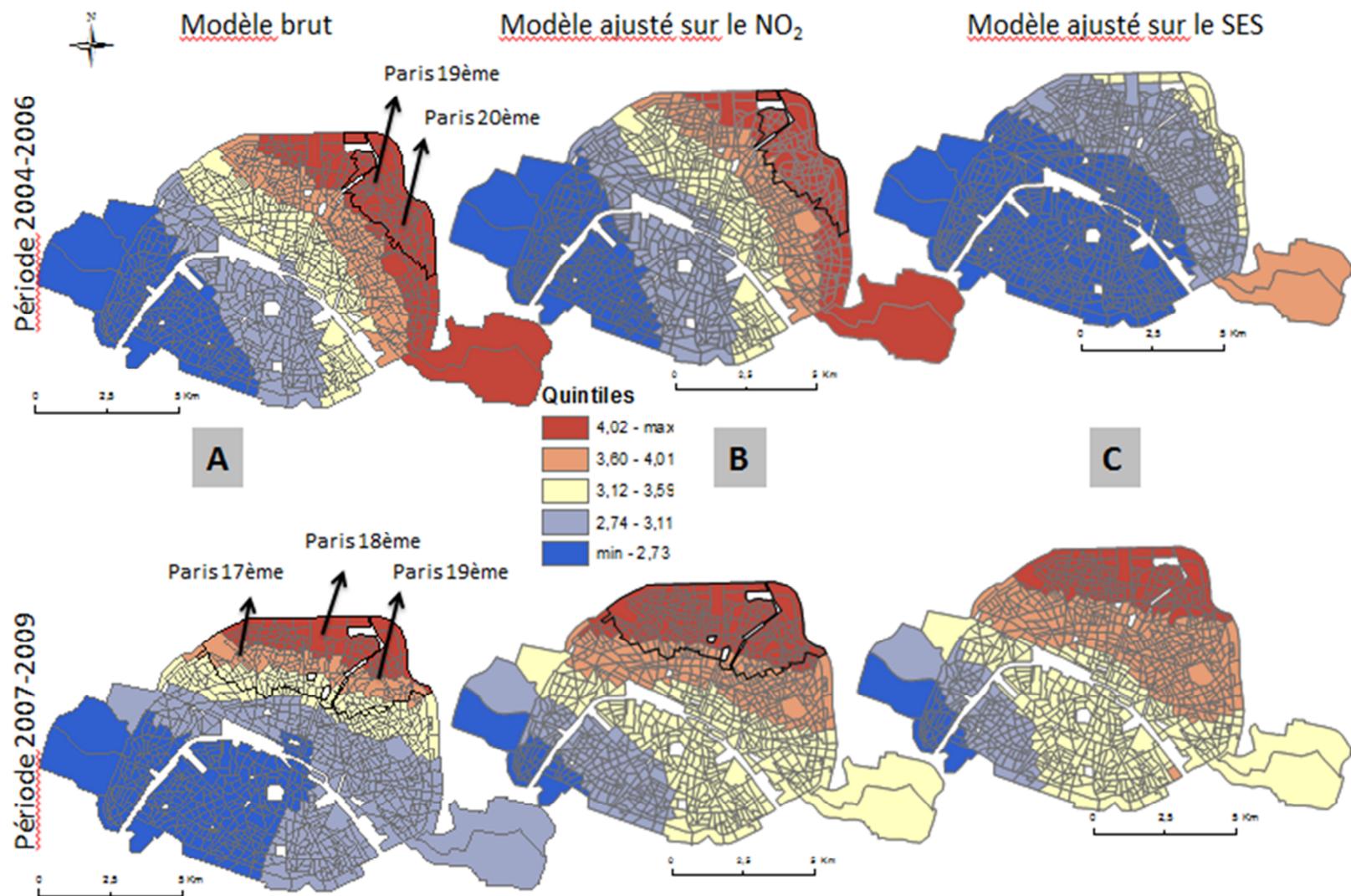


Figure 16 : Prévalence de la mortalité infantile estimée par le modèle GAM pour l'agglomération de Paris, issu du modèle Brut A), ajusté sur l'exposition au NO₂ B), ajusté sur la défaveur sociale C) sur les deux périodes d'études. La ligne noire et solide identifie les IRIS à risque significativement plus élevés de mortalité infantile.

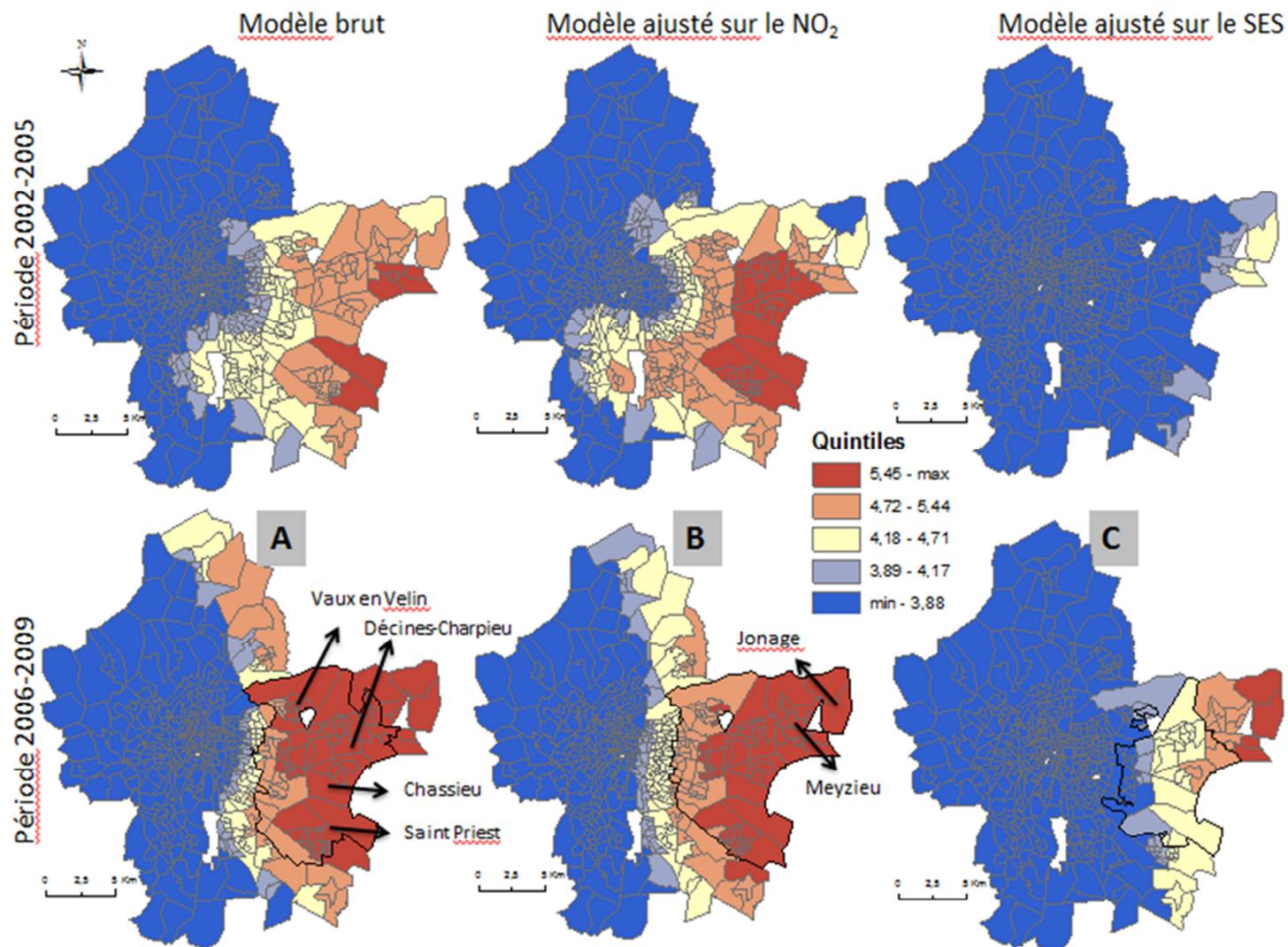


Figure 17 : Prévalence de la mortalité infantile estimée par le modèle GAM pour l'agglomération de Lyon, issu du modèle Brut A), ajusté sur l'exposition au NO₂ B), ajusté sur la défaveur sociale C) sur les deux périodes d'études. La ligne noire et solide identifie les IRIS à risque significativement plus élevés de mortalité infantile.

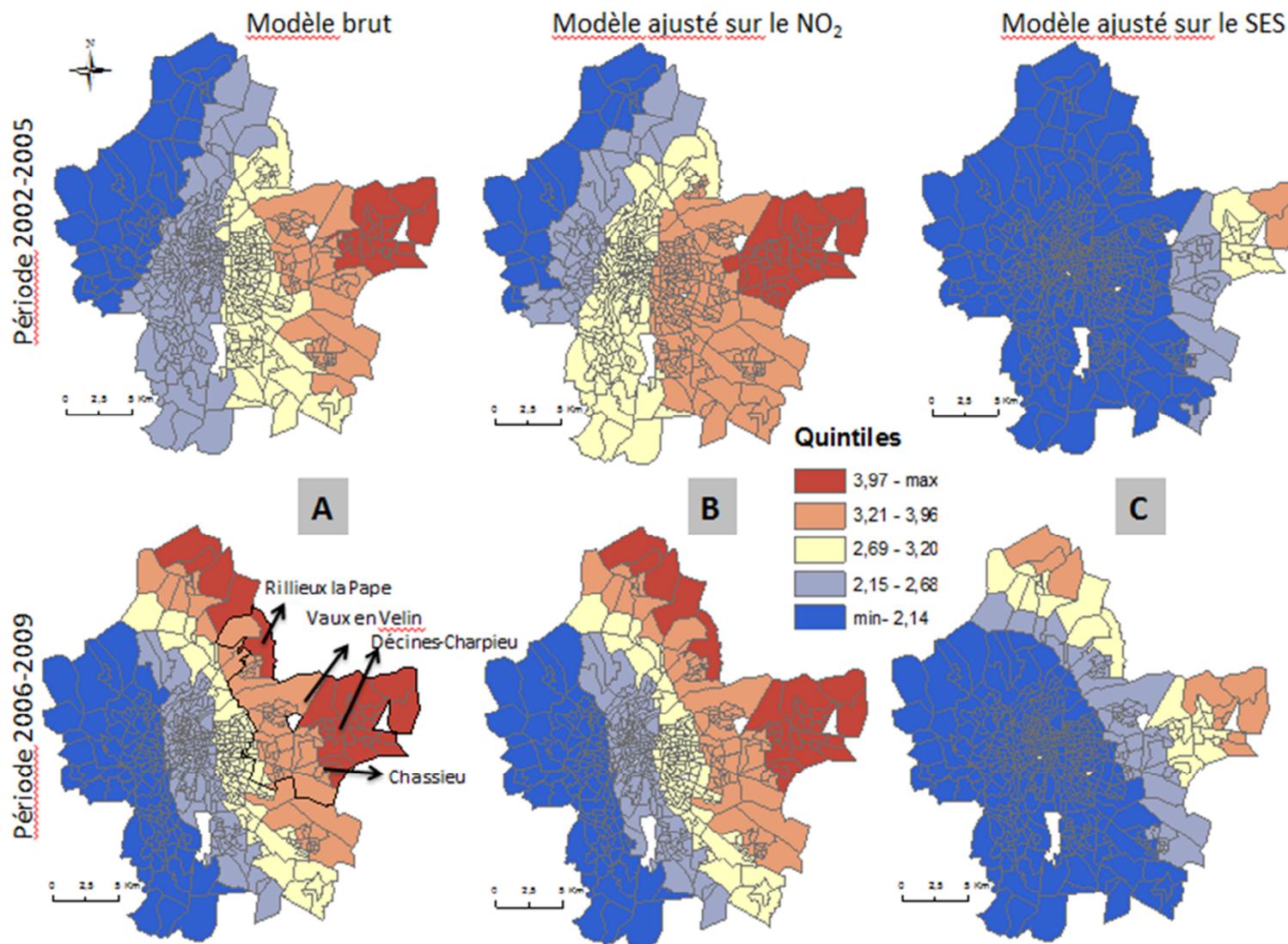


Figure 18 : Prévalence de la mortalité néonatale estimée par le modèle GAM pour l’agglomération de Lyon, issu du modèle Brut A), ajusté sur l’exposition au NO₂ B), ajusté sur la défaveur sociale C) sur les deux périodes d’études. La ligne noire et solide identifie les IRIS à risque significativement plus élevés de mortalité néonatale.

C. Discussion

Dans notre étude, l'évolution de la mortalité infantile et néonatale est sensible aux groupes de défaveur. Malgré une évolution plus importante du nombre de cas de décès néonataux, les schémas d'évolution sont les mêmes pour la mortalité infantile et néonatale. Tout d'abord, on constate que l'évolution des taux est différente selon les agglomérations. Alors que Lyon ne montre pas de progrès, Lille et Paris ont vu les taux diminuer entre les 2 périodes. De même les résultats démontrent une augmentation des inégalités sociales liés à la mortalité infantile et néonatale à Lyon, avec une augmentation plus importante dans les quartiers déjà défavorisés au début de la période d'observations. Les présentes cartes montrent ces évolutions importantes, différentes et spécifiques à chaque agglomération en termes de disparités géographiques du risque de mortalité infantile. En termes de mortalité infantile, Lille est caractérisée par une réduction du nombre de clusters de risque élevé de mortalité infantile en deuxième période, à l'inverse Lyon voit apparaître un cluster et Paris agrandit un cluster déjà présent en première période. En termes de mortalité néonatale, seule Lyon fait apparaître un cluster de risque de mortalité néonatale sensible à la défaveur socio-économique en deuxième période alors que les autres agglomérations sont caractérisées par une homogénéité du risque sur le territoire. Alors que la défaveur socio-économique semble être le déterminant le plus responsable de la variabilité spatiale de la mortalité infantile à Paris et à Lille en deuxième période, à Lyon c'est la combinaison de la défaveur et de l'exposition aux NO₂ qui explique une partie de l'hétérogénéité spatiale du risque (test global $p=0,11$). D'autres facteurs semblent nécessaires de prendre en compte dans les futures analyses des disparités géographiques du risque de mortalité infantile à Lyon.

Références : Chapitre 3

1. Shonkoff, JP, Phillips, DA. From Neurons to Neighborhoods: The Science of Early Childhood Development. Washington, D.C.: NATIONAL ACADEMY PRESS; 2000.
2. Blaxter M, Social Science Research Council (Great Britain), Great Britain. Dept. of Health and Social Security. The health of children a review of research on the place of health in cycles of disadvantage. 1981. Available from: <http://catalog.hathitrust.org/api/volumes/oclc/11260413.html>
3. Daguerre, A, Nativel, C. Les maternités précoces dans les pays développés : problèmes, dispositifs, enjeux politiques. CAF; 2004 p. 51. Report No.: 53.
4. Zeitlin J, Combiér E, Levaillant M, Lasbeur L, Pilkington H, Charreire H, et al. Neighbourhood socio-economic characteristics and the risk of preterm birth for migrant and non-migrant women: a study in a French district. *Paediatr Perinat Epidemiol*. 2011 Jul;25(4):347–56.
5. Blondel, B, Kermarrec, M. Les naissances en 2010 et leur évolution depuis 2003. Enquête nationale périnatale 2010. Paris: INSERM, DGS, DREES; 2011 Mai p. 117.
6. Laurent, E. Vers l'égalité des territoires. Dynamiques, mesures et politiques. France; 2013 Février p. 534 pages.
7. Quang Chi Dinh. Les inégalités sociales de la mortalité infantile s'estompent. 1998;314(1):89–106.
8. Leclerc, A, Kaminski, M, Lang, T. Inégaux face à la santé - Du constat à l'action. INSERM; 2008.
9. Moleux, M, Schaetzel, F, Scotton, C. Les inégalités sociales de santé : Déterminants sociaux et modèles d'action. IGAS, Inspection générale des affaires sociales; 2011 mai. Report No.: N°RM2011-061P.
10. Haut conseil de la santé publique. Les inégalités sociales de santé : sortir de la fatalité. HCSP; 2009.
11. Laumbach RJ. Outdoor air pollutants and patient health. *Am Fam Physician*. 2010 Jan 15;81(2):175–80.
12. Selevan SG, Kimmel CA, Mendola P. Identifying critical windows of exposure for children's health. *Environ Health Perspect*. 2000 Jun;108 Suppl 3:451–5.
13. Ritz B, Wilhelm M. Ambient Air Pollution and Adverse Birth Outcomes: Methodologic Issues in an Emerging Field. *Basic Clin Pharmacol Toxicol*. 2008 Feb 1;102(2):182–90.
14. Shah PS, Balkhair T. Air pollution and birth outcomes: A systematic review. *Environ Int*. 2011 Feb;37(2):498–516.
15. Walters DM, Breyse PN, Wills-Karp M. Ambient urban Baltimore particulate-induced airway hyperresponsiveness and inflammation in mice. *Am J Respir Crit Care Med*. 2001 Oct 15;164(8 Pt 1):1438–43.

