

UNIVERSITE PARIS DESCARTES

Master Ingénierie Mathématique pour les sciences du vivant  
*Spécialité Statistiques appliquées aux sciences sociales et à l'épidémiologie*

---

**Environnement de résidence et mobilité douce : une  
étude à partir de la cohorte RECORD**

---

UMR-S 707 : Epidémiologie, Systèmes d'Information, Modélisation

Rapport de stage

**BAUDET Nathalie**

## TABLE DES MATIERES

PRESENTATION DU LABORATOIRE .....	4
UMR-S 707.....	4
Equipe DS3 (« Déterminants sociaux de la santé et du recours aux soins ») .....	5
Groupe RECORD .....	6
INTRODUCTION.....	7
Etat de l'art .....	7
Objectifs .....	8
METHODES.....	8
Population .....	8
Variables .....	9
Analyses statistiques .....	12
RESULTATS .....	14
Sélection des participants de l'étude.....	14
Environnement et activité de marche.....	16
DISCUSSION.....	18
Sélection des participants de l'étude.....	18
Environnement et activité de marche.....	19
CONCLUSION .....	20
ANNEXES .....	31
REFERENCES.....	36

## REMERCIEMENTS

Je tiens à remercier le Directeur de l'UMR-S 707, Guy Thomas, pour m'avoir accueillie au sein de son unité.

Un remerciement tout particulier va à Basile Chaix, mon responsable au sein de l'équipe DS3, pour son encadrement, son soutien et son enseignement tout au long de ce stage. Merci aussi à l'ensemble du groupe RECORD pour son accueil chaleureux.

Merci à Brigitte Le Roux, responsable du Master, pour son suivi et son encadrement régulier.

## PRESENTATION DU LABORATOIRE

Le présent stage a été effectué au sein du groupe RECORD, qui appartient à l'équipe DS3, elle-même intégrée à l'UMR-S 707. Nous décrivons ici successivement ces trois entités.

### UMR-S 707

L'UMR-S 707 (Epidémiologie, Systèmes d'information, Modélisation) avec sa double tutelle Inserm et Université Pierre et Marie Curie a été créée en 2005. Elle regroupe aujourd'hui six chercheurs Inserm, un chercheur CNRS, 13 enseignants-chercheurs HdR, 6 ITA et 3 IATOS, 22 doctorants et 4 post-doctorants. La spécificité de l'unité demeure le recours prépondérant à la modélisation (mathématique, statistique et informatique) et aux technologies des systèmes d'information pour conduire des travaux dans le domaine de l'épidémiologie.

L'unité comporte deux équipes :

- Equipe 1 : *Epidémiologie des maladies infectieuses et modélisation*, dirigée par Guy Thomas. Cette équipe développe des systèmes d'information en épidémiologie, réalise des enquêtes épidémiologiques notamment sur le risque infectieux, et développe des outils d'analyse et d'interprétation des données basés sur la modélisation. L'équipe regroupe 11 chercheurs ou enseignants chercheurs (dont 10 HDR), 1 post-doc et 10 doctorants.

Le groupe d'Isabella Annesi-Maesano "Epidémiologie des maladies allergiques et respiratoires" a rejoint cette équipe lors du renouvellement quadriennal.

- Equipe 2 : *Déterminants sociaux de la santé et du recours aux soins*, dirigée par Pierre Chauvin. Cette équipe, qui existait dans le projet de création, a été renforcée par le recrutement CR2 Inserm de Basile Chaix (2007) avec qui le présent stage a été effectué, le recrutement CR1 CNRS d'Isabelle Parizot, et l'accueil du groupe de Bernard Larouzé, dont la thématique épidémiologique, orientée vers les pays en développement et le dépistage, présente une forte problématique sociale.

Depuis la création de l'unité, l'attention portée par le public, les scientifiques et les politiques aux maladies transmissibles émergentes, aux déterminants sociaux de la santé, et à l'environnement (pollution, allergies) ne s'est pas démentie. Sur le plan technologique, les progrès dans le domaine des systèmes d'information ont abouti à la disponibilité de bases de données contenant des descriptions détaillées des populations et de leurs environnements de vie. Dans ce contexte, la complémentarité des mathématiques et de la biostatistique, de l'informatique et de l'épidémiologie pour l'étude des maladies transmissibles, de l'impact de l'environnement et des facteurs sociaux des maladies apparaît de plus en plus indiscutable,

l'objectif étant de construire un cadre intégratif pour l'analyse et la prise de décision en santé publique.

### **Equipe DS3 (« Déterminants sociaux de la santé et du recours aux soins »)**

L'équipe de recherche sur les déterminants sociaux de la santé et du recours aux soins (DS3) a été créée en 2002 grâce au soutien initial du programme Avenir de l'Inserm octroyé à son responsable (P. Chauvin). Elle poursuit depuis sa création, et son intégration dans le projet de l'Unité 707, des recherches en santé publique sur les interrelations entre les situations sociales des individus, leurs perceptions de santé, et leurs stratégies et modes d'utilisation du système de soins. Ces recherches sont volontairement pluridisciplinaires, conjuguant des approches sociologiques et épidémiologiques, et des méthodes qualitatives et statistiques.

L'équipe compte dorénavant 5 chercheurs titulaires (1 DR1, 3 CR1 et 1 CR2, dont 1 CR1 CNRS associée) et 1 MCU-PH ; 2 praticiens hospitaliers, 4 doctorants et une allocataire de recherche ; 2 chercheuses contractuelles à temps partiel ; 2 ingénieurs titulaires (IE) et 1,5 ETP ingénieurs contractuels.

L'objectif est de réunir dans l'équipe de multiples disciplines dans une perspective partagée de recherche en santé publique : épidémiologie et sociologie bien sûr mais aussi clinique (médecine interne, infectiologie, psychiatrie), géographie et sciences politiques.

Trois hypothèses sont à l'origine du projet de l'équipe :

- la population ne peut pas être divisée entre « inclus » et « exclus ». Il s'agit là de deux idéaux-types qui constituent les bornes d'un continuum où s'inscrivent les situations sanitaires et sociales (et les besoins de santé) de chacun ;
- les déterminants de la santé et du recours aux soins décrivent un gradient social qui n'est pas totalement superposable aux gradients socioprofessionnels et économiques classiquement décrits. Au-delà de ces déterminants classiques, il est nécessaire de suivre trois approches novatrices : l'étude des liens sociaux, l'étude des caractéristiques contextuelles, le tout autant que possible dans une perspective biographique ;
- les inégalités sociales d'accès et de recours aux soins (curatifs, préventifs, sociosanitaires) existantes en France sont probablement à l'origine d'une part non négligeable des inégalités de santé constatées, et/ou les entretiennent et les accentuent.

## **Groupe RECORD**

Au sein de l'équipe DS3, Basile Chaix travaille plus particulièrement sur les effets du contexte physique et social de résidence sur la santé et sur leurs mécanismes d'influence, avec un intérêt particulier pour les pathologies coronaires et leurs facteurs de risques.

Pour pouvoir étudier de façon précise l'ensemble de ces dimensions, un projet de cohorte spécifique a été développé : la cohorte RECORD (Residential Environment and CORonary heart Disease). Cette étude apparie les données cliniques, biologiques, psychologiques et médicales recueillies lors de bilans de santé effectués par le Centre d'Investigations Préventives et Cliniques dans ses différents sites d'Île-de-France avec des données sociales et résidentielles spécifiquement recueillies pour la cohorte RECORD par auto-questionnaire et des données contextuelles issues d'un système d'information géographique.

L'objectif principal du projet RECORD est d'étudier les disparités sociales de santé qui existent en Île-de-France, avec un intérêt particulier pour les différences observées entre quartiers favorisés et défavorisés. L'étude part de l'idée qu'il est important de porter ces disparités territoriales de santé à la connaissance de la société et des pouvoirs publics.

Au-delà, il s'agit de comprendre comment le contexte géographique de vie dans ses multiples dimensions (environnement physique, services disponibles, relations sociales au sein des quartiers) influe sur la santé.

Cette étude concerne principalement les maladies coronaires et leurs facteurs de risque mais s'intéresse également à d'autres problèmes de santé, aux comportements de recours aux soins, etc.

Cette étude prévoit en outre de suivre de façon prospective la morbi-mortalité cardiovasculaire des participants à partir du PMSI et du système d'information des causes de décès du Cépi-DC ainsi que leurs consommations de soins à partir du SNIIR-AM.

Le projet RECORD est soutenu financièrement par plusieurs institutions : L'Institut de Recherche en Santé Publique (IRESP), l'Institut National de Prévention et d'Education à la Santé (INPES), l'Institut de Veille Sanitaire (InVS), les Ministères de la Santé et de la Recherche, la Caisse Nationale d'Assurance Maladie des Travailleurs Salariés, l'Agence Nationale de la Recherche (ANR), le Groupement Régional de Santé Publique d'Île-de-France, la Direction Régionale des Affaires Sanitaires et Sociales d'Île-de-France (DRASSIF), la Direction Régionale de la Jeunesse et des Sports d'Île-de-France (DRDJS), le Conseil Régional d'Île-de-France et l'École des Hautes Etudes en Santé Publique (EHESP).

