

Inserm



Institut national
de la santé et de la recherche médicale

Unité de recherche en épidémiologies et
sciences de l'information (U444)

Equipe de recherche sur les déterminants
sociaux de la santé et du recours aux soins

Les déterminants contextuels et spatiaux des modes de recours aux soins

Basile Chaix, Pierre Chauvin

Rapport scientifique à l'Action Concertée Incitative
« Terrains, techniques, théories ; travail interdisciplinaire en
Sciences humaines et sociales »

Décembre 2004

Résumé du travail

Depuis près de dix ans, l'épidémiologie sociale s'intéresse à l'impact que les caractéristiques du contexte de résidence peuvent avoir sur la santé des individus et sur leurs modes de recours aux soins, au-delà des effets imputables à leurs caractéristiques socio-économiques personnelles. Dans le cadre de ce projet, nous cherchons à avancer dans la connaissance des déterminants contextuels de la santé et du recours aux soins, qui ont reçu nettement moins d'attention en France qu'en Europe du Nord, en Angleterre, ou aux Etats-Unis.

Un de nos objectifs était de réfléchir sur les outils à mettre en œuvre pour décrire et expliquer les variations spatiales des phénomènes de santé et de recours aux soins, et de développer de nouvelles approches d'analyse permettant de combler les lacunes des méthodes actuellement utilisées dans ce champ de l'épidémiologie sociale. Ces développements méthodologiques ont été à la fois suscités par et mis en application dans deux études appliquées, l'une conduite à partir de données Suédoises, l'autre à partir de données Françaises.

Dans un premier temps, étudiant les variations spatiales de troubles mentaux liés à la consommation de substances psycho-actives dans la ville de Malmö en Suède à partir de données issues du Registre de Population, les modèles multiniveaux nous ont permis de voir que les variations d'un quartier à l'autre dans la prévalence des troubles étaient importantes. Se démarquant des pratiques d'analyse suivies par beaucoup d'auteurs, nous avons souligné l'intérêt qu'il y a à quantifier et modéliser les variations inter-quartiers des phénomènes de santé lorsque l'on cherche à évaluer l'importance du contexte pour ces phénomènes. Nous avons trouvé qu'une grande partie des variations spatiales des troubles mentaux considérés était liée aux variations du niveau socio-économique, tant mesuré au niveau individuel qu'au niveau des quartiers de résidence : au-delà de l'effet imputable au niveau socio-économique individuel, la prévalence des troubles était supérieure dans les quartiers pauvres.

Nous en sommes finalement venu à mettre en doute la pertinence de l'approche d'analyse multiniveau utilisée de façon quasi-hégémonique dans la littérature d'analyse contextuelle. En effet, fragmentant le territoire en une multitude de zones administratives et négligeant les connections spatiales qui existent entre ces zones, l'approche multiniveau ne fournit souvent que

des informations incomplètes sur la distribution spatiale des phénomènes de santé. Au-delà, mesurant les facteurs explicatifs du contexte de résidence au niveau de zones administratives arbitraires, elle s'avère souvent incapable de capter adéquatement les effets du contexte sur la santé. A partir des données Françaises de l'enquête SPS, nous avons étudié les variations des modes de recours aux soins sur le territoire métropolitain Français, et avons trouvé qu'une approche d'analyse qui tient compte de l'espace dans sa continuité intrinsèque permettait mieux de décrire et d'expliquer les variations spatiales des phénomènes.

Introduction

1) Utilité de l'analyse contextuelle en santé publique

Le champ de l'épidémiologie s'est longtemps inscrit dans le paradigme de l'individualisme méthodologique, qui postule que les facteurs influant sur la santé des personnes appartiennent au registre des caractéristiques individuelles.^{1,2} En suivant cette orientation d'analyse, on serait capable d'appréhender l'ensemble des processus agissant sur la santé des individus en tenant compte de leurs caractéristiques démographiques, sociales, psychologiques, anatomiques, biologiques, etc.

Au contraire, de nombreux travaux issus des sciences sociales ont cherché à mettre en évidence l'influence que le contexte de vie des individus peut avoir sur la santé.^{2,3} L'idée s'est ainsi progressivement formée dans le champ de l'épidémiologie sociale que les déterminants sociaux de la santé ont par nature une structure à niveaux (ou multiniveau), appartenant au niveau individuel, mais également au niveau du ménage, du lieu de résidence, ou du lieu de travail ou d'étude.⁴⁻⁸ En conséquence, il est aujourd'hui largement reconnu qu'une voie importante à suivre en épidémiologie sociale pour avancer dans la connaissance des mécanismes à l'origine des disparités sociales de santé est de s'intéresser aux effets du contexte, et notamment à ceux du contexte résidentiel.^{4,9}

Au-delà des objectifs de connaissance, il est important d'un point de vue de santé publique de tenir compte des relations qui existent entre le contexte de vie des individus et leur santé. En effet, les études d'analyse contextuelle offrent des perspectives nouvelles dans le champ de la santé publique, d'une part en décrivant les variations géographiques des phénomènes de santé, et d'autre part en affinant la compréhension que l'on a des mécanismes à l'origine des disparités de santé.

A – La description des variations géographiques des phénomènes de santé

Prendre en compte la dimension contextuelle des phénomènes de santé consiste d'abord à examiner si ceux-ci présentent des variations sur le territoire d'étude. Etant incapable d'identifier

des variations spatiales aux différentes échelles d'analyse considérées, on serait amené à conclure que le phénomène étudié ne présente pas de dimension contextuelle, et que sa variabilité est imputable à des facteurs mesurables au niveau individuel.^{10,11} Au contraire, si les méthodes mises en œuvre indiquent une variabilité géographique importante, le phénomène devient un objet d'intérêt en analyse contextuelle, qui cherche alors à en décrire et expliquer la distribution spatiale.¹² D'un point de vue de santé publique, la quantification des variations contextuelles des phénomènes indique si d'éventuels programmes d'information ou d'intervention doivent intégrer cette dimension contextuelle, ou si ces programmes peuvent être mis en œuvre de façon complètement invariante sur le territoire.¹² La description cartographique des disparités de santé ou de comportements relatifs à la santé aide également à identifier les zones d'intervention prioritaires et à répartir les ressources sur le territoire en tenant compte des besoins différenciés d'un endroit à l'autre.

B – La compréhension des mécanismes à l'origine des disparités géographiques de santé

Au-delà de la simple description des disparités territoriales de santé, l'objectif est d'avancer dans la compréhension des mécanismes qui les produisent. L'approche contextuelle s'est développée à partir du constat qu'il est nécessaire d'utiliser des données collectées au niveau individuel pour avancer dans la compréhension des déterminants sociaux de la santé.¹³ L'objectif de ce genre d'analyses est d'examiner si les variations géographiques identifiées sont intégralement liées à la composition variable des zones considérées en terme de caractéristiques individuelles, ou si elles résultent également d'effets proprement contextuels qui ne sauraient être captés au niveau individuel.^{14,15}

Ainsi que de nombreux auteurs l'ont indiqué, il est utile en santé publique d'examiner si les facteurs du contexte de résidence sont associés aux problèmes de santé après avoir tenu compte des facteurs démographiques et sociaux au niveau individuel.¹⁶ L'intérêt est de voir si l'on peut se contenter de cibler les programmes d'intervention sur la base des caractéristiques des individus, ou si l'on doit au-delà également tenir compte des caractéristiques des zones de résidence. L'idée avancée est qu'en cas d'effets directs des caractéristiques du contexte sur la santé des individus, la cible des programmes de santé publique manquerait d'inclure un nombre

important d'individus à risque si elle n'était définie que sur la base des facteurs de risque individuels.

Au-delà de la distinction entre effets de composition et effets contextuels, l'objectif est d'examiner quelles dimensions du contexte de résidence jouent sur la santé des individus.^{17,18} Cela permet d'adapter au mieux les programmes d'intervention aux mécanismes causaux identifiés. Concernant par exemple la pratique d'activités sportives, que l'on sait être liée au niveau socio-économique des individus,^{19,20} diverses études ont mis en évidence des variations significatives d'un quartier de résidence à l'autre.²¹ Les auteurs ont cherché à voir si ces variations spatiales étaient simplement dues à la composition variable des zones sur le plan des caractéristiques socio-économiques individuelles. Au-delà, ils ont trouvé que ces variations étaient en partie imputables au niveau socio-économique du quartier de résidence, mesuré en agrégeant les caractéristiques des individus.^{21,22} Un tel effet pourrait être dû au fait que les valeurs et habitudes comportementales d'un groupe social donné tendent à prévaloir dans les endroits où il est majoritaire, affectant ainsi l'ensemble des résidents, même si ils n'appartiennent pas à ce groupe social. Enfin, prenant en compte d'autres facteurs contextuels, les auteurs ont également pu montrer que la présence d'installations sportives et d'endroits où la marche ou la course peuvent être pratiquées en toute sécurité avait une influence sur la pratique sportive.²³

Quantifier les variations contextuelles des phénomènes, chercher à les expliquer en distinguant effets de composition et effets proprement contextuels, et avancer dans la connaissance des différents processus par lesquels le contexte influe sur la santé présentent donc un intérêt en santé publique.

2) Evaluer l'importance des effets du contexte sur la santé : l'importance de la question méthodologique

A – L'approche multiniveau en analyse contextuelle

Puisque les déterminants sociaux de la santé appartiennent à différents niveaux, la variabilité des phénomènes de santé présente une structure hiérarchique : au-delà de la variabilité qui existe entre individus d'un même groupe, une partie des variations survient d'une unité contextuelle à l'autre, l'individu et son contexte constituant des sources de variabilité distinctes et

hiérarchiquement organisées.²⁴ Concernant les méthodes d'analyses, les approches qui ne tiennent pas compte de cette structure complexe de la variabilité peuvent s'avérer en partie inefficaces. Afin de décrire et d'expliquer les variations de phénomènes qui opèrent à différents niveaux, la littérature d'épidémiologie sociale recourt aujourd'hui aux modèles multiniveaux (incluant des effets aléatoires au-delà des effets fixes²⁴⁻²⁷) ou dans une moindre mesure à des modèles basés sur l'équation d'estimation généralisée.²⁸⁻³¹ L'approche d'analyse multiniveau telle qu'elle est mise en œuvre aujourd'hui en épidémiologie sociale, s'est formée au cours des années 1980 par la réunion du courant d'analyse contextuelle et de la tradition statistique d'utilisation des modèles mixtes.²⁴ L'approche d'analyse multiniveau a d'abord été utilisée dans le champ des sciences de l'éducation.³² Cette approche n'a été utilisée dans le champ de l'épidémiologie sociale dans l'étude des effets du contexte de résidence sur la santé qu'à partir des années 1990, commençant vraiment à s'y établir au milieu de la décennie.^{2,33-35}

Cependant, diverses tendances se font jour dans la littérature dans l'utilisation qui est faite des modèles multiniveaux. Une première approche, suivie dans les études pionnières de la littérature d'analyse contextuelle ainsi que plus récemment, consiste à s'intéresser exclusivement aux mesures d'association entre facteurs contextuels et variables de santé individuelles.^{3,36,37} Dans ce cas, l'intérêt des modèles multiniveaux est de tenir compte de la structure hiérarchique des données (individus regroupés au sein de zones de résidence) lors de la procédure d'estimation des paramètres, et d'aboutir ainsi à des écart-types des forces d'association qui prennent en compte la corrélation intra-zone de la variable réponse.

Une telle utilisation des modèles multiniveaux apparaît en fait restrictive. Au cours d'une revue de littérature publiée dans la Revue d'Epidémiologie et de Santé Publique,³⁸ nous avons montré que la prise en compte des effets aléatoires des modèles fournit des informations utiles à l'interprétation des associations entre facteurs contextuels et phénomènes de santé.^{5,14,15,39,40}

Toutefois, une telle utilisation des effets aléatoires comme simples appuis dans l'interprétation des associations entre facteurs explicatifs et phénomènes de santé peut encore apparaître limitée. L'objet d'un travail coordonné par Juan Merlo de l'Hôpital Universitaire de Malmö en Suède auquel nous avons participé est de souligner que de tels effets aléatoires fournissent en eux-mêmes des informations importantes en santé publique sur les variations géographiques des phénomènes, sous la forme d'indicateurs que l'on appelle « mesures de

variation » par opposition aux « mesures d'association » classiques.^{12,41,42} Il apparaît que l'utilité de ces « mesures de variation » (telles que le coefficient de corrélation intraclasse ou coefficient de partition de la variance) a été sous-estimée dans la littérature. Nous avons participé de façon significative à une série de trois articles didactiques publiés dans le *Journal of Epidemiology and Community Health*, dont l'objectif était de clarifier l'intérêt respectif des mesures d'association (issues des effets fixes du modèle multiniveau) et des mesures de variation (issues des effets aléatoires). Ces deux types d'indicateurs fournissent des informations complémentaires permettant de juger de l'importance réelle du contexte pour la santé.^{12,41,42} C'est à ces questions que s'est intéressée la première des analyses que nous avons conduite (rapportée dans le premier chapitre de ce document), qui s'est attachée à l'étude des variations spatiales de troubles liés à la consommation de substances psycho-actives dans la ville de Malmö en Suède.

B – Comparaison de l'approche multiniveau et d'une perspective d'analyse spatiale dans l'étude des effets du contexte

Sur le plan méthodologique, nos travaux se sont toutefois orientés dans le sens d'une critique de l'approche multiniveau utilisée de façon quasi-hégémonique dans la littérature d'analyse contextuelle en épidémiologie sociale.³ Au travers de l'étude des variations géographiques des modes de recours aux soins en France, nous en sommes venu à l'idée que l'approche d'analyse multiniveau, du fait de sa conception de l'espace, ne fournit pas des informations optimales sur la variabilité spatiale des phénomènes de santé. L'approche multiniveau conçoit en effet l'espace comme fragmenté en zones distinctes le plus souvent définies à partir des limites administratives. La littérature géographique sur le «modifiable areal unit problem» a depuis longtemps montré que les résultats des analyses qui s'appuient sur un zonage administratif du territoire sont largement dépendants du découpage utilisé.⁴³⁻⁴⁶ L'effet d'agrégation qui intervient est d'une part dû à des phénomènes d'échelle, puisque les zones peuvent être définies à un niveau plus ou moins local («scale effect»). D'autre part, à une échelle donnée, les frontières considérées peuvent grouper les individus d'une multitude de façons différentes («zoning effect»). En conséquence, tant les indicateurs qui quantifient les variations d'une zone à l'autre que les mesures des effets du contexte sont dépendants du découpage en zones utilisé, et des différences

importantes dans les résultats peuvent être observées si d'autres découpages du territoire sont utilisés.⁴⁷

Au-delà de cette dépendance des indicateurs au découpage utilisé, une limite plus importante des modèles multiniveaux est de ne pas tenir compte des relations spatiales entre les zones, et de supposer que des individus provenant de zones différentes sont complètement indépendants même si ces zones sont adjacentes ou proches. En négligeant cette possible corrélation entre zones proches sur le territoire, les modèles multiniveaux ne permettent pas d'obtenir des informations optimales sur la distribution spatiale des phénomènes : ils ne renseignent que sur la force de la corrélation des phénomènes de santé à l'intérieur des zones, mais pas sur la portée de cette corrélation dans l'espace.

Au-delà de l'insuffisance des indicateurs qui décrivent les variations spatiales des phénomènes, une autre limite de l'approche multiniveau est de systématiquement définir les facteurs du contexte au niveau des zones administratives de résidence des individus. Or, rien n'assure *a priori* que les différents effets contextuels opèrent réellement au niveau des zones administratives considérées.¹⁷ Dans bien des cas, il se pourrait que les individus soient également affectés par les caractéristiques du contexte au-delà des limites administratives de leur zone de résidence, puisque leurs activités quotidiennes les amènent probablement à se déplacer dans cet espace élargi.

Ces différentes limites de l'approche d'analyse multiniveau sont liées à sa conception d'un espace fragmenté en zones administratives arbitraires déconnectées les unes des autres. Du fait de cette définition de l'espace, tant les mesures de variation dont il a été question ci-dessus que les mesures d'association entre facteurs contextuels et phénomènes de santé s'avèrent en partie inefficaces à rendre compte de la distribution spatiale des phénomènes de santé. Dans notre étude des variations géographiques des phénomènes de santé, nous avons cherché à développer une perspective spatiale d'investigation, qui s'appuie sur une conception continue de l'espace. Une première caractéristique de cette approche est de s'appuyer sur des modèles de régression spatiaux, qui quoique différents les uns des autres, ont pour point commun de ne pas fragmenter l'espace en zones déconnectées les unes des autres.⁴⁸⁻⁵⁴ En appliquant la notion de « mesure de variation » définie dans le cadre du modèle multiniveau, un de nos objectifs est de souligner que les modèles de régression spatiaux aboutissent à des indicateurs qui fournissent plus

d'informations sur la distribution spatiale des phénomènes que ceux que l'on obtient à partir des modèles multiniveaux.

Le second intérêt d'une approche spatiale en analyse contextuelle est de s'affranchir des limites administratives lors de la définition des facteurs contextuels explicatifs.⁵⁵⁻⁵⁷ Nous avons développé des méthodes de mesure de l'exposition aux caractéristiques du contexte qui tiennent compte de l'information contextuelle dans un espace continu centré sur le lieu de résidence des individus. Un avantage de ces approches est qu'elles parviennent certainement mieux à capter les effets du contexte environnant que les mesures réalisées au niveau des zones administratives pour les individus qui résident sur les marges de ces zones.⁵⁸

Sur un plan méthodologique, l'objectif de cette étude des variations géographiques des modes de recours aux soins était de comparer l'approche multiniveau couramment utilisée dans la littérature à cette perspective d'analyse spatiale. Nous avons cherché à voir si le fait de tenir compte de l'espace dans sa continuité intrinsèque permettait d'obtenir des informations sur la distribution spatiale des phénomènes, tant à partir des mesures classiques d'association qu'à partir des mesures de variation, qui resteraient inaccessibles dans le cadre de l'approche multiniveau qui fragmente le territoire en zones administratives déconnectées les unes des autres. Sur le plan des résultats, nous avons cherché à voir si les caractéristiques du milieu de résidence (en terme de niveau socio-économique ou d'accès aux médecins) étaient associées aux modes de recours aux soins, au-delà des effets imputables aux facteurs individuels (nous rapportons cette étude dans le second chapitre de ce document).

Chapitre 1 : Disparités géographiques des troubles liés à la consommation de substances psycho-actives dans la ville de Malmö en Suède

Introduction

Jusqu'à une période récente, la plupart des études en épidémiologie sociale se sont intéressées aux déterminants sociaux de la santé mesurés au niveau individuel.^{1,14} De telles études fournissent des informations utiles aux décideurs de santé publique, qui permettent de cibler les différents programmes d'intervention sur la base des caractéristiques sociales des individus. Toutefois, ainsi que le montre la littérature épidémiologique d'analyse contextuelle,^{2,3,59-61} un certain nombre de problèmes de santé pourrait aussi être liés aux caractéristiques du milieu de résidence, souvent appréhendé au niveau du quartier.⁴ En conséquence, différents auteurs ont souligné que l'étude des effets du contexte local de résidence constitue une des voies importantes par laquelle il est possible d'accroître notre connaissance des déterminants sociaux de la santé.

La plupart des études qui s'intéressent aux relations entre quartier de résidence et santé ont tenu compte du niveau socio-économique du quartier. Des études Nord Européennes et Nord Américaines ont par exemple montré qu'au-delà des effets liés au niveau socio-économique individuel, le fait de résider dans un quartier défavorisé était associé à différents problèmes de santé tels qu'avoir un faible poids à la naissance,^{18,49} présenter un risque accru de se blesser durant l'enfance,^{13,62} avoir des comportements qui endommagent la santé,^{34,59,61,63-65} être en surpoids,⁶⁶⁻⁶⁸ contracter des maladies cardiovasculaires,⁶⁹ et finalement mourir.⁷⁰ Dans la présente étude, nous avons cherché à voir si les troubles mentaux avaient eux-mêmes une dimension contextuelle, et si les éventuelles variations de prévalence entre quartiers étaient liées au niveau socio-économique du quartier.