16. Rahmalia A, Giorgis-Allemand L, Lepeule J, Philippat C, Galineau J, Hulin A, et al. Pregnancy exposure to atmospheric pollutants and placental weight: An approach relying on a dispersion model. *Environ Int.* 2012 Nov 1;48C:47–55.
17. Slama R, Darrow L, Parker J, Woodruff TJ, Strickland M, Nieuwenhuijsen M, et al. Meeting report: atmospheric pollution and human reproduction. *Environ Health Perspect.* 2008 Jun;116(6):791–8.
18. Slama R, Thiebaugeorges O, Goua V, Aussel L, Sacco P, Bohet A, et al. Maternal personal exposure to airborne benzene and intrauterine growth. *Environ Health Perspect.* 2009 Aug;117(8):1313–21.
19. Tabacova S, Baird DD, Balabaeva L. Exposure to oxidized nitrogen: lipid peroxidation and neonatal health risk. *Arch Environ Health.* 1998 Jun;53(3):214–21.
20. Tabacova S, Balabaeva L, Little RE. Maternal exposure to exogenous nitrogen compounds and complications of pregnancy. *Arch Environ Health.* 1997 Oct;52(5):341–7.
21. Marozienne L, Grazuleviciene R. Maternal exposure to low-level air pollution and pregnancy outcomes: a population-based study. *Environ Heal Glob Access Sci Source.* 2002 Dec 9;1(1):6.
22. Lee BE, Ha EH, Park HS, Kim YJ, Hong YC, Kim H, et al. Exposure to air pollution during different gestational phases contributes to risks of low birth weight. *Hum Reprod.* 2003 Mar 1;18(3):638–643.
23. Ponce NA, Hoggatt KJ, Wilhelm M, Ritz B. Preterm birth: the interaction of traffic-related air pollution with economic hardship in Los Angeles neighborhoods. *Am J Epidemiol.* 2005 Jul 15;162(2):140–8.
24. Brauer M, Lencar C, Tamburic L, Koehoorn M, Demers P, Karr C. A cohort study of traffic-related air pollution impacts on birth outcomes. *Environ Health Perspect.* 2008 May;116(5):680–6.
25. Darrow LA, Klein M, Flanders WD, Waller LA, Correa A, Marcus M, et al. Ambient air pollution and preterm birth: a time-series analysis. *Epidemiol Camb Mass.* 2009 Sep;20(5):689–98.
26. Stieb DM, Chen L, Eshoul M, Judek S. Ambient air pollution, birth weight and preterm birth: a systematic review and meta-analysis. *Environ Res.* 2012 Aug;117:100–11.
27. Woodruff TJ, Darrow LA, Parker JD. Air pollution and postneonatal infant mortality in the United States, 1999-2002. *Environ Health Perspect.* 2008 Jan;116(1):110–5.
28. Ritz B, Wilhelm M, Zhao Y. Air pollution and infant death in southern California, 1989-2000. *Pediatrics.* 2006 Aug;118(2):493–502.
29. Hajat S, Armstrong B, Wilkinson P, Busby A, Dolk H. Outdoor air pollution and infant mortality: analysis of daily time-series data in 10 English cities. *J Epidemiol Community Health.* 2007 Aug;61(8):719–22.

30. Scheers H, Mwalili SM, Faes C, Fierens F, Nemery B, Nawrot TS. Does air pollution trigger infant mortality in Western Europe? A case-crossover study. *Environ Health Perspect.* 2011 Jul;119(7):1017–22.
31. Romieu I, Ramírez-Aguilar M, Moreno-Macias H, Barraza-Villarreal A, Miller P, Hernández-Cadena L, et al. Infant mortality and air pollution: modifying effect by social class. *J Occup Environ Med Am Coll Occup Environ Med.* 2004 Dec;46(12):1210–6.
32. Weck RL, Paulose T, Flaws JA. Impact of environmental factors and poverty on pregnancy outcomes. *ClinObstetGynecol.* 2008 Jun;51(2):349–59.
33. Zeka A, Melly S, Schwartz J. The effects of socioeconomic status and indices of physical environment on reduced birth weight and preterm births in Eastern Massachusetts. *Environ Heal Glob Access Sci Source.* 2008;7:60.
34. Miranda ML, Maxson P, Edwards S. Environmental Contributions to Disparities in Pregnancy Outcomes. *Epidemiol Rev.* 2009 Nov 1;31(1):67–83.
35. Jerrett M, Burnett RT, Kanaroglou P, Eyles J, Finkelstein N, Giovis C, et al. A GIS - environmental justice analysis of particulate air pollution in Hamilton, Canada. *Environ Plan.* 2001;33(6):955–73.
36. Mortensen LH, Helweg-Larsen K, Andersen A-MN. Socioeconomic differences in perinatal health and disease. *Scand J Public Health.* 2011 Jul;39(7 Suppl):110–4.
37. Gray R, Bonellie SR, Chalmers J, Greer I, Jarvis S, Kurinczuk JJ, et al. Contribution of smoking during pregnancy to inequalities in stillbirth and infant death in Scotland 1994-2003: retrospective population based study using hospital maternity records. *BMJ.* 2009 Oct 1;339:b3754. doi: 10.1136/bmj.b3754.b3754.
38. O’Leary CM, Nassar N, Kurinczuk JJ, Bower C. The effect of maternal alcohol consumption on fetal growth and preterm birth. *BJOG Int J Obstet Gynaecol.* 2009 Feb;116(3):390–400.
39. Minnes S, Lang A, Singer L. Prenatal tobacco, marijuana, stimulant, and opiate exposure: outcomes and practice implications. *Addict Sci Clin Pr.* 2011 Jul;6(1):57–70.
40. Northstone K, Emmett P, Rogers I. Dietary patterns in pregnancy and associations with socio-demographic and lifestyle factors. *Eur J Clin Nutr.* 2008 Apr;62(4):471–9.
41. Macinko J, Starfield B, Shi L. The contribution of primary care systems to health outcomes within Organization for Economic Cooperation and Development (OECD) countries, 1970-1998. *Health Serv Res.* 2003 Jun;38(3):831–65.
42. Pilkington H, Blondel B, Papiernik E, Cuttini M, Charreire H, Maier RF, et al. Distribution of maternity units and spatial access to specialised care for women delivering before 32 weeks of gestation in Europe. *Health Place.* 2010 May;16(3):531–8.
43. Garlantézec R, Monfort C, Rouget F, Cordier S. Maternal occupational exposure to solvents and congenital malformations: a prospective study in the general population. *Occup Environ Med.* 2009 Jul;66(7):456–63.
44. Deuxième plan National Santé Environnement 2009-2013 (PNSE 2). France: Ministères Français; 2009 Juin p. 84.

45. LOI constitutionnelle n° 2005-205 du 1er mars 2005 relative à la Charte de l'environnement.
46. Edition de la DIV. Observatoire national des zones urbaines sensibles. 2004.
47. Laigle, L, Oehler, V. Les enjeux sociaux et environnementaux du développement urbain : la question des inégalités écologiques. Centre technique et scientifique du bâtiment (CSTB); 2004 Février.
40. Walker G. Beyond Distribution and Proximity: Exploring the Multiple Spatialities of Environmental Justice. In: Holifield R, Porter M, Executive GWC of EJS, editors. 2010 p. 23–46.
49. Kruize H, Driessen PPJ, Glasbergen P, van Egmond KND. Environmental equity and the role of public policy: experiences in the Rijnmond region. *Environ Manage.* 2007 Oct;40(4):578–95.
50. Arntzen A, Mortensen L, Schnor O, Cnattingius S, Gissler M, Andersen A-MN. Neonatal and postneonatal mortality by maternal education--a population-based study of trends in the Nordic countries, 1981-2000. *Eur J Public Health.* 2008 Jun;18(3):245–51.
51. Arntzen A, Nybo Andersen AM. Social determinants for infant mortality in the Nordic countries, 1980-2001. *Scand J Public Health.* 2004;32(5):381–9.
52. Wood AM, Pasupathy D, Pell JP, Fleming M, Smith GCS. Trends in socioeconomic inequalities in risk of sudden infant death syndrome, other causes of infant mortality, and stillbirth in Scotland: population based study. *BMJ.* 2012;344:e1552.
53. Smith LK, Manktelow BN, Draper ES, Springett A, Field DJ. Nature of socioeconomic inequalities in neonatal mortality: population based study. *BMJ.* 2010;341:c6654.
54. Sovio U, Dibden A, Koupil I. Social determinants of infant mortality in a historical Swedish cohort. *Paediatr Perinat Epidemiol.* 2012 Sep;26(5):408–20.
55. Singh GK, Kogan MD. Persistent socioeconomic disparities in infant, neonatal, and postneonatal mortality rates in the United States, 1969-2001. *Pediatrics.* 2007 Apr;119(4):e928–939.
56. Dinh, QC. Le relief social de la mortalité infantile : des écarts qui s'atténuent. *INSEE.* 1990;(245).
57. Padilla CM, Deguen S, Lalloue B, Blanchard O, Beaugard C, Troude F, et al. Cluster analysis of social and environment inequalities of infant mortality. A spatial study in small areas revealed by local disease mapping in France. *Sci Total Environ.* 2013 Jun 1;454-455:433–41.

CHAPITRE 4 :

DISCUSSION GÉNÉRALE ET PERSPECTIVES

Dans ce chapitre, nous dressons un bilan du travail accompli dans son ensemble. Nous présentons tout d'abord les originalités, les limites de l'étude, puis l'intérêt de l'approche spatiale dans la compréhension des inégalités sociales et environnementales de santé. Ce chapitre suggère ensuite des pistes de recherche qui pourraient faire suite à ce travail.

I. CONCLUSION GENERALE

A. Les mérites de l'étude

Le caractère original de ce travail repose principalement sur 3 points : une échelle d'analyse fine, la pluridisciplinarité de l'étude, et une approche spatiale innovante permettant de créer des outils d'aide à la décision.

Echelle d'analyse

La fine échelle d'analyse de ce travail permet :

- i) une compréhension plus fine des processus qui déterminent la répartition spatiale de la mortalité infantile et néonatale sur le territoire. Les analyses mises en œuvre permettent de comprendre la distribution inégale du « fardeau » environnemental selon le profil socio-économique des unités territoriales, et leur contribution sur les inégalités spatiales de santé.
- ii) comprendre ces inégalités à une échelle aussi précise alors que peu de travaux existent sur le sujet en Europe (1–3).

La pluridisciplinarité

La pluridisciplinarité combinant épidémiologie sociale (Chapitre 3, partie I), épidémiologie environnementale (Chapitre 3, partie III), biostatistique (Chapitre 2, partie V) et santé

publique en générale, mise en place dans ce travail a permis de répondre à la question des inégalités de santé sous plusieurs angles.

Approche Spatiale

L'approche spatiale est une approche innovante particulièrement dans l'étude des inégalités sociales de santé. La méthode GAM de détection de cluster, basée sur un modèle de risque spatialisé, permet de :

- i) combiner deux domaines importants : l'analyse classique des facteurs de risque, et l'analyse spatiale de la maladie ;
- ii) identifier des clusters à une échelle locale, pour mieux cibler les leviers d'action pour les professionnels ou les élus dans le cadre de la mise en place de politiques d'aménagement du territoire ou l'orientation des politiques sociales et de santé ;
- iii) présenter des cartes visant à communiquer l'information de manière efficace et directement utilisable pour la prise de décision des pouvoirs publics.

B. Les limites de l'étude

Il existe des limites à notre étude : liées au design d'étude écologique choisi (biais écologiques, inaccessibilité des données individuelles) et liées aux données d'exposition.

Limites liées au design de l'étude

Les limites propres au design de l'étude sont :

- i) les biais écologiques ; ils sont propres aux études réalisées à partir de données agrégées. Ils résultent du biais d'agrégation (les individus qui ont présenté l'effet considéré ne sont pas forcément ceux qui ont été exposés) et du « biais de spécification » ou biais de confusion écologique (4), par exemple un risque plus élevé de mortalité infantile dans une zone polluée peut être lié à un autre facteur spécifique à cette zone et non pris en compte.

Cependant, dans le cadre de ce travail nous avons travaillé à une résolution géographique qui soit la plus homogène possible en termes d'effectif de population, de caractéristiques socio-économiques, et d'utilisation de l'espace urbain. L'homogénéité des entités assure une

minimisation des biais écologiques qui permet de comparer plus facilement les associations observées à un niveau géographique à celles observées à un niveau individuel (5,6).

ii) l'inaccessibilité aux données individuelles est un facteur limitant dans l'étude des inégalités sociales de santé et la compréhension de ces mécanismes. Dans le temps imparti de cette thèse, il était impossible d'avoir accès aux données sanitaires, socio-économiques et comportementales individuelles pour les raisons suivantes.

- Dans un souci de rigueur méthodologique, nous avons contacté le CepiDC afin de recueillir ces données sanitaires individuelles disponibles sur le deuxième volet du bulletin n°7 (Bulletin en annexe) (poids à la naissance, âge de la mère, âge gestationnel, nombre de grossesse). Nous nous sommes heurtés à un problème d'identifiant afin de relier nos cas avec les caractéristiques individuelles qui leur correspondent. La profession des parents normalement disponibles sur les actes de décès étaient trop souvent absente. Nous n'avons donc pas utilisé cette variable.

- Nous n'avons pas d'informations sur les expositions environnementales individuelles qui influent sur la santé de la mère et du nouveau-né. En termes d'exposition à la pollution de l'air, une alternative l'approche écologique aurait été d'étudier l'exposition individuelle avec des capteurs portatifs. Cependant le coût et la lourdeur organisationnelle de telles études les limitent à de petits échantillons de la population et à de courtes durées de suivi. Or pour l'instant, l'évènement sanitaire analysé (mortalité infantile et néonatale) est rare et nécessite une taille de population et une durée de suivi à même de fournir la puissance statistique nécessaire aux analyses.

- De plus, nous ne disposons pas d'informations sur la santé de la mère (le diabète, l'obésité), la qualité de vie (l'exercice physique, la sédentarité, bonne qualité alimentaire, etc...)(7), les comportements à risque telle que la prise de substances toxiques (nicotine, caféine, cocaïne ou alcool (8–10), le tabagisme maternel pendant la grossesse ou le tabagisme passif (10,11) ou la régularité et la qualité des soins prénatals reçus (7).

- Enfin, il faut noter que la mobilité de la population au sein des IRIS ou entres IRIS n'a pas été prise en compte. Ce point a longtemps fait l'objet de réflexions lors de nos

travaux ; cependant les données à l'échelle de l'IRIS du recensement de 1999 (le seul disponible au début de l'étude) ne comportaient pas d'information à ce sujet. Les données du recensement de 2006 comportent quelques variables sur la mobilité résidentielle et la mobilité professionnelle, qui auraient pu être utilisées pour modéliser l'exposition, en recourant à des approches pondérées par le temps passé au travail ou au domicile ; néanmoins le recueil n'a pas été exhaustif pour ces données.

Limites liées au choix du polluant

Notre analyse de la pollution est basée sur la pollution de l'air et singulièrement le dioxyde d'azote.

i) Seul le NO₂ a été considéré dans ce travail de thèse. Nous voulions analyser les variations spatiales de la pollution de l'air entre les quartiers. Le NO₂ est connu pour être un marqueur de pollution de la circulation intra-urbaine (12,13) et pour être associé à des effets sur la santé à l'échelle locale. En outre, nous savons que le NO₂ présente une forte corrélation avec les particules ultrafines (PM₁₀, PM_{2.5}) (14). Ces particules inhalées peuvent avoir des effets néfastes sur la mortalité infantile (15–17). Au démarrage du programme Equit'Area, il n'était possible de modéliser qu'un seul polluant au vu de l'important travail de modélisation des AASQA en termes de temps et d'argent. Notre choix s'est porté sur le NO₂, néanmoins dans les prochains développements la modélisation des PM_{2.5} sera réalisée, et les concentrations en PM_{2.5} estimées à l'IRIS par les AASQAs. Ils seront intégrés au projet Equit'Area et analysés suivant le même plan d'analyse statistique développé dans cette thèse

ii) Les études sur les effets de la pollution de l'air ambiant sur la santé humaine sont compliquées par le fait que la pollution est à un mélange complexe qui varie dans l'espace et dans le temps (18). Ces expositions dépendent aussi de l'exposition professionnelle, de l'exposition à la fumée de tabac (active et passive), et de l'exposition à l'intérieur de la maison ainsi qu'au cours des déplacements (19). Pour prendre en compte le fait que les gens sont mobiles tout au long de la journée et s'exposent à un mélange de polluants qui varie selon la zone et les environnements, ce travail aurait nécessité un important recueil de données que le temps imparti pour ma thèse ne permettait pas. De même, notre indice de défaveur ainsi les zones études ne comportent que des zones urbaines. Il aurait été intéressant de comparer nos

résultats à l'aide d'un indice de défaveur et des polluants plus adaptés à la question des inégalités sociales de santé en zone rurale. Ce travail est en cours de réflexion au sein de l'équipe.