A part Basile Chaix, deux doctorants, une post-doctorante, une statisticienne et un stagiaire de Master chaque année travaillent sur le projet RECORD. Cinira Leal, doctorante, travaille sur les déterminants contextuels des facteurs de risque métaboliques et des recours aux soins associés. David Evans, doctorant, étudie la relation entre l'environnement géographique et l'activité physique et Sabrina Havard, post-doctorante travaille sur l'association entre la pollution atmosphérique, l'environnement socio-économique et la pression artérielle.

## **INTRODUCTION**

Des études montrent qu'une activité physique régulière est bénéfique pour la santé<sup>1</sup>. Cependant, en France, seulement 60% des adultes ont un niveau d'activité physique conforme aux recommandations. Par ailleurs, il n'est pas nécessaire d'effectuer une activité physique intense pour obtenir des effets positifs sur la santé. Ainsi, les discours de prévention s'orientent maintenant de plus en plus vers la promotion d'un mode de vie actif qui permet d'inclure une activité physique dans la vie courante. De ce fait, la mobilité douce est mise en avant et en particulier la marche. Pratiquée à un certain niveau et à une certaine vitesse, la marche semble avoir un effet bénéfique sur la santé<sup>2</sup>. De plus, la marche est une activité physique facilement praticable au quotidien et sans coût. En France, le Programme National Nutrition-Santé (PNNS)<sup>3</sup> fait la recommandation suivante : « L'activité physique minimum conseillée chez l'adulte correspond à la pratique de la marche à un pas soutenu 30 minutes par jour, la plupart, et si possible tous les jours de la semaine ».

### **Etat de l'art**

De plus en plus d'articles<sup>4</sup> s'intéressent aux effets de l'environnement physique<sup>5</sup> et de l'environnement de service sur la mobilité douce. Des associations ont ainsi été identifiées entre le temps de marche et les qualités esthétiques des quartiers, la diversité des destinations immédiatement accessibles<sup>6</sup> (commerces, parcs<sup>7</sup>, etc.) ou encore les caractéristiques du réseau de rues. Certaines études ont également pris en compte l'environnement social<sup>8</sup>, montrant par exemple que le sentiment d'insécurité dans un quartier diminue la probabilité de marcher.

Cependant, toutes ces études sont rarement ajustées sur le niveau socio-économique des quartiers, qui intervient probablement comme facteurs de confusion dans la multitude d'associations estimées. De plus, beaucoup d'études se restreignent à une seule catégorie d'individus (femmes<sup>9</sup>, personnes âgées<sup>10</sup>, etc.).

Une autre limite de ces études est l'intérêt quasi-exclusif pour l'environnement physique. En effet, les interactions sociales ont été très rarement prises en compte hormis sous la forme d'une variable d'insécurité.

Enfin, la plupart des études se sont intéressées à la marche en général<sup>11</sup>, sans assez distinguer entre la marche dite utilitaire et la marche récréative, alors qu'il a été suggéré que des déterminants différents existent pour ces deux types de marche.

## **Objectifs**

Nous avons conduit notre étude en deux temps. Tout d'abord nous avons étudié les facteurs qui peuvent influencer sur la participation à la cohorte RECORD. En effet, le recrutement des participants de la cohorte a été effectué lors d'exams de santé sans échantillonnage préalable. Il est donc important d'identifier les facteurs individuels et contextuels qui influent sur les chances de participation et ainsi d'éviter au mieux, par exemple par ajustement, les biais<sup>12</sup> qu'ils peuvent introduire.

Dans un second temps, nous nous sommes intéressés aux données sur la marche. Tout d'abord de façon descriptive, nous avons essayé de répondre aux questions suivantes : combien de temps marche t-on ? Quelles sont les caractéristiques individuelles qui influent sur les temps de marche ? Ensuite nous avons étudié l'influence des facteurs environnementaux sur la marche en nous intéressant particulièrement à la marche récréative. Nous avons cherché à identifier les caractéristiques du milieu de résidence qui, au-delà des effets individuels, interviennent comme obstacles à la mobilité douce ou, au contraire, encourageant à marcher.

## **METHODES**

### **Population**

L'étude des facteurs qui influent sur les chances de participation à la cohorte reposait sur deux bases de données différentes : la cohorte RECORD et le recensement de 1999.

La cohorte RECORD (Residential Environment and CORonary heart Disease) a été mise en place par l'INSERM U707 en collaboration avec le Centre d'Investigations Préventives et Cliniques (IPC). Les individus ont été recrutés lors de bilans de santé de la Sécurité Sociale réalisés dans différents sites du Centre IPC. Quatre centres de région parisienne ont été choisis pour l'étude : Paris, Argenteuil, Mantes-la-Jolie et Trappes. Les individus avaient entre 30 et 79 ans et résidaient dans 10 arrondissements de Paris et 111



autres communes d'Île-de-France qui avaient été sélectionnées a priori de façon à obtenir une grande diversité socio-territoriale dans l'échantillon. Le recrutement s'est échelonné entre mars 2007 et fin février 2008 et 7293 participants ont été inclus dans l'étude.

En plus de leur bilan de santé, les participants ont rempli trois questionnaires qui fournissent de nombreuses informations sur leur état de santé et leur situation sociale. Ils ont également complété le questionnaire RECORD qui permet d'obtenir des informations sur leur situation socio-économique, sur leurs modes de vie, sur leurs connaissances et attitudes de santé, et sur leur vécu au sein de leur quartier.

Les participants ont été géocodés avec précision. Nous leur avons attribué les coordonnées spatiales de leur adresse de résidence. Grâce à ce géocodage, nous avons pu identifier leur quartier de résidence. Dans toute l'étude, les quartiers ont pu être défini de deux façons : soit en utilisant la notion d'IRIS, soit en utilisant des zones circulaires autour de l'habitation des participants. Un IRIS est une unité géographique créée par l'INSEE qui correspond à un « petit quartier » et se définit comme un ensemble d'îlots contigus.

La base du Recensement de 1999 de l'INSEE nous a permis d'obtenir le nombre d'habitants par IRIS en fonction de l'âge, du sexe et du niveau d'instruction. Nous avons retenu tous les IRIS appartenant à un des 121 arrondissements parisiens ou communes présélectionnées pour l'étude RECORD, ce qui correspond à 2207 IRIS.

L'étude sur la marche est basée uniquement sur la cohorte RECORD. Lors d'un travail préliminaire sur la base de données, nous avons remarqué une baisse importante et significative de la proportion de valeurs manquantes et de réponses « 0h00 » pour les variables de temps de marche entre le début et la fin de l'étude due à une amélioration de la qualité du recueil. Ceci entraîne une augmentation significative du temps de marche global moyen entre les deux périodes. Après qu'un grand nombre de tests aient été réalisés, il a été décidé de supprimer tous les questionnaires qui avaient au moins quatre valeurs manquantes sur les huit possibles concernant la marche et de mettre les valeurs manquantes restantes à « 0h00 ». Ainsi sur les 7019 questionnaires restant, il n'existait plus de différence significative du temps de marche moyen entre le début et la fin de l'étude. Toute l'étude sur la marche a donc été effectuée sur ces 7019 participants.

## **Variables**

Pour l'étude sur les facteurs influant sur les chances de participation, les variables individuelles ont été recodées de telle sorte que l'on puisse les comparer entre les deux bases de données. L'âge a été divisé en trois classes (30-39 ans, 40-59 ans et 60 ans et

plus). Le niveau d'instruction a été divisé en trois classes : aucun diplôme, entre BEPC et BAC+2 et supérieur à BAC+2.

Pour obtenir les temps de marche des individus, un instrument de recueil du temps marché au cours des sept derniers jours a été mis en place. Il a été demandé aux participants d'indiquer le temps marché en fonction de la destination (travail, course, etc.) afin de différencier la marche utilitaire de la marche récréative. Par ailleurs, pour chaque destination, il a été demandé aux personnes d'indiquer le temps total marché au sein et en dehors de leur quartier. Nous avons obtenu ainsi six variables de temps de marche : temps de marche global (total et au sein du quartier), temps de marche récréative (total et au sein du quartier) et temps de marche utilitaire (total et au sein du quartier). A l'aide de ces variables nous avons créé quatre variables de proportion : proportion de marche totale au sein du quartier, proportion de la marche totale effectuée à des fins récréatives, proportion de marche récréative au sein du quartier et proportion de marche utilitaire au sein du quartier.

Nous avons utilisé des variables socio-démographiques individuelles telles que le sexe, l'âge, le statut matrimonial, le niveau d'instruction et le revenu du ménage par unité de consommation. Toutes ces variables ont été divisées en classes. L'âge est divisé en trois classes (30-45 ans, 45-65 ans et 65-79 ans). Pour le statut matrimonial, deux catégories ont été retenues : vivant seul ou en couple. Le niveau d'instruction a été divisé en quatre groupes : les personnes ne sachant ni lire ni écrire le français et/ou sans diplôme, les personnes ayant un CAP, un BEPC ou un brevet des collèges, les personnes ayant un BAC, un BTS ou un BAC+2 et les personnes ayant un diplôme supérieur à BAC+2. Quatre classes ont été créées pour le revenu par unité de consommation allant de moins de 833 euros par mois à plus de 2250 euros.