Dans la littérature, un certain nombre d'études ont cherché à voir si il existait des variations d'un quartier à l'autre dans l'utilisation de médicaments psychiatriques. Par exemple, une étude Hollandaise a rapporté des différences entre quartiers dans l'utilisation de benzodiazépines.⁷¹ Plus récemment, une étude Suédoise conduite dans la ville de Malmö a trouvé qu'il existait des variations d'importance modeste entre quartiers dans la propension à recourir à des médicaments anxiolytiques-hypnotiques.⁷² Cette analyse a indiqué qu'une partie des variations inter-quartiers

pouvait être due au niveau de participation sociale qui prévaut dans chaque quartier, au-delà de l'effet associé à diverses caractéristiques socio-économiques individuelles. Au-delà d'analyses s'intéressant à l'utilisation de médicaments, un certain nombre d'études ont exploré les variations de prévalence de troubles mentaux entre quartiers.^{33,73-76} Par exemple, une étude Anglaise basée sur des données d'enquête a rapporté des variations de morbidité psychiatrique faibles mais statistiquement significatives entre quartiers.³³ Toutefois, il est apparu que ces variations étaient en fait imputables à la composition variable des quartiers en terme de caractéristiques individuelles. Une autre étude, conduite en Hollande dans la ville d'Amsterdam, a abouti à des conclusions similaires :⁷⁵ les troubles mentaux étaient plus fréquents dans les quartiers défavorisés, mais cela était en fait dû à la concentration d'individus au statut socio-économique défavorisé dans de tels quartiers plutôt qu'à des effets proprement contextuels. Au contraire, une étude conduite dans la ville de Maastricht dans le même pays a trouvé que la prévalence de troubles mentaux non-psychotiques et non-organiques était légèrement plus élevée dans les quartiers défavorisés, après que l'on ait tenu compte dans le modèle de facteurs socio-économiques individuels.⁷⁷ De même, une étude conduite en Angleterre a indiqué que les résidents des quartiers défavorisés avaient une santé mentale légèrement plus dégradée que les habitants des quartiers plus aisés.⁷⁶

Il apparaît que l'on peut résumer cette littérature en concluant qu'il existe en général des variations faibles entre quartiers dans la prévalence de troubles mentaux, une large part de ces variations étant due à la composition variable des quartiers en terme de caractéristiques socio-économiques individuelles. Par ailleurs, dans ces études passées, les caractéristiques socio-économiques du quartier de résidence n'avaient aucune utilité ou seulement une utilité modérée pour comprendre les variations de santé mentale entre individus. Au total, les preuves réunies dans la littérature ne permettent pas réellement de conclure que le contexte de résidence a une importance cruciale sur la santé mentale.

Néanmoins, dans la présente étude, nous avons fait l'hypothèse qu'il pourrait y avoir une certaine hétérogénéité dans les troubles mentaux par rapport au fait de savoir si ils dépendent ou non des caractéristiques du quartier de résidence. Par exemple, les troubles mentaux organiques pourraient être complètement indépendants du contexte de résidence, alors que les problèmes mentaux liés à la consommation de substances psycho-actives dépendraient fortement des caractéristiques du quartier. Les résultats selon lesquels il n'existe qu'une influence modérée du

quartier de résidence sur les troubles mentaux pourraient en fait dissimuler une forte dépendance au contexte pour certaines catégories de troubles. Par conséquent, dans la présente étude, nous nous sommes intéressés à la distribution spatiale des troubles mentaux et comportementaux liés à la consommation de substances psycho-actives dans la ville de Malmö en Suède, et avons cherché à voir si le quartier de résidence avait un impact sur cette catégorie spécifique de troubles mentaux.

Une force de l'étude est que nous avons utilisé des données sur l'ensemble des individus âgés de 40 à 59 ans résidant dans la ville de Malmö en 2001. Ces données incluaient des informations de source administrative sur l'utilisation des soins, et indiquaient notamment si un trouble mental lié à la consommation de substances psycho-actives avait été diagnostiqué en 2001 pour chaque individu. Ces données étaient toutefois de nature transversale. En conséquence, dans la présente étude, nous n'étions pas en mesure d'étudier les facteurs causaux d'incidence de troubles mentaux, mais étions simplement capables d'identifier des facteurs associés à une prévalence accrue de troubles. Toutefois, ainsi que cela a été souligné ci-dessus, de telles études sont malgré tout utiles en santé publique, permettant aux décideurs de santé publique d'identifier des groupes sociaux ou des zones où la prévalence de problèmes de santé est accrue.

De façon plus spécifique, nos objectifs de recherche étaient les suivants :

1) Nous avons d'abord examiné si il existait des variations entre quartiers dans la prévalence de troubles mentaux et comportementaux liés à la consommation de substances psycho-actives. Nous avons utilisé des modèles multiniveaux afin de quantifier l'amplitude des variations entre quartiers,^{14,24,27} ainsi qu'un indicateur récemment proposé, l'odds ratio médian. L'avantage de cet indicateur est qu'il quantifie les variations géographiques des phénomènes de santé mesurés au niveau individuel sur l'échelle des odds ratios, qui est largement utilisée en épidémiologie, permettant ainsi d'évaluer plus facilement l'importance qu'il y a à tenir compte du quartier de résidence pour comprendre les variations de santé au niveau individuel.^{78,79} Un de nos objectifs ici est de souligner qu'au-delà des mesures classiques d'association entre facteurs du contexte et variables de santé, la mesure des variations entre quartiers fournit des informations utiles quand on s'intéresse aux relations entre contexte et santé.^{12,41,42,80,81} Au-delà, dans la présente étude, nous avons tracé des cartes de la ville de Malmö afin d'identifier visuellement les zones où la prévalence des troubles mentaux liés à la consommation de substances psycho-actives était accrue.

2) Une seconde étape a consisté à examiner si ces variations spatiales de prévalence de troubles mentaux étaient liées à des effets de composition, soit au fait que les individus présentent en moyenne des caractéristiques variables d'un endroit à l'autre.^{61,82}

3) Troisièmement, afin d'aller plus avant dans la connaissance de ces variations inter-quartiers, nous avons cherché à voir si le niveau socio-économique du quartier de résidence était associé à une prévalence accrue de troubles, après que l'on ait tenu compte de différentes caractéristiques des individus qui y vivaient. Nous avons eu recours à un indicateur qui a été récemment proposé pour le modèle logistique, l'interval odds ratio, ainsi qu'un indicateur associé que nous proposons (sorting out index), afin de voir si le facteur socio-économique contextuel permet réellement d'identifier les quartiers où la prévalence de troubles liés à la consommation de substances psycho-actives est accrue.^{78,79}

4) Enfin, nous avons cherché à voir si les individus qui cumulaient divers désavantages sociaux, mesurés au niveau individuel ou au niveau du quartier (solitude, pauvreté individuelle, résidence dans un quartier pauvre), présentaient des prévalences particulièrement accrues de troubles mentaux liés à la consommation de substances psycho-actives.

Méthodes

Sources de données

Le bureau régional du Comté de Scanie en Suède nous a fourni des données sur l'ensemble des 65 830 individus âgés de 40 à 59 ans dans la ville de Malmö en 2001. Ces données ont été extraites de la base de données d'Allocation de Ressources. La base de données originale comprenait des informations sur l'ensemble des contacts réalisés auprès de médecins privés ou publics en 2001, incluant certaines informations sur les diagnostics et les coûts. En tenant compte des trois premiers diagnostics posés à chaque séance (codés selon la Classification Internationale des Maladies, 10^{ème} version, CIM-10), le bureau régional de Scanie a défini des variables binaires indiquant si chaque individu avait ou non reçu des diagnostics dans différents groupes de problèmes de santé.

Mesures

Dans la présente étude, la variable réponse binaire considérée renvoie à la présence ou à l'absence de troubles mentaux ou comportementaux liés à la consommation de substances psycho-actives (CIM-10, codes F10-F19). Les substances potentiellement en jeu incluent l'alcool,

les opiacés, les cannabinoïdes, les sédatifs ou les hypnotiques, la cocaïne, les hallucinogènes, et le tabac. Les problèmes cliniques pouvant donner lieu à un diagnostic dans cette catégorie sont une intoxication aiguë, une utilisation nocive pour la santé, un syndrome de dépendance, un syndrome de sevrage, des troubles psychotiques, et un syndrome amnésique. Dans la base principale utilisée pour les analyses, nous n'avons pas d'autre information que la seule variable binaire indiquant la présence ou l'absence d'un diagnostic dans la catégorie F10-F19. Par contre, nous avons accès à ces informations détaillées dans une base séparée qui ne pouvait pas être liée à la base principale pour des raisons de confidentialité. Nous avons utilisé ces données détaillées pour obtenir des informations sur la distribution des troubles mentaux liés à la consommation de substances psycho-actives en diverses sous-catégories.

Comme variables individuelles, nous avons tenu compte du sexe, de l'âge, de l'état matrimonial, du niveau d'instruction, et du revenu des individus. L'âge a été divisé en quatre classes (40–44, 45–49, 50–54, et 55–59). Quant à l'état matrimonial, nous avons distingué les individus mariés ou en cohabitation d'une part, et les personnes célibataires, divorcées, ou veuves d'autre part. Le niveau d'instruction a été codé en deux groupes (neuf années d'instruction ou moins, plus de neuf années). Le revenu du ménage n'était pas disponible dans les données ; à la place, nous avons tenu compte du revenu individuel, codé en deux classes en utilisant le revenu médian comme valeur seuil. Le niveau d'instruction et le revenu individuel ont été pris en compte pour définir de façon approximative la position sociale des individus.⁸³

La ville de Malmö est divisée en 100 quartiers administratifs. La surface médiane de ces quartiers est de 0,5 kilomètre carré. En 2001, le nombre médian d'habitants dans un quartier était égal à 2046, et le nombre médian d'habitants âgés de 40 à 59 ans était de 510. Afin d'estimer le niveau socio-économique des quartiers,⁸⁴ nous avons calculé le revenu moyen des individus de 25 ans et plus dans les 100 quartiers de la ville. Il est en général recommandé de définir le niveau socio-économique du contexte à partir d'une variable qui est également prise en compte dans le modèle au niveau individuel, afin de capturer un effet socio-économique qui opère au niveau des quartiers. La variable socio-économique contextuelle a été divisée en quatre classes, en utilisant les quartiles de la distribution.

Analyses statistiques

Afin d'examiner si il existait des associations entre les variables individuelles divisées en classes et la prévalence des troubles mentaux liés à la consommation de substances psycho-actives, nous

avons utilisé le test habituel du chi-2. Afin de voir si la prévalence des troubles augmentait régulièrement avec le niveau de pauvreté du quartier, nous avons eu recours au test de tendance de Cochran-Armitage.⁸⁵

Nous avons ensuite conduit des analyses de régression. Comme souvent dans les études contextuelles, nos données avaient une structure hiérarchique, les individus étant groupés au sein des 100 quartiers administratifs de la ville de Malmö. Dans de telles études, l'hypothèse habituelle d'indépendance entre individus requise par les techniques de régression classiques est violée. En effet, il est courant que les individus qui résident dans un même quartier présentent en moyenne des caractéristiques de santé plus similaires que ne le font des individus qui proviennent de quartiers différents.^{40,61,72,86,87} Les modèles qui ne tiennent pas compte de cette corrélation des individus à l'intérieur des quartiers aboutissent à des mesures d'incertitude biaisées pour les effets que l'on cherche à mesurer. En particulier, il est bien connu que de tels modèles tendent à sous-estimer les écart-types des paramètres des effets contextuels, et conduisent souvent à conclure à l'existence d'un effet du contexte sur la santé alors qu'il n'y en a en fait pas. Au contraire, les modèles multiniveaux tiennent compte de la corrélation des individus à l'intérieur des quartiers au cours de la procédure d'estimation, et aboutissent comme tels à des écart-types non biaisés des paramètres,^{24-27,38,88} ce qui constitue une première raison de recourir à ces modèles. Toutefois, ainsi que discuté en détail dans d'autres publications^{12,41,42,80} et expliqué ci-dessous, une seconde utilité des modèles multiniveaux en analyse contextuelle est qu'ils permettent de quantifier les variations entre quartiers des phénomènes de santé, et d'évaluer ainsi l'importance que le contexte a pour les phénomènes étudiés.

Dans la présente étude, la variable réponse était de nature binaire. Nous avons par conséquent eu recours à un modèle logistique multiniveau. Supposons que Y_{ij} soit une variable binaire indiquant si chaque individu a reçu ou non un diagnostic de trouble mental ou comportemental lié à la consommation de substances psycho-actives, que X_{ij} soit un vecteur de variables explicatives, et que B soit le vecteur des coefficients correspondants. Afin d'étudier les variations de la probabilité p_{ij} d'avoir un trouble mental, nous avons estimé un modèle logistique multiniveau :

$$\text{logit}(p_{ij}) = \beta_0 + BX_{ij} + u_{0j}$$

$$u_{0j} \sim N(0; \sigma_{u0}^2)$$

Dans cette équation, γ_0 est l'intercept du modèle. La spécificité du modèle multiniveau est d'inclure un terme résiduel supplémentaire (u_{0j}), qui fournit des informations sur la prévalence spécifique des troubles liés à la consommation de substances psycho-actives dans le quartier j .⁸⁹ Toutefois, dans ce modèle, la procédure d'estimation n'estime pas directement un terme u_{0j} pour chaque quartier, mais suppose que les u_{0j} s sont normalement distribués et fournit la variance $\sigma_{u_0}^2$ de leur distribution. Ainsi que nous l'avons noté ci-dessus, au-delà de l'intérêt qu'il y a à obtenir des écart-types des paramètres qui tiennent compte de la dépendance des observations à l'intérieur des groupes, les modèles multiniveaux fournissent des informations pertinentes en analyse contextuelle.^{12,41,80} la variance $\sigma_{u_0}^2$ des résidus de niveau quartier va nous permettre de mesurer l'amplitude des variations entre quartiers dans la prévalence des troubles mentaux. Cette information aide à évaluer si le contexte résidentiel a ou non de l'importance pour le phénomène étudié. Le modèle multiniveau permet également d'estimer si ces variations entre quartiers sont statistiquement supérieures à 0, ce qui aide à conclure si le phénomène de santé étudié présente ou non une dimension contextuelle.¹²

Toutefois, dans le cas du modèle logistique, la variance inter-quartier $\sigma_{u_0}^2$ est exprimée sur l'échelle logistique. En conséquence, ce paramètre est difficile à interpréter sur le plan épidémiologique.⁹⁰ De ce fait, des approches ont été proposées dans la littérature statistique pour exprimer ces variations entre quartiers d'une façon qui soit plus facilement interprétable. D'une part, certains auteurs ont suggéré de recourir aux modèles de l'équation d'estimation généralisée pour mesurer le degré de similitude qui existe entre personnes d'un même quartier. En utilisant le modèle ALR (alternating logistic regression), il est en effet possible de quantifier la tendance des phénomènes à survenir en grappe à l'intérieur des quartiers («clustering») sur l'échelle des odds ratios, à l'aide du pairwise odds ratio.^{28,30,91} Dans la présente étude, puisque nous avons eu recours au modèle logistique multiniveau plutôt qu'à l'équation d'estimation généralisée, nous avons utilisé une autre option, qui consiste à convertir la variance inter-quartier sus-mentionnée sur l'échelle des odds ratio. Cela peut être fait en utilisant l'odds ratio médian qui a été proposé par Larsen⁷⁹ et a été appliqué au contexte épidémiologique par Larsen et Merlo.⁷⁸

Afin de définir l'odds ratio médian, considérons l'ensemble des paires d'individus qui présentent des caractéristiques similaires dans lesquelles les deux individus résident dans des quartiers différents. Pour chacune des paires d'individus, utilisant les résidus u_{0j} du modèle multiniveau, nous calculons l'odds ratio entre l'individu au risque moindre et celui au risque le

plus élevé (cet odds ratio dépend donc uniquement du résidu u_{0j} estimé dans chacun des deux quartiers). Considérant la distribution de cet odds ratio pour l'ensemble des paires d'individus de quartiers différents qui peuvent être formées avec nos données, l'odds ratio médian est simplement défini comme la valeur médiane de la distribution. Toutefois, en pratique, cet indicateur peut être directement calculé à partir de la variance inter-quartier σ_{u0}^2 du modèle multiniveau, à l'aide de la formule suivante (l'origine de cette formule étant mieux expliquée dans d'autres publications^{78,79}):

$$\text{MOR} = \exp \left[\left(2 \sigma_{u0}^2 \right) \cdot 0.6745 \right] \sim \exp(0.95 \sigma_{u0}^2)$$

L'odds ratio médian s'interprète de la façon suivante : lorsque l'on choisit aléatoirement deux individus dans des quartiers différents, l'odds ratio pour le risque de présenter un trouble mental entre l'individu au risque moindre et celui au risque le plus élevé est supérieur à la valeur de l'odds ratio médian dans 50% des cas. L'odds ratio médian fournit donc des informations sur l'amplitude des variations entre quartiers qui doivent être interprétées au niveau individuel : il mesure l'hétérogénéité qui existe entre individus provenant de quartiers différents. Un avantage de l'odds ratio médian est qu'il exprime la variance entre quartiers du modèle logistique multiniveau sur l'échelle des odds ratios, qui est largement utilisée en épidémiologie. Il permet donc de comparer l'amplitude des variations résiduelles entre quartiers à l'impact des facteurs pris en compte dans le modèle, qui sont également habituellement exprimés sur l'échelle des odds ratios. En rupture avec beaucoup d'auteurs qui concluent trop facilement et d'une façon excessive à l'importance du contexte pour la santé, cet outils conduit à mieux appréhender son influence exacte, en rapportant les variations inter-quartiers à l'effet des autres variables considérées.

Nous avons eu recours à la stratégie suivante pour étudier les variations de prévalence des troubles liés à la consommation de substances psycho-actives dans la ville de Malmö : nous avons d'abord estimé un modèle logistique multiniveau vide (c'est-à-dire ne contenant aucune variable explicative), qui nous a permis de quantifier les variations de prévalence des troubles sur l'échelle des odds ratios à l'aide de l'odds ratio médian. Afin d'examiner si cette variabilité inter-quartier était liée à des effets de composition, nous avons ensuite examiné si la variance inter-quartier σ_{u0}^2 diminuait lorsque les caractéristiques des individus ont été introduites dans le modèle. Il est en effet bien connu en analyse contextuelle qu'une partie des variations entre quartiers des phénomènes de santé est due à la composition variable des zones sur le plan des caractéristiques individuelles. Nous avons donc introduit de façon successive les différentes

caractéristiques individuelles dans le modèle afin d'identifier précisément les variables qui étaient à l'origine d'effets de composition.^{61,82} Nous avons alors cherché à voir si ces variations inter-quartiers demeuraient statistiquement significatives après inclusion des facteurs individuels dans le modèle. En l'absence d'effets de composition résiduels, des variations inter-quartiers significativement supérieures à zéro pourraient indiquer que certaines caractéristiques du quartier de résidence sont associées à la prévalence des troubles mentaux étudiés. Afin de tester cette hypothèse, nous avons finalement également inclus le niveau socio-économique du quartier de résidence dans le modèle (revenu moyen des habitants de 25 ans et plus). Nous avons cherché à voir si une part supplémentaire de la variabilité entre quartiers était expliquée à l'aide de ce facteur contextuel, et si les variations inter-quartiers demeuraient statistiquement supérieures à zéro dans ce modèle final.

Un des objectifs était d'identifier les quartiers où la prévalence des troubles liés à la consommation de substances psycho-actives était élevée. Par conséquent, il était important de savoir si la variable socio-économique contextuelle permet réellement de distinguer les quartiers à faible prévalence des quartiers à forte prévalence, après ajustement sur les facteurs individuels. Que la variable contextuelle soit utile ou non à cet effet ne dépend en fait pas seulement de l'importance de l'effet du niveau socio-économique du contexte, mais également de l'amplitude des variations résiduelles entre quartiers qui demeurent dans le modèle final (σ^2). En effet, même si la prévalence *moyenne* est plus élevée dans les quartiers défavorisés, savoir qu'un quartier est ou non défavorisé n'est pas très informatif si les variations résiduelles entre quartiers sont considérables : dans un tel cas, du fait de ces variations résiduelles, même un quartier défavorisé pourrait présenter une prévalence faible de troubles mentaux, et des quartiers favorisés pourraient avoir des prévalences plus importantes que d'autres quartiers qui le sont moins. Au contraire, si les variations résiduelles entre zones sont très faibles, savoir qu'un quartier est défavorisé est très informatif : dans ce cas, l'association positive entre niveau de pauvreté du quartier et prévalence des troubles est certainement assez forte pour faire que la plupart des quartiers défavorisés présentent une prévalence plus importante que les quartiers socialement favorisés.

Afin de tenir compte de cet aspect, nous avons eu recours à l'interval odds ratio proposé par Larsen⁷⁹ et introduit dans le champ de l'épidémiologie par Larsen et Merlo.⁷⁸ Afin de définir cet indicateur, considérons l'ensemble des paires d'individus de notre base de données dans

lesquelles un individu réside dans un quartier défavorisé et l'autre dans un quartier favorisé. Pour chacune de ces paires, utilisant le paramètre associé à l'effet socio-économique du quartier de résidence et les résidus u_{0j} de niveau quartier estimés à partir du modèle multiniveau, nous calculons l'odds ratio entre l'individu dans le quartier favorisé et celui dans le quartier défavorisé. De cette façon, nous obtenons un odds ratio composite qui intègre l'effet socio-économique contextuel et les variations résiduelles entre quartiers. Afin de comparer la force de l'effet du niveau socio-économique du quartier à l'amplitude des variations résiduelles entre quartiers, nous considérons la distribution de cet odds ratio composite pour l'ensemble des paires dans lesquelles un individu réside dans un quartier défavorisé et l'autre dans un quartier favorisé. Pour obtenir l'information qui nous intéresse, Larsen a proposé de définir l'interval odds ratio comme l'intervalle centré sur la médiane de la distribution qui comprend 80% des valeurs de l'odds ratio composite. Notons que les bornes inférieure et supérieure de l'interval odds ratio peuvent être calculées à partir des formules suivantes (dont l'origine est expliquée dans d'autres publications^{78,79}):

$$IOR_{\text{inférieur}} = \exp[\beta + \beta(2\sigma^2 u_0^2) - 1.2816] \sim \exp(\beta - 1.81\sigma^2 u_0^2)$$

$$IOR_{\text{supérieur}} = \exp[\beta + \beta(2\sigma^2 u_0^2) + 1.2816] \sim \exp(\beta + 1.81\sigma^2 u_0^2)$$

où β est le paramètre de régression pour l'effet socio-économique contextuel et $\sigma^2 u_0^2$ est la variance inter-quartier.