C. Apports scientifiques

Nos résultats ont permis de mettre en évidence :

- 1) l'existence d'inégalités socio-spatiales de mortalité infantile et néonatale : un gradient du risque de mortalité infantile à Lille et un gradient du risque de mortalité néonatale à Lille et Lyon.
- 2) l'existence d'inégalités environnementales : les quartiers les plus exposés à la pollution de l'air sont les quartiers les plus défavorisés à Lille et Marseille, la ville de Paris et les classes moyennes à Lyon
- 3) les facteurs déterminant le sur-risque de mortalité différent selon l'évènement sanitaire (infantile vs néonatale) et la zone géographique.
- 4) l'importance de la défaveur sociale comme déterminant du risque de mortalité infantile (Lille, Paris), et la contribution de la pollution de l'air dans les inégalités socio-spatiales du risque de mortalité néonatale (Lille).
- 5) l'intérêt de la mise en place d'un suivi dans le temps de ces inégalités comme outil d'évaluation des politiques sociales et/ou d'aménagement urbain.

L'apport clé de cette thèse est de mettre en évidence des inégalités différentes d'une agglomération à l'autre, d'un évènement sanitaire à l'autre et d'une période à l'autre.

II. **PERSPECTIVES ET AUTRES CONTRIBUTIONS DE LA RECHERCHE**

Cette thèse a été l'occasion de nombreux échanges scientifiques avec l'équipe qui ont ouvert de nouvelles pistes de réflexion originales. Certaines pistes que nous allons présenter ont déjà été explorées et certains de ces travaux sont en cours de mise en œuvre et de publications.

A. **Autres contributions de recherche**

Dans le cadre du projet Equit'Area, nous avons examiné le rôle d'autres expositions environnementales sur les inégalités socio spatiales de santé. Ces expositions environnementales (indépendantes ou combinées) peuvent avoir un effet négatif ou positif sur la santé.

1. **Expositions environnementales à « négatif effect »**

Les travaux réalisés sur les nuisances environnementales portent sur les industries polluantes et le bruit.

i) Dans un premier temps, nous nous sommes intéressés aux industries polluantes. L'objectif de l'étude est d'explorer la relation entre mortalité infantile, défaveur et proximité aux industries polluantes à Lille entre 2000 et 2009. Les industries polluantes de l'air ont été sélectionnées à partir de la base européenne des industries polluantes (EPER, European Pollutant Emission Register / Registre européen des émissions de polluants). La proximité aux industries polluantes a été analysée à l'aide de buffers⁵ et d'indicateurs tel que la présence de buffers dans l'IRIS. Malgré que la distance euclidienne des IRIS aux industries soit inversement proportionnelle et la défaveur de l'IRIS, aucune association significative n'a pu être mise en évidence avec la proximité aux industries polluantes. De ce travail, j'ai écrit un

⁵ Un buffer est une zone tampon (ici un cercle avec un rayon prédéfini) autour de l'industrie localisée sur une carte. Les populations vivant dans la zone recouverte par le buffer circulaire sont considérées comme exposées.

article publié dans *Environnement, Risques & Santé* (Article en annexe) (20), intitulé « Mortalité infantile, défaveur et proximité aux industries polluantes : une analyse spatiale conduite à fine échelle (agglomération de Lille, France) ».

ii) Dans un second temps, nous nous sommes penchés sur le bruit. L'objectif était d'étudier l'impact des caractéristiques du quartier (statut socio-économique et niveau de bruit ambiant) sur la répartition spatiale de la mortalité infantile dans l'agglomération de Lyon (seule agglomération sur laquelle nous disposons d'indicateur de bruit). A chaque IRIS a été attribué un indice socio-économique et un niveau Lden. Le Lden est un standard européen qui décrit l'exposition journalière aux nuisances sonores sur une période de 24 heures. Nos résultats montrent un impact du bruit sur la répartition spatiale de la mortalité infantile après ajustement sur les caractéristiques socio-économiques du quartier. Un article écrit par Wahida Kihal porte sur une analyse spatiale (modèle SatScan) de la relation les nuisances sonores, la défaveur et la mortalité infantile (actuellement en 2^{de} révision dans la revue *Environmental Health*) (Abstract en annexe).

2. Expositions environnementales à « positive effect »

Nous avons porté un intérêt aux expositions environnementales dites « positives » en débutant par la proximité des populations aux espaces verts et son rôle dans les inégalités sociales de santé. L'objectif a été d'étudier la relation entre les espaces verts et la distribution spatiale de la mortalité infantile en tenant compte des niveaux de privation des quartiers. Des indicateurs ont été créés tels que le nombre total d'espaces verts dans chaque IRIS, le pourcentage de la surface de l'IRIS occupé par un espace vert. Les résultats démontrent que les deux facteurs expliquent une part de la variabilité spatiale de la mortalité infantile à Lyon. Un article écrit par Wahida Kihal porte sur l'analyse de la relation entre la présence d'espaces verts et la mortalité infantile interprétée dans le cadre d'un modèle conceptuel des différents schémas pouvant relier les espaces verts aux issus défavorables de grossesse. (Actuellement en révision dans la revue *BMC Pediatric and child births*) (Abstract en annexe).

3. Multi-exposition

Suite à ces travaux c'est tout naturellement que notre intérêt s'est porté sur l'exposition à de multiples nuisances. Un indice composite d'exposition a été créé, à partir des concentrations de NO₂ annuels, de niveaux de bruit, de proximité aux espaces verts, de proximité aux installations industrielles, et aux trafics routiers en utilisant une analyse factorielle multiple, suivi d'une classification hiérarchique. Cet indice permet de synthétiser pour chaque IRIS, la « pression environnementale » qu'exerce le cadre de vie physique sur les habitants. Ces résultats sont présentés dans l'article écrit par Benoit Lalloue (Soumis au Journal of Exposure Science and Environmental Health) (Abstract en annexe).

B. Perspectives en santé publique

Les effets relativement importants observés à l'échelle intra-urbaine dans notre étude suggèrent un besoin pour de futures études et les politiques publiques de se concentrer sur la compréhension des effets sur la santé spécifiques à chaque agglomération en termes de mobilité, accès aux soins et aux comportements de la population.

Si la démarche ordonnée que nous avons suivie dans le cadre de cette thèse nous a permis d'avancer dans la représentation spatiale et l'évolution temporelle de la contribution de la pollution de l'air aux inégalités sociales de santé, leur interprétation demeure néanmoins contrainte par l'absence totale d'informations sur l'aménagement urbain (en terme d'accessibilité aux soins, de mobilité des populations) et les facteurs de risque individuels. Dans les perspectives du projet Equit'Area, l'intégration de données individuelles (sanitaires et socio-économique) aux données écologiques est en cours, ce qui permettra de dégager, par des analyses multi-niveaux, le rôle des caractéristiques socio-économiques des parents de celles subies sur leur lieu de résidence et donc d'approfondir les mécanismes sous-jacents aux inégalités sociales de santé.

Alors que l'étude des inégalités sociales de santé ou des effets des facteurs de risque environnementaux sur la santé est maintenant courante dans la littérature de santé publique, l'évaluation de leur effet conjoint ou de leur évolution par une approche spatiale reste à exploiter. Ce domaine encore vierge pourrait éclairer les politiques publiques sur les zones à prioriser dans le but de réduire les risques et les inégalités.

Références : Chapitre 4

1. Blumenshine P, Egerter S, Barclay CJ, Cubbin C, Braveman PA. Socioeconomic disparities in adverse birth outcomes: a systematic review. *Am J Prev Med.* 2010 Sep;39(3):263–72.
2. Guildea ZE, Fone DL, Dunstan FD, Sibert JR, Cartlidge PH. Social deprivation and the causes of stillbirth and infant mortality. *Arch Dis Child.* 2001 Apr;84(4):307–10.
3. Arntzen A, Nybo Andersen AM. Social determinants for infant mortality in the Nordic countries, 1980-2001. *Scand J Public Health.* 2004;32(5):381–9.
4. Gorla S, Stempfelet M, Crouy-Chanel P. Introduction aux statistiques spatiales et aux systèmes d'information géographique en santé environnement- Application aux études écologiques. [Internet]. Institut de veille sanitaire (InVS); 2011 p. 65 p. Available from: <http://www.invs.sante.fr>
5. Greenland S. Ecologic versus individual-level sources of bias in ecologic estimates of contextual health effects. *Int J Epidemiol.* 2001 Dec;30(6):1343–50.
6. Krieger N, Williams DR, Moss NE. Measuring social class in US public health research: concepts, methodologies, and guidelines. *Annu Rev Public Health.* 1997;18:341–78.
7. Blondel B, Kermarrec M. Les naissances en 2010 et leur évolution depuis 2003. Enquête nationale périnatale 2010. Paris: INSERM, DGS, DREES; 2011 Mai p. 117.
8. Patra J, Bakker R, Irving H, Jaddoe VWV, Malini S, Rehm J. Dose-response relationship between alcohol consumption before and during pregnancy and the risks of low birthweight, preterm birth and small for gestational age (SGA)-a systematic review and meta-analyses. *BJOG Int J Obstet Gynaecol.* 2011 Nov;118(12):1411–21.
9. O'Leary CM, Nassar N, Kurinczuk JJ, Bower C. The effect of maternal alcohol consumption on fetal growth and preterm birth. *BJOG Int J Obstet Gynaecol.* 2009 Feb;116(3):390–400.
10. Crane JMG, Keough M, Murphy P, Burrage L, Hutchens D. Effects of environmental tobacco smoke on perinatal outcomes: a retrospective cohort study. *BJOG Int J Obstet Gynaecol.* 2011 Jun;118(7):865–71.
11. Elsenbruch S, Benson S, Rucke M, Rose M, Dudenhausen J, Pincus-Knackstedt MK, et al. Social support during pregnancy: effects on maternal depressive symptoms, smoking and pregnancy outcome. *HumReprod.* 2007 Mar;22(3):869–77.
12. Chaix B, Gustafsson S, Jerrett M, Kristersson H, Lithman T, Boalt A, et al. Children's exposure to nitrogen dioxide in Sweden: investigating environmental injustice in an egalitarian country. *J Epidemiol Community Health.* 2006 Mar;60(3):234–41.

13. Yanosky JD, Schwartz J, Suh HH. Associations between measures of socioeconomic position and chronic nitrogen dioxide exposure in Worcester, Massachusetts. *J Toxicol Environ Health A*. 2008;71(24):1593–602.
14. Seaton A, Dennekamp M. Hypothesis: Ill health associated with low concentrations of nitrogen dioxide--an effect of ultrafine particles? *Thorax*. 2003 Dec;58(12):1012–5.
15. Woodruff TJ, Darrow LA, Parker JD. Air pollution and postneonatal infant mortality in the United States, 1999-2002. *Environ Health Perspect*. 2008 Jan;116(1):110–5.
16. Scheers H, Mwalili SM, Faes C, Fierens F, Nemery B, Nawrot TS. Does air pollution trigger infant mortality in Western Europe? A case-crossover study. *Environ Health Perspect*. 2011 Jul;119(7):1017–22.
17. Romieu I, Ramírez-Aguilar M, Moreno-Macias H, Barraza-Villarreal A, Miller P, Hernández-Cadena L, et al. Infant mortality and air pollution: modifying effect by social class. *J Occup Environ Med Am Coll Occup Environ Med*. 2004 Dec;46(12):1210–6.
18. Goldberg MS. On the interpretation of epidemiological studies of ambient air pollution. *J Expo Sci Environ Epidemiol*. 2007 Dec;17 Suppl 2:S66–70.
19. Crouse DL, Ross NA, Goldberg MS. Double burden of deprivation and high concentrations of ambient air pollution at the neighbourhood scale in Montreal, Canada. *Soc Sci Med* 1982. 2009 Sep;69(6):971–81.
20. Padilla C, Lalloué B, Zmirou-Navier D, Séverine D. Mortalité infantile, défaveur et proximité aux industries polluantes: une analyse spatiale conduite à fine échelle (agglomération de Lille, France). *Environ Risques Santé*. 2011 Jun 1;10(3):216–21.

ANNEXES

- **Article publié** : Lalloué B, Monnez JM, **Padilla C**, LeMeur N, Zmirou-Navier D, Deguen S. A statistical procedure to create a neighborhood socioeconomic index for health inequalities analysis. Int J Equity Health. 2013 Mar 28;12:21

- **Article publié** : **Padilla C**, Benoit Lalloué, Denis Zmirou-Navier & Séverine Deguen. Infant mortality, deprivation and proximity to polluting industrial facilities – A small-scale spatial analysis with census block data (Lille Metropolitan Area, France). Environnement Risques et Santé. Vol 10, n°3 ; 216-21 – Déc2010.

- **Abstract** : Kihal W, **Padilla C**, Lalloué B, Gelormini M, Rey G, Zmirou-Navier D, Deguen S. Green space and Social inequalities among a cohort of neonatal mortality in France. Waiting editor in IJERPH (International Journal of environmental Research and public Health)

- **Abstract** : Kihal W, **Padilla CM**, Lalloué B, Rougier C, Defrance J, Zmirou-Navier D, Deguen S. An exploratory spatial analysis to assess the relationship between deprivation, noise and infant mortality. Waiting editor in Pediatric and perinatal epidemiology journal.

_ **Abstract** : Lalloué B, Monnez JM, **Padilla C**, Kihal W, Zmirou-Navier D, Deguen S. Data Analysis technics, A tool for Cumulative Exposure Assessment. Submitted in Exposure Science and Environmental Epidemiology journal

- Bulletin n°7 : Acte de décès

- Acte de décès néonatal

RESEARCH

Open Access

A statistical procedure to create a neighborhood socioeconomic index for health inequalities analysis

Benoît Lalloué^{1,2,3*}, Jean-Marie Monnez³, Cindy Padilla^{1,2}, Wahida Kihal¹, Nolwenn Le Meur^{1,4}, Denis Zmirou-Navier^{1,2,5} and Séverine Deguen^{1,2}

Abstract

Introduction: In order to study social health inequalities, contextual (or ecologic) data may constitute an appropriate alternative to individual socioeconomic characteristics. Indices can be used to summarize the multiple dimensions of the neighborhood socioeconomic status. This work proposes a statistical procedure to create a neighborhood socioeconomic index.

Methods: The study setting is composed of three French urban areas. Socioeconomic data at the census block scale come from the 1999 census. Successive principal components analyses are used to select variables and create the index. Both metropolitan area-specific and global indices are tested and compared. Socioeconomic categories are drawn with hierarchical clustering as a reference to determine “optimal” thresholds able to create categories along a one-dimensional index.

Results: Among the twenty variables finally selected in the index, 15 are common to the three metropolitan areas. The index explains at least 57% of the variance of these variables in each metropolitan area, with a contribution of more than 80% of the 15 common variables.

Conclusions: The proposed procedure is statistically justified and robust. It can be applied to multiple geographical areas or socioeconomic variables and provides meaningful information to public health bodies. We highlight the importance of the classification method. We propose an R package in order to use this procedure.

Keywords: Socioeconomic status, Multidimensional index, Principal component analysis, Hierarchical classification, Small-area analysis

Social health inequalities are well documented in the epidemiological literature. Studies show that, for a wide array of health outcomes (infant mortality and pregnancy [1-3], cardiovascular and respiratory diseases [4-8], mental health [9,10], etc.), the burden of disease is different between deprived and affluent populations [11]. Most studies on social health inequalities use personal socioeconomic characteristics [1,6,7] which are often difficult and long to obtain, especially if an individual-based epidemiological study has to be set in order to collect them. Although

contextual data cannot be used and interpreted as individual data (due among other issues to the ecological fallacy), it is easier to retrieve aggregate data from existing databases. Further, when the spatial units are small, as in our case, the ecological bias is reduced [12]. Besides, it is in some cases relevant, very convenient or even necessary to use this aggregate socioeconomic data as an alternative source of information for public health research [2,5,8-10,13-15].

Moreover, even when personal information is available, studies have shown that measures of neighborhood socioeconomic status (SES) explain significant variations in health status, even after adjustment for individual socioeconomic characteristics, suggesting that neighborhood SES may be by itself a risk factor [4,16,17].