Nous avons également utilisé des variables d'expérience individuelle (soutien ou cohésion sociale au sein du quartier, hostilité ou méfiance envers les habitants du quartier, etc.). Des scores ont été créés à partir de différents items du questionnaire.

Pour les deux parties de l'étude (sélection des participants et activité de marche) nous disposons aussi d'un certain nombre de variables environnementales. Des variables sur la structure socio-démographique des quartiers ont été définies à l'échelle des IRIS : densité de population, revenu médian, proportion de demandeurs d'emploi, proportion de RMIstes, prix immobiliers, proportion de personnes de plus de 15 ans fortement éduquées, proportion de personnes non imposables. Des variables sur l'environnement physique ont aussi été créés : variables de bâti (hauteur moyenne des bâtiments, pourcentage de la surface totale qui est bâtie, etc.), densité de destinations (commerces, services, etc.), caractéristiques du réseau de rues (nombre d'intersections, taille des tronçons, etc.),

présence de transports en commun, présence de parcs et d'espaces verts, qualités esthétiques de l'environnement. Ces variables ont été définies au niveau IRIS pour l'étude sur la sélection des participants, et dans des zones circulaires d'un rayon de 500 mètres autour du lieu de résidence de chaque participant pour l'étude sur la marche.

Nous avons créé des variables écométriques<sup>13</sup> qui s'appuient sur des modèles multiniveaux à effets aléatoires afin de caractériser l'environnement physique et d'interactions sociales. Dans le questionnaire RECORD, différentes questions relatives à chacune des dimensions environnementales qui nous intéressent ont été posées. Nous disposons ainsi des réponses de différentes personnes d'un même quartier à ces questions. Les réponses à ces questions ont été prises en compte comme variables à expliquer dans des modèles multiniveaux à trois niveaux (questions, individus, quartiers). Dans ces modèles, les quartiers ont été définis au niveau TRIRIS (regroupement de plusieurs IRIS). Un intercept aléatoire a été introduit à la fois au niveau des individus et au niveau des quartiers. L'approche écométrique consiste à extraire l'effet aléatoire de niveau quartier de ces modèles, et à s'en servir comme d'une variable explicative dans nos modèles où les variables de marche sont les variables réponses. Ainsi, six variables ont été définies au moyen de cette technique :

- un score indiquant si l'environnement est propice ou non à un mode de vie actif basé sur les items suivants : « environnement déplaisant pour se promener à pieds », « manque d'espaces pour faire du sport en extérieur », « manque d'espaces verts à proximité » ;
- un score de dégradation de l'environnement physique basé sur les problèmes rapportés suivants : « façades de bâtiments en mauvais état », « quartier mal entretenu », « ordures et déchets sur la voie publique », « vandalisme » ;
- un score de dégradation de l'environnement social basé sur les problèmes rapportés suivants : « victimisation au cours des douze derniers mois », « incivilités de certains résidents », « sentiment d'insécurité », « impolitesse des résidents », « bruit excessif des gens du quartier » ;
- un score de cohésion sociale au sein du quartier basé sur les items suivants : « services mutuels rendus », « sentiment d'appartenance à une communauté », « action collective au sein du quartier », « existence de liens entre habitants du quartier » ;
- un score de stigmatisation basé sur les items suivants : « j'ai déjà été jugé négativement parce que je vis dans ce quartier », « ce quartier vous a déjà fait honte à vous personnellement » ;

- un score de soutien basé sur les items suivants : « possibilité d'être hébergé dans le quartier en cas de problème dans le domicile », « aide au sein du quartier dans les tâches quotidiennes en cas de maladie », « quelqu'un de confiance dans le quartier pour discuter en cas de décision à prendre », « quelqu'un au sein du quartier pour obtenir du réconfort lorsque triste ou déprimé ».

Les analyses de ce rapport ont été réalisées à partir des données obtenues à l'inclusion des individus dans l'étude.

## **Analyses statistiques**

### *Sélection des participants de l'étude*

Pour étudier les facteurs qui influent sur les chances de participer à la cohorte, nous avons utilisé des modèles multiniveaux ainsi que des modèles de régression spatiale. Dans les deux cas nous avons spécifié une distribution de Poisson pour les résidus et une fonction de lien logarithme. La variable réponse était le nombre de participants de l'étude RECORD dans chaque IRIS présélectionné pour un profil donné (en fonction de l'âge, du sexe et du niveau d'étude). Le logarithme du nombre d'habitants par IRIS issu du recensement de 1999 pour chaque profil a été spécifié comme offset dans les modèles.

Dans les modèles multiniveaux, nous avons défini un effet aléatoire au niveau IRIS. Afin de caractériser l'autocorrélation spatiale entre les IRIS, nous avons calculé des coefficients de Moran<sup>14</sup> à partir des résidus de ces modèles pour différentes distances entre les IRIS (moins de 2000 mètres, entre 2000 et 4000 mètres, etc.).

Dans les modèles spatiaux, nous avons spécifié deux effets aléatoires (toujours au niveau IRIS), un structuré spatialement et un deuxième sans structure spatiale. Pour l'effet structuré, nous avons spécifié un processus conditionnel autorégressif (CAR) qui modélise l'autocorrélation des résidus entre IRIS contigus. Nous avons estimé les variances de chacun des effets, ce qui nous a permis de déterminer la part de la variance géographique structurée spatialement.

Nous avons à chaque fois estimé trois modèles : un premier ajusté seulement sur l'âge et le sexe, un second où le niveau d'instruction individuel a été ajouté, et un dernier contenant les trois variables individuelles et les variables contextuelles retenues.

Les interactions entre les effets du niveau d'instruction individuel et des variables contextuelles ont ensuite été testées en utilisant des variables ordinales. Nous avons alors estimé des modèles multiniveaux avec des variables de 12 catégories<sup>15</sup> combinant les

modalités des deux variables de l'interaction pour les interactions retenues, ajustés sur les variables individuelles et les autres variables contextuelles.

La méthode de Monte Carlo par chaîne de Markov (MCMC)<sup>16</sup> a été utilisée pour estimer les modèles multiniveaux et les modèles CAR avec le logiciel Winbugs. Nous avons choisi des distributions *a priori* peu informatives. Le DIC (Deviance Information Criterion) a permis de comparer nos différents modèles. Plus le DIC est petit, plus l'adéquation du modèle aux données (pénalisée par le nombre de paramètres) est satisfaisante. L'estimation de l'erreur de Monte Carlo a permis de vérifier la convergence des algorithmes de MCMC. En pratique, lorsque cette erreur était inférieure à 5% de l'écart type estimé du paramètre, nous avons considéré que l'algorithme convergeait.

Afin d'appréhender de façon plus fine les facteurs qui influent sur les chances de participation à l'étude RECORD, nous avons cherché à différencier entre les facteurs qui influent sur les chances d'effectuer un bilan de santé dans un des sites du Centre IPC dans la population générale, et les facteurs qui influent sur les chances de participer à l'étude RECORD parmi les personnes qui avaient effectué un bilan de santé.

Deux analyses complémentaires ont ainsi été menées. La première utilisait le même modèle que pour l'analyse principale en remplaçant la variable réponse par le nombre d'individus par IRIS ayant passé un examen de santé dans un des 4 sites du Centre IPC aux dates de recrutement de l'étude (qu'ils aient ou non ensuite accepté de participer à l'étude RECORD). La deuxième analyse a estimé un modèle binomial, qui évaluait les chances de participer ou non à l'étude RECORD parmi les individus qui avaient passé un bilan de santé et qui étaient éligibles en termes d'âge et de commune de résidence.

Toutes les analyses complémentaires ont été réalisées avec des modèles multiniveaux (tableau 5). Ne connaissant pas les adresses des personnes qui n'ont pas été retenues ou qui n'ont pas accepté de participer à l'étude, ces analyses complémentaires n'ont pas pu être effectuées au niveau IRIS (impossible de géocoder les individus). Elles ont donc été réalisées au niveau communal. A des fins de comparaison, le modèle principal a été de nouveau estimé au niveau communal.

### *Environnement et activité de marche*

De façon descriptive, nous avons étudié les temps de marche pour les différents types de marche (globale, utilitaire et récréative) ainsi que les différentes proportions de temps de marche en fonction tout d'abord de caractéristiques individuelles (sexe, âge, revenus, niveau d'instruction, statut matrimonial, etc.). Nous avons utilisé différents tests statistiques (Jonckheere-Terpstra, Wilcoxon ou Kruskal-Wallis) pour déterminer s'il existait des différences ou des tendances entre les classes de ces variables individuelles quant aux

variables réponse. Le même travail a été effectué avec les variables d'expérience au sein du quartier ou les variables de perception par rapport au quartier.