L'interval odds ratio est un indicateur qui intègre ici l'effet socio-économique du contexte et les variations résiduelles entre quartiers. Supposons que le niveau de pauvreté du quartier soit positivement associé à la prévalence des troubles mentaux considérés. Si les variations résiduelles entre quartiers sont faibles, l'odds ratio composite (basé sur l'effet positif du niveau socio-économique contextuel et sur les résidus de niveau quartier) est supérieur à 1 dans la majorité des cas. En conséquence, l'interval odds ratio (ou intervalle centré comprenant 80% des valeurs de l'odds ratio composite) ne comprend pas la valeur 1. Au contraire, si les variations résiduelles entre quartiers sont importantes par rapport à l'effet socio-économique du quartier, il est alors très probable de trouver des quartiers défavorisés qui présentent une prévalence de troubles mentaux plus faible que celle que présentent certains quartiers favorisés. L'odds ratio composite étant ainsi fréquemment inférieur à 1, l'interval odds ratio comprend la valeur 1. Ainsi, pour une force donnée de l'effet socio-économique du contexte, que l'interval odds ratio comprenne ou non la valeur 1 dépend de l'amplitude des variations résiduelles entre quartiers. Par conséquent,

l'intervalle odds ratio va nous permettre de juger de l'importance de l'effet socio-économique du contexte par rapport à l'amplitude des variations inter-quartiers résiduelles. Il fournit donc des informations complémentaires à celles auxquelles conduit l'odds ratio classique.

Dans son travail, Larsen a indiqué qu'il existe d'autres options pour décrire la distribution de l'odds ratio composite.^{78,79} Un problème de l'intervalle odds ratio est qu'il est assez arbitraire de le définir comme un intervalle comprenant 80% plutôt que 70% ou 90% des valeurs. Un intervalle de 80% pourrait inclure la valeur 1 alors qu'un intervalle de 70% ne la comprendrait pas. Au contraire, un intervalle de 80% pourrait ne pas inclure la valeur 1 alors qu'un intervalle de 90% le ferait. En conséquence, pour pallier à ce problème, nous suggérons de rapporter le pourcentage des paires d'individus dans la distribution de l'odds ratio composite pour lesquelles cet odds ratio est supérieur à 1. Si le facteur contextuel pris en compte n'est d'absolument aucune utilité pour distinguer les quartiers à faible prévalence des quartiers à forte prévalence, l'odds ratio composite qui intègre l'effet positif du niveau socio-économique du contexte et les variations inter-quartiers résiduelles n'est supérieur à 1 que dans 50% des cas. Au contraire, si l'effet socio-économique contextuel est extrêmement fort par rapport aux variations résiduelles entre quartiers, l'odds ratio composite est supérieur à 1 dans 100% des cas. Nous utilisons ce pourcentage compris entre 50% et 100% comme un indicateur que nous appelons «*sorting out index*», qui permet de quantifier l'importance de l'effet contextuel par rapport aux variations inter-quartiers résiduelles.

Dans un modèle ajusté sur les facteurs individuels, cet indicateur nous permet de comparer l'effet socio-économique du contexte aux variations entre quartiers provenant d'autres sources qui persistent après que l'on ait tenu compte de la composition des quartiers en termes de caractéristiques individuelles. Toutefois, dans la présente étude, nous avons également calculé le *sorting out index* à partir d'un modèle logistique multiniveau qui n'incluait que le facteur socio-économique contextuel (pas les variables individuelles). Dans ce cas, cet indicateur nous permet d'évaluer si le facteur socio-économique contextuel peut être directement utilisé pour distinguer les quartiers qui ont une prévalence élevée de troubles de ceux qui présentent une faible prévalence.

Résultats

En 2001, 1,45% des individus de 40 à 59 ans dans la ville de Malmö ont reçu un diagnostic de trouble mental ou comportemental lié à la consommation de substances psycho-actives (n = 956 sur 65 830 individus). Parmi les personnes ayant reçu un tel diagnostic, l'alcool était en jeu chez

88% des individus, les sédatifs et les hypnotiques chez 9%, les stimulants chez 7%, les opioïdes chez 6%, et les cannabinoïdes chez moins de 3% des individus. Quatorze pour-cent des individus avaient reçu des diagnostics impliquant plusieurs substances psycho-actives. Sur le plan clinique, un syndrome de dépendance avait été diagnostiqué chez 91% de ces individus et une utilisation nocive pour la santé chez 13% d'entre eux. Des troubles psychotiques n'avaient été trouvés que pour 2% des individus qui avaient reçu un diagnostic de trouble lié à la consommation de substances psycho-actives. Toutefois, dans la présente étude, nous n'étions pas en mesure de distinguer ces différents sous-diagnostics dans la base de données principale. Par conséquent, les analyses qui suivent ne s'intéressent qu'à l'existence ou non chez les individus d'un diagnostic global de trouble lié à la consommation de substances psycho-actives.

Nous avons d'abord cherché à voir si la prévalence des troubles mentaux liés à la consommation de substances psycho-actives variait d'un quartier à l'autre, c'est-à-dire si ce phénomène de santé avait une dimension contextuelle. Le modèle multiniveau a indiqué qu'il existait de très fortes variations d'un quartier à l'autre, et que celles-ci étaient très nettement supérieures à 0 (tableau 1). Comme nous l'avons indiqué ci-dessus, il est en fait plus parlant d'exprimer ces variations entre quartiers sur l'échelle des odds ratio à l'aide de l'odds ratio médian. Dans le modèle vide, l'odds ratio médian était égal à 2,13 (tableau 1), ce que l'on peut interpréter de la façon suivante : lorsque l'on choisit aléatoirement deux individus dans des quartiers différents, l'odds ratio estimé entre l'individu au risque moindre et celui au risque le plus élevé est supérieur à 2,13 dans 50% des cas. L'odds ratio médian indique donc que des variations substantielles existaient entre quartiers dans la prévalence des troubles liés à la consommation de substances psycho-actives, montrant que ces troubles mentaux sont fortement contextualisés dans la ville de Malmö. Considérant les résidus de niveau quartier du modèle multiniveau, la carte rapportée en haut de la figure 1 permet de visualiser les variations de prévalence de ces troubles mentaux parmi les 100 quartiers de la ville. La figure indique qu'un certain nombre de quartiers au centre et au nord de la ville présentent une prévalence élevée de troubles mentaux et comportementaux liés à la consommation de substances psycho-actives.

Tableau 1 : Résultats des modèles multiniveaux sur les troubles mentaux liés à la consommation de substances psycho-actives

	Modèle vide		Modèle incluant les facteurs individuels		Modèle incluant les facteurs individuels et le revenu moyen du quartier		IOR ^a	Sorting out index
	Indice	IC 95% ^a	Indice ^b	IC 95%	Indice ^b	IC 95%		
Effets fixes (odds ratios)								
Age (réf. : 40–44 ans)								
45–49 ans			1,48	(1,23; 1,77)	1,48	(1,23; 1,77)		
50–54 ans			1,30	(1,07; 1,57)	1,30	(1,08; 1,57)		
55–59 ans			1,35	(1,11; 1,63)	1,36	(1,12; 1,65)		
Sexe: homme (réf. : femme)								
			2,37	(2,06; 2,72)	2,35	(2,04; 2,70)		
Etat matrim. : seul (réf. : marié)								
			4,46	(3,74; 5,31)	4,34	(3,64; 5,16)		
Instruction: faible (réf. : élevée)								
			1,38	(1,20; 1,59)	1,36	(1,18; 1,56)		
Revenu : faible (réf. : élevé)								
			3,65	(3,09; 4,30)	3,46	(2,94; 4,09)		
Revenu moyen du quartier (réf. : quatrième quartile)								
Troisième quartile					1,40	(1,00; 1,96)	[0,79; 2,48]	77%
Second quartile					2,04	(1,49; 2,80)	[1,15; 3,62]	94%
Premier quartile					2,13	(1,56; 2,91)	[1,20; 3,78]	95%
Effets aléatoires								
Variance inter-quartier	0,62	(0,38; 0,88)	0,17	(0,07; 0,26)	0,10	(0,03; 0,17)		
Odds ratio médian	2,13		1,49		1,36			

^a IC, intervalle de confiance ; IOR, interval odds ratio.

^b Dans les modèles incluant différents facteurs, l'ensemble des effets sont ajustés les uns par rapport aux autres.

Afin de déterminer quelle part des variations géographiques de prévalence était imputable à la composition variables des quartiers, nous avons d'abord cherché à voir si les différents facteurs individuels étaient associés à la prévalence de troubles mentaux. Ainsi que l'indique le tableau 2, les analyses univariées ont montré que la prévalence des troubles liés à la consommation de substances psycho-actives était significativement supérieure chez les hommes, chez les personnes célibataires, divorcées, ou veuves, chez les individus peu instruits, et chez ceux qui avaient un faible revenu. L'inclusion simultanée de ces facteurs dans le modèle multiniveau a confirmé que les troubles mentaux étaient plus fréquents chez ces différentes catégories d'individus (tableau 1).

Figure 1 : Résidus de niveau quartier des modèles multiniveaux sur les troubles liés à la consommation de substances psycho-actives pour les 100 quartiers de Malmö, estimés à partir du modèle multiniveau vide, du modèle individuel, et du modèle contextuel

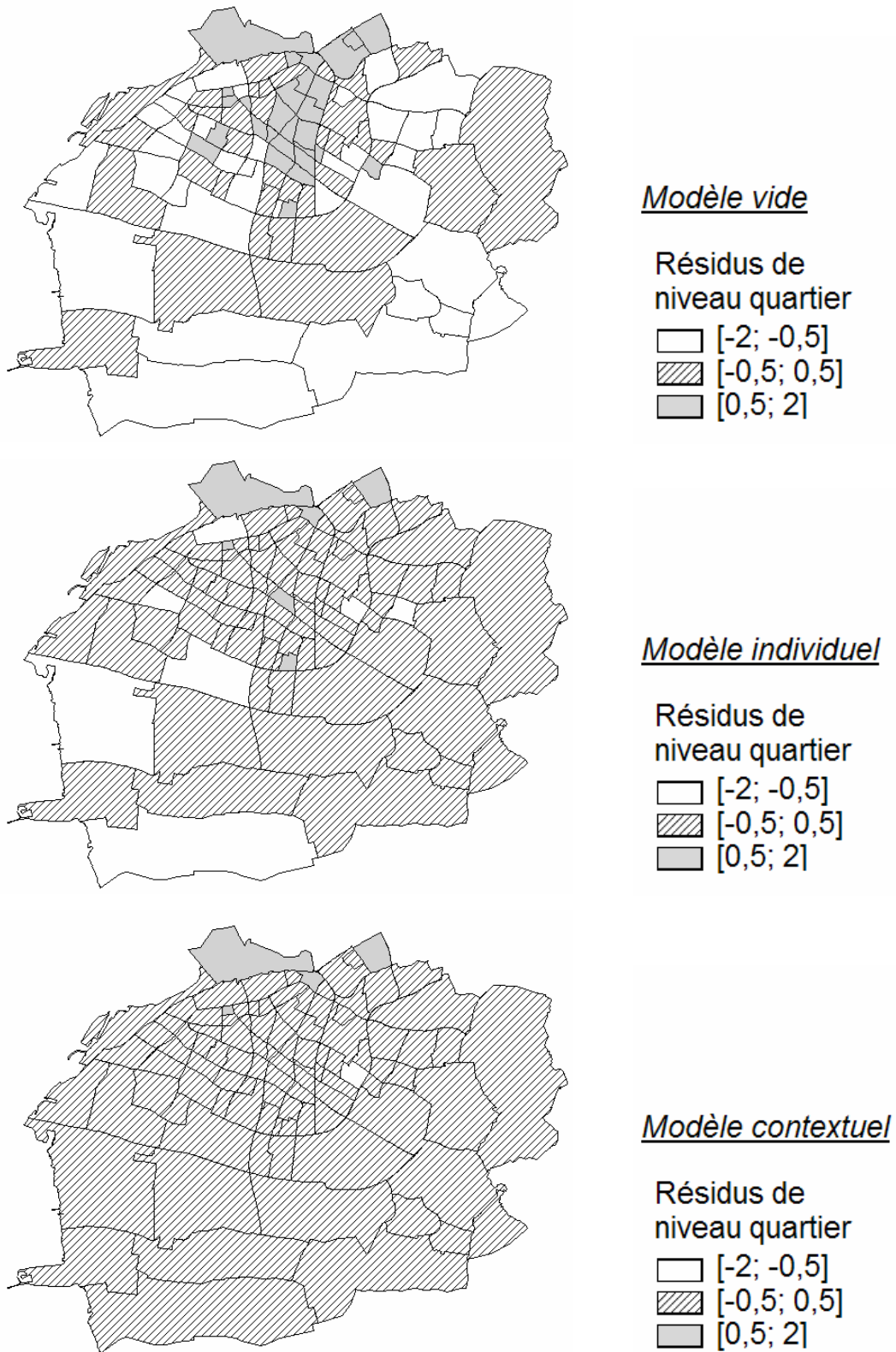


Tableau 2 : Prévalence de troubles mentaux liés à la consommation de substances psycho-actives, en fonction des caractéristiques des individus et du quartier

Variable (fréquence)	Prévalence (fréquence)	en %	p
Sexe			< 0,0001 ^a
Homme (33 306)	1,98 (661)		
Femme (32 524)	0,91 (295)		
Age			0,0124 ^a
40–44 ans (17 138)	1,27 (217)		
45–49 ans (16 686)	1,69 (282)		
50–54 ans (16 483)	1,43 (236)		
55–59 ans (15 523)	1,42 (221)		
Etat matrimonial			< 0,0001 ^a
Seul (31 321)	2,54 (794)		
Mariés, en cohabitation (34 509)	0,47 (162)		
Niveau d’instruction			< 0,0001 ^a
9 années ou moins (15 042)	2,21 (332)		
Plus que 9 années (49 246)	1,24 (613)		
Revenu individuel			< 0,0001 ^a
= à la valeur médiane ^b (32 915)	2,32 (762)		
> à la valeur médiane ^b (32 915)	0,59 (194)		
Revenu moyen du quartier			< 0,0001 ^c
Premier quartile (16 761)	2,29 (384)		
Second quartile (15 560)	1,97 (307)		
Troisième quartile (16,274)	1,06 (173)		
Quatrième quartile (17,235)	0,53 (92)		

^a Test du khi-deux

^b 189 523 Couronnes Suédoises (approximativement égal à 26 700 dollars)

^c Test de tendance de Cochran-Armitage

La variance entre quartiers estimée par le modèle multiniveau a diminué de 73% lorsque les facteurs individuels ont été introduits dans le modèle, et l’odds ratio médian a chuté de 2,13 à 1,49. Nous pouvons également visualiser cette réduction des variations non expliquées entre quartiers en comparant sur la figure 1 la carte du haut à la carte du milieu. Sur cette dernière, nous avons rapporté les résidus de niveau quartier du modèle multiniveau qui inclut les facteurs individuels. Pour cette seconde carte, nous avons eu recours aux mêmes valeurs seuil que sur la carte du haut pour répartir les quartiers en trois catégories différentes en fonction de la valeur du résidu de niveau quartier. Il est évident que beaucoup plus de quartiers appartiennent à la catégorie intermédiaire sur la carte du milieu que sur la carte du haut. Cette catégorie renvoyant aux quartiers qui ont une prévalence ni particulièrement faible

ni particulièrement élevée, on voit que beaucoup de résidus de niveau quartier qui étaient significativement négatifs ou positifs dans le modèle multiniveau vide ont été réduits en direction de 0 lorsque les facteurs individuels ont été introduits dans le modèle. Une telle chute de la variance inter-quartier indique qu'une portion importante de la variabilité géographique de prévalence des troubles mentaux était due aux différences dans la composition des quartiers en terme de caractéristiques individuelles.

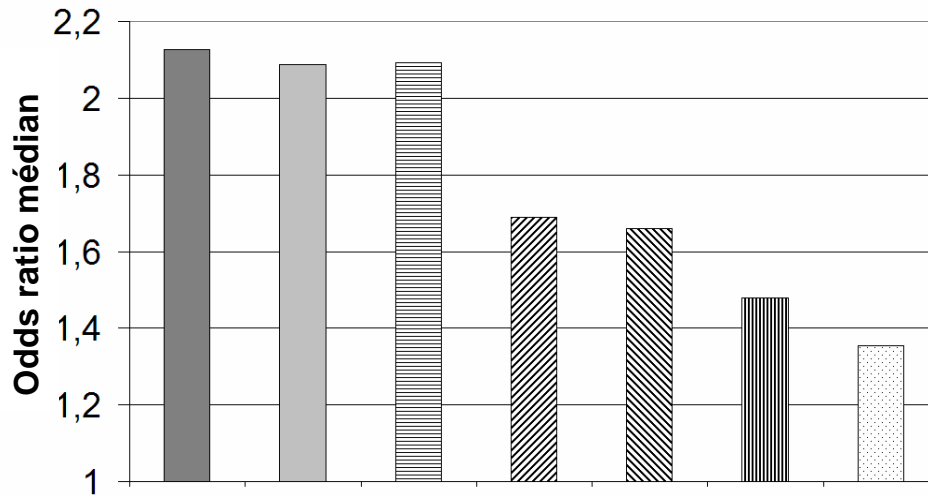
En incluant de façon successive les différents facteurs individuels dans le modèle, nous avons pu identifier de façon plus précise les caractéristiques qui étaient à l'origine de ces effets de composition importants. Ainsi que l'indique la figure 2, une large part des variations entre quartiers dans la prévalence des troubles mentaux liés à la consommation de substances psycho-actives était due à la composition variable des quartiers en terme d'état matrimonial. Puisque les personnes célibataires, divorcées, ou veuves avaient une prévalence accrue de troubles liés à la consommation de substances psycho-actives, la proportion variable de tels individus d'un quartier à l'autre était liée à des variations géographiques importantes de prévalence. Un effet de composition d'intensité moindre lié au revenu des individus a pu être identifié. Ainsi que l'indique la figure 2, l'introduction de ces différents facteurs a réduit de façon importante les variations de prévalence non expliquées entre quartiers.

Après inclusion des caractéristiques individuelles dans le modèle, la variabilité entre quartiers dans la prévalence des troubles liés à la consommation de substances psycho-actives était significativement supérieure à 0. Nous avons cherché à voir si le niveau socio-économique du quartier était associé à la prévalence des troubles, après que l'on ait tenu compte de l'effet des caractéristiques individuelles. De façon univariée (tableau 2), le test de Cochran-Armitage a indiqué que la prévalence des troubles mentaux augmentait de façon linéaire avec le niveau de pauvreté du quartier de résidence.

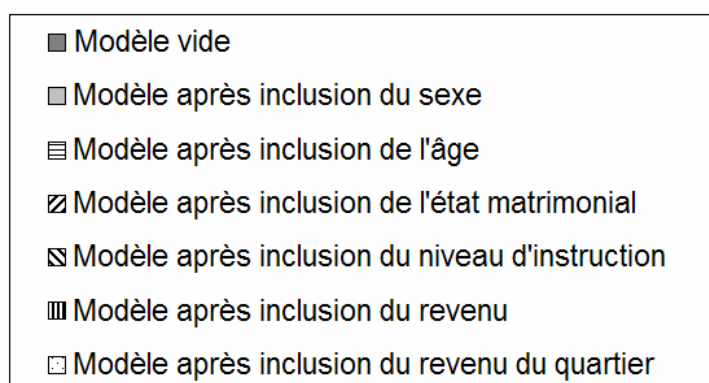
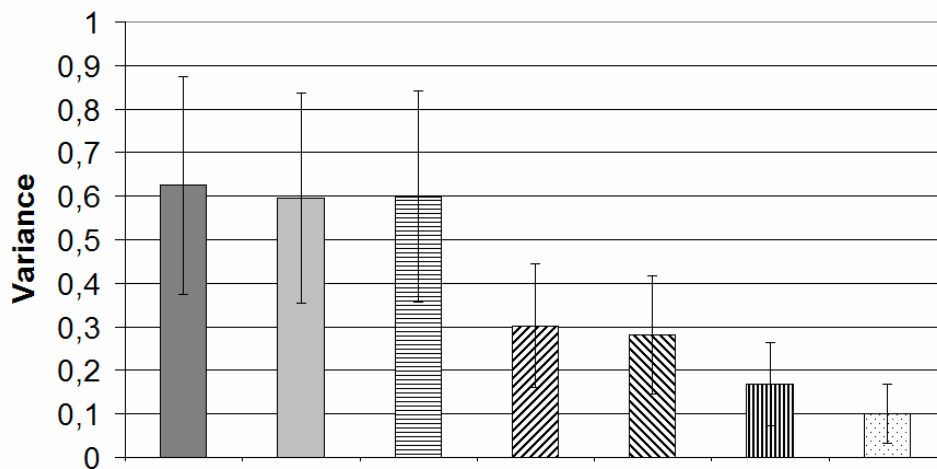
La variable socio-économique contextuelle a ensuite été introduite dans le modèle logistique multiniveau. Après ajustement sur les facteurs individuel, la prévalence des troubles mentaux liés à la consommation de substances psycho-actives était supérieure dans les quartiers défavorisés (tableau 1). Ainsi que le tableau 1 et la figure 2 l'indiquent, la variance résiduelle entre quartiers a à nouveau chuté de 41% lorsque la variable socio-économique contextuelle a été introduite dans le modèle. L'odds ratio médian a chuté de 1,41 à 1,36. En comparant la carte du milieu à la carte du bas sur la figure 1, il apparaît également qu'une portion additionnelle de la variabilité entre quartiers a pu être expliquée par le niveau socio-économique du quartier.