* Correspondence: benoit.lalloue@ehesp.fr

¹EHESP Rennes, Sorbonne Paris Cité, Rennes, France

²Inserm, UMR IRSET Institut de recherche sur la santé l'environnement et le travail - 1085, Rennes, France

Full list of author information is available at the end of the article

Neighborhood SES is a complex concept involving many aspects, such as employment, income, education, housing and social bonds [18-22]. Several studies have used only one variable to represent SES, a limitation that makes it difficult to take account of the multiple dimensions of neighborhood SES. Another possibility is to study the one-by-one association (for instance with simple regression models) between the outcome of interest and different variables simultaneously with the purpose to identify if a particular component of the SES is particularly associated with the outcome. However, this comparison between associations is not trivial (for instance comparison between non nested models), especially when variables are correlated, and should be carefully conducted in order to obtain sound conclusions. Eventually, it is possible to include several socioeconomic variables in the same model but this may lead to statistical problems when dealing with multicollinearity and the large number of parameters to be estimated.

To overcome these problems, a socioeconomic composite index may be set up at the level of a neighborhood, that may summarize the many aspects that encompass the concept of SES with a large variety of possible techniques: additive scores with different weighting approaches (Z-score, experts weightings) [22-27], principal component analysis or factor analysis [22,26,28-31], spatial or GIS-based analyses [32,33], or others methods [34,35]. This kind of index can then be used for different purposes such as reveal the existence of social health inequalities through an epidemiological study, giving an overview of the situation for decision makers, or identify particular extreme areas where it is needed to focus public action (social planning, urban planning, . . .). Townsend [36] and Carstairs [24] indices, which are commonly used in various countries, cover different topics. Because they are based on only 4 variables, these indices may not always give a comprehensive representation of SES that can be used by public health bodies, at local or national level, in order to determine where action might be justified and effective. Indices that incorporate a greater number of variables from different dimensions may be more appropriate for this purpose. Moreover, these indices or Jarman's UPA[25], were built 20 years ago and may not take into account possible modifications that occurred along time and across countries in the definitions of the variables or in the evolution of how these variables may contribute to the SES. Utilization of these indices across different countries may also be difficult due to cultural or historical distinctions (e.g. "social classes", when defined, can have very different definitions; also, "proportion of households not owner of their house" vary considerably across countries); comparisons within countries may also be hampered by demographic and urban policy factors (e.g. the "proportion of households not owner of a car" highly depend on the availability of public transport and sharply contrast central

urban areas and peri-urban or rural areas). Moreover, regardless of the creation procedure, the interpretation of the variables included in an index can be very different according to the implementation area. For instance, variables like ratio of individual houses or proportion of farmers do not have the same interpretation from a SES point of view in urban or rural areas. The interpretation of each variable included in an index must of course be done according to the context. In this setting, it could be helpful to have a versatile procedure which would allow taking into account these changes and selecting variables among a given set rather than fixing a mandatory list of variables.

A rigorous methodological approach is required to ensure that the index is statistically well founded and provides a good approximation of SES. Principal component analysis appears to be particularly suitable for developing composite indices because this statistical method creates non-correlated linear combinations of the variables with maximal variance, which allows the best contrast between statistical units. Furthermore, in ecologic epidemiological studies, mapping is a crucial step for showing the spatial distribution of deprivation in public health studies. While discretization of quantitative variables is frequently used for this purpose, mainly by using quantiles, it is an arbitrary technique which relies on the categories sample sizes rather than on similarities between units. Yet several other methods, such as hierarchical classification, are available to create homogeneous categories of similar spatial units.

In this context, this study presents a procedure based on statistical criteria and justification for selecting socioeconomic variables in order to create a neighborhood socioeconomic index meant to provide meaningful information to public health bodies and allow epidemiological assessment of social health inequalities. In view to assess its ability to be generalized, the procedure was applied in three contrasted French metropolitan areas to create both area-specific and global socioeconomic indices. Eventually, an R package was created in order to give an easy way to implement the procedure in a variety of contexts. This package contains the basic functions needed to run the procedure, obtain the corresponding SES index and create categories.

Material and methods

Study setting and small area level

The study was carried out in three large metropolitan areas in France: Lille (Nord Pas de Calais region, northern France), Lyon (Rhône-Alpes region, central and eastern France) and Marseille (Provence-Alpes-Côte d'Azur region, south eastern France) with a total population of around 3.8 million. These urban areas are the three largest in France after Paris. They differ in some important socio-demographic features. For instance, the Lille Métropole has a higher population of people under age 25, more blue-collar workers and individual houses than the other

cities; Grand Lyon has a higher rate of foreigners and white collar workers, and fewer people with no qualifications; while the Aix-Marseille urban unit has a higher rate of single parent families, higher unemployment, a lower rate of people with steady jobs and less social housing (see in Additional file 1 the detailed socioeconomic characteristics of the metropolitan areas).

The statistical units were the sub-municipal French census block groups (called IRIS for “Îlots Regroupés pour l’Information Statistique”) defined by the National Institute of Statistics and Economic Studies (INSEE). These are the smallest units for which socioeconomic and demographic information is available from the French national census (Table 1). These units have an average of 2,000 inhabitants and are constructed in collaboration with local actors (municipalities and communities) to be as homogeneous as possible in terms of socio-demographic characteristics and land use. They also take account of physical obstacles that may break up urban landscapes, such as arterial roads, green spaces, bodies of water, and must have unambiguously identifiable contours, stable over time. Census block groups (BGs) are divided into three distinct categories in order to take into account some special cases: 1) “housing” BGs represent the vast majority of BGs (92% for all France in 2008) and count generally between 1800 and 5000 inhabitants; 2) “activity” BGs include at least 1,000 employees and at least twice as many employees as residents (e.g. industrial estates or business districts); and 3) “miscellaneous” BGs are specific wide areas sparsely populated (leisure parks, port areas, forest, etc.). As activity and miscellaneous BGs have some particular profiles due to the way they are defined, this study only considered housing BGs for the creation of the socioeconomic index. Housing BGs were treated in our statistical analysis as active units while activity and miscellaneous BGs were treated as supplementary units (meaning that they were not part of the construction of the socioeconomic index but will have an index value).

Socioeconomic data

Socioeconomic data were taken from the 1999 national census (provided by the INSEE) and provided counts of

population, households and residences at BG scale covering all the social, economic and demographic aspects. Median income (for 2001) at the BG scale was taken from another national study also provided by the INSEE (“Revenus fiscaux des ménages”, INSEE – DGI). Using this raw data, 48 variables were defined at the BG scale based on the INSEE definitions. These variables were chosen to be representative of the theoretical concepts of SES and in line with the variables most often used in the literature [21,22,24-31,33,36,37]. We also introduced some variables which were not totally part of the SES concept but that could be considered as linked with it (and could also give insight about the neighborhood environment). For instance we included the proportion of people over the age of 65, which could reflect the proportion of retired people; or the proportion of people who have moved from their municipality since the last census which could give indications about the “residential instability” of the neighborhood population.

All variables were associated with family structure, household type, immigration status, mobility, employment, income, education and housing (a detailed list of these variables can be found in Table 2). The aim of introducing a spectrum of variables broader than just the variables known to be indicators of the SES was to offer the possibility to examine the utility of taking into account some “proxies” and also to have a data driven approach in order to determine the variables which maximize the index’s variance.

Some of the variables were intentionally redundant and represented the same notion, in view to determine which best represented this notion. There were two such groups: 7 variables of unemployment (ID 12 to ID 18 in Table 2) and 3 variables of labor force (ID 9 to ID 11 in Table 2). There were an unexpectedly high number of missing values for median income (see Table 1) but, willing to keep this variable in the analysis, we filled missing values with the average value of the adjacent BGs.

Creation of the socioeconomic index

The socioeconomic index was created by improving and expanding a procedure previously developed by our

Table 1 Description of the three study urban areas

	Lille Métropole	Grand Lyon	Aix-Marseille
Population in 1999 (inhabitants)	1,091,438	1,193,384	1,349,772
Population in 2007 (inhabitants)	1,106,885	1,260,348	1,434,845
Number of municipalities	85	58	38
Number of census block groups	506	510	630
Number of housing blocks (% of census block groups)	475 (94%)	465 (91%)	563 (89%)
Number of census block groups with missing median income (% of census blocks)	119 (24%)	95 (19%)	106 (17%)
Area (km ²)	611.45	527.15	1289.59

Table 2 Description of the first selection of 48 socioeconomic variables at the census block group scale

<i>Unless stated otherwise, variables are proportions expressed in %</i>		Var. Id.
Family and household	People under the age of 25 in the total population	1
	People over the age of 65 in the total population	2
	People living outside the household (boarder students, soldier in garrison, people in jail, people in nursing home or in hospital, etc.) in the total population	3
	SINGLE-PARENT FAMILIES IN THE TOTAL POPULATION	4
	Householders living alone in the total population	5
Immigration and mobility	FOREIGN PEOPLE IN THE TOTAL POPULATION	6
	FOREIGN IMMIGRANTS (SINCE THE LAST CENSUS) IN THE TOTAL POPULATION	7
	People who have moved from their municipality since the last census in the total population	8
Employment and income	<i>People in the labor force in the total population</i> ^a	9
	Men in the labor force in the total male population ^a	10
	Women in the labor force in the total female population ^a	11
	<i>Unemployed people in the labor force</i> ^b	12
	Unemployed foreigners in the labor force ^b	13
	Unemployed people in the 15-24 years old labor force ^b	14
	Over 50 years old unemployed people in the labor force ^b	15
	Unemployed people in the male labor force ^b	16
	Unemployed people in the female labor force ^b	17
	People unemployed for more than 1 year in the labor force ^b	18
	SELF-EMPLOYED (INDEPENDENT WORKERS, EMPLOYERS, ETC.) IN THE LABOR FORCE	19
	PEOPLE WITH UNSTABLE JOBS IN THE LABOR FORCE (APPRENTICES, TRAINEES, TEMPORARY JOBS, ETC.)	20
	PEOPLE WITH STEADY JOBS IN THE LABOR FORCE	21
	Farmers in the labor force	22
	Managers in the labor force	23
	Blue-collar workers in the labor force	24
	MEDIAN INCOME PER CONSUMPTION UNIT (IN EUROS PER YEAR) ^c	25
Education	People 6-15 years old attending school in the 6-15 years old population	26
	PEOPLE WITH NO SCHOOL GRADUATION (AND NOT STUDYING) IN THE 15 YEARS OLD AND MORE POPULATION	27
	PEOPLE WITH BASIC OR INTERMEDIATE GENERAL OR VOCATION QUALIFICATIONS (AND NOT STUDYING) IN THE 15 YEARS OLD AND MORE POPULATION	28
	PEOPLE WITH GENERAL OR VOCATIONAL MATURITY CERTIFICATES (AND NOT STUDYING) IN THE 15 YEARS OLD AND MORE POPULATION	29
	People with at least a lower tertiary education (and not studying) in the 15 years old and more population	30
	People with a higher educational degree (and not studying) in the 15 years old and more population	31
Housing	Students in the 15 years old and more population	32
	Individual houses in the main residences	33
	Multiple dwelling units in the main residences	34
	NON-OWNER-OCCUPIED IN THE MAIN RESIDENCES	35
	Subsidized housing in the main residences	36
	Main residences built before 1968	37
	Main residences built after 1990	38
	Main residences less than 40 m ²	39
	Main residences larger than 150 m ²	40

Table 2 Description of the first selection of 48 socioeconomic variables at the census block group scale (Continued)

Main residences without bathtub or shower	41
Main residences without toilet	42
Main residences without heating	43
Main residences with a parking space (garage or other)	44
MAIN RESIDENCES WITH MORE THAN ONE PERSON PER ROOM	45
AVERAGE NUMBER OF PEOPLE PER ROOM ^c	46
HOUSEHOLDS WITHOUT A CAR	47
HOUSEHOLDS WITH 2 OR MORE CARS	48

^aRedundant group "labor force".

^bRedundant group "unemployment".

^cNot a proportion.

UPPERCASE : variables selected commonly for each metropolitan areas and global analysis.

Italic: variables selected during the "reduction of redundant groups" step for the global analysis.

team [38]. The three steps (Figure 1) described below were used:

Step 1 - Study of the redundant variables (in our case variables ID 9-11 and ID 12-18 aforementioned). To avoid issues due to redundant covariates (correlation over 0.8 for most of them in all the study areas) one variable was selected from each group by applying principal component analysis (PCA, see Additional file 2) to each of the two groups of redundant variables. The first component was a good representation of the group of variables, strongly correlated with all of the variables (if not and if a variable was not well represented by the first component, - a situation that did not happen in our

case - this variable should actually not have been part of the redundant group and should not have been included in it). However, to ease interpretation, the variable with the largest correlation with the first component was selected. Reducing the two redundant groups to one variable each reduced, in our case, the number of variables to 40 (the two selected variables at this step are shown in italic font in Table 2).

Step 2 - Selection of the variables. PCA on these 40 variables (i.e. after selection of one variable per group of redundant variables) was used to select the variables with a contribution to the first component larger than the average one, i.e. variables that were best correlated with the first component.

Step 3-Construction of the final index. A final PCA was carried out including the variables selected in step 2. Provided that the first component of this PCA could be interpreted (according to the meaning of the variables in the given context) as a "SES component" (which was expected giving the variables selected and confirmed *a posteriori*), it was used to calculate the socioeconomic index as the reduced first component. This normalization gave an index with mean 0 and standard deviation 1. Since the purpose of the whole procedure was to create a single index, the second and subsequent components of the final PCA were not used (in practice, the variance explained by the second component was far below that explained by the first component and there was no clear interpretation of it).

The choice of PCA as the main technique in the procedure was done in order to use as little constraints and hypotheses as possible, as well as to keep a strongly data driven approach. This approach is not based on a model set a priori, like Factor Analysis. Moreover, it is known that a PCA where the last eigenvalues are near and close to 0 gives results very similar to those of a Factor Analysis.

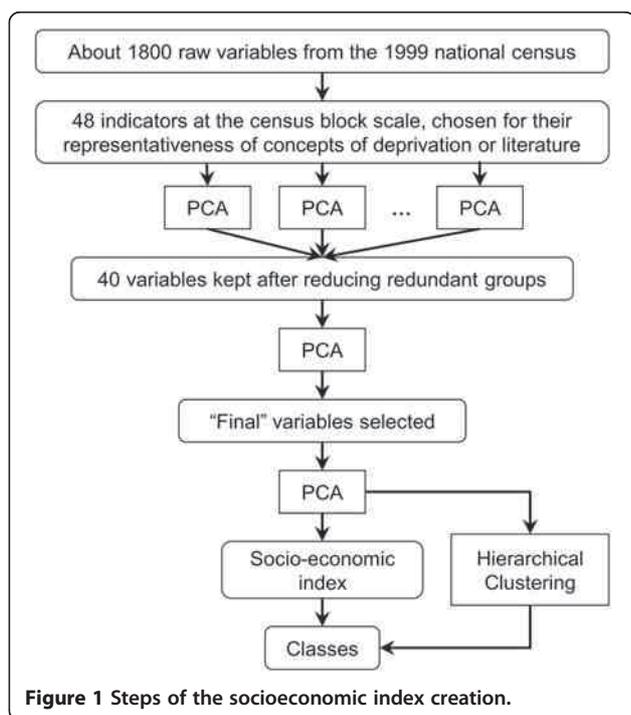


Figure 1 Steps of the socioeconomic index creation.

This procedure was applied to each metropolitan area independently (giving socioeconomic indices specific to each area) and to the three urban areas altogether. This produced four different indices that, as a result, could be compared. Step 1 was kept in the procedure and applied to each metropolitan area because differences between them in their socioeconomic make-up might lead to different choices.

Hierarchical clustering and optimal thresholds

Socioeconomic categories were created so that the index could be used as an explanatory variable to determine possible non-linear relationships in a variety of applications, and be used for mapping.

Defining socioeconomic index quantiles is a common technique but it is sometimes unsatisfactory. Indeed, classes created with quantiles are only based on the number of units and therefore may not correctly classify units according to their similarity if they are not distributed homogeneously. This can lead to merging in the same class very different groups of units, or to split into two classes a homogeneous group.

Hierarchical clustering (HC) is frequently used after PCA [39] in data mining to create meaningful categories. Given a set of p variables measured on n elements, each element is represented as a point in \mathbb{R}^p . A distance between elements d (usually the Euclidian distance) and a distance between categories Δ (based on d) are defined. HC algorithm creates a hierarchy of categories step by step by merging at each step the two categories which are the nearest according to Δ . When Δ is a particular distance (the Ward's distance), this algorithm allows to obtain categories homogeneous in their composition and heterogeneous between them (i.e. with a maximum between-categories inertia). The most appropriate partition is then selected from the hierarchy of categories. More methodological details about HC are available in Appendix 2.