Après quelques analyses préliminaires, nous nous sommes focalisés sur deux variables à expliquer : le temps de marche récréative total et le pourcentage de marche récréative effectuée au sein du quartier. Des modèles linéaires et log-linéaires n'étaient pas envisageables avec ces deux variables car l'hypothèse de normalité des résidus était très nettement rejetée par le test de Kolmogorov-Smirnov et leur distribution était manifestement non normale. Pour éviter une trop grande perte d'information liée à l'utilisation d'un modèle logistique, nous avons opté pour des modèles polytomiques ordonnés. Les deux variables réponses ont donc été divisées en 6 classes. Les modèles ont d'abord été estimés avec SAS<sup>17</sup> (PROC LOGISTIC et PROC GLIMMIX) puis avec Winbugs<sup>18</sup> (algorithme de MCMC). Nous avons estimé des modèles de régression spatiale au niveau TRIRIS. Comme pour l'étude sur la sélection des participants, nous avons spécifié deux effets aléatoires, un structuré spatialement (processus conditionnel autorégressif) et un autre non structuré spatialement.

Nous avons d'abord estimé des modèles vides (n'incluant aucune variable explicative). Nous avons ensuite estimé nos modèles en ajustant seulement sur des variables individuelles (âge, sexe, niveau d'étude, revenu, etc.). Nous avons alors testé nos différentes variables contextuelles afin de trouver d'éventuelles associations. Nous avons d'abord introduit ces variables une à une dans nos modèles, puis avons testé simultanément celles que nous avions retenues.

Les interactions entre les variables retenues ont ensuite été étudiées. Cela a d'abord été fait en testant des produits de variables ordinales afin de savoir lesquelles étaient significatives puis confirmé en créant des variables combinées de 16 catégories.

## **RESULTATS**

### **Sélection des participants de l'étude**

Les résultats du modèle multiniveau sur les chances de participation à l'étude RECORD sont présentés dans le tableau 1. La variance au niveau IRIS du modèle ajusté sur l'âge et le sexe indiquait l'existence de variations des chances de participation entre les différents quartiers (tableau 1). Les chances de participation étaient 3,57 fois plus élevées pour les 25% des habitants des IRIS dont les chances sont les plus élevées comparées au 25% des habitants des IRIS qui ont les chances les plus faibles. Comme montré sur la figure 1 avec les coefficients de Moran, l'autocorrélation spatiale entre les IRIS dans les chances de

participation diminuait à mesure que la distance entre les IRIS augmentait, et s'annulait à partir d'une distance supérieure à 12 km. L'estimation des modèles spatiaux a montré qu'il y avait une dépendance spatiale importante entre les quartiers voisins quant aux chances de participation : le modèle ajusté sur l'âge et le sexe a indiqué que 86% de la variance entre les IRIS était structurée spatialement. Les DIC étaient plus petit pour tous les modèles spatiaux que pour les modèles multiniveaux.

L'introduction du niveau d'instruction individuel dans les deux modèles (multiniveau et spatial) a montré qu'une faible part de la variabilité spatiale des chances de participation était lié à cette caractéristique individuelle (tableau 1).

Les modèles ont ensuite été estimés avec toutes les variables environnementales retenues. Les coefficients de corrélation de Pearson entre les variables environnementales sont représentés dans le tableau 4. Comme montré dans le tableau 2, les chances de participer diminuaient avec l'augmentation de la distance entre le centroïde de l'IRIS de résidence et le site du Centre IPC le plus proche.

Par rapport au niveau socio-économique des quartiers, les chances de participer à l'étude augmentaient à mesure que le revenu médian du quartier et les prix immobiliers augmentaient. Cependant, nous remarquons aussi qu'elles augmentaient avec la proportion de demandeurs d'emploi.

Quant aux variables sur l'environnement physique et de services, nous observons que les chances de participer augmentaient à mesure que diminuait la hauteur moyenne du bâti et la densité de service.

Nous observons sur la figure 1 et sur le tableau 1 qu'une grande partie de l'autocorrélation spatiale était expliquée par les variables contextuelles introduites dans notre modèle, mais qu'il restait cependant une part non négligeable de variabilité spatiale non expliquée.

Deux interactions avec le niveau d'instruction individuel ont été mises en évidence, d'une part avec la distance au site du Centre IPC le plus proche et d'autre part avec les prix immobiliers du quartier. Les interactions exprimées au moyen de variables de catégories combinées sur le tableau 3 indiquaient qu'un niveau d'instruction individuel élevé augmentait l'effet négatif de la distance au site du centre IPC et l'effet positif des prix immobiliers sur les chances de participation.

Les analyses complémentaires nous ont permis de montrer que toutes ces variables environnementales étaient principalement associées au fait d'effectuer un examen de santé au Centre IPC parmi la population de ces quartiers plutôt qu'au fait de participer à l'étude RECORD parmi les personnes venues à l'examen de santé. Seul les prix immobiliers

restaient légèrement associés à la participation à la cohorte parmi les personnes qui avaient effectué un bilan de santé. En revanche, le niveau d'instruction individuel était très fortement associé à la participation : parmi les patients présents au Centre IPC, plus une personne avait un niveau d'instruction faible, moins ses chances de participer à l'étude RECORD étaient élevées.

## **Environnement et activité de marche**

### *Résultats descriptifs*

Tous les résultats descriptifs sont rapportés dans les tableaux en annexe.

L'ensemble des participants était composé à 65% d'homme et à 40% de personnes âgées entre 45 et 59 ans. Le niveau d'étude était globalement assez élevé, puisque environ 68% avaient un diplôme de niveau baccalauréat ou supérieur.

En moyenne, les participants ont rapporté un peu plus de 8h de marche au cours des sept jours précédant l'enquête. Vingt-sept pour-cent de ce temps avait été consacré à la marche récréative. Cependant, 30% des enquêtés n'avaient effectué aucune marche récréative durant cette période.

De façon descriptive, l'âge avait un effet positif sur toutes les variables de temps de marche (totale, récréative et utilitaire) : plus une personne était âgée plus ses temps de marche avaient été importants. Les hommes avaient marché significativement plus que les femmes pour leurs loisirs, alors que les femmes avaient eu un temps de marche utilitaire plus important que les hommes.

Les personnes ne sachant ni lire ni écrire le français et/ou qui n'avaient aucun diplôme avaient eu des temps de marche inférieurs aux personnes possédant un diplôme.

Nous avons aussi observé que les personnes aisées avaient marché en proportion moins dans leur environnement proche que les individus moins favorisés.

En ce qui concerne les variables d'expérience, le soutien social dont bénéficiait une personne au sein de son quartier tendait à augmenter ses temps de marche. Nous avons remarqué une différence de près de 50 minutes de marche total sur 7 jours entre une personne qui bénéficiait ou ne bénéficiait pas de soutien. De même, la perception d'un environnement extérieur propice à un mode de vie actif augmentait la proportion de marche récréative effectuée au sein du quartier (de 42% à 53%).



### *Résultats des modèles polytomiques ordonnés*

En ce qui concerne le temps total de marche récréative, nous avons remarqué un effet du sexe, de l'âge et de la profession. En effet, les hommes, les personnes de plus de 60 ans et les individus inactifs, chômeurs ou retraités marchaient plus pour leur loisir (tableau 6).

Pour les variables contextuelles, le nombre de destinations accessibles dans un rayon de 500 mètres autour du lieu de résidence du participant et le pourcentage de personnes fortement éduquées dans le quartier avaient un effet positif sur le temps de marche récréative total (tableau 6).

La part de la variance structurée spatialement a diminué à mesure que nous avons ajouté les variables dans le modèle (tableau 7). Elle est devenue quasiment nulle lorsque le modèle était ajusté sur les variables individuelles et les variables contextuelles. De plus, le DIC encourageait à retenir le modèle incluant toutes les variables malgré sa plus grande complexité. Les variables contextuelles semblaient donc expliquer la quasi-totalité des variations spatiales entre TRIRIS pour le temps de marche récréative total.

D'après l'estimation de l'erreur de Monte Carlo, l'algorithme ne convergait pas pour nos trois modèles spatiaux (vide, individuel et contextuel) sur la proportion de marche récréative effectuée au sein du quartier. Nous avons donc estimé des modèles multiniveaux (tableau 8) au niveau TRIRIS sans effet structuré spatialement pour lesquels l'algorithme convergait.

Nous avons observé (tableau 9) une influence du niveau d'instruction individuel sur le pourcentage de marche récréative effectué au sein du quartier : plus une personne avait un niveau d'instruction faible, plus elle avait de chances d'effectuer une part importante de sa marche récréative dans son environnement proche. Nous avons aussi vu un effet du sexe, les hommes ayant plus de chances d'effectuer une part importante de leur marche récréative dans leur quartier que les femmes.

Nous avons remarqué une association négative assez forte entre le revenu individuel par unité de consommation et la proportion de temps marché pour son loisir au sein du quartier lors des estimations préliminaires des modèles à effets fixes ajusté sur l'âge, le sexe et le niveau d'instruction. L'estimation du modèle multiniveau au niveau TRIRIS a fait entièrement disparaître cette association.

Au niveau des variables contextuelles (tableau 9), la surface de parc dans une zone circulaire de 500 mètres autour de l'habitation du participant et la variable écométrique d'environnement propice à un mode de vie actif, étaient associées positivement avec le pourcentage de marche récréative effectué au sein du quartier.