Figure 2 : Variations résiduelles entre quartiers dans les modèles multiniveaux consécutifs, exprimées à l'aide de l'odds ratio médian (haut) et à l'aide de la variance inter-quartier (bas)

Odds ratio médian



Variance inter-quartier



Ainsi qu'on l'a noté ci-dessus, une question importante en santé publique est de savoir si un facteur contextuel donné – dans notre cas le niveau socio-économique du quartier – permet d'identifier les zones où la prévalence des problèmes de santé est accrue. Nous avons trouvé que la prévalence des troubles mentaux liés à la consommation de substances psycho-actives était accrue dans les quartiers défavorisés, mais l'utilité de ce facteur socio-économique contextuel pour identifier les quartiers à forte prévalence dépend également de l'importance des variations résiduelles entre quartiers. Nous avons utilisé l'interval odds ratio^{78,79} afin de comparer la force de l'effet socio-économique contextuel à l'amplitude des variations résiduelles entre quartiers. Ainsi que l'indique le tableau 1, l'interval odds ratio pour le fait de résider dans un quartier défavorisé plutôt que favorisé était égal à [1.20; 3.78], et ne comprenait pas la valeur 1. Cela signifie que l'odds ratio composite pour l'effet contextuel du niveau de pauvreté est supérieur à 1 dans la plupart des cas, même quand on tient compte des variations résiduelles entre quartiers dans le calcul de l'odds ratio.

Au-delà de l'interval odds ratio, nous avons cherché à exprimer la distribution de l'odds ratio composite (qui intègre l'effet socio-économique du contexte et les variations résiduelles entre quartiers) d'une façon peut-être plus aisément interprétable. Nous suggérons de rapporter le pourcentage des paires d'individus pour lesquelles l'odds ratio composite est supérieur à 1. Dans notre cas, cet indicateur que nous appelons *sorting out index* était égal à 95% (tableau 1) : lorsque l'on tenait compte des variations résiduelles entre quartiers, la prévalence ajustée était plus forte pour un individu vivant dans un quartier défavorisé que pour un autre vivant dans un quartier favorisé dans 95% des cas. Cela indique que le niveau socio-économique du contexte de résidence permet largement d'identifier les quartiers où la prévalence ajustée de troubles mentaux est élevée, ce qui s'explique par le fait que l'effet socio-économique contextuel était assez important alors que les variations résiduelles entre quartiers étaient modestes.

Toutefois, afin de savoir si le facteur socio-économique contextuel peut être directement utilisé pour distinguer les quartiers à faible et à forte prévalence, l'interval odds ratio et le *sorting out index* doivent être estimés à partir d'un modèle logistique multiniveau qui ne contient que le facteur contextuel (mais pas les variables individuelles d'ajustement). Les résultats de ce modèle sont rapportés dans le tableau 3. L'odds ratio lié au fait de résider dans les quartiers les plus défavorisés plutôt que dans les quartiers les plus favorisés était égal à 4,99. Cet odds ratio est nettement plus fort que ce qu'il était dans un modèle ajusté sur les facteurs individuels, car il intègre une partie des effets du niveau socio-économique individuel. Ainsi que l'indique le tableau 3, le *sorting out index* était égal à 87%, 97%, et 99%

pour le fait de résider dans les second, troisième, et quatrième quartiles de pauvreté du quartier de résidence. Quand on tient compte des variations résiduelles entre quartiers, l'odds ratio estimé entre un individu des quartiers les plus riches et un individu des quartiers les plus pauvres est supérieur à 1 dans 99% des cas. Cela montre que le facteur socio-économique contextuel peut être directement utilisé afin de distinguer les quartiers à faible et à forte prévalence.

Tableau 3: Résultats des modèles multiniveaux sur les troubles mentaux liés à la consommation de substances psycho-actives : modèle ne contenant que le facteur socio-économique contextuel (et pas de variables individuelles)

	Indice ^a	IC 95% ^b	IOR ^b	Sorting out index
Effets fixes (odds ratios)				
Revenu moyen du quartier (réf. : quatrième quartile)				
Troisième quartile	2,21	(1,48; 3,30)	[0,89; 5,51]	87%
Second quartile	3,80	(2,60; 5,57)	[1,53; 9,48]	97%
Premier quartile	4,99	(3,44; 7,24)	[2,00; 12,43]	99%
Effets aléatoires				
Variance inter-quartier	0,25	(0,14; 0,37)		

^a L'ensemble des effets sont ajustés les uns par rapport aux autres.

^b IC, intervalle de confiance; IOR, interval odds ratio.

Enfin, nous avons cherché à voir si le fait de combiner des désavantages sociaux au niveau individuel et au niveau du quartier de résidence était associé à une prévalence particulièrement accrue de troubles liés à la consommation de substances psycho-actives. Pour ce faire, nous avons testé les interactions entre effets des différents facteurs. Ces analyses ont indiqué que l'état matrimonial était un puissant modificateur pour les effets du revenu. Nous avons donc estimé le modèle multiniveau séparément pour les personnes célibataires, divorcées, ou veuves d'un côté, et pour les personnes mariées ou en cohabitation de l'autre. Ainsi que nous le rapportons dans le tableau 4, ces modèles ont d'abord indiqué que l'association entre revenu individuel et prévalence de troubles liés à la consommation de substances psycho-actives était nettement plus forte chez les individus qui vivaient seuls que chez les personnes qui vivaient en couple. Deuxièmement, l'association entre pauvreté du quartier et prévalence des troubles était également plus forte pour les individus qui vivaient seuls. Par conséquent, ces modèles ont indiqué que l'état matrimonial était un modificateur de l'effet du revenu, à la fois au niveau individuel et au niveau du quartier. Une telle

modification d'effet indique que les individus qui vivaient seuls *et* qui avaient un faible revenu ou vivaient dans un quartier pauvre présentaient une prévalence particulièrement accrue de troubles mentaux liés à la consommation de substances psycho-actives. Ainsi que nous nous y attendions, nous avons observé une plus grande réduction de la variance entre quartiers en introduisant les facteurs individuels et contextuel dans le modèle pour les personnes vivant seules que pour les personnes mariées ou en cohabitation (figure 3).

Table 4: Résultats des modèles multiniveaux sur les troubles mentaux liés à la consommation de substances psycho-actives, stratifié sur l'état matrimonial

	Parmi les individus seuls		Parmi les individus mariés, en cohabitation	
	Indice ^a	IC 95% ^b	Indice ^a	IC 95%
Effets fixes (odds ratios)				
Age (réf. : 40–44 ans)				
45–49 ans	1,45	(1,19; 1,77)	1,69	(1,04; 2,75)
50–54 ans	1,27	(1,04; 1,56)	1,47	(0,89; 2,44)
55–59 ans	1,22	(0,99; 1,52)	2,09	(1,29; 3,37)
Sexe: homme (réf. : femme)	2,37	(2,03; 2,77)	2,09	(1,50; 2,92)
Instruction: faible (réf. : élevée)	1,36	(1,17; 1,58)	1,30	(0,92; 1,84)
Revenu : faible (réf. : élevé)	4,04	(3,33; 4,90)	2,11	(1,49; 3,00)
Revenu moyen du quartier (réf. : quatrième quartile)				
Troisième quartile	1,62	(1,09; 2,41)	1,16	(0,65; 2,07)
Second quartile	2,29	(1,58; 3,32)	1,97	(1,15; 3,40)
Premier quartile	2,65	(1,83; 3,82)	1,20	(0,67; 2,14)
Effets aléatoires				
Variance inter-quartier	0,10	(0,02; 0,17)	0,27	(0,02; 0,52)

^a L'ensemble des effets sont ajustés les uns par rapport aux autres.

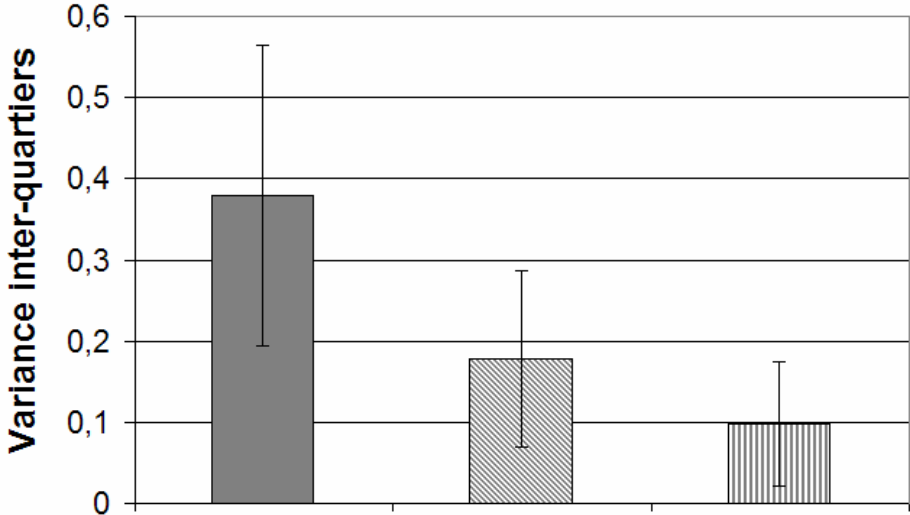
^b IC, intervalle de confiance.

Discussion

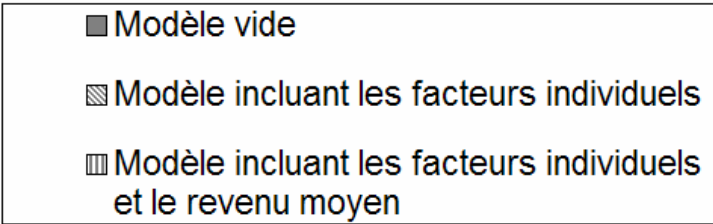
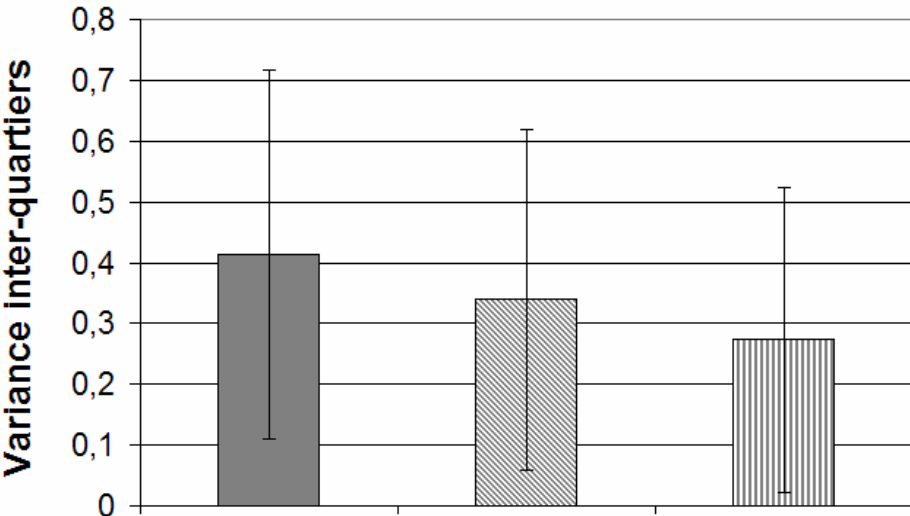
Dans la présente étude, notre but était d'étudier la distribution spatiale des troubles mentaux et comportementaux liés à la consommation de substances psycho-actives dans la ville de Malmö en Suède, et d'identifier les quartiers où la prévalence de troubles était élevée. Un des points forts de notre étude est que nous avons utilisé des données sur l'ensemble des individus de 40 à 59 ans qui résidaient dans la ville de Malmö en 2001. En conséquence, la puissance statistique était importante, et largement suffisante pour explorer et décrire les variations spatiales de prévalence des troubles mentaux considérés.

Figure 3 : Variance des résidus de niveau quartier dans les modèles multiniveaux consécutifs stratifiés sur l'état matrimonial

Pour les individus célibataires, divorcés, ou veufs



Pour les individus mariés ou en cohabitation



Quant aux faiblesses de notre étude, une première limite est que nous n'étions pas capables de distinguer entre les différents sous-groupes de diagnostics dans la catégorie globale des troubles mentaux liés à la consommation de substances psycho-actives (F10-F19). Un tel groupe de diagnostics inclut des tableaux cliniques divers, et renvoie à l'utilisation de différents types de substances psycho-actives. Toutefois, les statistiques descriptives que nous avons fournies indiquent que la plupart des individus qui avaient reçu un diagnostic dans le groupe F10-F19 présentaient un syndrome de dépendance à l'alcool. Nos futures études sur ce thème dans la ville de Malmö gagneront à distinguer les différents sous-groupes de diagnostics.

D'autre part, puisque notre étude est basée sur des données transversales, il est impossible de conclure à l'existence de relations causales entre les facteurs de risque étudiés et l'incidence de troubles liés à la consommation de substances psycho-actives. Afin d'étudier les facteurs qui influent sur l'incidence de ces troubles, il est nécessaire de répéter ces analyses à partir de données longitudinales. De telles analyses longitudinales fourniront des informations utiles pour mettre en place des programmes de prévention ou d'intervention.

De plus, les données sur l'existence de troubles liés à la consommation de substances psycho-actives proviennent de diagnostics réalisés par des médecins au cours de contacts avec le système de soins en 2001. Un certain nombre d'individus souffrant de troubles mentaux liés à la consommation de substances psycho-actives pourraient n'avoir eu aucun contact avec le système de soins en 2001. La mesure dans laquelle des individus avec de tels troubles mentaux ont ou non reçu un diagnostic dans le groupe F10-F19 en 2001 pourrait dépendre des caractéristiques des individus et de leur quartier de résidence, mais nous n'étions absolument pas en mesure d'examiner si cela a ou non affecté les associations que nous rapportons.

Concernant la distribution spatiale des troubles liés à la consommation de substances psycho-actives, le modèle multiniveau nous a permis de montrer qu'il existait des variations de prévalence importantes entre quartiers. On peut en conclure que ces troubles mentaux spécifiques présentent une dimension contextuelle notable dans la ville de Malmö. Nous avons observé que la prévalence des troubles était particulièrement accrue au centre et au nord de la ville. Nous avons eu recours à l'odds ratio médian pour quantifier l'amplitude des variations entre quartiers sur l'échelle des odds ratios. Ces variations inter-quartiers étaient plus fortes que celles rapportées dans nombre d'études précédentes.^{33,73-76} Alors que ces analyses se sont souvent appuyées sur une définition générique des troubles mentaux, nous nous sommes intéressés à une catégorie spécifique de troubles, ceux liés à la consommation de substances psycho-actives. Nos résultats suggèrent que l'importance modeste du contexte

pour les troubles mentaux en général pourrait cacher des situations particulières de forte influence du contexte résidentiel sur des catégories plus spécifiques de troubles. En particulier, il est apparu que les troubles dont l'étiologie est liée à la consommation de substances psycho-actives dépendaient fortement des caractéristiques sociales du contexte.

Dans un second temps, nous avons introduit des variables individuelles dans le modèle. Le fait d'être un homme, de vivre seul, et d'avoir un faible revenu étaient associés une prévalence supérieure de troubles mentaux liés à la consommation de substances psycho-actives. Ainsi que cela a été expliqué ci-dessus, ces associations au niveau individuel ne nous permettent pas de tirer des conclusions sur les mécanismes causaux qui sont à l'œuvre. Par exemple, pour l'état matrimonial, il se pourrait que les personnes célibataires, divorcées, ou veuves aient plus de facilité à consommer des substances psycho-actives, n'étant pas encouragées par un partenaire à réduire leur consommation. Cela correspondrait à un effet causal de l'état matrimonial, par lequel le couple aurait un effet protecteur. Toutefois, la relation causale peut évidemment jouer en sens contraire, puisque les personnes avec une consommation importante de substances psycho-actives pourraient avoir plus de difficultés à trouver un partenaire, ou être soumises à un risque accru de divorcer ou de se séparer quand elles sont en couple.

Toutefois, même si les associations identifiées dans notre étude ne permettent pas de tirer des conclusions définitives sur les relations causales en jeu, elles fournissent néanmoins des informations utiles pour les décideurs de santé publique. En effet, nos résultats peuvent servir si l'on cherche à identifier les populations pour lesquelles la prévalence de troubles liés à la consommation de substances psycho-actives est supérieure.

Nous avons trouvé qu'une part importante des variations géographiques dans la prévalence des troubles liés à la consommation de substances psycho-actives était imputable à la composition variable des quartiers en terme d'état matrimonial et de revenu individuel. Toutefois, après ajustement sur ces facteurs, la variance inter-quartier demeurait supérieure à 0. Afin d'expliquer ces variations résiduelles, nous avons tenu compte d'un facteur mesuré au niveau des quartiers administratifs, et avons trouvé que le fait de résider dans un quartier pauvre était associé à une prévalence accrue de troubles liés à la consommation de substances psycho-actives, après ajustement sur les facteurs individuels. De même qu'avec les associations au niveau individuel, notre schéma d'étude ne nous permet pas de conclure à l'existence d'un effet causal du niveau socio-économique du quartier sur la santé mentale. En effet, dans notre étude transversale, les mécanismes causaux qui sous-tendent l'association entre pauvreté du contexte et prévalence des troubles pourraient jouer dans les deux directions

opposées. D'une part, la pauvreté pourrait avoir un impact sur la santé mentale des individus.⁷⁷ D'autre part, même si les études entreprises jusqu'à aujourd'hui n'ont pas réellement réussi à mettre en évidence des processus de migration sélective liés à des variables de santé,⁹² les migrations pourraient aussi dans notre cas contribuer au phénomène de concentration des troubles liés à la consommation de substances psycho-actives dans les quartiers les plus pauvres de la ville. Par conséquent, l'hypothèse d'un effet causal de la pauvreté du contexte de résidence sur l'incidence de troubles liés à la consommation de substances psycho-actives ne peut être validée ici et mérite d'être étudiée à partir de données longitudinales. Toutefois, malgré cette incertitude sur les mécanismes causaux en jeu, nos résultats montrent que des interventions destinées aux personnes présentant des troubles liés à la consommation de substances psycho-actives pourraient être plus particulièrement utiles dans les quartiers les plus défavorisés de la ville.

D'un point de vue de santé publique, il est particulièrement utile d'examiner si le niveau socio-économique du quartier permet réellement d'identifier les quartiers où la prévalence des troubles mentaux considérés est la plus importante. Pour ce faire, il est nécessaire de comparer la force de l'effet du niveau de pauvreté du quartier à l'amplitude des variations résiduelles entre quartiers. En cas de variations résiduelles importantes entre quartiers, il serait hautement probable de trouver un quartier favorisé qui présente une plus forte prévalence de troubles qu'un quartier défavorisé. Dans ce cas, la variable socio-économique contextuelle ne permettrait pas de distinguer efficacement les quartiers à forte prévalence des quartiers à faible prévalence. Afin d'obtenir des informations à ce sujet, nous avons utilisé l'interval odds ratio proposé par Larsen.⁷⁹ Suivant ensuite l'idée de cet auteur selon laquelle d'autres options que l'interval odds ratio pourraient être utilisées pour décrire la distribution de l'odds ratio composite, nous avons proposé un indicateur associé, le sorting out index, qui fournit les informations souhaitées sous une forme peut-être plus intuitive et facilement interprétable. L'interval odds ratio comme le sorting out index ont indiqué que l'effet du niveau socio-économique du quartier était important par rapport aux variations résiduelles entre quartiers, que les variables individuelles d'ajustement soient ou non incluses dans le modèle. Par exemple, le sorting out index a indiqué que, même lorsque l'on tenait compte des variations non expliquées entre quartiers, l'odds ratio pour le fait de résider dans les quartiers les plus pauvres plutôt que dans les quartiers les plus riches était supérieur à 1 dans 99% des cas (95% des cas dans un modèle ajusté sur les facteurs individuels).

Nous avons ensuite cherché à voir si la prévalence des troubles mentaux était particulièrement accrue chez les individus qui combinaient différents désavantages sociaux

mesurés au niveau individuel et au niveau du quartier. Nous avons trouvé que l'accroissement de prévalence chez les personnes qui avaient un faible revenu ou qui résidaient dans des quartiers défavorisés était d'amplitude nettement plus forte quand ces personnes étaient également célibataires, divorcées, ou veuves. En d'autres termes, l'état matrimonial intervenait comme modificateur de l'effet du revenu, à la fois au niveau individuel et au niveau du quartier. Par conséquent, d'éventuels programmes d'intervention pourraient être plus particulièrement utiles si ils sont ciblés sur des individus qui combinent plusieurs désavantages sociaux comme le fait de vivre seul, d'être pauvre, ou de résider dans un quartier pauvre.

En conclusion, notre étude contribue à enrichir le corps des connaissances sur les déterminants sociaux de la santé. Nos résultats suggèrent que de tels déterminants ont une structure multiniveau, et qu'ils ne peuvent par conséquent pas être simplement pris en compte et mesurés au niveau individuel. Lorsque l'on cherche à identifier des populations qui présentent une santé dégradée, il apparaît nécessaire de tenir compte des facteurs du contexte en plus des caractéristiques habituelles au niveau individuel.