However, HC is a multidimensional technique which uses several components of a PCA (often 5 or 10). Now, we wanted here to create the categories from a one-dimensional index (it is also possible, but not for the same purpose, to keep directly the categories created by HC in order to have a qualitative index). Then, we used HC as a reference to determine "optimal" thresholds able to construct a one-dimensional classification as close as possible to the HC. There were two possible cases, depending on the number of categories:

- either the categories constructed with HC were not distributed along the first principal axis of the PCA (the second and subsequent axes affecting classification) in which case it was not possible to determine thresholds along the first axis that would be able to correctly approximate HC categories. The

index could not be used by itself and the number of categories was therefore reduced.

- Or the categories were distributed along the first component of the PCA (i.e. our socioeconomic index). In this case, optimal thresholds were determined using a simple iterative algorithm: at each step, categories were defined with new thresholds along the index values and the concordance rate between this classification and the clustering using HC were calculated. Only values with the best concordance percentage were kept. It created socioeconomic categories using the socioeconomic index by itself.

Comparison of indices and classifications

Pearson's coefficient of correlation was used to compare the area-specific indices to the global one, which encompasses the 3 metropolitan areas, and also the Carstairs' and Townsend indices with ours.

Carstairs' index [24] was constructed as the sum of the standardized proportions of total unemployment, of households without a car, of households with more than one person per room, and of blue-collar workers (since French census do not use "social classes"). Townsend's index [36] was constructed as the sum of the standardized log-proportions of total unemployment and households with more than one person per room, and the proportions of households without a car, and of non-owner-occupied main residences.

The concordance percentage was used to compare pairs of classifications. This is the percentage of BGs in the same class in both classifications (the diagonal of the confusion matrix). The R software [40], with the FactoMineR package [39] and the SesIndexCreaToR package, was used to create the indices and clustering, to determine thresholds and draw comparisons.

R package

Since the procedure described here is more complex than for some other SES indices, we specifically developed the SesIndexCreaToR package. The version 1.0 of this package (currently freely available on the website of the EquitArea project: http://www.equitarea.org/documents/packages_1.0-0/) contains the basic functions needed to run the procedure (in its entirety or only in some steps) and to obtain the corresponding SES index. The user may also create categories of this index with different methods (hierarchical clustering with or without k-nearest neighbors, quantiles, or intervals). We project to extend the package in the future and among other improvements we foresee to add tools to help the interpretation of these categories and the visualization of the results.

Results

Constructing indices, selecting variables and determining contributions

The four socioeconomic indices were built as the first component of PCA using the data for each urban area separately and for an overall analysis of the combined sets of data. In each case, this first component of the PCA was positively correlated with variables of low SES (unemployment, single-parent families, overcrowding, etc.) and negatively correlated with variables of high SES (income, steady jobs, high level of education, etc.); therefore, it was interpreted as a true SES component (see Figure 2). It always explained a large proportion of the total variance (Table 3). By contrast, the second component had never a clear interpretation and explained less than 17% of the total variance.

Fifteen of the 20 variables selected by the procedure (step 1 to 3) were common to all four indices and accounted for more than 77% of the construction of the first component (see below). This result showed that the procedure was robust and that, despite the substantial socio-demographic differences between the urban areas, the same common variables could explain a large part of the socioeconomic variability. Now, the ranks of the contributions of the variables to the index were different in each urban area (Additional file 3) although none was far from the overall average contribution.

For Lille Métropole, the procedure selected 21 variables for the index (for Grand Lyon, Marseille urban area and global analysis it selected 20, 20 and 19, respectively). The first component of the final PCA explained 61% (resp. 58%, 57% and 57%) of the total variance while the second 12% (resp. 17%, 15% and 11%). The variables common to

all four indices contributed 80% (resp. 84%, 78% and 88%) to the index. In each metropolitan area, local experts associated with the project confirmed that our SES index globally well-represent the socioeconomic true profile of the neighborhoods.

Comparisons between indices

The indices in each metropolitan area were compared with the global index restricted respectively to the BGs of each area (the global index was constructed on the BGs of the three areas altogether but, to allow comparison with the area-specific index, only BGs of this area were considered), as well as with those proposed by Carstairs and Townsend.

In general, all the correlations between the area-specific and the overall index restricted to the BGs of each metropolitan area were above 0.9 (Table 4) with a clear linear association (Additional files 4, 5 and 6). There were very good correlations between our indices (constructed either for each urban area or for the three areas altogether) and the Carstairs and Townsend indices (always larger than 0.91) suggesting that the socioeconomic dimension measured by our index is very close to that measured by the well-known and often used Carstairs and Townsend indices.

Comparisons between classifications

The initial number of categories we tested for classification was five because it is a usual number of categories used in spatial epidemiology, especially for mapping. For all four indices, the categories obtained through HC depended both on the first and second axes of the PCA and the categories were not distributed solely

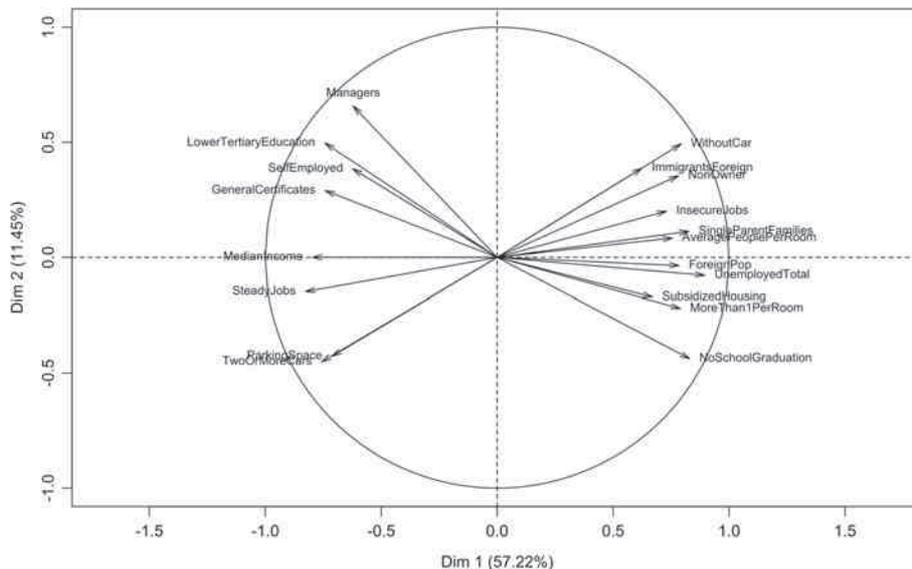


Figure 2 Circle of correlation for final step, global analysis.

Table 3 Percentage of variance explained by the two first components of the final PCA, by area

	Lille Métropole	Grand Lyon	Marseille urban unit	Global
1 st component	60.73%	57.79%	57.29%	57.22%
2 nd component	12.13%	16.71%	14.66%	11.45%

along the first component. This observation suggests that 5 categories were too many should the objective be to use only the first component (i.e. our index). The analysis showed that classification with quintiles was not optimal, since the hierarchical clustering gave very different results. The number of categories was, therefore, reduced until categories could be constructed only using the first axis of the final PCA. In each case, the largest number of categories was three. This number creates the optimal classification according to the HC. Average values of the common variables for the global analysis can be seen in Additional file 7.

Concordance rates between the different techniques are shown in Table 5. Very low concordance was found for the HC with 5 categories and quintile classifications, with less than 50% of similar classifications for the three metropolitan areas and about 60% for the metropolitan areas taken together. HC was unable to construct 5 categories using only the first PCA axis. The concordance rate between HC with 3 categories and terciles was between 69% and 78%. This could be explained by the smaller number of categories. Even such concordance rates mean that about one quarter of the BGs had a different class across the two classification methods. When comparing the SES index categories (created by HC or by optimal threshold) with Carstairs' and Townsend's indices categorized through quantiles, the concordance rates remained low.

By contrast, the concordance rates between HC in 3 categories and optimal thresholds were between 93% and 97%, confirming that the classification in three categories was fully along the first axis. Additional files 8, 9 and 10 show for each metropolitan area maps of the three socioeconomic categories created either by tertiles or optimal thresholds, and the range of the SES index for each category (the higher the index, the most deprived the area)..

Table 4 Correlation coefficients between indices

	Lille Métropole	Grand Lyon	Marseille urban unit	Global
Global ^a	0.99	1	0.99	/
Carstairs	0.92	0.96	0.91	0.94
Townsend	0.98	0.94	0.96	0.96

^aWhen comparing the global index with a city index, global index is restricted to that particular city's census block groups.

Discussion

This study developed a rigorous statistical procedure to create socioeconomic indices in urban contexts, improving and extending a previous work.[38] The procedure was applied and validated on three different French urban areas and proved its robustness in different socio-demographic settings. An R package was developed in order to help applying this procedure in other contexts.

As for most studies developing new methodologies to construct a neighborhood socioeconomic index, the preliminary selection of variables was based on a literature review [10,15,26,27,29]. The social and material deprivation index developed by Pampalon et al [28] included people without qualifications, employment ratio, average income, individuals living alone, individuals divorced, separated or widowed, and single parent families. They chose these variables according to four criteria: well documented health links, variables previously used as "geographic proxies" in social health inequality studies, variables belonging to the material or social dimension of deprivation and the availability of data for their study area. Carstairs and Townsend followed a similar procedure for selecting 4 variables characterizing neighborhood deprivation in their indices. However, this approach was only a preliminary step in the construction of our index. One originality of our procedure lays in selecting the final variables for the index by usage of data mining techniques rather than only information gleaned from a literature review, allowing to discard part of the subjectivity that may influence the choice of the variables. This data driven approach allows the data "speak by themselves". Although it was what we expected, it was not sure, before the PCA was implemented that the first component would be a good socioeconomic index. This appeared *a posteriori* as the PCA explored the data and revealed their underlying structure.

About 20 variables, a number not defined *a priori*, were selected for each metropolitan area, encompassing the various domains of SES. This allowed to determine the common determinants of SES in the various areas and also to select determinants which are more specific in each area. The larger number of variables compared with other indices gives room for a finer spatial description of SES and of specific characteristics of each metropolitan area, providing information which public health bodies might find helpful in determining key targets for local actions. Indeed, once the index is constructed and used to identify BGs with the lowest SES, it is possible to return to the variables that compose the index in order to see which ones could be a leverage for action, a property that more simple indexes lack. Using this method (use such an index, in a quantitative or qualitative way, to identify lowest SES areas and then go back to the individual variables to have more details) in an epidemiological study to describe the spatial distribution of some disease or cause of mortality in

Table 5 Concordance rates between different clustering techniques and between indexes

	HC (5) ^a vs. quintiles ^b	HC (3) ^a vs. tertiles ^b	HC (3) ^a vs. optimal thresholds ^b	Optimal thresholds vs. tertiles ^b	Carstairs		Townsend	
					HC (5) ^a vs. quintiles ^b	Optimal thresholds vs. tertiles ^b	HC (5) ^a vs. quintiles ^b	Optimal thresholds vs. tertiles ^b
Lille Métropole	41%	78%	98%	79%	38%	70%	42%	78%
Grand Lyon	48%	74%	93%	78%	47%	77%	40%	75%
Marseille urban unit	48%	69%	97%	67%	51%	67%	50%	69%
Global	63%	71%	97%	72%	57%	70%	55%	71%

^aHierarchical Clustering using Principal Components (in parenthesis, the number of categories chosen).

^bConcordance rate (percent of census block groups categorized into the same class using the two different clustering schemes).

a metropolitan area will not only allow to flag communities where the risk is highest, but will also provide information on the social and economic characteristics of these communities upon which appropriate and focused preventive policies can be devised and implemented.

The large number of common variables (15 of the 20 variables) across the metropolitan areas shows the stability of the results and the good representation of the underlying concept of SES conveyed by the index. These variables reveal the common determinants of SES in different French metropolitan areas, at BG level, which is the smallest administrative unit for which census data is available. The specific SES patterns in each area can be assessed in two different ways: through the variables which are specific to each area, and through the relative contribution of each variable to the final index. As a result, the procedure proposed in this study can be used alternatively to build a city-specific index which can be applied locally, for instance to determine priority BGs for local action, or a global index to compare a set of cities with the same metric.

However, one should remember that data and indices used here are area-based and not person-based. Indeed, although BGs are constructed in order to be as homogeneous as possible, there is still individual variability within them which cannot be assessed by aggregated data. Therefore, as it is now well-known, inference at the individual scale from indices created at the BG scale can be tricky due to the ecological fallacy. SES indices presented here are neighborhood SES indices and should be used as a way to assess the contextual socioeconomic setting in which people live rather than a way to approximate the individual SES.

When socioeconomic indices were first constructed, categories were delineated to show the spatial distribution of SES on maps and to investigate the existence of non-linear social relationships with some outcome of interest. So far, to our knowledge, most of the studies classifying deprivation scales have used quantiles [2,10,13,15,27,29,33] without questioning the validity of this classification method from a statistical point of view. This simple approach should be used with caution; our study suggests

that it might put dissimilar geographical units in the same class and separate similar units, according to HC.

Using HC, the first dimension alone of the final PCA was not sufficient to create 5 socioeconomic categories. Although we could have kept the results of the HC as a qualitative index, this would have contradicted with our aim to have a one-dimensional index. In this study, but without possible generalization to other data, it was preferable to use a 3-categories classification built only with the first component of the final PCA.

Despite its statistical justification, this study has some limitations. Some are induced by the very nature of an index. Since indices are composite syntheses of several variables, they have no unit. This can reduce the interpretability of their application, especially regression models, the meaning of an increase or decrease of one unit of the SES index being difficult to express. From a public policy point of view, an index alone cannot give indications on how to operate to change the situation. Although the indices created by the procedure we propose share these limitations, we think they are interesting as first indicators of 'global' neighborhood SES and as a synthetic tool to point out the situation to policy makers. Eventually, one may return, as aforementioned, to the variables composing the index to have a better insight of the actual situation of the identified neighborhoods and the variables that most contribute to this signal.

Secondly, median income had to be estimated where the data was missing. Because BGs with incomplete information on median income were a minority (maximum 24% for the Lille metropolitan area) and because only one variable among the 20 used in the indices had such missing data, incompleteness has probably little effect. A perspective for improvement could be to use more advanced techniques to handle missing data.

Thirdly, utilization of a large amount of data requires preparation and calculation before applying the procedure, which is time consuming. It also calls for technical know-how. This procedure is clearly more complex than number of other indices. We think this is the price to pay for a deeper analysis of SES and its determinants and a more

detailed interpretation of the results. While our index showed a high correlation with the Carstairs and Townsend indices, we think it allows more in depth analysis, when needed, and overcomes some of the limitations faced by between and within countries comparisons due to the low number and the nature of the variables than compose these well-known indices. Similar studies in other countries that allow usage of detailed socioeconomic information at BG levels would help assess the robustness of the procedure in other social contexts.

Fourthly, HC has no criteria regarding the size of the categories and so it can yield categories with very different sizes, which can be a limitation when linking their distribution with other attributes such as the prevalence of some health condition or of some exposure factor.

As a summary, a major strength of the procedure presented in this article is its versatility: it is not restricted to a particular set of data or type of study, and can be used for a large variety of contexts such as social epidemiology, environmental justice assessment, public health studies or urban and social planning. The application of this procedure on three large metropolitan areas shows high correlations with well-known indices like Townsend's and Carstairs', which appears to confirm that the created index represents the same socioeconomic notion. Although this procedure is more complicated than these other methods to create a SES index, the variables included in the final SES index allows a wider representation of the dimensions of SES, both to identify the best variables to distinguish BGs at the metropolitan area scale and to have better information on the particularities of the BGs. Then, it allows finer analysis of key determinants of health inequalities and reflection on local policies that would aim to cope with these inequalities. Another innovation in this study is the use of HC to constitute SES categories and compare them to the classically used quantiles. This approach allows having categories with more homogeneous compositions and which can consequently increase contrasts between them. Finally, we provide an R package able to reproduce the procedure easily. In conclusion, this procedure can be used to produce a SES index with a strong statistical basis and great scope for interpretation and relevance to public health bodies. The set of selected variables had a high proportion of common determinants of SES; they could also identify some features more specific to each area. Comparison of clustering methods showed that care should be taken to derive homogeneous categories.

Additional files

Additional file 1: Base Socioeconomic Characteristics of the Three Study Urban Areas.

Additional file 2: Principal component analysis and Hierarchical clustering.

Additional file 3: Correlations and Contributions of Variables to the First Component and Variance Explained by the First Component, According to the Study Area.

Additional file 4: Plot of city SES index vs. global index restricted to each city (housing census block groups only).