L'interaction entre ces deux variables environnementales prises sous formes ordinales était significative et positive. Nous avons observé un effet synergique entre les deux variables. Les résultats du modèle avec la combinaison des catégories de ces deux variables confirment bien cet effet (tableau 10). Un environnement propice à un mode de vie actif augmentait l'effet positif de la surface des parcs sur la proportion de marche récréative effectuée au sein du quartier.

## **DISCUSSION**

### **Sélection des participants de l'étude**

Nous avons d'abord cherché à étudier les effets des caractéristiques individuelles et contextuelles des participants de l'étude RECORD sur les chances de participation à l'étude. Les résultats trouvés dans cette partie sont très importants pour la suite des études à conduire à partir de la cohorte. En effet, dans les travaux futurs sur les effets de l'environnement sur la santé basés sur cette cohorte, il faudra prendre en compte ces dimensions contextuelles associées à la participation.

#### *Limite*

Une limite de cette étude est la différence de période et de champ de population entre les données du recensement de 1999 et les données issues de l'étude RECORD (2007-2008), qui peut entraîner une légère différence d'effectifs de population entre les deux bases de données.

#### *Interprétations*

Nous avons trouvé que les chances de participer à l'étude RECORD décroissaient à mesure qu'augmentait la distance entre le centroïde de l'IRIS et le site du Centre IPC le plus proche et ce encore plus pour les personnes avec un niveau d'instruction faible. Cet effet jouait plus sur les chances de participer à un examen de santé que sur les chances de participer à l'étude parmi les personnes présentes à l'examen. Cette association peut peut-être s'expliquer par un manque de moyens ou de motivation pour se rendre dans un des sites d'examen de la part des personnes moins favorisées.

Les différentes mesures de densités (proportion de surface bâtie, densité de service, hauteur moyenne du bâti ou densité de population communale) indiquaient qu'une forte densité était associée à une diminution des chances de participer à la cohorte RECORD. Même si le nombre d'individus vus pour un examen de santé est plus élevé pour le site

principal (Paris) que dans les autres, la population plus élevée de Paris et sa proche banlieue peut avoir comme résultat la baisse des chances de participer dans les zones plus denses.

Le niveau socio-économique des quartiers influait très fortement sur le fait d'aller passer un examen de santé mais très faiblement sur le fait de participer à l'étude parmi les personnes ayant effectué un bilan. Cependant, une proportion élevée de demandeurs d'emploi augmentait les chances de participer à la cohorte. Une explication possible pour cet effet est que 3 des 4 sites où ont eu lieu les recrutements se situent dans des zones géographiques plutôt défavorisées spécialement pour recruter des participants qui sont en situation de précarité.

Le facteur qui influait le plus sur la participation à la cohorte RECORD parmi les personnes ayant passé un bilan de santé est le niveau d'instruction individuel. Ceci peut s'expliquer par le fait qu'un des critères d'inclusion était que la personne ait les capacités de lire et de comprendre le français afin de répondre au questionnaire par elle-même, mais aussi par des différences de motivation à participer à une étude scientifique.

## **Environnement et activité de marche**

### *Limites et perspectives*

Nous nous sommes limités dans le présent document à l'étude de la marche récréative. Il reste encore beaucoup à explorer à partir des données recueillies pour appréhender toutes les relations qui peuvent exister entre la marche et l'environnement résidentiel.

Dans cette première vague de l'étude, la notion de quartier était implicite. Il n'avait pas été demandé aux participants de définir les contours de leur quartier. Dans une perspective de réinterrogation de la cohorte, des recherches sont faites pour trouver un moyen pour que les participants puissent définir avec précision les limites de leur quartier (encadrement de rues, cartes, etc.).

Par ailleurs, les temps de marche, ayant été rapportés plutôt que mesurés, sont sujets à des biais de remémoration, qui pourraient être différentiels entre les classes sociales.

### *Interprétations*

L'algorithme de MCMC ne convergait pas pour les modèles spatiaux sur la proportion de marche récréative effectuée au sein du quartier. L'effet spatial est peut être plus difficile à

estimer du fait de la taille d'échantillon moindre (4804 participants) pour ce modèle car 30% des participants n'ont effectué aucune marche récréative.

Les participants de l'étude avaient plus tendance à marcher pour leur loisir si leur quartier comportait de nombreuses destinations accessibles (services, commerces, etc.). Les chances de marcher longtemps pour son plaisir sont plus importantes pour les résidents des quartiers plus favorisés.

Nous avons trouvé un effet synergique entre la surface de parcs et l'environnement propice à un mode de vie actif sur la proportion de marche récréative effectuée dans son environnement proche. L'un des items sur lequel est basé la variable écométrique d'environnement propice à un mode de vie actif est : « l'environnement est déplaisant pour se promener à pieds ». L'aspect esthétique du quartier perçu par ses habitants augmente donc l'effet positif de la surface des parcs d'incitation à la marche.

## **CONCLUSION**

Notre étude a montré qu'il était possible d'identifier et de quantifier les facteurs individuels et environnementaux qui peuvent influencer sur la participation à une cohorte dont le recrutement a été effectué sans échantillonnage.

Au-delà des caractéristiques individuelles, nous avons montré que des caractéristiques environnementales des quartiers peuvent avoir un rôle notable sur l'activité de marche récréative. Afin de promouvoir l'activité de marche, les pouvoirs publics devraient donc porter une attention particulière à l'accessibilité des populations à des espaces verts et services de divers types ainsi qu'à l'esthétique générale des quartiers.

**Tableau 1.** Variations spatiales et autocorrélation de la participation à la cohorte RECORD

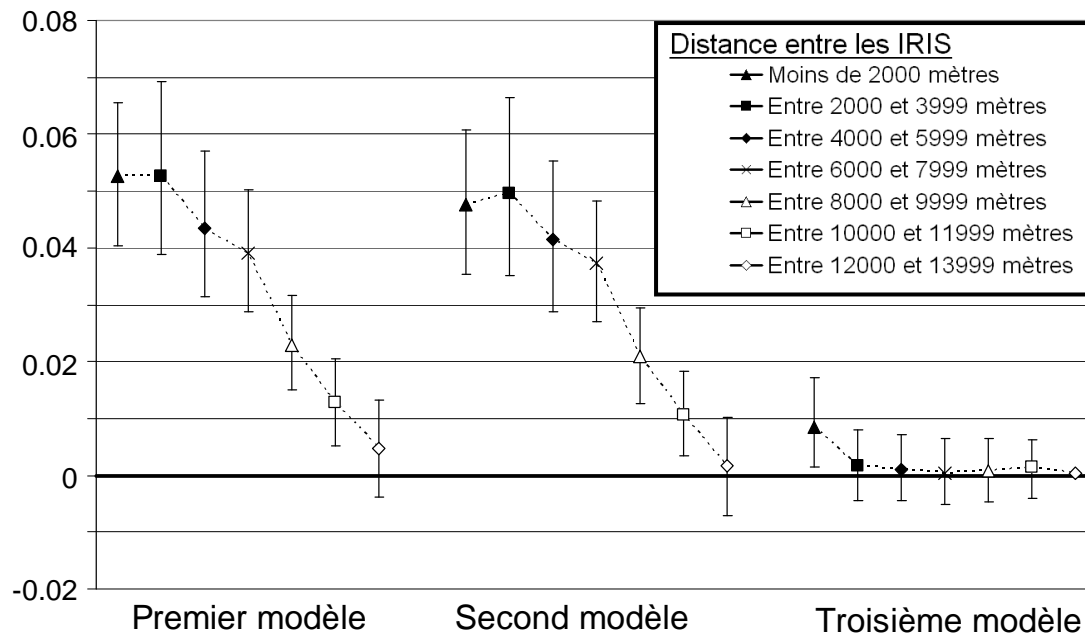
	<b>Modèle ajusté sur l'âge et le sexe</b>	<b>Modèle ajusté sur l'âge, le sexe et le niveau d'instruction</b>	<b>Modèle complet</b>
<b>Modèle multiniveaux</b>			
Variance au niveau IRIS	0,31 (0,27 ; 0,35)	0,28 (0,25 ; 0,32)	0,15 (0,12 ; 0,18)
Odds ratio interquartile	3,57 (3,28 ; 3,88)	3,41 (3,14 ; 3,70)	2,44 (2,24 ; 2,65)
DIC	34524	34286	33980
<b>Modèle CAR</b>			
Variance non structurée spatialement	0,06 (0,03 ; 0,09)	0,06 (0,03 ; 0,10)	0,05 (0,02 ; 0,08)
Variance structurée spatialement	0,37 (0,26 ; 0,49)	0,33 (0,22 ; 0,45)	0,24 (0,14 ; 0,35)
Paramètre de corrélation	0,998 (0,997 ; 0,999)	0,998 (0,996 ; 0,999)	0,996 (0,989 ; 0,998)
Pourcentage de variance structurée	86 (75, 94)	84 (71, 92)	82 (66 ; 93)
DIC	34157	33959	33861

**Tableau 2.** Effet des variables individuelles et environnementales sur la participation à la cohorte RECORD, estimé par un modèle de régression spatiale.