Remerciements

Nous souhaitons remercier la Région de Scanie en Suède pour son soutien dans cette étude. Le projet de recherche a été financé par une subvention de Ministère Français de la Recherche dans le cadre de l'Action Concertée Incitative «Terrains, techniques, et théories en Sciences humaines et sociales » (TTT027), par le programme « Avenir » de l'INSERM attribué à Pierre Chauvin et à son équipe de recherche, et par une subvention du FAS (le Conseil Suédois de Recherche sur les Conditions de Travail et les Questions Sociales) attribuée à Juan Merlo.

Chapitre 2 : Les variations des modes de recours aux soins sur le territoire métropolitain Français

Introduction

Beaucoup d'études conduites en France ou dans d'autres pays industrialisés ont montré que les comportements de recours aux soins des individus varient sur le territoire. D'un point de vue de santé publique, il est important d'examiner si ces variations spatiales sont dues au fait que les individus présentent des caractéristiques différentes d'un endroit à l'autre, ou à proprement parler aux caractéristiques du contexte de résidence. Le contexte de résidence influence les modes de recours aux soins au travers de l'offre de soins disponible à proximité. Mais joue-t-il également au travers de son niveau socio-économique, en redoublant ainsi les effets du statut socio-économique au niveau individuel ?

Pour répondre à ces questions, il n'est pas possible de recourir à l'approche écologique, qui se contente de mettre en relation des mesures agrégées d'utilisation des soins et des variables explicatives elles-mêmes agrégées à un niveau plus ou moins local. Pour distinguer des effets du contexte d'effets opérant au niveau individuel, il est indispensable de s'appuyer sur des données collectées au niveau individuel, orientation qui est au centre de l'analyse contextuelle.^{1,3,14,93}

La plupart des études contextuelles menées à partir de données individuelles se sont appuyées sur le paradigme de l'analyse multiniveau. Dans cette approche, le territoire d'étude est fragmenté en zones, qui n'ont souvent qu'une origine administrative.⁹⁴ Les données utilisées dans les analyses ont ainsi une structure hiérarchique, puisque les individus sont regroupés au sein de zones géographiques. Le traitement de ces données hiérarchiques nécessite des méthodes d'analyse spécifiques telles que les modèles multiniveaux.^{25,27}

En fait, l'approche multiniveau en analyse contextuelle présente certaines limites, liées au fait qu'elle étudie les variations géographiques des phénomènes en s'appuyant sur un territoire fragmenté en zones administratives. Sur le plan statistique, une limite majeure des modèles multiniveaux est qu'ils postulent que deux individus résidant dans des zones différentes sont absolument indépendants (corrélation nulle), même si les zones sont proches ou adjacentes.²⁴ En d'autres termes, les modèles multiniveaux tiennent compte du fait que les individus proviennent de zones différentes, mais ils négligent complètement les relations spatiales entre ces zones.

Il apparaît donc nécessaire d'analyser les variations spatiales des comportements de recours aux soins dans un espace continu plutôt que fragmenté en zones administratives. De telles techniques d'analyse spatiales ont été utilisées dans le champ de l'analyse écologique, à partir de données agrégées.^{48-53,95-97} Toutefois, dans le champ de l'analyse contextuelle (à partir de données collectées au niveau individuel), très peu d'études ont essayé de mettre en œuvre ces méthodes d'analyse spatiale.^{98,99} Afin d'avancer dans cette perspective, une option est de recourir aux modèles spatiaux mixtes, dont l'intérêt est de tenir compte de la distribution spatiale des phénomènes sans fragmenter le territoire en zones aux limites arbitraires.¹⁰⁰ En effet, les modèles spatiaux mixtes s'appuient sur des structures de corrélation spatiale : ils postulent que la corrélation entre individus peut être exprimée à l'aide d'une fonction décroissante de la distance spatiale qui les sépare.

Habituellement en analyse contextuelle, on tente d'expliquer pourquoi les individus résidant au même endroit ont des comportements plus similaires que des individus vivant à des endroits différents. On tient compte pour ce faire des caractéristiques des individus et de leur contexte de résidence.¹⁴ Dans notre cas où la corrélation géographique est capturée au moyen d'une structure de corrélation spatiale, on ne cherche pas seulement à expliquer les similitudes de comportement entre individus vivant dans la même zone mais également les similitudes de comportements entre individus vivant dans des zones proches ou adjacentes. Plusieurs hypothèses complémentaires peuvent être avancées pour expliquer cette corrélation entre zones adjacentes. L'une d'entre elles est que les individus ne sont pas seulement affectés par les caractéristiques de leur zone locale de résidence, mais également par le contexte au-delà de ses limites, puisque leurs activités sociales ne s'arrêtent certainement pas à ces frontières administratives.¹⁷ Toutefois, la meilleure façon de tenir compte de cet état de fait n'est certainement pas de mesurer les caractéristiques du contexte de résidence au niveau de zones administratives plus larges. De telles mesures sont en effet particulièrement inadéquates pour les individus qui résident sur les marges de ces zones. Par conséquent, nous proposons de mesurer les caractéristiques du contexte de résidence dans un espace continu autour du lieu de résidence des individus, plutôt qu'au niveau des zones administratives de résidence aux limites arbitraires.^{55-57,101}

Dans notre analyse des variations spatiales des modes de recours aux soins en France, nous avons cherché à comparer l'approche multiniveau à cette approche spatiale. Il est important pour les décideurs de santé publique de mieux identifier les caractéristiques individuelles et contextuelles associées avec les différents modes de recours aux systèmes de soins ambulatoires. D'une part, la sous-utilisation des soins de spécialistes peut conduire à

restreindre les options de traitement offertes aux patients.^{102,103} Mais d'autre part, au contraire, une sur-utilisation des soins de spécialistes qui ne serait pas accompagnée d'un recours à un médecin généraliste régulièrement suivi par la personne peut conduire à un manque de coordination dans les soins.^{104,105} Utilisant un échantillon d'individus répartis sur le territoire métropolitain français, (a) nous avons entrepris une analyse multiniveau des modes de recours aux soins, et avons examiné si les modèles multiniveaux permettaient de tenir adéquatement compte de la corrélation spatiale des comportements ; (b) nous avons cherché à voir si les modèles mixtes spatiaux constituaient un moyen plus adéquat de rendre compte de la distribution dans l'espace des comportements de recours aux soins ; enfin (c), nous nous sommes demandés si l'on parvenait mieux à expliquer les variations spatiales des modes de recours aux soins en mesurant les facteurs explicatifs du contexte de résidence de façon continue autour du lieu de résidence des individus plutôt qu'au niveau des zones administratives.

Données et méthodes

Bases de données, et variables étudiées

Nous avons utilisé les données des enquêtes sur la Santé et la Protection Sociale conduites par l'Institut de Recherche, d'Etude et de Documentation en Economie de la Santé (IRDES) en 1998 et 2000.¹⁰⁶ La moitié de l'échantillon a été enquêtée en 1998, l'autre moitié en 2000. L'échantillon national métropolitain utilisé dans les analyses est représentatif des assurés des organismes de Sécurité Sociale des salariés, des agriculteurs et des indépendants (à l'exclusion des ayants-droits). Soixante-quatre pour cent des individus contactés ont accepté de participer à l'enquête. Les données d'enquête ont été appariées aux fichiers administratifs des organismes de Sécurité Sociale, qui rapportent les consultations de médecins réalisées par chaque individu sur une période annuelle.

Deux variables-réponse différentes ont été définies pour les analyses. La première variable, définie à partir d'une question d'enquête, indique si les individus avaient ou non un médecin généraliste régulier. L'analyse de cette variable-réponse a été entreprise dans le sous-groupe des individus enquêtés en 2000 ($n = 5\,227$), puisque la question d'enquête correspondante avait seulement été posée cette année-là. La seconde variable-réponse prise en compte indique si plus de la moitié des consultations réalisées par chaque individu au cours d'une année l'avaient été auprès de médecins spécialistes plutôt que de médecins généralistes. Cette variable a été définie à partir des données administratives des organismes de Sécurité Sociale, qui n'ont pu être appariées aux données d'enquête que pour 9309 individus sur

10 955. Nous avons étudié cette variable-réponse parmi les individus qui avaient eu au moins une consultation sur la période annuelle considérée (n = 8 102). Après avoir exclu les individus de moins de 18 ans, les échantillons d'étude étaient de 5 217 pour l'analyse du fait de ne pas avoir de médecin généraliste régulier, et de 8 093 pour l'analyse du pourcentage de consultations de spécialistes.

Concernant les données contextuelles, nous disposons d'informations agrégées au niveau des municipalités de résidence des individus. Il s'agissait d'une part de données socio-économiques issues du recensement de la population. D'autre part, nous disposons du nombre de lieux de consultation de médecins de chaque spécialité dans chaque municipalité (plusieurs médecins de spécialité identique localisés à la même adresse comptant pour un unique lieu de consultation). Les données de localisation des médecins nous ont été fournies par le Ministère de la Santé (Pascale Breuil, Bureau des Professions Médicales).

Variables explicatives

Définition des caractéristiques du contexte au niveau des zones administratives

Le territoire métropolitain Français peut être subdivisé en 36 500 municipalités, ou de façon alternative en 348 zones d'emploi définies par l'Institut National de la Statistique et des Etudes Economiques en agrégeant les municipalités adjacentes entre lesquelles existent des flux importants de déplacement domicile-travail.¹⁰⁷ Dans la base de données réservée à l'analyse du fait de ne pas avoir de médecin généraliste régulier, 3 233 municipalités et 338 zones d'emploi étaient représentées. Dans la base de données pour l'analyse du pourcentage de consultations de spécialistes, 4 421 municipalités et 340 zones d'emploi étaient représentées. Considérant les zones dans cette dernière base de données, le nombre médian d'habitants dans les municipalités était de 2 185 (écart inter-quartile : 794 – 6 533), et de 98 495 dans les zones d'emploi (61 818 – 178 720).

Au niveau des zones administratives, nous avons tenu compte du pourcentage d'habitants avec un faible niveau d'instruction (niveau inférieur au BEPC), et des densités de généralistes et de spécialistes (nombres de lieux de consultations par kilomètre carré). Pour ce qui concerne les spécialistes, nous avons d'abord calculé la densité pour chaque spécialité séparément. Nous avons ensuite déterminé un indicateur de densité commun à l'ensemble des spécialités en calculant la moyenne pondérée des densités calculées pour chacune des spécialités (et préalablement standardisées). Dans le calcul de cette moyenne, nous avons utilisé des pondérations égales aux proportions de consultations annuelles dédiées à chacune

des spécialités. Les différentes variables contextuelles ont été définies au niveau des municipalités et des zones d'emploi.

Une approche spatiale pour définir les indicateurs du contexte

Les trois facteurs contextuels pris en compte dans les analyses (pourcentage d'individus avec un niveau d'instruction faible, offre en médecins généralistes et spécialistes) ont également été mesurés dans un espace continu autour du lieu de résidence des individus.¹⁰¹ Pour mettre en œuvre cette approche, nous avons utilisé un système d'information géographique, dans lequel les municipalités sont géoréférencées en tant que polygones. L'adresse exacte de résidence des individus enquêtés n'étant pas connue, ceux-ci ont été positionnés au centroïde de leur municipalité de résidence.

Le pourcentage d'habitants avec un faible niveau d'instruction était accessible au niveau des municipalités. Toutefois, afin d'éviter de compter sur un espace fragmenté en zones arbitraires, nous avons régulièrement positionné des points tous les kilomètres sur le territoire Français (soit environ 540 000 points au total). Nous avons alors attribué à chacun de ces points le pourcentage d'individus faiblement instruits de la municipalité dans laquelle le point était situé.

Il est nécessaire d'utiliser des poids afin de tenir compte du fait que les points situés à une plus grande distance des individus ont un impact moindre sur eux que des points plus proches. On définit en général ces poids à partir d'une fonction décroissante de la distance arbitrairement choisie.¹⁰¹ Nous avons cherché à les définir de façon moins arbitraire, en tenant compte des distances que les individus parcourent pour aller travailler. Nous supposons que l'impact sur les individus d'un endroit situé à une certaine distance est fonction de leur propension à se rendre régulièrement à cette distance, et avons estimé cette propension à partir de la probabilité que les individus actifs ont de couvrir cette distance pour aller travailler. Nous avons utilisé le recensement de 1999, qui fournit la municipalité de résidence et la municipalité de travail pour les 22 millions d'actifs en France métropolitaine. Pour chaque actif, nous avons calculé la distance en ligne droite du centroïde de sa municipalité de résidence au centroïde de sa municipalité de travail (cette distance étant définie comme égale à 0 pour les actifs qui travaillaient dans leur municipalité de résidence). Nous avons observé que cette distance suivait une distribution exponentielle pour les 22 millions d'actifs. La densité de probabilité estimée de cette distribution est $h(d) = 0.0799 \cdot \exp(-0.0799 \cdot d)$ où d est la distance mesurée en kilomètres. Nous avons utilisé les valeurs de cette fonction

décroissante de la distance comme poids dans le calcul du niveau moyen d'instruction dans le contexte de résidence.

Dans ce calcul, nous considérons l'ensemble des points autour du lieu de résidence des individus ayant un poids au moins égal à 5% de la valeur du poids pour un point situé à 0 kilomètre de l'individu (le poids attribué à chaque point étant calculé à partir de la fonction décroissante de la distance décrite ci-dessus, en tenant compte de la distance du point à l'individu). Puisque $h(d) = 0.05 \cdot h(0)$ pour $d = 37,5$ kilomètres, le niveau d'instruction moyen dans le contexte de résidence a été calculé comme la moyenne pondérée des valeurs aux points situés dans un rayon de 37,5 kilomètres autour du lieu de résidence des individus.

Nous avons également cherché à quantifier le niveau d'offre de soins dans un espace continu autour du lieu de résidence des individus plutôt qu'au niveau des zones administratives. Chaque lieu de consultation a été positionné aléatoirement au sein de sa municipalité (nous ne pouvions localiser plus précisément les lieux d'exercice des médecins dans le système d'information géographique). Pour chaque individu, nous avons déterminé le nombre de lieux de consultation de chaque spécialité dans un rayon de 50 kilomètres. Dans ce calcul, nous avons utilisé les poids issus de la fonction décroissante de la distance décrite ci-dessus, afin de tenir compte du fait que les médecins situés à une plus grande distance des individus sont certainement moins accessibles que les médecins plus proches. Nous avons calculé un tel indicateur spatial pour les médecins généralistes. Pour l'accès aux médecins spécialistes, nous avons calculé un indicateur synthétique en faisant la moyenne pondérée des indicateurs calculés séparément pour chaque spécialité (nous avons utilisé dans ce calcul des poids correspondant aux proportions de consultations réalisées auprès de chaque spécialité dans notre échantillon).

Facteurs d'ajustement au niveau individuel

Les modèles de régression ont été ajustés sur différentes caractéristiques démographiques, socio-économiques, et de santé des individus. Le détail de ces variables d'ajustement individuelles figure dans le tableau 1.

Analyse statistique

Présentation du modèle multiniveau et du modèle spatial mixte

Puisque que nous nous intéressons à des variables réponse binaires, nous considérons le modèle logistique. Soit y_{ij} la valeur de la variable binaire pour l'individu i dans la zone j . Nous pouvons estimer un modèle logistique multiniveau vide (c'est-à-dire ne contenant aucune variable explicative) :²⁷

$$y_{ij} = \mu_j + e_{ij} \quad (\text{équation 1})$$

$$\text{logit}(\mu_j) = \mu_0 + u_j$$

$$u_j \sim N(0, \sigma_u^2)$$

où u_j est le résidu de la zone j associé à l'intercept μ_0 . Le résidu u_j permet de tenir compte du niveau spécifique du phénomène dans la zone j . En estimant la variance σ_u^2 des résidus de niveau zone, le modèle est capable de tenir compte de la structure hiérarchique des données (individus regroupés au sein de zones).

TABLEAU 1. Variables individuelles utilisées comme facteurs d'ajustement dans les modèles de régression sur l'utilisation des soins, France, 1998 et 2000.

Variables	Catégories
Age	Moins de 30 ans*; 30–44 ans; 45–59 ans; 60–74 ans; 75 ans et plus
Sexe	Homme*; femme
Statut matrimonial	Marrié ou en couple*; célibataire; divorcé; veuf
Santé perçue (sur une échelle de 0 à 10)	Score faible (de 0 à 6)*; score moyen-faible (égal à 7); score moyen-élevé (égal à 8); score élevé (égal à 9 ou 10)
Nombre de maladies†	0*; 1 ou 2; 3 ou 4; plus de 4
Niveau d'instruction	Niveau primaire ou moins*; niveau secondaire; université; en cours d'étude
Catégorie socioprofessionnelle‡	Ouvrier non qualifié*; ouvrier qualifié; employé; profession intermédiaire; cadre
Position par rapport à l'emploi	Activité professionnelle*; chômage; autre
Revenu du ménage par unité de consommation	Premier quartile*; second quartile; troisième quartile; quatrième quartile
Assurance santé	Assurance de base seulement*; assurance complémentaire; autre

* Cette catégorie est la catégorie de référence dans le modèle.

† Une liste de maladies a été fournie aux individus pour les aider à rapporter leurs problèmes de santé. Les médecins du CREDES ont complété la liste de maladies de chaque individu en tenant compte de son utilisation de médicaments et de ses consultations de professionnels de santé. Les maladies dentaires ne sont pas prises en compte.

‡ La catégorie socioprofessionnelle a été définie conformément à la Nomenclature des Professions et Catégories Sociales élaborée par l'Institut National de la Statistique et des Etudes Economiques.

Au contraire du modèle multiniveau, le modèle spatial mixte ne s'appuie pas sur un territoire fragmenté en zones arbitraires, mais tient compte de la position de chaque observation dans l'espace.¹⁰⁰ Considérons le modèle logistique suivant :

$$y_i = \beta_i + e_i \quad (\text{équation 2})$$

$$\text{logit}(\beta_i) = \beta_0$$

Soit d_{ij} la distance spatiale entre les lieux de résidence de deux individus i et j . Dans le modèle spatial mixte, la corrélation entre les résidus e_i et e_j pour les individus i et j est une fonction décroissante de la distance d_{ij} .

En géostatistique d'où ces modèles sont issus, il existe des cas dans lesquels deux mesures géologiques répétées au même endroit conduiront à des résultats exactement similaires. On peut tenir compte de cette situation en spécifiant la structure de corrélation suivante pour les résidus du modèle logistique :

$$\text{Corr}(e_i, e_j) = \beta^2 [\exp(-d_{ij}/\beta)] / \beta^2 \quad (\text{équation 3})$$

où $\exp(-d_{ij}/\beta)$ indique que la corrélation entre observations est proportionnelle à la distance spatiale qui les sépare. Avec cette spécification, deux observations localisées au même endroit ($d_{ij} = 0$) sont parfaitement corrélées (corrélation égale à 1).

Toutefois, dans d'autres applications géostatistiques, il peut exister des variations géologiques à une échelle trop fine pour que l'on puisse en tenir compte lors des mesures. Cela conduit à la possibilité d'obtenir des résultats différents si l'on répète une mesure géologique au même endroit. On peut tenir compte d'un tel « effet de pépité » en ajoutant un paramètre supplémentaire à la structure de corrélation :

$$\text{Corr}(e_i, e_j) = \beta^2 [\exp(-d_{ij}/\beta)] / (\beta^2 + \sigma_1^2) \quad (\text{équation 4})$$

Avec cette spécification, deux observations localisées au même endroit peuvent avoir une corrélation qui vaut moins que 1. Ce modèle convient certainement mieux à notre analyse épidémiologique, puisque deux individus résidant au même endroit n'ont pas nécessairement des comportements de recours aux soins identiques.

Analyse des données

L'analyse multiniveau a été entreprise de deux façons différentes. Dans un premier temps, nous avons tenu compte des municipalités comme second niveau du modèle (l'analyse prend donc en compte la municipalité d'appartenance de chaque individu). Dans un second temps, nous avons eu recours aux zones d'emploi comme second niveau du modèle multiniveau. Afin d'obtenir des estimations précises des variations inter-municipalités ou inter-zones d'emploi dans les comportements de recours aux soins, les paramètres des modèles

multiniveaux ont d'abord été estimés au moyen d'une procédure de Monte Carlo par Chaîne de Markov (MLwiN 1.2., Institut d'Education, Londres). Puisque le coefficient de corrélation intraclasse manque de cohérence sur le plan statistique dans le modèle logistique multiniveau,^{108,109} nous avons exprimé les variations inter-zones à l'aide de l'odds ratio médian proposé par Klaus Larsen.⁷⁹ Supposons que nous choissions aléatoirement deux individus avec des caractéristiques identiques dans deux zones différentes. A partir des résidus du modèle multiniveau estimés pour ces deux zones, on calcule l'odds ratio entre l'individu au risque moindre et l'individu au risque le plus élevé. Considérant que cet odds ratio a une certaine distribution (on obtiendrait d'autres valeurs en choisissant des individus dans des zones différentes), l'odds ratio médian se définit comme la valeur médiane de la distribution.