Additional file 5: Plot of SES index vs. Carstairs' index, according to the study area (housing census block groups only).

Additional file 6: Plot of SES index vs. Townsend's index, according to the study area (housing census block groups only).

Additional file 7: Average Values of the Common Variables per Category Created With HC for the Global Analysis.

Additional file 8: Maps of the socioeconomic index for Lille Metropole, in three categories by tertiles or optimal thresholds.

Additional file 9: Maps of the socioeconomic index for Grand Lyon, in three categories by tertiles or optimal thresholds.

Additional file 10: Maps of the socioeconomic index for Aix-Marseille urban area, in three categories by tertiles or optimal thresholds.

Abbreviations

BG: Census block group; HC: Hierarchical clustering; PCA: Principal component analysis; SES: Socio-economic status.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

BL has contributed to the creation of the method, implemented the procedure, created the R package, performed and interpreted the application, drafted the article and conducted the literature review. JMM has contributed to the creation of the method, provided statistical rigor, contributed to the interpretation of the results and helped to draft the manuscript. CP has collected socioeconomic data and helped to draft the manuscript. WK has constructed some socioeconomic variables, filled the missing data for median income and helped to draft the manuscript. NLM has given insight and expertise for the implementation and creation of the R package and helped to draft the manuscript. DZN guarantees quality assurance and helped to draft the manuscript. SD, head of Project Equit'Area research examining the role of environmental exposures on health inequalities, has followed up the general labor, has contributed to the definition of the method, interpretation of results, writing section and its finalization. All authors read and approved the final manuscript.

Acknowledgments

This work and the Equit'Area project are supported by the French National Research Agency (ANR, contract-2010-PRSP-002-01) and the EHESP School of Public Health. This research was also jointly supported by the Direction Générale de la Santé (DGS), the Caisse Nationale d'Assurance Maladie des Travailleurs Salariés (CNAMTS), the Régime Social des Indépendants (RSI), the Caisse Nationale de Solidarité pour l'Autonomie (CNSA), the Mission Recherche de la Direction de la Recherche, des Etudes, de l'Evaluation et des Statistiques (MiRe-DREES) and l'Institut national de prévention et de promotion de la santé (Inpes), under the research call launched by the French Institute of Public Health Research (IReSP) in 2010.

Author details

¹EHESP Rennes, Sorbonne Paris Cité, Rennes, France. ²Inserm, UMR IRSET Institut de recherche sur la santé l'environnement et le travail - 1085, Rennes, France. ³Lorraine University, CNRS, INRIA UMR 7502, Institut Elie Cartan, Lorraine, France. ⁴UMR936 INSERM, Université de Rennes 1, Rennes, France. ⁵Lorraine University, Medical School, Lorraine, France.

Received: 20 December 2012 Accepted: 17 March 2013

Published: 28 March 2013

References

1. Arntzen A, Samuelsen SO, Bakketeig LS, Stoltenberg C: Socioeconomic status and risk of infant death. A population-based study of trends in Norway, 1967-1998. *Int J Epidemiol* 2004, **33**:279-288.

2. Krieger N, Chen JT, Waterman PD, Soobader M-J, Subramanian SV, Carson R: **Choosing area based socioeconomic measures to monitor social inequalities in low birth weight and childhood lead poisoning: The Public Health Disparities Geocoding Project (US).** *J Epidemiol Community Health* 2003, **57**:186–199.
3. Singh GK, Kogan MD: **Persistent socioeconomic disparities in infant, neonatal, and postneonatal mortality rates in the United States, 1969–2001.** *Pediatrics* 2007, **119**:e928–e939.
4. Chaix B, Rosvall M, Merlo J: **Recent increase of neighborhood socioeconomic effects on ischemic heart disease mortality: a multilevel survival analysis of two large Swedish cohorts.** *Am J Epidemiol* 2007, **165**:22–26.
5. Deguen S, Lalloué B, Bard D, Havard S, Arveiler D, Zmirou-Navier D: **A small-area ecologic study of myocardial infarction, neighborhood deprivation, and sex: a Bayesian modeling approach.** *Epidemiology* 2010, **21**:459–466.
6. Ellison-Loschmann L, Sunyer J, Plana E, Pearce N, Zock J-P, Jarvis D, Janson C, Antó JM, Kogevinas M: **Socioeconomic status, asthma and chronic bronchitis in a large community-based study.** *Eur Respir J* 2007, **29**:897–905.
7. Prescott E, Godtfredsen N, Vestbo J, Osler M: **Social position and mortality from respiratory diseases in males and females.** *Eur Respir J* 2003, **21**:821–826.
8. Tassone EC, Waller LA, Casper ML: **Small-Area Racial Disparity in Stroke Mortality.** *Epidemiology* 2009, **20**:234–241.
9. Curtis S, Copeland A, Fagg J, Congdon P, Almog M, Fitzpatrick J: **The ecological relationship between deprivation, social isolation and rates of hospital admission for acute psychiatric care: a comparison of London and New York City.** *Health Place* 2006, **12**:19–37.
10. Tello JE, Jones J, Bonizzato P, Mazzi M, Amadeo F, Tansella M: **A census-based socio-economic status (SES) index as a tool to examine the relationship between mental health services use and deprivation.** *Soc Sci Med* 2005, **61**:2096–2105.
11. Marmot M: **Social determinants of health inequalities.** *Lancet* 2005, **365**:1099–1104.
12. Elliott P, Wakefield J, Best N, Briggs D: *Spatial epidemiology: methods and applications.* Oxford; New York: Oxford University Press; 2001.
13. Barceló MA, Saez M, Saurina C: **Spatial variability in mortality inequalities, socioeconomic deprivation, and air pollution in small areas of the Barcelona Metropolitan Region, Spain.** *Sci Total Environ* 2009, **407**:5501–5523.
14. Borrell C, Mari-Dell'Olmo M, Serral G, Martínez-Beneito M, Gotsens M: **Inequalities in mortality in small areas of eleven Spanish cities (the multicenter MEDEA project).** *Health Place* 2010, **16**:703–711.
15. Singh GK: **Area deprivation and widening inequalities in US mortality, 1969–1998.** *Am J Public Health* 2003, **93**:1137–1143.
16. Zeka A, Melly SJ, Schwartz J: **The effects of socioeconomic status and indices of physical environment on reduced birth weight and preterm births in Eastern Massachusetts.** *Environ Health* 2008, **7**:60.
17. Meijer M, Röhl J, Bloomfield K, Grittner U: **Do neighborhoods affect individual mortality? A systematic review and meta-analysis of multilevel studies.** *Soc Sci Med* 2012, **74**:1204–1212.
18. Braveman PA, Cubbin C, Egerter S, Chideya S, Marchi KS, Metzler M, Posner S: **Socioeconomic status in health research: one size does not fit all.** *JAMA* 2005, **294**:2879–2888.
19. Gordon D: **Census based deprivation indices: their weighting and validation.** *J Epidemiol Community Health* 1995, **49**(Suppl 2):S39–S44.
20. Krieger N, Williams DR, Moss NE: **Measuring social class in US public health research: concepts, methodologies, and guidelines.** *Annu Rev Public Health* 1997, **18**:341–378.
21. Galobardes B, Shaw M, Lawlor DA, Lynch JW, Davey Smith G: **Indicators of socioeconomic position (part 2).** *J Epidemiol Community Health* 2006, **60**:95–101.
22. Morris R, Carstairs V: **Which deprivation? A comparison of selected deprivation indexes.** *J Public Health Med* 1991, **13**:318–326.
23. Carstairs V: **Deprivation indices: their interpretation and use in relation to health.** *J Epidemiol Community Health* 1995, **49**(Suppl 2):S3–S8.
24. Carstairs V, Morris R: **Deprivation: explaining differences in mortality between Scotland and England and Wales.** *BMJ* 1989, **299**:886–889.
25. Jarman B: **Identification of underprivileged areas.** *Br Med J (Clin Res Ed)* 1983, **286**:1705–1709.
26. Eibner C, Sturm R: **US-based indices of area-level deprivation: results from HealthCare for Communities.** *Soc Sci Med* 2006, **62**:348–359.
27. Fukuda Y, Nakamura K, Takano T: **Higher mortality in areas of lower socioeconomic position measured by a single index of deprivation in Japan.** *Public Health* 2007, **121**:163–173.
28. Pampalon R, Hamel D, Gamache P, Raymond G: **A deprivation index for health planning in Canada.** *Chronic Dis Can* 2009, **29**:178–191.
29. Rey G, Jouglé E, Fouillet A, Hémon D: **Ecological association between a deprivation index and mortality in France over the period 1997–2001: variations with spatial scale, degree of urbanicity, age, gender and cause of death.** *BMC Public Health* 2009, **9**:33.
30. Salmond C, Crampton P, Sutton F: **NZDep91: A New Zealand index of deprivation.** *Aust N Z J Public Health* 1998, **22**:835–837.
31. Messer LC, Laraia BA, Kaufman JS, Eyster J, Holzman C, Culhane J, Elo I, Burke JG, O'Campo P: **The development of a standardized neighborhood deprivation index.** *J Urban Health* 2006, **83**:1041–1062.
32. Anthopoulos R, James SA, Gelfand AE, Miranda ML: **A spatial measure of neighborhood level racial isolation applied to low birthweight, preterm birth, and birthweight in North Carolina.** *Spat Spatiotemporal Epidemiol* 2011, **2**:235–246.
33. Bell N, Schuurman N, Hayes MV: **Using GIS-based methods of multicriteria analysis to construct socio-economic deprivation indices.** *Int J Health Geogr* 2007, **6**:17.
34. Pernet C, Delpierre C, Dejardin O, Grosclaude P, Launay L, Guittet L, Lang T, Launoy G: **Construction of an adaptable European transnational ecological deprivation index: the French version.** *J Epidemiol Community Health* 2012, **66**(11):982–989.
35. Mari-Dell'Olmo M, Martínez-Beneito MA, Borrell C, Zurriaga O, Nolasco A, Domínguez-Berjón MF: **Bayesian factor analysis to calculate a deprivation index and its uncertainty.** *Epidemiology* 2011, **22**:356–364.
36. Townsend P, Phillimore P, Beattie A: *Health and deprivation: inequality and the North.* London: Routledge; 1988.
37. Carstairs V, Morris R: **Deprivation and mortality: an alternative to social class?** *J Public Health* 1989, **11**:210–219.
38. Havard S, Deguen S, Bodin J, Louis K, Laurent O, Bard D: **A small-area index of socioeconomic deprivation to capture health inequalities in France.** *Soc Sci Med* 2008, **67**:2007–2016.
39. Lê S, Josse J, Husson F: **FactoMineR: An R package for multivariate analysis.** *J Stat Software* 2008, **25**:1–18.
40. R Development Core Team: *R: A language and environment for statistical computing.* Vienna Austria: R Foundation for Statistical Computing; 2011.

doi:10.1186/1475-9276-12-21

Cite this article as: Lalloué *et al.*: A statistical procedure to create a neighborhood socioeconomic index for health inequalities analysis. *International Journal for Equity in Health* 2013 **12**:21.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



Mortalité infantile, défaveur et proximité aux industries polluantes : une analyse spatiale conduite à fine échelle (agglomération de Lille, France)

CINDY PADILLA^{1,2}

BENOIT LALLOUÉ^{1,2}

DENIS ZMIROU-NAVIER^{1,2,3,4}

DEGUEN SÉVERINE^{1,2}

¹ EHESP (École des hautes études en santé publique), Rennes, Sorbonne Paris Cité
Avenue du Pr Léon-Bernard
35043 Rennes cedex, France
<cindy.padilla@ehesp.fr>
<sylvaine.cordier@univ-rennes1.fr>

² Institut de recherche en santé environnement travail
Rennes
France

³ Inserm U954 Vandœuvre-les-Nancy
France
<Denis.Zmirou@nancy.inserm.fr>

⁴ Université de Nancy, Vandœuvre-les-Nancy
France
<helene.primerano@uhp-nancy.fr>

Tirés à part :
C. Padilla

Article reçu le 14 décembre 2010,
accepté le 21 mars 2011

Résumé. L'objectif de l'étude est d'explorer la relation entre mortalité infantile, défaveur et proximité aux industries polluantes à Lille entre 2000 et 2009. L'unité géographique est l'Iris (îlots regroupés pour l'information statistique), la plus petite unité pour laquelle des données de population sont disponibles. Les informations sur les cas de mortalité infantile ont été recueillies auprès des mairies. L'adresse de résidence a permis de leur affecter leur Iris de résidence. Les industries polluantes ont été sélectionnées à partir de la base européenne des industries polluantes (EPER, *European Pollutant Emission Register/Registre européen des émissions de polluants*). Un indice de défaveur a été estimé à l'échelle de l'Iris à partir des données de l'Insee du recensement de 1999. Une analyse bayésienne prenant en compte à la fois la forte variabilité des taux de mortalité dans chaque Iris ainsi que l'autocorrélation spatiale a permis de déterminer la relation avec la déprivation et la proximité aux industries. Les résultats montrent que le risque de mortalité infantile est multiplié par 1,53 (IC 95 % [1,33-1,76]) pour une augmentation d'une unité de la défaveur. Aucune association significative n'a pu être mise en évidence avec la proximité aux industries polluantes. Notre résultat confirme, à une échelle géographique fine, l'influence du niveau socio-économique sur la mortalité infantile en France. La finesse de notre unité géographique n'apparaît pas adaptée à l'indicateur de proximité aux industries polluantes élaboré.

Mots clés : défaveur ; disparités d'état sanitaire ; indicateurs d'état sanitaire ; inégalités ; mortalité infantile ; théorème de Bayes.

Abstract

Infant mortality, deprivation and proximity to polluting industrial facilities – A small-scale spatial analysis with census data (Lille Metropolitan Area, France)

The objective of the study was to investigate the association between infant mortality, deprivation and proximity to polluting industries in Lille between 2000 and 2009. The geographical unit used was the smallest French census block: the Iris. Information on infant deaths was collected from local authorities. These data include the family address, so that they can be assigned to their Iris of residence. Polluting industries were identified from the European database of polluting industries (EPER). A deprivation index was calculated across the Iris, from the Insee census data of 1999. Bayesian analysis, taking both the high variability in mortality rates in each Iris and the spatial autocorrelation into account, made it possible to determine the relation with deprivation and industries. The results show that the risk of infant mortality was multiplied by 1.53 (95% CI [1.33-1.76]) for a unit increase of deprivation, no significant association was observed with proximity to polluting industries. Our results confirmed the influence of socioeconomic status on infant mortality across the smallest scale area for which data were available in France. No association was observed with proximity to polluting industries, probably because the small scale of our units was not appropriate for the index of proximity to the polluting facilities that we used.

Key words: Bayes theorem; deprivation; health status disparities; health status indicators; inequalities; infant mortality.

Pour citer cet article : Padilla C, Lalloué B, Zmirou-Navier D, Séverine D. Mortalité infantile, défaveur et proximité aux industries polluantes : une analyse spatiale conduite à fine échelle (agglomération de Lille, France). *Env Risque Sante* 2011 ; 10 : 216-21. doi : 10.1684/ers.2011.0455

Le constat des inégalités sociales de santé (ISS) est aujourd'hui solidement établi dans la plupart des pays industrialisés. Globalement, les populations socio-économiquement défavorisées sont davantage touchées que les populations les plus aisées par de nombreux et divers problèmes de santé (cancers, asthme, maladies cardiovasculaires [1, 2]). Récemment, ce même constat d'inégalités sociales relatif à la mortalité ou la mortalité périnatale a été rapporté dans la littérature [3, 4].

Selon les territoires d'étude, ces ISS peuvent présenter des caractéristiques très différentes. La mortalité infantile, reconnue pour être un indicateur de l'état de santé globale des populations, présente des variations importantes entre les territoires français, et ce à toutes les échelles : régionale, départementale et infra-départementale. Seule, une analyse à une résolution spatiale fine conduit à une compréhension approfondie de la répartition géographique des inégalités de santé, élément essentiel pour révéler les inégalités au niveau local souvent masqué par des estimations à grande échelle comme la ville ou la région [5].

Parmi les facteurs susceptibles d'expliquer ces inégalités, plusieurs déterminants sociaux ont été proposés. Les facteurs de risques individuels (alcoolisme, tabagisme, sédentarité, obésité, diabète...), connus de la littérature, seraient plus prévalents parmi les populations les plus défavorisées. N'expliquant cependant qu'une partie des ISS observées, d'autres facteurs tels que les facteurs contextuels [6, 7] (c'est-à-dire qui ne sont pas dus à l'individu mais que l'individu subit) tels que l'accès inégal au système de soins, à la prise en charge et aux traitements les plus adaptés, l'environnement de vie, ont été avancés.