	<b>Exp(<math>\beta</math>)</b>	<b>IC à 95%</b>
Age (vs. 60 ans et plus)		
50 – 59 ans	2,60	(2,46 ; 2,75)
30 – 49 ans	1,13	(1,05 ; 1,21)
Femme (vs. homme)	0,47	(0,45 ; 0,50)
Niveau d'instruction (vs. élevé)		
Moyen	0,73	(0,69 ; 0,77)
Faible	0,57	(0,52 ; 0,63)
Distance au centre IPC (vs. élevé)		
Moyen-élevé	1,12	(0,96 ; 1,30)
Moyen-faible	1,27	(1,07 ; 1,50)
Faible	1,49	(1,24 ; 1,80)
Revenu médian (vs. faible)		
Moyen-faible	1,21	(1,08 ; 1,35)
Moyen-élevé	1,31	(1,14 ; 1,50)
Elevé	1,60	(1,33 ; 1,88)
Prix immobiliers (vs. faible)		
Moyen-faible	1,06	(0,96 ; 1,18)
Moyen-élevé	1,16	(1,03 ; 1,30)
Elevé	1,50	(1,31 ; 1,71)
Proportion de demandeurs d'emploi (vs. faible)		
Moyen-faible	1,02	(0,93 ; 1,12)
Moyen-élevé	1,27	(1,12 ; 1,43)
Elevé	1,41	(1,21 ; 1,64)
Proportion de surface bâtie (vs. élevé)		
Moyen-élevé	1,17	(1,05 ; 1,30)
Moyen-faible	1,23	(1,07 ; 1,41)
Faible	1,22	(1,05 ; 1,42)
Hauteur moyenne du bâti (vs. élevé)		
Moyen-élevé	0,97	(0,89 ; 1,06)
Moyen-faible	1,08	(0,97 ; 1,21)
Faible	1,21	(1,07 ; 1,37)
Densité de service (vs. élevé)		
Moyen-élevé	1,08	(0,97 ; 1,20)
Moyen-faible	1,23	(1,07 ; 1,40)
Faible	1,30	(1,11 ; 1,52)

Figure 1

### Coefficients de Moran (intervalle de crédibilité à 95%)



**Tableau 3.** Association entre, d'une part, une variable de catégories combinées du niveau d'instruction et (i) de la distance au centre IPC ou (ii) des prix immobiliers, et d'autre part, la participation à l'étude RECORD, ajustée sur l'âge, le sexe et les variables environnementales (deux modèles séparés)\*

	<b>Exp(<math>\beta</math>)</b>	<b>IC 95%</b>
<b>Catégories combinées du niveau d'instruction et de la distance au centre IPC</b>		
Instruction individuelle faible		
Distance au centre élevé	Ref.	
Distance au centre moyen-élevé	1,15	(0,85 ; 1,56)
Distance au centre moyen-faible	1,39	(1,02 ; 1,89)
Distance au centre faible	2,40	(1,78 ; 3,21)
Instruction individuelle moyenne		
Distance au centre élevé	1,54	(1,25 ; 1,89)
Distance au centre moyen-élevé	1,66	(1,31 ; 2,12)
Distance au centre moyen-faible	1,99	(1,54 ; 2,56)
Distance au centre faible	2,38	(1,81 ; 3,11)
Instruction individuelle élevée		
Distance au centre élevé	2,32	(1,86 ; 2,92)
Distance au centre moyen-élevé	2,62	(2,04 ; 3,36)
Distance au centre moyen-faible	2,74	(2,12 ; 3,55)
Distance au centre faible	2,81	(2,13 ; 3,69)
<b>Catégories combinées du niveau d'instruction des prix immobiliers</b>		
Instruction individuelle faible		
Prix immobiliers faible	Ref,	
Prix immobiliers moyen-faible	0,91	(0,73 ; 1,14)
Prix immobiliers moyen-élevé	0,92	(0,71 ; 1,17)
Prix immobiliers élevé	0,99	(0,74 ; 1,31)
Instruction individuelle moyenne		
Prix immobiliers faible	0,99	(0,85 ; 1,16)
Prix immobiliers moyen-faible	1,14	(0,96 ; 1,36)
Prix immobiliers moyen-élevé	1,32	(1,10 ; 1,59)
Prix immobiliers élevé	1,62	(1,33 ; 1,97)
Instruction individuelle élevée		
Prix immobiliers faible	1,72	(1,42 ; 2,07)
Prix immobiliers moyen-faible	1,58	(1,31 ; 1,91)
Prix immobiliers moyen-élevé	1,60	(1,32 ; 1,94)
Prix immobiliers élevé	2,25	(1,85 ; 2,75)

\*Avant d'estimer les modèles avec les variables combinées ; on estime un modèle avec des variables ordinales pour le niveau d'instruction individuel ; la distance au centre et les prix immobiliers et les termes produits entre le niveau d'instruction et les deux variables environnementales, Les paramètres sont - 0,05 (IC à 95%: -0,19 ; 0,10) pour l'effet du niveau d'instruction ; -0,40 (IC à 95%: -0,50 ; -0,29) pour l'effet de la distance ; 0,07 (IC à 95%: -0,02 ; 0,17) pour l'effet des prix immobiliers ; 0,12 (IC à 95%: 0,08 ; 0,15) pour le niveau d'instruction  $\times$  le terme de distance ; et 0,03 (IC à 95%: -0,01 ; 0,07) pour le niveau d'instruction  $\times$  terme des prix immobiliers.



**Tableau 4.** Corrélation de Pearson (et intervalle de confiance à 95%) entre les variables environnementales retenues dans le modèle final de régression spatiale (n = 2207 iris)

	Distance au centre	Revenu médian	Proportion de demandeurs d'emploi	Densité de services	Prix immobiliers	Proportion de surface bâtie	Hauteur moyenne du bâti
Distance au centre	1	-0,22 (-0,23 ; -0,21)	+0,02 (+0,01 ; +0,03)	-0,34 (-0,35 ; -0,33)	-0,20 (-0,21 ; -0,19)	-0,38 (-0,39 ; -0,37)	-0,31 (-0,30 ; -0,29)
Revenu médian		1	-0,51 (-0,52 ; -0,51)	+0,19 (+0,18 ; +0,20)	+0,78 (+0,77 ; +0,78)	+0,13 (+0,12 ; +0,14)	+0,03 (+0,02 ; +0,04)
Proportion de demandeurs d'emploi			1	+0,08 (+0,07 ; +0,09)	-0,43 (-0,44 ; -0,42)	+ ,13 (+0,12 ; +0,14)	+0,10 (+0,09 ; +0,11)
Densité de services				1	+0,11 (+0,10 ; +0,12)	+0,80 (+0,79 ; +0,80)	+ ,44 (+0,43 ; +0,45)
Prix immobiliers					1	+0,04 (+0,03 ; +0,05)	+0,07 (+0,06 ; +0,08)
Proportion de surface bâtie						1	+0,40 (+0,39 ; +0,41)
Hauteur moyenne du bâti							1

**Tableau 5.** Effet des variables environnementales communales sur les chances de (i) participer à l'étude RECORD (dans la population générale), (ii) effectuer un bilan de santé (dans la population générale), et (iii) participer à la cohorte RECORD (parmi les personnes ayant effectué un examen de santé), estimé par un modèle multiniveau ajusté sur les variables individuelles.

	<b>Participation à l'étude parmi la population générale</b>	<b>Effectué un bilan de santé parmi la population générale</b>	<b>Participation à l'étude parmi les personnes qui ont fait un bilan de santé</b>
	<b>Exp(<math>\beta</math>) IC à 95%</b>	<b>Exp(<math>\beta</math>) IC à 95%</b>	<b>OR IC à 95%</b>
Niveau d'instruction (vs. élevé)			
Moyen	0,72 (0,68 ; 0,75)	1,55 (1,47 ; 1,64)	0,77 (0,70 ; 0,85)
Faible	0,56 (0,51 ; 0,62)	0,77 (0,74 ; 0,81)	0,13 (0,12 ; 0,15)
Distance au centre (vs. élevé)			
Moyen-élevé	1,12 (0,91 ; 1,39)	1,04 (0,85 ; 1,28)	-
Moyen-faible	1,35 (1,10 ; 1,67)	1,35 (1,09 ; 1,65)	-
Faible	1,67 (1,37 ; 2,04)	1,74 (1,42 ; 2,13)	-
Prix immobiliers (vs. faible)			
Moyen-faible	1,47 (1,10 ; 1,99)	1,12 (0,83 ; 1,53)	1,14 (0,96 ; 1,34)
Moyen- élevé	1,70 (1,21 ; 2,44)	1,33 (0,95 ; 1,93)	1,12 (0,93 ; 1,34)
Elevé	2,26 (1,56 ; 3,39)	1,76 (1,20 ; 2,66)	1,20 (1,01 ; 1,43)
Proportion de RMIstes (vs. élevé)			
Moyen-élevé	1,03 (0,80 ; 1,31)	0,99 (0,78 ; 1,26)	-
Moyen-faible	1,25 (0,91 ; 1,76)	1,21 (0,89 ; 1,66)	-
Faible	1,46 (1,02 ; 2,16)	1,33 (0,94 ; 1,91)	-
Proportion de demandeurs d'emploi (vs. faible)			
Moyen-faible	1,06 (0,84 ; 1,33)	1,09 (0,86 ; 1,37)	-
Moyen-élevé	1,23 (0,88 ; 1,76)	1,30 (0,93 ; 1,82)	-
Elevé	1,49 (0,95 ; 2,40)	1,18 (0,76 ; 1,86)	-
Densité de population (vs. élevé)			
Moyen-élevé	1,21 (1,00 ; 1,46)	1,23 (1,01 ; 1,49)	-
Moyen-faible	1,34 (1,06 ; 1,67)	1,45 (1,15 ; 1,82)	-
Faible	1,36 (1,01 ; 1,83)	1,29 (0,95 ; 1,71)	-
Hauteur moyenne du bâti (vs. élevé)			
Moyen-élevé	1,21 (1,01 ; 1,44)	1,23 (1,03 ; 1,48)	-
Moyen-faible	1,26 (1,02 ; 1,56)	1,27 (1,03 ; 1,57)	-
Faible	1,27 (1,01 ; 1,60)	1,31 (1,05 ; 1,65)	-

**Tableau 6.** Effet des variables individuelles et environnementales sur le temps de marche récréative total, estimé par un modèle de régression spatiale.