Afin d'examiner si il y a de l'autocorrélation spatiale que les modèles multiniveaux ne parviennent pas à prendre en compte, nous avons calculé la statistique I de Moran¹¹⁰ à partir de fonctions implémentées sous le logiciel R¹¹¹ Tel que nous l'avons calculé, le I de Moran indique si deux zones adjacentes ont des résidus (issus du modèle multiniveau) plus similaires que ce à quoi on s'attendrait si les résidus étaient distribués dans l'espace de façon complètement aléatoire. I est égal à 0 quand il n'y a pas d'autocorrélation spatiale, et est positif lorsque les zones adjacentes présentent des valeurs plus similaires que ce à quoi on s'attendrait sous l'hypothèse d'une distribution spatiale aléatoire.

Les modèles spatiaux mixtes ont été estimés à l'aide de la macro SAS GLIMMIX (version 8.02, Institut SAS, Caroline du Nord, Etats-Unis d'Amérique). Pour modéliser de façon spatiale les données, les individus ont été aléatoirement localisés au sein de leur municipalité de résidence. Afin de voir si le modèle spatial permettait mieux de rendre compte de la variabilité dans les données que les modèles multiniveaux, les modèles multiniveaux ont été ré-estimés à l'aide la macro GLIMMIX. Du fait de la procédure d'estimation utilisée, le critère d'information d'Akaïke ne pouvait pas être utilisé. Nous nous sommes donc servis de la déviance pour comparer les différents modèles.

Tout en tenant compte des variables individuelles dans les modèles, nous avons introduit les variables contextuelles. Celles-ci n'ont été retenues dans les modèles que si elles étaient significativement associées aux variables réponse dans le modèle multiniveau ou dans le modèle spatial mixte. Nous avons successivement estimé les effets du contexte au niveau des municipalités, au niveau des zones d'emploi, ou à l'aide de la méthode spatiale que nous avons présentée. Afin de mieux comparer ces différentes formulations des effets du contexte, les indicateurs ont tous été divisés en quartiles.

TABLEAU 2. Résultats des modèles multiniveaux et des modèles mixtes spatiaux vides sur l'utilisation des soins, France, 1998 et 2000

	Ne pas avoir de médecin généraliste régulier	Avoir un pourcentage élevé de consultations de spécialistes
Modèle multiniveau (individus, municipalités)†		
Variance inter-zone σ_u^2 (écart-type)	0,382 (0,133)**	0,175 (0,059)**
Odds ratio médian	1,80	1,49
I de Moran pour les résidus (écart-type)	0,33 (0,02)***	0,20 (0,02)***
Déviance (scaled deviance)	4738,9	8841,2
Modèle multiniveau (individus, zones d'emploi)†		
Variance inter-zone σ_u^2 (écart-type)	0,249 (0,068)***	0,140 (0,030)***
Odds ratio médian	1,61	1,43
I de Moran pour les résidus (écart-type)	0,24 (0,03)***	0,32 (0,03)***
Déviance (scaled deviance)	4056,3	8625,6
Modèle mixte spatial§		
σ^2 (écart-type)	0,032 (0,008)***	0,033 (0,010)***
σ_1^2 (écart-type)	1,084 (0,023)***	1,116 (0,018)***
σ (écart-type)	16,40 (9,64)*	115,5 (64,7)*
Déviance (scaled deviance)	3603,2	7840,6

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ (Les p sont bilatéraux).

† Les paramètres des modèles multiniveaux ont été estimés par la méthode de Monte Carlo par Chaîne de Markov (MLwiN). Pour calculer le I de Moran, nous avons utilisé les résidus de niveau zone de 67% des municipalités ($n = 2167$) pour le fait de ne pas avoir de généraliste régulier et de 73% des municipalités ($n = 3227$) pour le fait d'avoir un pourcentage élevé de consultations de spécialistes (les autres municipalités n'avaient pas de municipalité adjacente dans l'échantillon). Tous les résidus de niveau zone d'emploi ont été utilisés pour calculer le I de Moran. Du fait de la normalité des résidus, le I de Moran est normal sous l'hypothèse nulle, avec une moyenne égale à 0 et une variance connue. Nous avons calculé un p bilatéral pour le I de Moran. La déviance provient de modèles multiniveaux estimés à l'aide de GLIMMIX.

§ Les paramètres des modèles spatiaux ont été estimés avec GLIMMIX. Le test Z de Wald a été utilisé pour les paramètres de covariance.

Résultats

Douze pour cent des individus ont indiqué qu'ils n'avaient pas de médecin généraliste habituel, et 23% des individus avaient réalisé plus de la moitié de leurs consultations annuelles auprès de spécialistes plutôt que de généralistes. Les modèles multiniveaux ont indiqué que des variations existaient pour chacune des deux variables réponse au niveau des municipalités ou au niveau des zones d'emploi ($p < 0.001$) (tableau 2). Par exemple, l'odds ratio médian était égal à 1,80 pour le fait de ne pas avoir de médecin généraliste régulier dans le modèle multiniveau vide dans lequel les municipalités forment le second niveau du modèle : lorsque l'on choisit aléatoirement deux individus dans des municipalités différentes, l'odds ratio calculé à partir des résidus de niveau municipalité entre l'individu au risque moindre et l'individu au risque le plus élevé était supérieur à 1,80 dans la moitié des cas. La statistique I de Moran était significativement positive dans l'ensemble des modèles multiniveaux ($p < 0.001$). Tenant compte de la corrélation à l'intérieur des zones mais pas de la corrélation existant entre zones adjacentes, le modèle multiniveau est donc incapable de tenir compte de l'ensemble de la corrélation géographique existant dans les données (tableau 2).

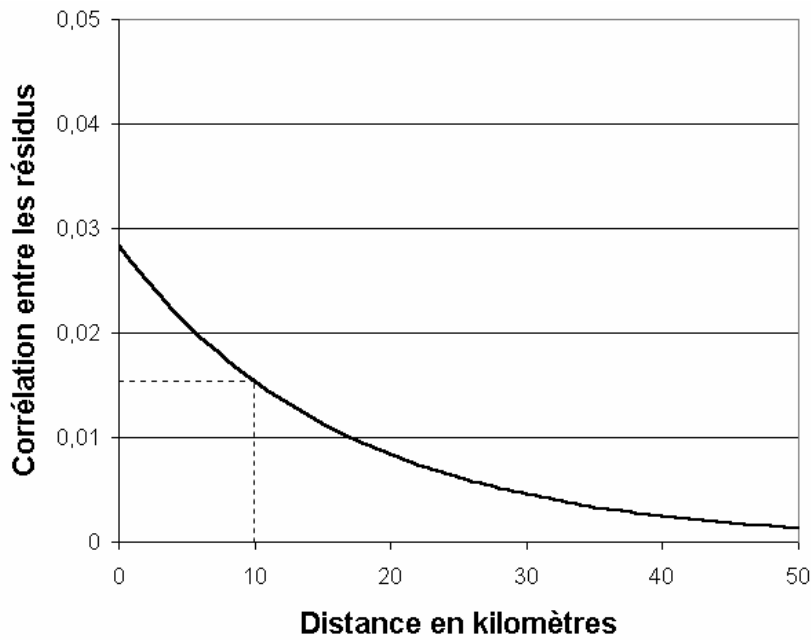
A partir des modèles spatiaux mixtes vides (tableau 2), on a par exemple trouvé que deux individus localisés au même endroit avaient une corrélation égale à 0,028 dans le fait de ne pas avoir de médecin généraliste régulier (figure 1). Cette corrélation entre individus diminuait à mesure qu'augmentait la distance spatiale entre eux. Par exemple, la corrélation n'était plus que de 0,015 pour deux individus distants de 10 kilomètres. Pour chacune des deux variables-réponse, nous avons trouvé que la déviance était largement plus faible dans le modèle spatial mixte vide que dans les modèles multiniveaux vides. Cela indique que la structure de corrélation spatiale permet mieux de rendre compte de la variabilité géographique inhérente aux données que le modèle multiniveau qui fragmente le territoire en zones arbitraires (tableau 2).

Le modèle spatial ajusté sur les facteurs individuels a indiqué qu'un faible pourcentage d'individus faiblement instruits dans le contexte de résidence était associé à des chances accrues de ne pas avoir de médecin généraliste régulier (tableau 3). Toutefois, l'offre en médecins n'était pas associée à cette variable réponse. Un faible pourcentage d'individus faiblement instruits ainsi qu'une offre importante en médecins spécialistes étaient indépendamment associés à des chances plus élevées d'avoir réalisé plus de la moitié de ses consultations annuelles auprès de spécialistes. De façon indépendante, l'offre en médecins généralistes semblait associée de façon négative à cette variable réponse (nous avons retenu

cet effet dans le modèle spatial en dépit de sa non-significativité, parce qu'il était significatif dans les modèles multiniveaux ; c.f. ci-dessous).

FIGURE 1. Corrélation entre les comportements individuels définie à l'aide d'une fonction décroissante de la distance spatiale entre eux, estimée à partir des modèles spatiaux vides pour le fait de ne pas avoir de médecin généraliste régulier et pour le pourcentage de consultations de spécialistes, France, 1998 et 2000.

Ne pas avoir de médecin généraliste régulier



Pourcentage de consultations de spécialistes

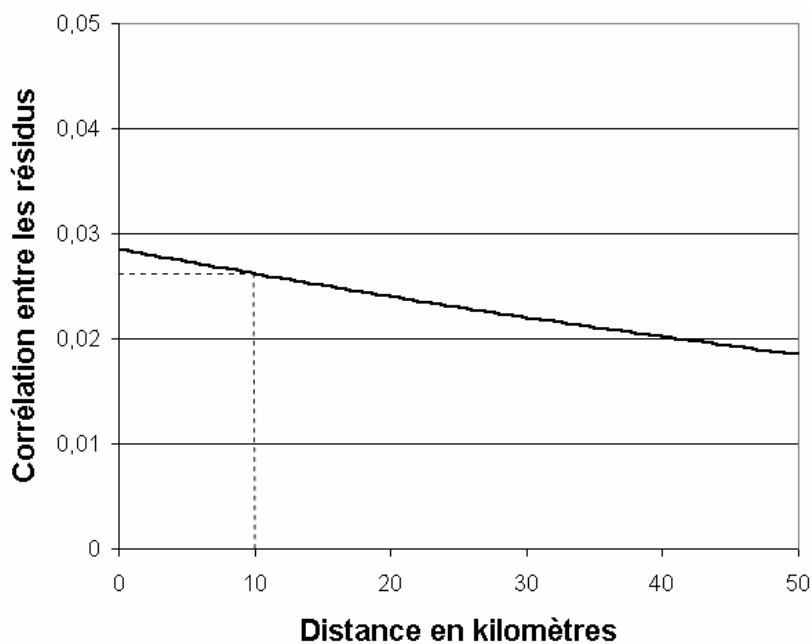


TABLEAU 3. Effets contextuels sur l'utilisation des soins, estimés à partir de modèles mixtes spatiaux ajustés sur les facteurs individuels (les effets de contexte ont été mesurés au niveau des municipalités, au niveau des zones d'emploi, et avec notre approche spatiale), France, 1998 et 2000

	Effets au niveau des municipalités*		Effets au niveau des zones d'emploi*		Effets mesurés à l'aide de l'approche spatiale*	
	OR	IC 95%	OR	IC 95%	OR	IC 95%
Variable réponse : Ne pas avoir de généraliste régulier‡						
Pourcentage d'individus faiblement instruits (réf. : "élevé", 4 ^{ème} quartile)						
Moyen-élevé (troisième quartile)	1,03	0,80 - 1,33	0,97	0,75 - 1,27	1,02	0,75 - 1,37
Moyen-faible (second quartile)	1,27	0,99 - 1,63	1,22	0,93 - 1,59	1,49	1,10 - 2,00
Faible (premier quartile)	1,79	1,38 - 2,31	1,86	1,40 - 2,46	2,24	1,61 - 3,13
Variable réponse : Avoir un pourcentage élevé de consultations de spécialistes‡						
Pourcentage d'individus faiblement instruits (réf. : "élevé", 4 ^{ème} quartile)						
Moyen-élevé (troisième quartile)	1,09	0,92 - 1,30	1,10	0,91 - 1,34	1,18	0,97 - 1,45
Moyen-faible (second quartile)	1,30	1,08 - 1,57	1,20	0,97 - 1,49	1,38	1,09 - 1,74
Faible (premier quartile)	1,50	1,23 - 1,84	1,17	0,90 - 1,53	1,62	1,15 - 2,28
Offre en spécialistes (réf. : "faible", 1 ^{er} quartile)						
Moyen-faible (second quartile)	1,02	0,86 - 1,21	1,20	0,96 - 1,50	1,40	1,05 - 1,88
Moyen-élevé (troisième quartile)	1,19	0,93 - 1,53	0,97	0,71 - 1,33	1,70	1,08 - 2,67
Elevé (quatrième quartile)	1,48	1,05 - 2,09	1,95	1,14 - 3,33	2,03	1,13 - 3,67
Offre en généralistes (réf. : "faible", 1 ^{er} quartile)						
Moyen-faible (second quartile)	0,92	0,77 - 1,09	0,89	0,72 - 1,11	0,88	0,66 - 1,18
Moyen-élevé (troisième quartile)	0,87	0,67 - 1,12	1,05	0,78 - 1,42	0,69	0,44 - 1,09
Elevé (quatrième quartile)	0,82	0,57 - 1,17	0,66	0,39 - 1,10	0,59	0,33 - 1,07

* Les odds ratios ont été ajustés sur l'ensemble des facteurs individuels rapportés dans le tableau 1.

‡ Dans le modèle sur le fait de ne pas avoir de généraliste régulier, la seule variable contextuelle retenue était le pourcentage d'individus peu instruits. Dans le modèle sur le pourcentage de consultations de spécialistes, les trois facteurs contextuels ont été inclus (leurs effets sont mutuellement ajustés).

Bien que les intervalles de confiance soient larges, il est apparu que les effets contextuels étaient plus importants lorsque les variables contextuelles étaient mesurées de façon spatiale plutôt qu'au niveau des zones administratives (tableau 3). Dans le modèle sur le pourcentage de consultations de spécialistes, le niveau d'instruction moyen et l'offre en spécialistes n'étaient associés simultanément et de façon notable avec la variable réponse que lorsque ces facteurs étaient mesurés à l'aide de notre approche spatiale (il convient malgré tout de noter que les intervalles de confiance étaient larges).

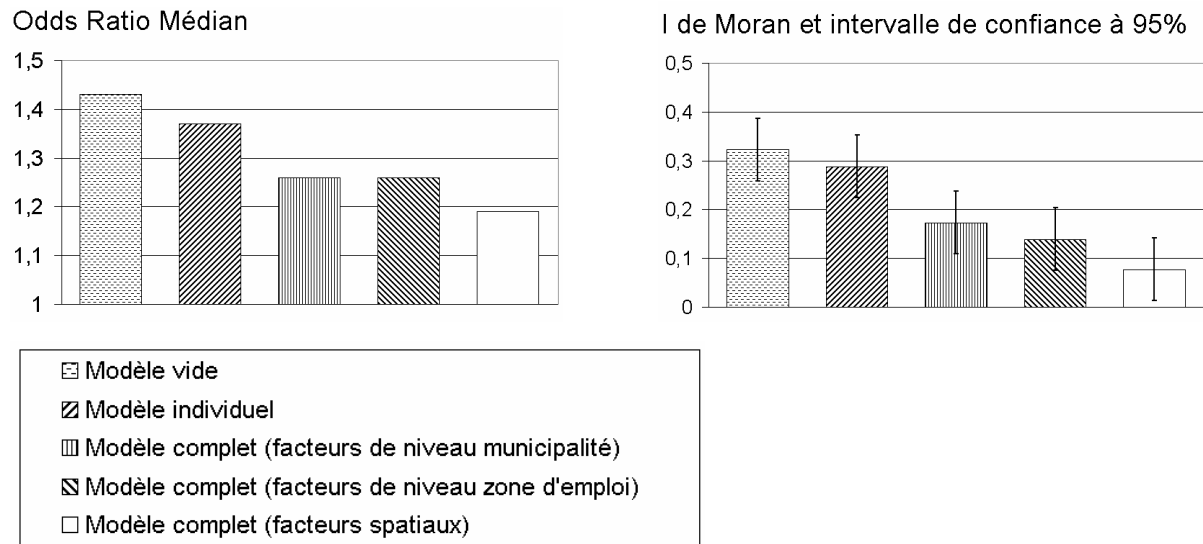
Nous avons calculé l'odds ratio médian et le I de Moran dans les modèles multiniveaux vides, dans les modèles multiniveaux incluant les variables individuelles, et dans les modèles multiniveaux incluant les caractéristiques du contexte mesurées des différentes manières. Nous représentons ces indicateurs en haut de la figure 2 pour le modèle multiniveau sur le pourcentage de consultations de spécialistes dans lequel les individus sont groupés par zones d'emploi. Il est apparu que l'hétérogénéité entre zones d'emploi (traduite au moyen de l'odds ratio médian) diminuait lorsque les variables individuelles et les variables contextuelles ont été introduites dans le modèle. L'odds ratio médian était à son plus faible niveau lorsque les variables contextuelles ont été mesurées à l'aide de notre approche spatiale. De même, le I de Moran a chuté lorsque les caractéristiques individuelles et contextuelles ont été prises en compte dans le modèle, et était à son niveau le plus faible lorsque les variables du contexte étaient mesurées de façon spatiale. Les mêmes résultats ont été obtenus pour le fait de ne pas avoir de médecin généraliste régulier, et pour les modèles multiniveaux tenant compte des municipalités de résidence au second niveau du modèle. Dans tous les modèles multiniveaux à l'exception de celui pour la proportion de consultations de spécialistes dans lequel les individus étaient regroupés par municipalités de résidence, le I de Moran demeurait significativement différent de 0 ($p < 0.05$) après que les facteurs contextuels aient été pris en compte. Même dans les modèles les mieux spécifiés, l'approche multiniveau s'avérait donc incapable de tenir entièrement compte de la corrélation géographique dans les données.

Dans les modèles spatiaux vides, dans les modèles spatiaux incluant les variables individuelles, et dans les modèles spatiaux incluant les facteurs contextuels, nous avons calculé la fonction de corrélation des résidus (qui décroît lorsque augmente la distance spatiale entre les individus). Ces fonctions de corrélation sont représentées au bas de la figure 2 pour les modèles sur la proportion de consultations de spécialistes. On voit encore que la corrélation résiduelle était à son plus faible niveau lorsque les facteurs du contexte ont été

mesurés à l'aide de notre approche spatiale. Un résultat identique a été obtenu à partir des modèles sur le fait de ne pas avoir de médecin généraliste régulier.

FIGURE 2. Explication des variations géographiques des chances d'avoir un pourcentage élevé de consultations de spécialistes, à l'aide des variables individuelles et contextuelles, France, 1998 et 2000. Haut de la figure : odds ratio médian calculé à partir du modèle multiniveau de niveau zone d'emploi, et statistique I de Moran calculée à partir des résidus du modèle (intervalle de confiance à 95%). Bas de la figure : corrélation résiduelle entre individus en fonction de la distance spatiale entre eux, à partir du modèle mixte spatial.