En dépit de nombreux facteurs de risques déjà identifiés, une partie des inégalités sociales de santé demeure, à ce jour, inexpliquée. Dans ce contexte, les nuisances environnementales ont été suspectées comme pouvant contribuer à ces inégalités [8, 9]. En effet, parmi la population générale, les groupes socio-économiquement défavorisés pourraient être à la fois plus exposés à des sources polluantes et plus susceptibles aux effets sanitaires en résultant, en raison d'un état sanitaire comparativement dégradé et d'un moindre accès aux soins précoces. À travers ces deux mécanismes pouvant agir de manière indépendante ou combinée, les populations défavorisées pourraient souffrir de plus forts effets sanitaires de l'exposition à des facteurs de risque environnementaux.

L'objectif de cette étude est d'étudier la relation entre les variabilités géographiques des inégalités de mortalité infantile et la défaveur socioéconomique et environnementale à une échelle fine de l'agglomération de Lille, en France, au cours de la période 2000-2009.

Matériel et méthode

Cette étude s'intègre dans le programme de recherche EQUITAREA que nous développons au sein de l'École

des hautes études en santé publique (EHESP). Son objectif général est d'étudier la contribution des expositions environnementales et d'un cumul de celles-ci aux inégalités sociales de santé en France. Il s'agit d'une étude écologique conduite à l'échelle des îlots regroupés pour l'information statistique (Iris, regroupant en moyenne 2 000 habitants). L'agglomération lilloise est composée de 85 communes et de 506 Iris.

Les décès d'enfants de moins d'1 an ont été recueillis auprès des mairies de l'agglomération de Lille et l'adresse de résidence recueillie a permis de leur affecter leur Iris de résidence.

Le niveau socio-économique des Iris a été estimé par l'indicateur de défaveur socio-économique [10, 11]. Il a été construit à partir des données du recensement de l'Insee de 1999. Il est composé de 19 variables qui reflètent la dimension matérielle et sociale de la défaveur. L'indice a été introduit dans les modèles statistiques sous deux formes : en quantitatif et en cinq classes approximativement équivalentes aux quintiles de la distribution (la catégorie 1 représentant les Iris les plus favorisés et la catégorie 5 les Iris les plus défavorisés). La répartition spatiale de l'indicateur de défaveur montre un fort contraste à l'échelle de l'Iris. Ces contrastes opposent les Iris les plus démunis qui sont concentrés dans les grandes villes de l'agglomération, notamment Lille, Roubaix et Tourcoing. Les Iris les plus privilégiés sont au contraire en périphérie (figure 1).

La proximité des industries polluantes a été déterminée à partir des données du registre européen des industries polluantes (EPER, *European Pollutant Emission Register/Registre européen des émissions de polluants*). La base EPER 2004 est le premier registre européen des émissions rejetées dans l'air et l'eau par les grandes et moyennes installations industrielles dépassant un seuil d'émissions par type de polluants. Il couvre 50 polluants différents et contient des données provenant de tous les États membres (<http://www.eper.cec.eu.int/>). Sur l'agglomération de Lille, au total, 52 industries polluantes dépassant le seuil autorisé ont été enregistrées. Un indice de proximité a été estimé pour chaque Iris en cumulant les distances entre les industries localisées dans l'agglomération et le centroïde de chaque Iris.

Pour estimer les risques relatifs de la mortalité infantile liée au statut socio-économique à l'échelle du quartier, nous avons utilisé un modèle hiérarchique bayésien proposé par Besag *et al.* [12]. Ce modèle nous permet de prendre en compte la variabilité due au faible nombre de cas dans chaque unité géographique et de considérer l'information fournie par les valeurs des unités géographiques voisines. En effet, ce modèle contrôle à la fois l'hétérogénéité spatiale et la composante d'auto-corrélation spatiale de nos données. Les modèles ont été réalisés à l'aide du logiciel WinBUGS version 1.4.1 (MRC Biostatistics Unit Cambridge, Royaume Uni) Nous avons introduit le niveau de pauvreté du quartier dans le modèle de différentes façons : directement auprès de l'indice quanti-

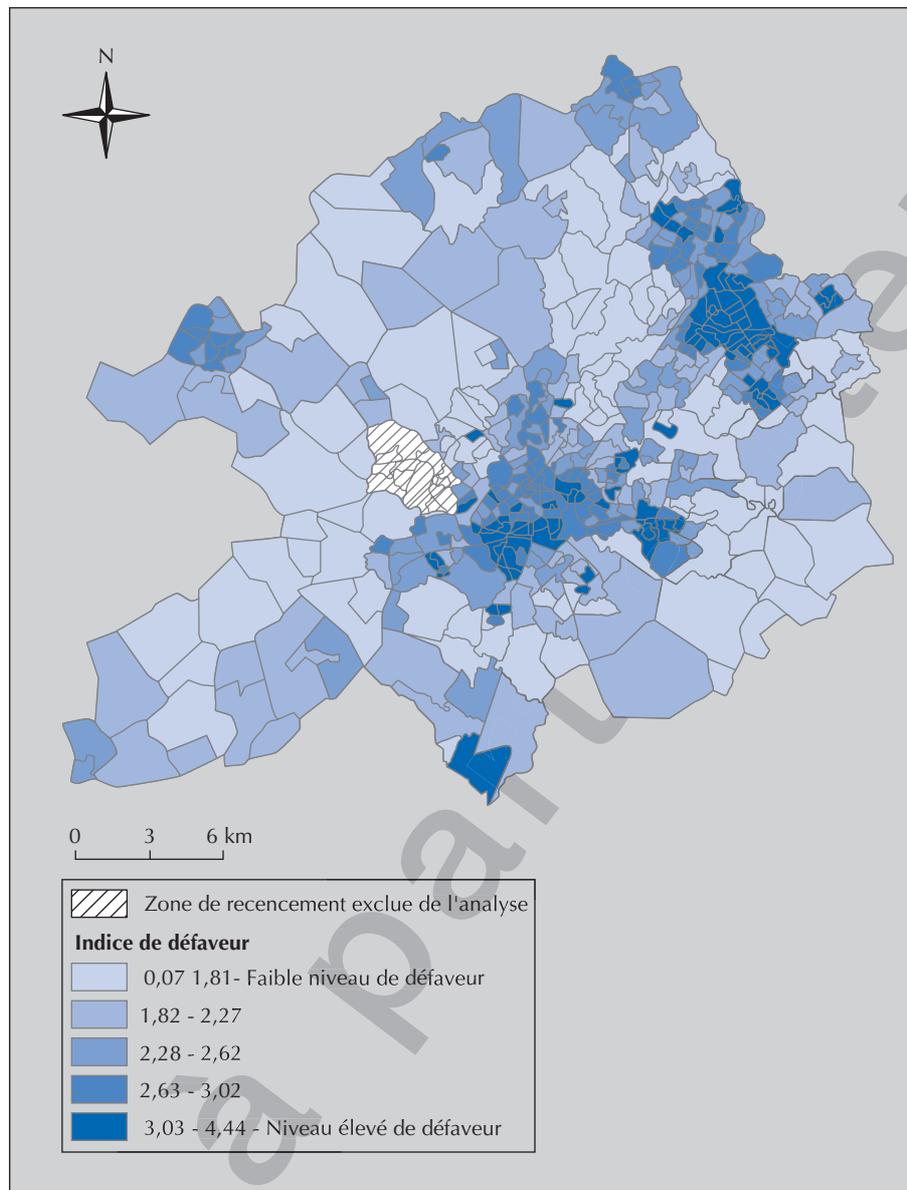


Figure 1. Indice de défaveur de l'agglomération de Lille à l'échelle des îlots regroupés pour l'information statistique.

Figure 1. Deprivation Index in the Lille metropolitan area, by French census block.

tatif de statut économique et social (SES) et comme covariables correspondant aux catégories de défaveur exceptée la première (la plus privilégiée) utilisée comme référence.

Résultats

L'agglomération de Lille est caractérisée par un taux de mortalité infantile moyen pour la période 2000-2009 de

5,1 décès d'enfants de moins de 1 an pour 1 000 naissances vivantes (écart type = 7,2) et une médiane de 3,8 décès pour 1 000 naissances vivantes (figure 2).

Quel que soit le modèle mis en œuvre, les résultats montrent une influence significative de la défaveur sur la mortalité infantile. Plus précisément, il existe un clair gradient du risque de mortalité infantile de la catégorie la plus privilégiée (catégorie 1, servant de référence) à la catégorie des plus démunis : catégorie 5. En effet, pour les femmes résidant dans les Iris les plus défavorisés, le risque de mortalité infantile est multiplié

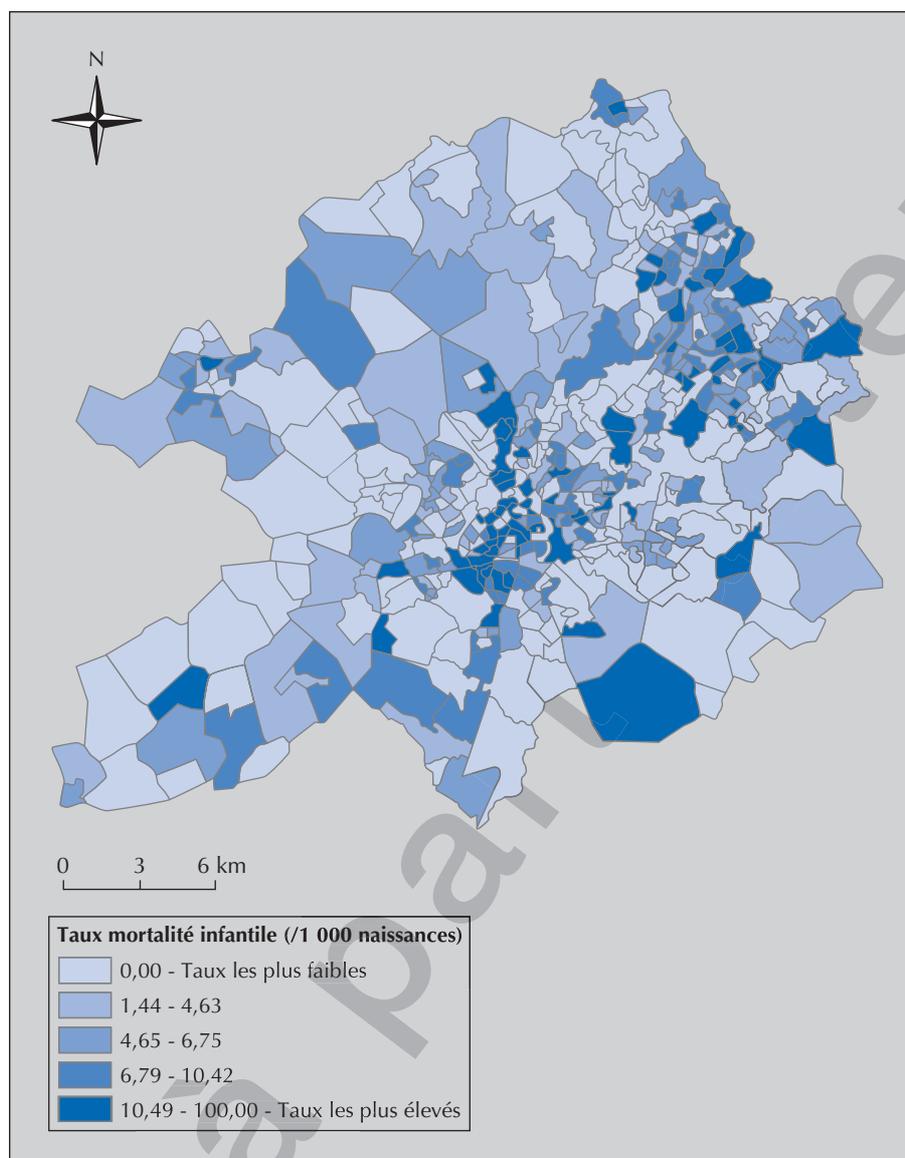


Figure 2. Taux de mortalité infantile de l'agglomération de Lille à l'échelle des îlots regroupés pour l'information statistique.

Figure 2. Infant mortality rate in the Lille metropolitan area by French census block.

par 2,8 par rapport aux femmes résidant dans les Iris les plus favorisés (OR = 2,83; IC 95 % = [2,02-3,99]) (figure 1).

En revanche, aucune association significative avec la proximité aux industries polluantes, y compris en tenant compte de possibles interactions avec le niveau socio-économique n'a été mise en évidence (tableau 1) Cependant, on peut constater que les populations les plus défavorisées seraient localisées à plus grande proximité des industries polluantes que les populations plus aisées (tableau 2).

Discussion et conclusion

Notre étude confirme l'existence d'inégalités sociales de santé à l'échelle géographique fine. En effet, nous montrons que le risque de mortalité infantile est inversement proportionnel au statut socio-économique de l'Iris de résidence des mères. En revanche, nous n'avons pas démontré de lien significatif avec la proximité aux industries polluantes à l'échelle de l'Iris malgré qu'il a été constaté que les populations les plus défavorisées seraient

Tableau 1. Risque relatif (RR) de la mortalité infantile selon l'indice de défaveur.*Table 1. Relative risk (RR) of infant mortality by deprivation index.*

Modèles	Catégories de défaveur	RR [95 % IC]
Indice de défaveur	SES	1,55 [1,33-1,78]
Niveau de défaveur	C1 : classe moins défavorisée	1,00 ^a
	C2	1,41 [0,99-2,00]
	C3	1,82 [1,31-2,57]
	C4	2,39 [1,70-3,37]
	C5 : classe plus défavorisée	2,59 [1,88-3,61]

^a Catégorie de référence ; SES : statut économique et social.

Tableau 2. La population la plus défavorisée vit plus près des industries polluantes que la population plus aisée.*Table 2. The most deprived population lives closer to industrial facilities than the more affluent ones.*

Classes de l'indice de défaveur	Moyenne géométrique (km)	km [95 % IC]
Classe 1 : [0,07-1,81]	559	[533-586]
Classe 2 : [1,82-2,27]	557	[527-582]
Classe 3 : [2,28-2,62]	561	[536-586]
Classe 4 : [2,63-3,02]	544	[519-571]
Classe 5 : [3,03-4,44]	499	[483-516]

localisées à plus grande proximité des industries polluantes que les populations plus aisées.

La principale limite de notre étude tient à l'indicateur utilisé pour mesurer la proximité des populations aux industries polluantes. En effet, celui-ci n'est probablement pas adapté à la finesse de l'unité géographique. Une synthèse de la littérature [13] a permis d'identifier trois grandes catégories d'indicateurs de proximité aux industries. La proximité des populations aux installations industrielles émettrices de polluants ou classées « à risque industriel » sera dorénavant analysée à l'aide de la méthode des *buffers*.

Un autre inconvénient de notre étude est le manque d'informations notamment pour les émissions des industries polluantes. En effet, la base de données EPER n'enregistre que les industries polluantes dépassant un seuil fixé par la réglementation européenne et ce seuil est déjà très élevé. De nombreuses industries polluantes ne sont donc pas prises en compte dans notre étude. Dans le cadre d'une collaboration avec l'Institut national de l'environnement industriel et des risques (Ineris), les données collectées par l'Inventaire national spatialisé, qui regroupent toutes les industries quel que soit leur seuil ainsi que les émissions polluantes par ces industries, seront prises en compte dans une prochaine analyse.

En revanche, notre étude a été réalisée à l'échelle la plus fine possible compte tenu de la disponibilité des données et utilise un indice de défaveur validé par d'autres études à Strasbourg [10, 11, 14, 15]. À notre connaissance, aucune étude française n'a été réalisée à l'échelle de l'Iris pour étudier les variations spatiales de la

mortalité infantile. De plus, nous avons utilisé une méthodologie statistique adaptée au « design » et à la rareté de l'événement sanitaire étudié. La prise en compte de l'auto-corrélation spatiale est importante et renforce notre résultat. Cette méthode a été rarement utilisée dans les études épidémiologiques en santé publique.

En conclusion, notre étude a confirmé l'existence de disparités socio-spatiales du risque de mortalité infantile. De nouveaux développements méthodologiques sont actuellement en cours pour construire un indicateur de proximité des populations aux industries polluantes qui prendrait en compte les émissions de celles-ci, plus adapté à la finesse de notre analyse.

Ce projet, dans sa globalité, a pour but d'explorer les interactions entre trois familles de variables environnementales, socio-économiques et événements de santé, qui ne sont pas souvent considérées simultanément dans les études épidémiologiques. Alors que l'étude des inégalités sociales de santé ou des effets des facteurs de risque environnementaux sur la santé est maintenant courante dans la littérature de santé publique, l'évaluation de leur effet conjoint ne l'est pas et pourrait éclairer les politiques publiques visant à réduire les risques et les inégalités. ■

Remerciements et autres mentions

Financement : Agence nationale de la recherche (ANR), direction générale de la santé (DGS), École des hautes études en santé publique (EHESP), Agence de l'environnement et de la maîtrise de l'énergie (Ademe), Région Nord Pas-de-Calais ; **conflits d'intérêts :** aucun.