	<b>OR</b>	<b>IC à 95%</b>
Age (vs. 30 – 44 ans)		
45 – 59 ans	1,16	(1,10 ; 1,27)
60 ans et plus	1,23	(1,00 ; 1,47)
Homme (vs. femme)	1,30	(1,19 ; 1,42)
Niveau d'instruction (vs. élevé)		
Moyen-élevé	1,16	(0,65 ; 0,94)
Moyen-faible	1,13	(1,01 ; 1,27)
Faible	0,78	(1,04 ; 1,29)
Profession (vs. actifs)		
Chômeurs	1,40	(1,24 ; 1,58)
Retraités	2,17	(1,79 ; 2,63)
Inactifs	1,53	(1,26 ; 1,87)
Nombre de destinations (vs. faible)		
Moyen-faible	1,10	(0,96 ; 1,24)
Moyen-élevé	1,10	(0,94 ; 1,25)
Elevé	1,33	(1,13 ; 1,54)
Pourcentage de personnes fortement éduquées (vs. faible)		
Moyen-faible	1,23	(1,10 ; 1,41)
Moyen-élevé	1,26	(1,10 ; 1,48)
Elevé	1,40	(1,20 ; 1,65)

**Tableau 7.** Variation spatiale du temps de marche récréative total.

	<b>Modèle vide</b>	<b>Modèle individuel</b>	<b>Modèle complet</b>
Variance non structurée spatialement	0,032 (0,017 ; 0,059)	0,032 (0,016 ; 0,061)	0,035 (0,017 ; 0,063)
Variance structurée spatialement	0,036 (0,010 ; 0,087)	0,033 (0,007 ; 0,081)	0,001 (0,002 ; 0,014)
Pourcentage de variance structurée	53 (20 ; 79)	50 (16 ; 78)	4 (0,5 ; 33)
DIC	24057	23790	23767

**Tableau 8.** Modèle multiniveau sur la proportion de marche récréative effectuée au sein du quartier

	<b>Modèle vide</b>	<b>Modèle individuel</b>	<b>Modèle complet</b>
Variance au niveau TRIRIS	0,096 (0,05 ; 0,16)	0,097 (0,05 ; 0,16)	0,066 (0,029 ; 0,12)
DIC	16205	16164	16122

**Tableau 9.** Modèle multiniveaux sur la proportion de marche récréative effectuée au sein du quartier en fonction des variables individuelles et contextuelles

	<b>OR</b>	<b>IC à 95%</b>
Age (vs. 30 – 44 ans)		
45 – 59 ans	0,82	(0,73 ; 0,93)
60 ans et plus	0,77	(0,67 ; 0,88)
Homme (vs. femme)	1,15	(1,03 ; 1,29)
Niveau d'instruction (vs. élevé)		
Moyen-élevé	0,98	(0,86 ; 1,11)
Moyen-faible	1,16	(1,02 ; 1,33)
Faible	2,00	(1,60 ; 2,48)
Surface de parc (vs. faible)		
Moyen-faible	0,98	(0,85 ; 1,14)
Moyen-élevé	1,18	(1,01 ; 1,37)
Elevé	1,28	(1,09 ; 1,50)
Environnement propice à un mode de vie actif (vs. faible)		
Moyen-faible	0,97	(0,83 ; 1,14)
Moyen-élevé	1,26	(1,07 ; 1,48)
Elevé	1,48	(1,25 ; 1,74)

**Tableau 10.** Association entre une variable combinant les catégories de la surface de parc et d'un environnement propice à un mode de vie actif, et la proportion de marche récréative au sein du quartier (ajustée sur les variables individuelles)\*

	OR	IC 95%
Catégories combinées de la surface de parc et de l'environnement propice à un mode de vie actif		
Surface de parc faible		
Environnement faible	Ref,	
Environnement moyen-faible	0,96	(0,71 ; 1,30)
Environnement moyen-élevé	1,25	(0,92 ; 1,69)
Environnement élevé	1,15	(0,85 ; 1,56)
Surface de parc moyen-faible		
Environnement faible	0,92	(0,67 ; 1,19)
Environnement moyen-faible	0,88	(0,67 ; 1,16)
Environnement moyen-élevé	1,18	(0,89 ; 1,59)
Environnement élevé	1,34	(0,98 ; 1,83)
Surface de parc moyen-élevé		
Environnement faible	1,08	(0,83 ; 1,41)
Environnement moyen-faible	1,03	(0,79 ; 1,36)
Environnement moyen-élevé	1,42	(1,06 ; 1,91)
Environnement élevé	1,71	(1,27 ; 2,33)
Surface de parc moyen-élevé		
Environnement faible	1,03	(0,67 ; 1,57)
Environnement moyen-faible	1,18	(0,86 ; 1,60)
Environnement moyen-élevé	1,41	(1,08 ; 1,87)
Environnement élevé	1,91	(1,48 ; 2,49)

\*Avant d'estimer les modèles avec les variables combinées ; on estime un modèle avec des variables ordinales pour la surface de parc et l'environnement propice à un mode de vie actif et les termes produits entre les deux variables environnementales, Les paramètres sont -0,03 (IC à 95%:-0,17 ; 0,10) pour l'effet de la surface des parcs ; 0,04 (IC à 95%:-0,10 ; 0,17) pour l'effet de l'environnement propice à un mode de vie actif ; 0,04 (IC à 95%:0,001 ; 0,10) pour la surface des parcs × le terme d'environnement propice à un mode de vie actif.

## ANNEXES

---

### Annexe 1. Caractéristiques socio-démographiques des participants de la cohorte RECORD

---

	<b>Pourcentage</b>
<b>Genre</b>	
Homme	34,25
Femme	65,75
<b>Age</b>	
Moins de 45 ans	35,59
Entre 45 et 65 ans	41,62
Plus de 65 ans	22,80
<b>Niveau d'instruction</b>	
Ne sais pas lire ou écrire le français ou sans diplôme	7,53
CAP, BEPC, brevet des collèges	24,15
BAC, BTS ou BAC+2	29,62
Licence, maîtrise et +	38,71

---

**Annexe 2.** Temps de marche et proportion de marche récréative en fonction des variables socio-démographiques individuelles

	Globale	Utilitaire	Récréative	Proportion récréative
<b>Genre</b>				
Homme	497	349	148	28%
Femme	518	375	144	25%
<i>p-value (Wilcoxon)</i>	0,0522	0,0002	0,0002	<0,0001
<b>Age</b>				
Moins de 45 ans	456	345	112	24%
Entre 45 et 65 ans	504	363	140	26%
Plus de 65 ans	581	369	212	33%
<i>Statistique (Jonckherre-Tepstra)</i>	10,41	4,13	12,42	10,55
<i>p-value</i>	<0,0002	<0,0001	<0,0001	<0,0001
<b>Niveau d'instruction</b>				
Ne sais pas lire ou écrire le français ou sans diplôme	414	299	115	27%
CAP, BEPC, brevet des collèges	546	379	167	28%
BAC, BTS ou BAC+2	528	376	152	27%
Licence, maîtrise et +	481	345	136	26%
<i>p-value (Kruskal-Wallis)</i>	<0,0001	<0,0001	<0,0001	0,2828
<b>Revenus (euros)</b>				
Inférieur à 833	517	378	139	25%
Compris entre 833 et 1400	517	371	146	27%
Compris entre 1400 et 2250	504	348	156	28%
Supérieur à 2250	478	335	142	27%
<i>Statistique (Jonckherre-Tepstra)</i>	2,50	1,47	3,63	3,15
<i>p-value</i>	0,0061	0,0703	0,0001	0,0008



**Annexe 3.** Temps de marche au sein des quartiers et proportions de marche fonction des variables socio-démographiques individuelles