A partir du modèle multiniveau



A partir du modèle spatial

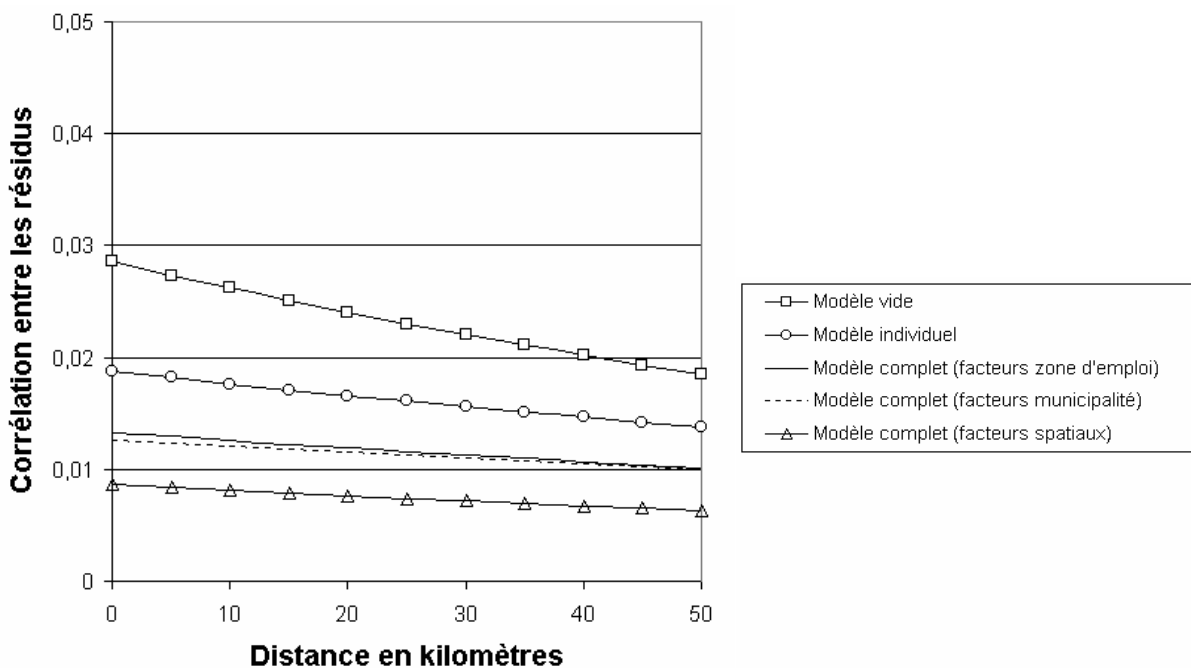
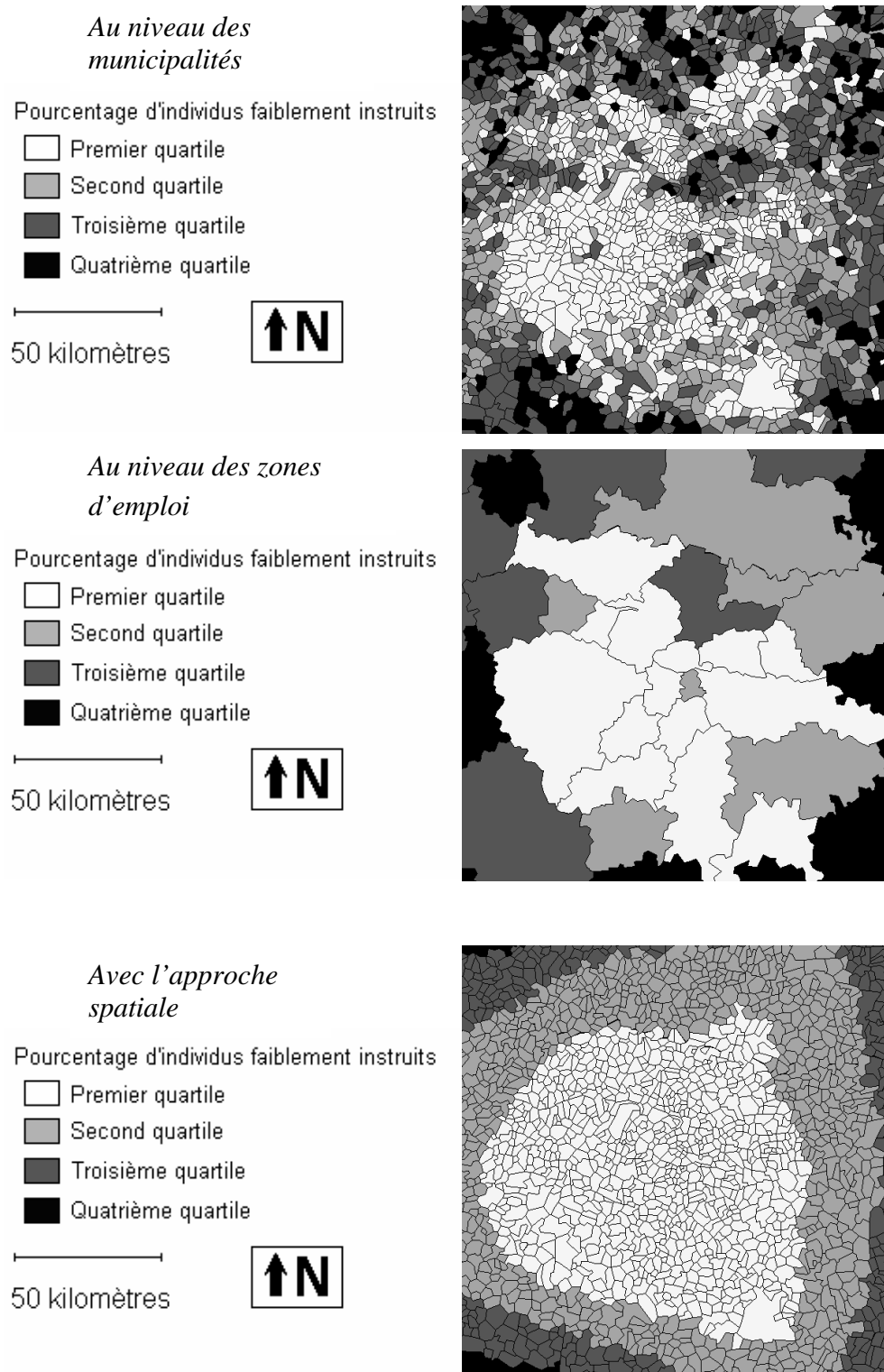


FIGURE 3. Variations du niveau d’instruction moyen dans le contexte de résidence, mesuré au niveau des municipalités, des zones d’emploi, et à l’aide de l’approche spatiale dans une zone rectangulaire autour de Paris, France, 1998 et 2000.



Il est important de noter que les différentes approches de mesure des facteurs du contexte conduisent à des représentations cartographiques différentes lorsque l'on cherche à identifier les zones avec un niveau d'exposition faible ou important à ces caractéristiques contextuelles. Nous illustrons cet état de fait sur la figure 3, où le pourcentage d'individus faiblement instruits dans le contexte de résidence, défini à partir des trois approches différentes, est représenté sur la carte d'une zone qui s'étend autour de la ville de Paris.

Enfin, nous avons cherché à voir si les modèles multiniveaux surestimaient le niveau de significativité des effets contextuels, du fait de leur incapacité à tenir compte de la corrélation spatiale des observations. Considérant divers modèles incluant les facteurs contextuels (modèle logistique classique, modèles multiniveaux, modèle spatial mixte), nous avons divisé le paramètre associé à chaque effet contextuel par son écart-type, afin de quantifier son degré de significativité. Nous avons trouvé que le modèle multiniveau surestimait de façon systématique le degré de significativité des effets contextuels par rapport au modèle spatial mixte (tableau 4). Par exemple, l'effet de l'offre en médecins généralistes sur le pourcentage de consultations de spécialistes était significativement différent de 0 dans les modèles multiniveaux, mais ne l'était plus dans le modèle spatial.

TABLEAU 4. Niveau de significativité des effets du contexte sur l'utilisation des soins (défini en divisant les paramètres par leurs écart-types), estimé à partir des modèles non-multiniveaux, des modèles multiniveaux, et des modèles mixtes spatiaux, France, 1998 et 2000

	Modèle non- multiniveau	Modèle multiniveau (municipalités)	Modèle multiniveau (zones d'emploi)	Modèle mixte spatial
Variable réponse : Ne pas avoir de généraliste régulier				
Pourcentage d'individus faiblement instruits, premier quartile	7,0	7,1	6,0	4,8
Variable réponse : Avoir un pourcentage élevé de consultations de spécialistes				
Pourcentage d'individus faiblement instruits, premier quartile	5,5	5,7	4,7	2,8
Offre en spécialistes, quatrième quartile	3,1	3,4	3,3	2,4
Offre en généralistes, quatrième quartile	3,1	3,3	3,0	1,7

Discussion

Nous avons trouvé qu'une approche qui recourt à des méthodes de modélisation spatiale et définit les effets du contexte dans un espace continu autour du lieu de résidence des individus plutôt qu'au niveau de leur zone administrative de résidence était plus à même de rendre compte des variations spatiales des modes de recours aux soins en France.

Dans notre exemple, les modèles multiniveaux n'ont pas permis de tenir adéquatement compte de la corrélation géographique des comportements de recours aux soins. Les modèles multiniveaux postulent que les zones à haut risque et les zones à faible risque sont aléatoirement réparties dans l'espace. Au contraire, la statistique I de Moran calculée à partir des résidus des modèles multiniveaux indique que des individus vivant dans des zones adjacentes tendaient à avoir des comportements plus similaires que ce à quoi on pourrait s'attendre sous l'hypothèse d'une distribution spatiale aléatoire. Les modèles mixtes spatiaux ont mieux permis de rendre compte de la variabilité spatiale des comportements de recours aux soins que les modèles multiniveaux. En définissant la corrélation géographique à l'aide d'une fonction décroissante de la distance entre individus, les modèles mixtes spatiaux étaient capables de tenir compte de l'autocorrélation spatiale que les modèles multiniveaux ne pouvaient capturer.

Par conséquent, de même que les modèles logistiques non-multiniveaux qui ne tiennent pas compte de la corrélation des comportements à l'intérieur des zones conduisent souvent à des inférences erronées,²⁴ les modèles multiniveaux qui ne prennent pas en compte l'autocorrélation spatiale existant dans les données peuvent surestimer le niveau de significativité des effets contextuels et aboutir à de fausses inférences. De façon plus fondamentale, afin de mieux comprendre la distribution spatiale des comportements de recours aux soins, il est intéressant de montrer qu'une structure de corrélation spatiale permet mieux de capter les variations du phénomène que ne le font les modèles multiniveaux.

Dans une perspective explicative, nous nous posons les deux questions suivantes. Premièrement, pourquoi les individus vivant dans une même municipalité ont-ils des comportements plus similaires que des individus vivant dans des municipalités différentes ? Deuxièmement, pourquoi les individus vivant dans des municipalités proches ou adjacentes présentent-ils des similitudes de comportements ? Dans nos analyses, le I de Moran (qui mesure la corrélation entre individus issus de municipalités adjacentes) a diminué lorsque les facteurs contextuels mesurés au niveau des municipalités ont été introduits dans les modèles. En effet, des municipalités proches l'une de l'autre ont en moyenne des caractéristiques plus similaires que des municipalités plus éloignées. En conséquence, l'inclusion des facteurs de

niveau municipalité dans les modèles a en partie permis d'expliquer l'autocorrélation spatiale des comportements de recours aux soins.

Une explication complémentaire à la corrélation des comportements observée entre municipalités adjacentes est peut-être que les individus ne sont pas seulement affectés par les caractéristiques de leur municipalité de résidence, mais également par les caractéristiques des municipalités adjacentes. Une option pour tenir compte de cet état de fait serait de définir les facteurs du contexte au niveau de zones administratives plus larges (comme les zones d'emploi que nous avons utilisées). Toutefois, cette solution n'est pas adéquate non plus, en particulier pour les individus résidant aux marges de ces zones. Par conséquent, nos indicateurs qui mesurent les caractéristiques du contexte dans un espace continu autour du lieu de résidence des individus pourraient constituer une alternative intéressante. Ceux-ci permettent de tenir compte du fait que les individus ne sont pas seulement affectés par les caractéristiques de leur lieu de résidence, mais également (quoique de façon moins intense) par les caractéristiques des lieux alentours. Nous avons observé que les différentes approches de mesure des effets du contexte conduisent à une identification cartographique différente des zones avec un risque élevé ou faible.

Dans les modèles multiniveaux, les variables contextuelles définies à l'aide de notre approche spatiale ont permis d'expliquer une plus grande part de la corrélation des comportements à l'intérieur des zones et entre zones adjacentes que les indicateurs construits à partir des zones administratives. Dans les modèles spatiaux, ces indicateurs définis dans un espace continu autour du lieu de résidence des individus ont également permis d'expliquer une plus grande part de l'autocorrélation spatiale.

Nous sommes conscients du fait que l'approche multiniveau est plus adaptée lorsque le contexte est défini de façon non géographique (comme lorsque l'on s'intéresse à des lieux de travail ou à des écoles), ou lorsque l'on étudie des processus qui opèrent spécifiquement à l'échelle des zones administratives (par exemple, liés aux politiques publiques). Toutefois, dans beaucoup d'analyses, il pourrait être plus approprié d'étudier les variations géographiques des phénomènes de santé dans un espace continu, en utilisant des techniques de modélisation spatiale et des indicateurs qui capturent l'espace de façon continue.

Conclusion générale et perspectives

L'analyse des effets du contexte sur la santé a connu un développement considérable au cours des dix dernières années,^{3,112} et est aujourd'hui considérée comme l'une des principales voies à suivre pour avancer dans la compréhension des disparités sociales de santé.⁴ Les premières générations d'analyses contextuelles ont confirmé qu'il y a un réel intérêt à tenir compte des effets du contexte de résidence lors de l'étude d'un certain nombre de phénomènes de santé. Toutefois, l'analyse contextuelle telle qu'elle est pratiquée aujourd'hui a certainement atteint ses limites, et est ainsi mise en demeure de réviser ses schémas d'analyse pour permettre de nouvelles avancées dans la compréhension des effets du contexte sur la santé.⁹ Les limites de la méthodologie utilisée renvoient : 1) à l'utilisation d'une distinction trop rigide entre effets individuels et effets contextuels,¹¹³ 2) à un intérêt peut-être trop exclusif pour un petit nombre de facteurs contextuels (souvent de type socio-économique), sans que des efforts suffisants n'aient encore été mis en œuvre pour identifier de façon plus précise les mécanismes d'influence du contexte sur la santé,¹⁷ 3) à l'utilisation de données transversales, qui s'ajoute au problème des biais de confusion pour compromettre définitivement l'identification d'effets causaux du contexte,¹¹⁴ et enfin 4) à l'utilisation de l'approche d'analyse multiniveau, qui ne permet peut-être pas de décrire et de rendre compte efficacement des variations inter-zones des phénomènes de santé.

La période initiale d'engouement pour cette thématique des effets du contexte sur la santé, qui correspond peut-être à l'enfance de ce champ d'analyse, doit maintenant faire place à une période à la fois critique et constructive, au cours de laquelle les différentes limites évoquées ci-dessus devront être discutées et résolues. Ce n'est qu'en s'attachant véritablement à surmonter ces difficultés qu'il sera possible d'accroître l'intérêt de ces analyses contextuelles pour le champ de la santé publique. Au cours de notre travail, nous avons pris conscience de la nécessité d'avancer dans cette voie, et avons commencé à proposer certaines solutions. Nous nous sommes tout particulièrement intéressés à la question des méthodes d'analyse des variations spatiales et des méthodes de mesure des facteurs du contexte.

L'objectif d'une série d'articles publiés dans le *Journal of Epidemiology and Community Health* sous la direction de Juan Merlo^{12,41} est de souligner que la modélisation de la variance inter-zone fournit elle-même des informations d'intérêt en santé publique, au-delà de la modélisation des associations entre facteurs contextuels et phénomènes de santé.^{42,109} Un tel message n'avait probablement pas été délivré de façon aussi claire ni aussi complète dans la

littérature d'épidémiologie sociale qu'au travers de la série d'articles que nous avons écrite de façon collaborative.

Cette distinction entre modélisation des associations et modélisation de la variance a servi de point de départ pour la suite de notre travail, puisque nous nous efforçons de montrer que dans beaucoup de cas d'analyses, à la fois les mesures d'association et les mesures de variation fournissent des informations incomplètes en santé publique lorsqu'elles sont mises en œuvre au travers de l'approche d'analyse multiniveau. Dans beaucoup de cas, ces deux types d'indicateurs permettent d'aller plus avant dans la description et la compréhension de la distribution spatiale des phénomènes de santé lorsque l'on s'appuie sur une vision plus continue de l'espace. Au travers d'une étude des variations des modes de recours aux soins sur le territoire métropolitain Français, nous avons donc cherché à proposer une perspective d'analyse spatiale des variations géographiques des phénomènes de santé.

Toutefois, plutôt que d'apporter des réponses définitives, de tels travaux laissent simplement entrevoir que les processus contextuels qui influent sur la santé des individus se distribuent de façon complexe dans l'espace. Dans le cadre de travaux futurs conduits dans l'Equipe Avenir « Déterminants Sociaux de la Santé et du Recours aux Soins » de l'unité 707 de l'INSERM, nous chercherons à aller plus avant dans l'étude des distributions spatiales des phénomènes de santé et des mécanismes contextuels qui en sont à l'origine.

Perspectives de recherche

Notre projet de recherche futur nous conduira tout particulièrement à nous interroger sur les méthodes d'analyse à mettre en œuvre et les données à utiliser pour avancer dans la compréhension des effets du contexte sur la santé. L'objectif de ces progrès méthodologiques est de fertiliser de nouvelles approches d'intervention dans le champ de la santé publique. Notre projet de recherche s'articulera autour des interrogations suivantes :

1) Quelles approches faut-il mettre en œuvre pour décrire la distribution spatiale des phénomènes de santé ?

Au cours de nos premiers travaux, nous avons utilisé différentes méthodes de régression spatiale afin d'obtenir des informations sur la distribution spatiale des phénomènes.^{52,98-100,115} D'autres approches statistiques existent que nous n'avons pas encore testées, et certaines méthodes d'analyse sont en cours de constitution. Nous projetons d'engager une comparaison générale des différentes approches d'analyse. L'objectif essentiel sera alors d'évaluer l'intérêt relatif des résultats qu'elles fournissent sur un plan de santé publique. Cherchant à voir si les différentes méthodes permettent de capter la spécificité des distributions spatiales des

phénomènes, il sera intéressant de comparer différents phénomènes de santé dont les distributions spatiales seraient largement différentes. La question sera alors de voir si les approches d'analyse envisagées parviennent à fournir des informations assez spécifiques sur chacune des distributions spatiales, qui puissent s'avérer utiles lors de la mise en place de programmes de santé publique adaptés à chaque problème de santé.

2) Quelles approches faut-il mettre en œuvre pour mesurer les facteurs du contexte qui influent sur la santé ?

Dans notre recherche future, nous nous efforcerons de travailler à une meilleure identification des processus par lesquels le contexte est susceptible d'influer sur la santé. Nous chercherons à voir si l'on parvient à clarifier le mode opératoire des effets contextuels en distinguant différents types de facteurs, et notamment différentes dimensions économiques et sociales du contexte de résidence. Toutefois, une condition primordiale pour avancer dans cette voie est de raffiner les méthodes de mesure des facteurs du contexte, à la fois dans l'espace et dans le temps :

2A – Mesure dans l'espace : Un second volet de notre travail a consisté à proposer différentes méthodes de mesure des facteurs du contexte dans un espace continu autour du lieu de résidence des individus. Une telle approche est complètement nouvelle et devra être perfectionnée. Des analyses de sensibilité devront être conduites à chaque fois, afin d'examiner si le surcroît de complexité dans les approches de mesure apparaît justifié au regard des résultats obtenus. Par exemple, est-il justifié d'introduire des pondérations dans le calcul des indicateurs contextuels, pondérations qui permettent de tenir compte du fait que les individus sont avant tout affectés par les localisations les plus proches et dans une moindre mesure par les localisations plus éloignées ? Dans cette voie, on peut même imaginer de mettre au point des procédures d'estimation qui permettent de sélectionner la pondération la plus appropriée, ce qui permettrait de savoir dans quelle mesure une localisation située à proximité a un impact plus important sur les individus qu'une localisation située deux fois plus loin.

Nos approches de mesure des facteurs du contexte tiennent compte d'un espace continu autour du lieu de résidence des individus. Cela constitue un progrès par rapport aux mesures classiques qui négligent l'espace au-delà des limites administratives de la zone de résidence des individus. Toutefois, des limites physiques et sociales existent dans l'espace, et l'hypothèse d'une parfaite continuité spatiale n'est certainement pas adéquate. Dans nos études à venir conduites à l'échelle d'une ville donnée (Malmö en Suède, Paris en France), nous proposons de considérer l'intégralité du réseau routier de ces villes, qui nous permettra

de découper l'espace en pâtés de maison au sein desquels nous calculerons les indicateurs contextuels. Nous chercherons à voir si ces mesures contextuelles qui réintroduisent des discontinuités dans un espace continu permettent encore mieux de rendre compte des variations spatiales des phénomènes de santé.

2B – Mesure dans le temps : Une limite importante de la littérature est que les associations entre facteurs contextuels et variables de santé ont souvent été mises en évidence à partir de données transversales.¹¹⁴ Dans la suite de nos travaux, ayant accès à des données longitudinales tant Françaises que Suédoises, nous chercherons à étudier l'impact que les facteurs du contexte peuvent avoir sur l'incidence ultérieure de problèmes de santé. Dans cette perspective, nous chercherons à voir si les effets du milieu de résidence jouent de façon cumulative dans le temps : une exposition prolongée à un facteur contextuel donné est-elle à l'origine d'un effet plus important sur les individus qu'une exposition moins durable ? Dans cette optique, connaissant l'histoire résidentielle des individus, il sera d'une part nécessaire de tenir compte des caractéristiques des lieux de résidence successifs des individus, mais également des changements de caractéristiques des différents lieux de résidence au cours du temps. L'objectif sera de voir si cette mesure cumulative de l'exposition au contexte permet mieux de rendre compte des variations des phénomènes de santé que l'approche classique qui la capture à un moment donné.

3) Dans quelle mesure la moindre exhaustivité et la moindre précision spatiale de certaines sources de données utilisées conduisent-elles à une perte d'informations pertinentes en santé publique ?

Dans les prochaines années, l'Equipe Avenir sur les « Déterminants Sociaux de la Santé et du Recours aux Soins » sera amenée à renforcer sa coopération avec le Département de Médecine Communautaire de l'Hôpital Universitaire de Malmö, au travers d'une collaboration avec Juan Merlo. Pour nos travaux futurs, nous aurons donc accès à des données Suédoises, caractérisées par l'exhaustivité quant à la population couverte, par une très grande précision spatiale, et par une dimension longitudinale. Concernant la situation Française, nous comptons notamment travailler à partir des données de l'Echantillon Démographique Permanent de l'INSEE, qui pour riches qu'elles soient, n'atteignent pas des degrés d'exhaustivité et de précision spatiale comparables. Cherchant à comparer diverses méthodes d'analyse des variations spatiales et de mesure des facteurs du contexte, nous examinerons à chaque fois si la moindre précision des données Françaises aboutit à une perte d'informations qui seraient utiles d'un point de vue de santé publique.

S'attachant à remplir ces objectifs de recherche, nos travaux futurs devraient contribuer à lever des limites inhérentes à l'analyse contextuelle telle qu'elle est pratiquée aujourd'hui en épidémiologie sociale, et permettre ainsi d'affirmer l'importance de son rôle en santé publique.