Références

1. Dalstra JA, Kunst AE, Borrell C, *et al.* Socioeconomic differences in the prevalence of common chronic diseases: an overview of eight European countries. *Int J Epidemiol* 2005 ; 34 : 316-26.
2. Menvielle G, Kunst AE, Stirbu I, *et al.* Socioeconomic inequalities in alcohol related cancer mortality among men: to what extent do they differ between Western European populations? *Int J Cancer* 2007 ; 121 : 649-55.
3. Singh GK, Kogan MD. Persistent socioeconomic disparities in infant, neonatal, and postneonatal mortality rates in the United States, 1969-2001. *Pediatrics* 2007 ; 119 : e928-39.
4. Smith GC, Fretts RC. Stillbirth. *Lancet* 2007 ; 370 : 1715-25.
5. Barcelo MA, Saez M, Saurina C. Spatial variability in mortality inequalities, socioeconomic deprivation, and air pollution in small areas of the Barcelona Metropolitan Region, Spain. *Sci Total Environ* 2009 ; 407 : 5501-23.
6. Costa G, D'errico A. Inequalities in health: do occupational risks matter? *Eur J Public Health* 2006 ; 16 : 340.
7. Siegrist J, Marmot M. Health inequalities and the psychosocial environment-two scientific challenges. *Soc Sci Med* 2004 ; 58 : 1463-73.
8. Evans GW, Kantrowitz E. Socioeconomic status and health: the potential role of environmental risk exposure. *Annu Rev Public Health* 2002 ; 23 : 303-31.
9. O'Neill MS, Jerrett M, Kawachi I, *et al.* Health, wealth, and air pollution: advancing theory and methods. *Environ Health Perspect* 2003 ; 111 : 1861-70.
10. Deguen S, Lalloue B, Bard D, Havard S, Arveiler D, Zmirou-Navier D. A small-area ecologic study of myocardial infarction, neighborhood deprivation, and sex: a Bayesian modeling approach. *Epidemiology* 2010 ; 21 : 459-66.
11. Havard S, Deguen S, Bodin J, Louis K, Laurent O, Bard D. A small-area index of socioeconomic deprivation to capture health inequalities in France. *Soc Sci Med* 2008 ; 67 : 2007-16.
12. Besag J, York J, Mollie A. Bayesian image restoration with two applications in spatial statistics. *Ann I Stat Math* 1991 ; 43 : 1-21.
13. Deguen S, Paulus H, Zmirou-Navier D. Indicators of residential proximity to industrial facilities, a literature review dealing with proximity index. *Environ Int* 2011.
14. Havard S, Deguen S, Zmirou-Navier D, Schillinger C, Bard D. Traffic-related air pollution and socioeconomic status: a spatial autocorrelation study to assess environmental equity on a small-area scale. *Epidemiology* 2009 ; 20 : 223-30.
15. Laurent O, Bard D, Filleul L, Segala C. Effect of socioeconomic status on the relationship between atmospheric pollution and mortality. *J Epidemiol Community Health* 2007 ; 61 : 665-75.

Abstract : An exploratory spatial analysis to assess the relationship between deprivation, noise and infant mortality (Soumis à Environmental Health)

Wahida Kihal^{1*}, Cindy.M Padilla^{1,2}, Benoit Lalloué^{1,2,3}, Christophe Rougier⁴, Jérôme Defrance⁴, Denis Zmirou-Navier^{1,2,3}, Séverine Deguen^{1,2}

¹ EHESP School of Public Health–Rennes-France

² INSERM U1085-IRSET – Research Institute of environmental and occupational health. Rennes, France

³ Lorraine University–Nancy-France

⁴ CSTB Scientific and technical Center for Building– Saint-Martin-d’Hères- France

Background: Few studies explored how noise could contribute to social health inequalities, and even less considered infant mortality or its risk factors as the health event of interest. We investigate in this paper the impact of neighbourhood characteristics, both the socioeconomic status and ambient noise levels, on the spatial distribution of infant mortality in the Lyon metropolitan area, France.

Methods: All infant deaths (n=715) that occurred between 2000 and 2009 were geocoded at the census block level. Each census block was assigned multi-component socioeconomic characteristics and Lden levels, a measure of exposure to noise. Using a spatial–scan statistic, we examined whether there were significant clusters of high risk of infant mortality according to these neighborhood characteristics.

Results: Our results highlight that infant mortality is not randomly spatially distributed, with clusters of high risk in the southeast of the Lyon metropolitan area (RR=1.44 ; p= 0.09). After adjustment on socioeconomic characteristics and noise levels, this cluster disappears or shifts according to different scenarios, suggesting that noise and socioeconomic characteristics explain part of the spatial distribution of infant mortality.

Conclusion: Our findings show an impact of noise on the spatial distribution of mortality after adjusting on socio-economic characteristics. Because of the scarcity of studies, it is difficult to compare our findings with others. The results seem plausible in view of 3 hypothetical not exclusive pathways we propose in a conceptual framework: (i) a psychological pathway, (ii) a physiological disruption process and (iii) an unhealthy behaviors pathway. These findings require further research for confirmation and interpretation.

Key words: noise exposure, neighborhood deprivation, infant mortality, spatial analysis

Abstract : Green space, social inequalities and neonatal mortality in France (Soumis à BMC Pregnancy and Childbirth)

Wahida Kihal W¹, Cindy Padilla C^{1,2}, Benoît Lalloué^{1,2,3}, Marcello Gelormini¹, Denis Zmirou-Navier^{1,2,3}, Séverine Deguen^{1,2}

¹ EHESP School of Public Health–Rennes, Sorbonne-Paris Cité, France

² INSERM U1085-IRSET – Research Institute of environmental and occupational health. Rennes, France

³ Lorraine University–Vandoeuvre-les-Nancy-France

Background. Few studies have considered using environmental amenities to explain social health inequalities.

Nevertheless, Green spaces that promote good health may have an effect on socioeconomic health inequalities. In developed countries, there is considerable evidence that green spaces have a beneficial effect on the health of urban populations and recent studies suggest they can have a positive effect on pregnancy outcomes.

Objective. To investigate the relationship between green spaces and the spatial distribution of infant mortality taking account neighborhood deprivation levels.

Methods. The study took place in Lyon metropolitan area, France. All infant deaths that occurred between 2000 and 2009 were geocoded at census block level. Each census block was assigned greenness and socioeconomic deprivation levels. The spatial–scan statistic was used to identify high risk cluster of infant mortality according to these neighborhood characteristics.

Results. The spatial distribution of infant mortality was not random with a high risk cluster in the south east of the Lyon metropolitan area ($p < 0.003$). This cluster disappeared ($p = 0.12$) after adjustment for greenness level and socioeconomic deprivation, suggesting that these factors explain part of the spatial distribution of infant mortality. These results are discussed using a conceptual framework with 3 hypothetical pathways by which green spaces may have a beneficial effect on adverse pregnancy outcomes: (i) a psychological pathway, (ii) a physiological disruption process and (iii) an environmental pathway.

Conclusions. These results add some evidence to the hypothesis that there is a relationship between access to green spaces and pregnancy outcomes but further research is required to confirm this.

Key words: Greenness level, neighborhood deprivation, infant mortality, spatial analysis

Abstract : Data Analysis technics, A tool for Cumulative Exposure Assessment (Soumis à Exposure Science and Environmental Epidemiology journal)

Benoît Lalloué^{1,2,3,4}, Jean-Marie Monnez^{3,4}, Cindy Padilla^{1,2}, Wahida Kihal¹, Denis Zmirou-Navier^{1,2,5}, Séverine Deguen^{1,2}

1 EHESP Rennes, Sorbonne Paris Cité, France

2 Inserm, UMR 1085-IRSET (Institut de recherche sur la santé l'environnement et le travail), France

3 Lorraine University, CNRS UMR 7502, Institut Elie Cartan, France

4 Lorraine University, INRIA, CNRS UMR7502, BIGS (INRIA Nancy - Grand Est / IECN), France

5 Lorraine University Medical School, France

Background: Everyone is subject to environmental exposures from various sources, with negative health impacts (air, water and soil contamination, noise...) or with positive effects (e.g. green space). Studies considering such complex environmental settings in a global manner are rare.

Aims: We propose to use data mining techniques to create a composite exposure index with a data-driven approach, in view to assess the environmental burden experienced by populations. We illustrate this approach on a large French metropolitan area.

Methods: The study was carried out in the Great Lyon area (France, 1,260,348 inhabitants in 2007, 527 km²) at the census block group (BG) scale. Indicators on NO₂ annual concentrations, noise levels, proximity to green spaces and to industrial plants, and road traffic were synthetized using Multiple Factor Analysis (MFA), which allows to explore the relations between exposures and BGs without a priori knowledge, and to synthetize indicators of different types. Hierarchical clustering was then used to create BG classes.

Results: The four first components of the MFA explained respectively 23, 14, 13 and 12% of the total variance. Clustering in 3 classes group: 1) BGs far from industries, greener and with less noise and air pollution than the average; 2) BGs close to industries with less green spaces and air pollution than the average; and 3) BGs far from industries but with higher levels of noise, air pollution and traffic. Greater numbers of classes were tested in order to assess a variety of clustering.

Conclusions: We present a data driven approach using data mining techniques, which seem overlooked for cumulative exposure assessment in complex environmental settings. Although it cannot be applied directly for risk or health effect assessment, the resulting index can help to identify hot spots of cumulative exposure, to prioritize urban policies or to compare the environmental burden across study areas in an epidemiological framework.

Certificat de décès (bulletin n°7)

**Ce bulletin doit être adressé
à la Direction Départementale des Affaires Sanitaires et Sociales**

RÉPUBLIQUE FRANÇAISE
INSTITUT NATIONAL DE LA STATISTIQUE
ET DES ÉTUDES ÉCONOMIQUES

BULLETIN DE DÉCÈS

N° de l'acte _____

N° d'ordre
de l'avis
de décès _____

Département _____ Commune (1) _____

Commune (2) _____

Section _____

Date du décès _____

 Jour Mois Année

7

A. RENSEIGNEMENTS RELATIFS AU DÉFUNT

3 **Sexe** Masculin 1 Féminin 2

4 **Né(e) le** Jour, mois, année _____

5 **À** Commune (2) _____

Département ou pays (3) _____

6 **Profession** Retraité(e) 1

Inactif(ve) 2

Actif(ve) (y compris demandeur d'emploi) .. 3 → préciser profession et situation

profession _____

(si demandeur d'emploi, indiquer l'ancienne profession)

situation (4) ... Salarié(e) de l'État Autre salarié(e) À son compte

ou des collectivités locales

7 **Nationalité** Française 1

Étrangère 2 → préciser _____

8 **Domicile** Commune (2) _____

Département ou pays (3) _____

9 **État matrimonial** Célibataire 1 Veuf(ve) 3

Marié(e) 2 Divorcé(e) 4

B. RENSEIGNEMENTS RELATIFS AU DÉCÈS

10 **Lieu du décès** Logement ou domicile 1 Hospice, maison de retraite 4

Établissement hospitalier 2 Voie ou lieu public 5

Clinique privée 3 Autre lieu 6

11 **Le certificat médical confidentiel de décès a-t-il été fourni ?** OUI NON

(1) Numéro de la commune au Code officiel géographique.

(2) Pour Paris, Lyon, Marseille, indiquer l'arrondissement.

(3) Département métropolitain : code ou nom en clair.
DOM, TOM ou pays étranger : nom en clair.

(4) Salarié(e) de l'État ou des collectivités locales :
- y compris les hôpitaux publics, les PTT, les arsenaux ... ;
- non compris les entreprises publiques (SNCF, EDF, banques, etc.).
Autre salarié(e) : y compris les entreprises publiques et les organismes de Sécurité sociale.

Fait le,
(date de création du bulletin)
L'officier de l'état civil.

NATIONALE 1 01206 1

Résumé

Inégalités sociales de santé et expositions environnementales. Une analyse spatio-temporelle du risque de mortalité infantile et néonatale dans 4 agglomérations françaises, 2000-2009.

En France, la réduction des inégalités sociales de santé est un objectif explicite des politiques de santé depuis une dizaine d'années. La mortalité infantile et la mortalité périnatale sont reconnues comme des indicateurs pertinents pour caractériser l'état de santé général d'une population. En dépit de taux globalement bas, des disparités spatiales persistent en termes de niveaux et d'évolutions récentes. Cette mortalité précoce est sensible aux effets de la pollution atmosphérique. Pourtant peu d'études ont exploré la contribution des nuisances environnementales aux inégalités socio-spatiales de santé. Dans ce contexte, l'objectif de la thèse est d'analyser par une approche spatio-temporelle, la contribution de l'exposition au dioxyde d'azote, indicateur de qualité de l'air, aux inégalités sociales de mortalité infantile et néonatale en France entre 2000 et 2009.

Nous avons effectué une étude épidémiologique de type écologique en utilisant l'IRIS comme unité géographique dans les agglomérations de Lille, Paris, Lyon, et Marseille. Tous les cas ont été recueillis à partir des registres de décès des mairies et ont été géocodés en utilisant l'adresse de résidence des parents. Les données socioéconomiques et démographiques ont été estimées à partir des recensements de 1999 et 2006. Un indice composite qui englobe de multiples dimensions a été utilisé pour analyser la défaveur socioéconomique globale à l'échelle du quartier. Les concentrations moyennes de dioxyde d'azote ont été modélisées par les réseaux de surveillance de la qualité de l'air. Des modèles statistiques additifs généralisés ont permis de prendre en compte l'autocorrélation spatiale et de générer des cartes à l'aide de lissage sur la longitude et la latitude tout en ajustant sur les variables d'intérêt.

Les résultats mettent en évidence l'existence d'inégalités socio-spatiales de mortalité infantile et néonatale et d'inégalité environnementale. Plus précisément, nous avons trouvé des zones de sur-risque de mortalité infantile et néonatale en relation avec la défaveur socio-économique et la qualité de l'air ambiant. La force et la direction des associations varient entre les villes, démontrant ainsi la difficulté de généraliser ces observations au niveau national ; ces résultats sont ville-spécifique. Ces associations varient également selon la période d'étude et l'évènement sanitaire étudié. Cette thèse illustre une approche innovante pour identifier les zones qui cumulent des conditions défavorables. Ce domaine encore vierge pourrait éclairer les politiques publiques sur les zones à prioriser dans le but de réduire les risques et les inégalités socio-territoriales de qualité des milieux de vie et de santé.

Mots-clés : mortalité infantile, dioxyde d'azote, inégalités, défaveur socio-économique, analyse spatiale, fine échelle

Abstract

Social inequalities in health and environmental exposures. A spatio-temporal analysis of the risk of infant and neonatal mortality in four French metropolitan areas, 2000-2009.

In France, reducing social health inequalities has become an explicit objective of health policy for the last decade. Infant and neonatal mortality are recognized as indicators of the health status of a population. Characterized by generally low rates, infant mortality presents large spatial variations that change over time. This early mortality is sensitive to air pollution insults. Yet, time and geographical patterns of this disproportionate distribution of environmental and social burden remain poorly assessed. In this context, the thesis objectives were to analyze by a spatial and temporal approach, the contribution of exposure to nitrogen dioxide to social inequalities in infant and neonatal mortality in France between 2000 and 2009.

We conducted an ecological type epidemiological study using the French census block as the geographical unit in the metropolitan areas of Lille, Paris, Lyon, and Marseille. All cases were collected in the cities deaths registries and were geocoded using address of parent's residence. Socioeconomic and demographic data were estimated from the 1999 and 2006 census. A composite index which encompasses multiple dimensions was used to analyze global deprivation at the neighborhood scale. Average nitrogen dioxide concentrations were modeled by the air quality monitoring networks. Generalized additive models allowed to take into account spatial autocorrelation and generate maps using smoothing on longitude and latitude while adjusting for covariates of interest.

The results highlight the existence of socio-spatial inequalities in infant and neonatal mortality and of environmental disparities. Specifically, we found areas of elevated risk depending on infant and neonatal mortality in relation with socioeconomic deprivation and ambient air quality. The strength and direction of the associations varied between cities demonstrating the difficulty to generalize these observations at the national level; these results are city-specific. These associations also vary according to the period and the health event. This thesis illustrates an innovative approach to identify areas which cumulate adverse conditions. This approach could be a tool for evaluation of social policies and / or urban development to reduce environmental and health inequalities.

Key words : infant mortality, nitrogen dioxide, inequality, deprivation, spatial analysis, small area level