	Temps total quartier	Proportion total quartier	Proportion utilitaire quartier	Proportion récréatif quartier
<b>Genre</b>				
Homme	224	49%	51%	49%
Femme	231	49%	51%	46%
<i>p-value (Wilcoxon)</i>	0,0821	0,1848	0,2857	0,0167
<b>Age</b>				
Moins de 45 ans	208	49%	51%	50%
Entre 45 et 65 ans	223	48%	50%	47%
Plus de 65 ans	265	50%	54%	46%
<i>Statistique (Jonckherre-Tepstra)</i>	7,79	0,04	2,78	-3,21
<i>p-value</i>	<0,0001	0,4839	0,00270	0,0006
<b>Niveau d'instruction</b>				
Ne sais pas lire ou écrire le français ou sans diplôme	205	56%	58%	59%
CAP, BEPC, brevet des collèges	247	49%	51%	48%
BAC, BTS ou BAC+2	234	48%	50%	56%
Licence, maîtrise et +	213	48%	51%	48%
<i>p-value (Kruskal-Wallis)</i>	0,0060	<0,0001	<0,0001	<0,0001
<b>Revenus (euros)</b>				
Inférieur à 833	245	52%	53%	53%
Compris entre 833 et 1400	231	49%	51%	49%
Compris entre 1400 et 2250	226	48%	51%	47%
Supérieur à 2250	205	46%	50%	44%
<i>Statistique (Jonckherre-Tepstra)</i>	-1,25	-5,13	-2,36	-5,87
<i>p-value</i>	0,1062	<0,0001	0,0089	<0,0001

**Annexe 4.** Temps de marche et proportion de marche récréative en fonction des variables d'expériences

	Globale	Utilitaire	Récréative	Proportion récréative
<b>Attachement</b>				
Faible	493	359	134	25%
Moyen	509	359	150	27%
Elevé	512	356	156	28%
<i>Statistique (Jonckherre-Tepstra)</i>	2,03	0,58	5,46	5,05
<i>p-value</i>	0,0211	0,2799	<0,0001	<0,0001
<b>Hostilité/méfiance</b>				
Faible	494	348	146	27%
Moyen	496	352	143	26%
Elevé	535	382	153	27%
<i>Statistique (Jonckherre-Tepstra)</i>	1,62	1,77	0,13	-1,39
<i>p-value</i>	0,0526	0,0379	0,4475	0,0816
<b>Environnement extérieur (Espace vert, équipements sportifs extérieurs, etc.)</b>				
Faible	516	382	134	24%
Moyen	507	359	148	27%
Elevé	490	335	156	29%
<i>Statistique (Jonckherre-Tepstra)</i>	-1,23	-4,06	4,03	6,73
<i>p-value</i>	0,1077	<0,0001	<0,0001	<0,0001
<b>Cohésion</b>				
Faible	495	353	142	27%
Moyen	511	367	144	26%
Elevé	534	372	162	28%
<i>Statistique (Jonckherre-Tepstra)</i>	1,96	1,27	2,50	2,08
<i>p-value</i>	0,0249	0,1008	0,0061	0,0186
<b>Soutien</b>				
Faible	474	341	133	25%
Moyen-faible	498	351	147	27%
Moyen-élevé	518	374	147	26%
Elevé	524	366	158	28%
<i>Statistique (Jonckherre-Tepstra)</i>	5,03	4,15	5,93	4,01
<i>p-value</i>	<0,0001	<0,0001	<0,0001	<0,0001

**Annexe 5.** Temps de marche au sein des quartiers et proportions de marche fonction des variables d'expériences

	Temps total quartier	Proportion total quartier	Proportion utilitaire quartier	Proportion récréatif quartier
<b>Attachement</b>				
Faible	205	46%	49%	42%
Moyen	232	49%	52%	47%
Elevé	243	51%	52%	54%
<i>Statistique (Jonckherre-Tepstra)</i>	7,21	7,33	4,73	8,96
<i>p-value</i>	<0,0001	<0,0001	<0,0001	<0,0001
<b>Hostilité/méfiance</b>				
Faible	225	50%	52%	49%
Moyen	225	49%	51%	48%
Elevé	233	47%	50%	46%
<i>Statistique (Jonckherre-Tepstra)</i>	-0,86	-3,43	-3,04	-1,81
<i>p-value</i>	0,1948	0,0003	0,0012	0,0345
<b>Environnement extérieur (Espace vert, équipements sportifs extérieurs, etc.)</b>				
Faible	222	48%	51%	42%
Moyen	228	48%	50%	48%
Elevé	228	50%	52%	53%
<i>Statistique (Jonckherre-Tepstra)</i>	0,92	3,40	0,93	6,73
<i>p-value</i>	0,1788	0,0003	0,1746	<0,0001
<b>Cohésion</b>				
Faible	223	49%	52%	47%
Moyen	227	47%	49%	46%
Elevé	247	50%	52%	52%
<i>Statistique (Jonckherre-Tepstra)</i>	2,52	1,34	-0,42	2,52
<i>p-value</i>	0,0058	0,0902	0,3367	<0,0057
<b>Soutien</b>				
Faible	200	47%	50%	45%
Moyen-faible	219	47%	50%	46%
Moyen-élevé	239	49%	51%	49%
Elevé	244	50%	53%	50%
<i>Statistique (Jonckherre-Tepstra)</i>	7,75	4,24	3,50	3,70
<i>p-value</i>	<0,0001	<0,0001	0,0002	0,0001

## REFERENCES

- 1 World Health Report Geneva. *World Health Organization*. 2002.
- 2 Hakim AA, Petrovitch H, Burchfiel CM, Ross GW, Rodriguez BL, White LR, *et al.* Effects of walking on mortality among nonsmoking retired men. *N Engl J Med*. 1998 ; 338: 94-9.
- 3 Oppert J, Simon C, Rivière D, Guezennec C. Activité physique et santé : arguments scientifiques, pistes pratiques. Les synthèses du Programme National Nutrition-Santé. *Paris : ministère de la santé et des solidarités* 2005.
- 4 Owen N, Humpel N, Leslie E, Bauman A, Sallis JF. Understanding environmental influences on walking ; Review and research agenda. *Am J Prev Med*. 2004 ; 27: 67-76.
- 5 Humpel N, Owen N, Iverson D, Leslie E, Bauman A. Perceived environment attributes, residential location, and walking for particular purposes. *Am J Prev Med*. 2004 ; 26: 119-25.
- 6 Boer R, Zheng Y, Overton A, Ridgeway GK, Cohen DA. Neighborhood design and walking trips in ten U.S. metropolitan areas. *Am J Prev Med*. 2007 ; 32: 298-304.
- 7 Giles-Corti B, Broomhall MH, Knuiaman M, Collins C, Douglas K, Lange A, *et al.* Increasing Walking. How Important Is Distance To, Attractiveness, and Size of Public Open Space? *American Journal of Preventive Medicine*. 2005 ; 28: 169-76.
- 8 Giles-Corti B, Donovan RJ. Relative influences of individual, social environmental, and physical environmental correlates of walking. *American Journal of public health*. 2003 ; 93: 1583-89.
- 9 Ball K, Timperio A, Salmon J, Giles-Corti B, Roberts R, Crawford D. Personal, social and environmental determinants of educational inequalities in walking : a multilevel study. *Journal of Epidemiology and Community Health* 2007 ; 61: 108-14.
- 10 Nagel CL, Carlson NE, Bosworth M, Michael YL. The relation between Neighborhood Built Environment and walking activity among older Adults. *American Journal of Epidemiology*. 2008 ; 168: 461-68.
- 11 Giles-Corti B, Donovan RJ. Socioeconomic Status Differences in Recreational Physical Activity Levels and Real and Perceived Access to a Supportive Physical Environment. *Preventive Medicine*. 2002 ; 35: 601-11.
- 12 Nohr EA, Frydenberg M, Henriksen TB, Olsen J. Does low participation in cohort studies induce bias? *Epidemiology*. 2006 ; 17: 413-8.
- 13 Chaix B, Lindstrom M, Rosvall M, Merlo J. Neighbourhood social interactions and risk of acute myocardial infarction. *J Epidemiol Community Health*. 2008 ; 62: 62-8.
- 14 Chaix B, Merlo J, Subramanian SV, Lynch J, Chauvin P. Comparison of a spatial perspective with the multilevel analytical approach in neighborhood studies: the case of mental and behavioral disorders due to psychoactive substance use in Malmo, Sweden, 2001. *Am J Epidemiol*. 2005 ; 162: 171-82.
- 15 Kaufman JS. Interaction reaction. *Epidemiology*. 2009 ; 20: 159-60.

- 16 Best N, Richardson S, Thomson A. A comparison of Bayesian spatial models for disease mapping. *Stat Methods Med Res.* 2005 ; 14: 35-59.
- 17 Gaineroff MJ. New York state Psychiatric Institute: Using the Proportional Odds Model for Health-Related Outcomes: Why, When, and How with Various SAS® Procedures. . *SAS institute, Statistics and data analysis.* 205-30.
- 18 Mwalili SM, Lesaffre E, Declerck D. A Bayesian ordinal logistic regression model to correct for interobserver measurement error in a geographical oral health study. *Applied Statistics.* 2005 ; 55: 77-93.