Liste de références générale

- 1 Diez-Roux AV. Bringing context back into epidemiology: variables and fallacies in multilevel analysis. *Am J Public Health* 1998;88:216-22.
- 2 Duncan C, Jones K, Moon G. Health-related behaviour in context: a multilevel modelling approach. *Soc Sci Med* 1996;42:817-830.
- 3 Pickett KE, Pearl M. Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. *J Epidemiol Community Health* 2001;55:111-22.
- 4 Kawachi I, Berkman LF, editors. *Neighborhoods and Health*. New York, NY: Oxford University Press, 2003.
- 5 Mooij T. Pupil-class determinants of aggressive and victim behaviour in pupils. *Br J Educ Psychol* 1998;68:373-385.
- 6 Palmer RF, Graham JW, White EL, Hansen WB. Applying multilevel analytic strategies in adolescent substance use prevention research. *Prev Med* 1998;27:328-336.
- 7 Johnson RA, Hoffmann JP. Adolescent cigarette smoking in U.S. racial/ethnic subgroups: findings from the National Education Longitudinal Study. *J Health Soc Behav* 2000;41:392-407.
- 8 Kivimaki M, Vahtera J, Pentti J, Ferrie JE. Factors underlying the effect of organisational downsizing on health of employees: longitudinal cohort study. *BMJ* 2000;320:971-975.
- 9 O'Campo P. Invited commentary: Advancing theory and methods for multilevel models of residential neighborhoods and health. *Am J Epidemiol* 2003;157:9-13.
- 10 Merlo J, Östergren PO, Hagberg O, Lindstrom M, Lindgren A, Melander A, et al. Diastolic blood pressure and area of residence: multilevel versus ecological analysis of social inequity. *J Epidemiol Community Health* 2001;55:791-8.
- 11 Merlo J, Lynch JW, Yang M, Lindstrom M, Ostergren PO, Rasmusen NK, et al. Effect of neighborhood social participation on individual use of hormone replacement therapy and antihypertensive medication: a multilevel analysis. *Am J Epidemiol* 2003;157:774-83.
- 12 Merlo J, Chaix B, Yang M, Lynch JW, Rastam L. A brief conceptual tutorial on multilevel analysis in social epidemiology – linking the statistical concept of clustering to the idea of contextual phenomenon. *J Epidemiol Community Health* 2004; in press.
- 13 O'Campo P, Rao RP, Gielen AC, Royalty W, Wilson M. Injury-producing events among children in low-income communities: the role of community characteristics. *J Urban Health* 2000;77:34-49.

- 14 Diez-Roux AV. Multilevel analysis in public health research. *Annu Rev Public Health* 2000;21:171-92.
- 15 Pampalon R, Duncan C, Subramanian SV, Jones K. Geographies of health perception in Quebec: a multilevel perspective. *Soc Sci Med* 1999;48:1483-1490.
- 16 Diez-Roux AV. Investigating neighborhood and area effects on health. *Am J Public Health* 2001;91:1783-9.
- 17 Morenoff JD. Neighborhood mechanisms and the spatial dynamics of birth weight. *AJS* 2003;108:976-1017.
- 18 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;87:1113-1118.
- 19 Karvonen S, Rimpela A. Socio-regional context as a determinant of adolescents' health behaviour in Finland. *Soc Sci Med* 1996;43:1467-1474.
- 20 Tuinstra J, Groothoff JW, van den Heuvel WJ, Post D. Socio-economic differences in health risk behavior in adolescence: do they exist? *Soc Sci Med* 1998;47:67-74.
- 21 Karvonen S, Rimpela AH. Urban small area variation in adolescents' health behaviour. *Soc Sci Med* 1997;45:1089-1098.
- 22 Sundquist J, Malmstrom M, Johansson SE. Cardiovascular risk factors and the neighbourhood environment: a multilevel analysis. *Int J Epidemiol* 1999;28:841-845.
- 23 Huston SL, Evenson KR, Bors P, Gizlice Z. Neighborhood environment, access to places for activity, and leisure-time physical activity in a diverse North Carolina population. *Am J Health Promot* 2003;18:58-69.
- 24 Snijders T, Bosker R. *Multilevel Analysis. An introduction to basic and advanced multilevel modelling*. London, England: Sage Publications, 1999.
- 25 Goldstein H, Browne W, Rasbash J. Multilevel modelling of medical data. *Stat Med* 2002;21:3291-315.
- 26 Goldstein H. *Multilevel Statistical Models*. 2nd ed. London: Edward Arnold, 1995.
- 27 Leyland AH, Goldstein H. *Multilevel modelling of health statistics*. Chichester, England: Wiley, 2001.
- 28 Bobashev GV, Anthony JC. Clusters of marijuana use in the United States. *Am J Epidemiol* 1998;148:1168-74.
- 29 Bobashev GV, Anthony JC. Use of alternating logistic regression in studies of drug-use clustering. *Subst Use Misuse* 2000;35:1051-73.

- 30 Preisser JS, Arcury TA, Quandt SA. Detecting patterns of occupational illness clustering with alternating logistic regressions applied to longitudinal data. *Am J Epidemiol* 2003;158:495-501.
- 31 Petronis KR, Anthony JC. A different kind of contextual effect: geographical clustering of cocaine incidence in the USA. *J Epidemiol Community Health* 2003;57:893-900.
- 32 Raudenbush SW, Bryk AS. A hierarchical model for studying school effects. *Sociology of Education* 1986;59:1-17.
- 33 Duncan C, Jones K, Moon G. Psychiatric morbidity: a multilevel approach to regional variations in the UK. *J Epidemiol Community Health* 1995;49:290-295.
- 34 O'Campo P, Gielen AC, Faden RR, Xue X, Kass N, Wang MC. Violence by male partners against women during the childbearing year: a contextual analysis. *Am J Public Health* 1995;85:1092-1097.
- 35 Carr-Hill RA, Rice N, Roland M. Socioeconomic determinants of rates of consultation in general practice based on fourth national morbidity survey of general practices. *BMJ* 1996;312:1008-1012.
- 36 Diez Roux AV, Merkin SS, Hannan P, Jacobs DR, Kiefe CI. Area characteristics, individual-level socioeconomic indicators, and smoking in young adults: the coronary artery disease risk development in young adults study. *Am J Epidemiol* 2003;157:315-26.
- 37 Kennedy BP, Kawachi I, Glass R, Prothrow-Stith D. Income distribution, socioeconomic status, and self rated health in the United States: multilevel analysis. *BMJ* 1998;317:917-921.
- 38 Chaix B, Chauvin P. L'apport des modèles multiniveau dans l'analyse contextuelle en épidémiologie sociale : une revue de la littérature. *Rev Epidemiol Sante Publique* 2002;50:489-499.
- 39 Merlo J, Östergren PO, Broms K, Bjorck-Linne A, Liedholm H. Survival after initial hospitalisation for heart failure: a multilevel analysis of patients in Swedish acute care hospitals. *J Epidemiol Community Health* 2001;55:323-329.
- 40 Wang J, Siegal HA, Falck RS, Carlson RG. Needle transfer among injection drug users: a multilevel analysis. *Am J Drug Alcohol Abuse* 1998;24:225-237.
- 41 Merlo J, Yang M, Chaix B, Lynch JW, Rastam L. A brief conceptual tutorial on multilevel analysis in social epidemiology – investigating contextual phenomena in different groups of individuals. *J Epidemiol Community Health* 2004; in press.

- 42 Merlo J. Multilevel analytical approaches in social epidemiology: measures of health variation compared with traditional measures of association. *J Epidemiol Community Health* 2003;57:550-2.
- 43 Fotheringham AS, Wong DWS. The modifiable areal unit problem in multivariate statistical analysis. *Environ Plan A* 1991;23:1025-1044.
- 44 Amrhein CG. Searching for the elusive aggregation effect: evidence from statistical simulations. *Environ Plan A* 1995;27:105-119.
- 45 Martin D. An assessment of surface and zonal models of population. *Int J Geographical Information Systems* 1996;10:973-989.
- 46 Holt D, Steel DG, Tranmer M. Area homogeneity and the modifiable areal unit problem. *Geographical Systems* 1996;3:181-200.
- 47 Mitchell R. Multilevel modeling might not be the answer. *Environ Plan A* 2001;33:1357-1360.
- 48 Kleinschmidt I, Sharp BL, Clarke GP, Curtis B, Fraser C. Use of generalized linear mixed models in the spatial analysis of small-area malaria incidence rates in Kwazulu Natal, South Africa. *Am J Epidemiol* 2001;153:1213-21.
- 49 English PB, Kharrazi M, Davies S, Scalf R, Waller L, Neutra R. Changes in the spatial pattern of low birth weight in a southern California county: the role of individual and neighborhood level factors. *Soc Sci Med* 2003;56:2073-88.
- 50 Kleinschmidt I, Sharp B, Mueller I, Vounatsou P. Rise in malaria incidence rates in South Africa: a small-area spatial analysis of variation in time trends. *Am J Epidemiol* 2002;155:257-64.
- 51 Joines JD, Hertz-Picciotto I, Carey TS, Gesler W, Suchindran C. A spatial analysis of county-level variation in hospitalization rates for low back problems in North Carolina. *Soc Sci Med* 2003;56:2541-53.
- 52 Werneck GL, Maguire JH. Spatial modeling using mixed models: an ecologic study of visceral leishmaniasis in Teresina, Piaui State, Brazil. *Cad Saude Publica* 2002;18:633-7.
- 53 Leyland AH, Langford IH, Rasbash J, Goldstein H. Multivariate spatial models for event data. *Stat Med* 2000;19:2469-78.
- 54 Langford IH, Leyland AH, Rasbash J, Goldstein H. Multilevel modelling of the geographical distributions of diseases. *J R Stat Soc Ser C Appl Stat* 1999;48:253-68.
- 55 Treno AJ, Gruenewald PJ, Johnson FW. Alcohol availability and injury: the role of local outlet densities. *Alcohol Clin Exp Res* 2001;25:1467-71.

- 56 Liu GC, Cunningham C, Downs SM, Marrero DG, Fineberg N. A spatial analysis of obesogenic environments for children. *Proc AMIA Symp* 2002;459-63.
- 57 Ali M, Emch M, Tofail F, Baqui AH. Implications of health care provision on acute lower respiratory infection mortality in Bangladeshi children. *Soc Sci Med* 2001;52:267-77.
- 58 Fotheringham AS, Charlton ME, Brunsdon C. Spatial variations in school performance: a local analysis using geographically weighted regression. *Geographical & Environmental Modelling* 2001;5:43-66.
- 59 Kleinschmidt I, Hills M, Elliott P. Smoking behaviour can be predicted by neighbourhood deprivation measures. *J Epidemiol Community Health* 1995;49 Suppl 2:S72-S77.
- 60 Hart C, Ecob R, Smith GD. People, places and coronary heart disease risk factors: a multilevel analysis of the Scottish Heart Health Study archive. *Soc Sci Med* 1997;45:893-902.
- 61 Duncan C, Jones K, Moon G. Smoking and deprivation: are there neighbourhood effects? *Soc Sci Med* 1999;48:497-505.
- 62 Reading R, Langford IH, Haynes R, Lovett A. Accidents to preschool children: comparing family and neighbourhood risk factors. *Soc Sci Med* 1999;48:321-330.
- 63 Kalff AC, Kroes M, Vles JS, Hendriksen JG, Feron FJ, Steyaert J, et al. Neighbourhood level and individual level SES effects on child problem behaviour: a multilevel analysis. *J Epidemiol Community Health* 2001;55:246-250.
- 64 Cunradi CB, Caetano R, Clark C, Schafer J. Neighborhood poverty as a predictor of intimate partner violence among White, Black, and Hispanic couples in the United States: a multilevel analysis. *Ann Epidemiol* 2000;10:297-308.
- 65 Ross CE. Walking, exercising, and smoking: does neighborhood matter? *Soc Sci Med* 2000;51:265-274.
- 66 Ellaway A, Anderson A, Macintyre S. Does area of residence affect body size and shape? *Int J Obes Relat Metab Disord* 1997;21:304-8.
- 67 van Lenthe FJ, Mackenbach JP. Neighbourhood deprivation and overweight: the GLOBE study. *Int J Obes Relat Metab Disord* 2002;26:234-40.
- 68 Robert SA, Reither EN. A multilevel analysis of race, community disadvantage, and body mass index among adults in the US. *Soc Sci Med* 2004;59:2421-34.
- 69 Diez-Roux AV, Nieto FJ, Muntaner C, Tyroler HA, Comstock GW, Shahar E, et al. Neighborhood environments and coronary heart disease: a multilevel analysis. *Am J Epidemiol* 1997;146:48-63.

- 70 Yen IH, Kaplan GA. Neighborhood social environment and risk of death: multilevel evidence from the Alameda County Study. *Am J Epidemiol* 1999;149:898-907.
- 71 Groenewegen PP, Leufkens HG, Spreeuwenberg P, Worm W. Neighbourhood characteristics and use of benzodiazepines in The Netherlands. *Soc Sci Med* 1999;48:1701-11.
- 72 Johnell K, Merlo J, Lynch J, Blennow G. Neighbourhood social participation and women's use of anxiolytic-hypnotic drugs: a multilevel analysis. *J Epidemiol Community Health* 2004;58:59-64.
- 73 Weich S, Holt G, Twigg L, Jones K, Lewis G. Geographic variation in the prevalence of common mental disorders in Britain: a multilevel investigation. *Am J Epidemiol* 2003;157:730-7.
- 74 Weich S, Twigg L, Holt G, Lewis G, Jones K. Contextual risk factors for the common mental disorders in Britain: a multilevel investigation of the effects of place. *J Epidemiol Community Health* 2003;57:616-21.
- 75 Reijneveld SA, Schene AH. Higher prevalence of mental disorders in socioeconomically deprived urban areas in The Netherlands: community or personal disadvantage? *J Epidemiol Community Health* 1998;52:2-7.
- 76 Wainwright NW, Surtees PG. Places, people, and their physical and mental functional health. *J Epidemiol Community Health* 2004;58:333-9.
- 77 Driessen G, Gunther N, Van Os J. Shared social environment and psychiatric disorder: a multilevel analysis of individual and ecological effects. *Soc Psychiatry Psychiatr Epidemiol* 1998;33:606-612.
- 78 Larsen K, Merlo J. Appropriate assessment of neighborhood effects on individual health – integrating random and fixed effects in multilevel logistic regression. *Am J Epidemiol* 2004; in press.
- 79 Larsen K, Petersen JH, Budtz-Jorgensen E, Endahl L. Interpreting parameters in the logistic regression model with random effects. *Biometrics* 2000;56:909-14.
- 80 Merlo J, Chaix B, Yang M, Lynch JW, Rastam L. A brief conceptual tutorial on multilevel analysis in social epidemiology – interpreting neighbourhood differences and the effects of neighbourhood characteristics on individual health. *J Epidemiol Community Health* 2004; in press.
- 81 Merlo J, Asplund K, Lynch JW, Rastam L, Dobson A. Population effects on individual systolic blood pressure – a multilevel analysis of WHO MONICA project. *Am J Epidemiol* 2004;159:1168-1179.

- 82 Shouls S, Congdon P, Curtis S. Modelling inequality in reported long term illness in the UK: combining individual and area characteristics. *J Epidemiol Community Health* 1996;50:366-376.
- 83 Merlo J, Gerdtham UG, Lynch J, Beckman A, Norlund A, Lithman T. Social inequalities in health – do they diminish with age? Revisiting the question in Sweden 1999. *Int J Equity Health* 2003;2:2.
- 84 Geronimus AT, Bound J. Use of census-based aggregate variables to proxy for socioeconomic group: evidence from national samples. *Am J Epidemiol* 1998;148:475-86.
- 85 Corcoran C, Mehta C, Senchaudhuri P. Power comparisons for tests of trend in dose-response studies. *Stat Med* 2000;19:3037-50.
- 86 Austin PC, Goel V, van Walraven C. An introduction to multilevel regression models. *Can J Public Health* 2001;92:150-154.
- 87 Jones K, Gould MI, Duncan C. Death and deprivation: an exploratory analysis of deaths in the health and lifestyle survey. *Soc Sci Med* 2000;50:1059-1079.
- 88 Rice N, Jones A. Multilevel models and health economics. *Health Econ* 1997;6:561-575.
- 89 Subramanian SV, Jones K, Duncan C. Multilevel methods for public health research. In: Kawachi I, Berkman LF, editors. *Neighborhoods and Health*. New York, NY: Oxford University Press, 2003:65-111.
- 90 Chaix B, Bobashev GV, Merlo J, Chauvin P. Re: "Detecting patterns of occupational illness clustering with alternating logistic regressions applied to longitudinal data". *Am J Epidemiol* 2004;160:505-506 (letter).
- 91 Katz J, Carey VJ, Zeger SL, Sommer A. Estimation of design effects and diarrhea clustering within households and villages. *Am J Epidemiol* 1993;138:994-1006.
- 92 Boyle P, Norman P, Rees P. Does migration exaggerate the relationship between deprivation and limiting long-term illness? A Scottish analysis. *Soc Sci Med* 2002;55:21-31.
- 93 Merlo J, Asplund K, Lynch JW, Rastam L, Dobson A. Population effects on individual systolic blood pressure - a multilevel analysis of WHO MONICA project. *Am J Epidemiol* (in press).
- 94 Boyle MH, Willms JD. Place effects for areas defined by administrative boundaries. *Am J Epidemiol* 1999;149:577-85.
- 95 Carrat F, Valleron AJ. Epidemiologic mapping using the "kriging" method: application to an influenza-like illness epidemic in France. *Am J Epidemiol* 1992;135:1293-300.

- 96 Sabel CE, Boyle PJ, Loytonen M, Gatrell AC, Jokelainen M, Flowerdew R, et al. Spatial clustering of amyotrophic lateral sclerosis in Finland at place of birth and place of death. *Am J Epidemiol* 2003;157:898-905.
- 97 Green C, Hoppa RD, Young TK, Blanchard JF. Geographic analysis of diabetes prevalence in an urban area. *Soc Sci Med* 2003;57:551-60.
- 98 Gemperli A, Vounatsou P, Kleinschmidt I, Bagayoko M, Lengeler C, Smith T. Spatial patterns of infant mortality in Mali: the effect of malaria endemicity. *Am J Epidemiol* 2004;159:64-72.
- 99 Banerjee S, Wall MM, Carlin BP. Frailty modeling for spatially correlated survival data, with application to infant mortality in Minnesota. *Biostatistics* 2003;4:123-42.
- 100 Littell RC, Milliken GA, Stroup WW, Wolfinger R. *SAS System for Mixed Models*. Cary, North Carolina, USA: SAS Institute, 1996.
- 101 Grasland C, Mathian H, Vincent JM. Multiscalar analysis and map generalisation of discrete social phenomena: Statistical problems and political consequences. *Stat J UN Econ Comm Eur* 2000;17:1-32.
- 102 Soloway B. Primary care and specialty care in the age of HAART. *AIDS Clin Care* 1997;9:37-9.
- 103 Baker DW, Hayes RP, Massie BM, Craig CA. Variations in family physicians' and cardiologists' care for patients with heart failure. *Am Heart J* 1999;138:826-34.
- 104 Grumbach K, Selby JV, Damberg C, Bindman AB, Quesenberry C, Jr., Truman A, et al. Resolving the gatekeeper conundrum: what patients value in primary care and referrals to specialists. *JAMA* 1999;282:261-6.
- 105 Bodenheimer T, Lo B, Casalino L. Primary care physicians should be coordinators, not gatekeepers. *JAMA* 1999;281:2045-9.
- 106 Auvray L, Dumesnil S, Le Fur P. Santé, soins et protection sociale en 2000 [Health, healthcare and insurance in 2000] (in French). Paris, France: CREDES, 2001.
- 107 Zonage d'Etudes [Geographic subdivisions of the territory] (in French). Paris, France: Institut National de la Statistique et des Etudes Economiques (http://www.insee.fr/fr/nom_def_met/nomenclatures/zonages_etudes/index.htm).
- 108 *"Partitioning variation in multilevel models"*. Goldstein H, Browne W, Rasbash J ed: (<http://www.ioe.ac.uk/hgpersonal/Variance-partitioning.pdf2002>).
- 109 Chaix B, Bobashev GV, Merlo J, Chauvin P. Re: "Detecting patterns of occupational illness clustering with alternating logistic regressions applied to longitudinal data". *Am J Epidemiol* (in press).

110 Walter SD. The analysis of regional patterns in health data. II. The power to detect environmental effects. *Am J Epidemiol* 1992;136:742-59.

111 Bivand R. Spatial dependence: weighting schemes, statistics and models. (<http://cran.r-project.org/src/contrib/PACKAGES.html#spded>).

112 Bingenheimer JB, Raudenbush SW. Statistical and substantive inferences in public health: issues in the application of multilevel models. *Annu Rev Public Health* 2004;25:53-77.

113 Oakes JM. The (mis)estimation of neighborhood effects: causal inference for a practicable social epidemiology. *Soc Sci Med* 2004;58:1929-52.

114 Curtis S, Southall H, Congdon P, Dodgeon B. Area effects on health variation over the life-course: analysis of the longitudinal study sample in England using new data on area of residence in childhood. *Soc Sci Med* 2004;58:57-74.

115 Diggle P, Moyeed R, Rowlingson B, Thomson M. Childhood malaria in the Gambia: a case-study in model-based geostatistics. *J R Stat Soc Ser C Appl Stat* 2002;51:493-506.