

Application des modèles multiniveaux à l'étude des effets du contexte de résidence sur la survenue de cardiopathies ischémiques.

Rapport de stage effectué dans l'UMR-S 707. Equipe de recherche sur les déterminants sociaux de la santé et du recours aux soins.

Introduction

Des disparités de santé liées à la catégorie socio-économique des individus ont été rapportées pour un grand nombre de pathologies. Ces déterminants sociaux peuvent être définis à plusieurs échelles : au niveau individuel (niveau d'instruction, profession, revenu, etc.) et au niveau du contexte de vie (caractéristiques de l'environnement de résidence, du lieu de travail, etc.) (1). L'analyse contextuelle, étudiant l'influence de l'environnement de résidence sur la santé, peut permettre d'augmenter l'efficacité d'éventuelles actions de santé publique en ciblant de façon plus adéquate les populations sur lesquelles intervenir. Son but est aussi d'améliorer la connaissance des mécanismes à l'origine des disparités de santé. Par exemple, des études ont montré que les caractéristiques socio-économiques du lieu de résidence influaient sur le poids à la naissance (2), le tabagisme (3, 4), l'activité physique (5) et la mortalité (6, 7), ceci indépendamment des effets liés aux caractéristiques socio-économiques des individus. Etudier et mesurer l'influence du contexte local sur la santé permet donc de progresser dans la compréhension des disparités géographiques décrites par les géographes de la santé.

Une association entre facteurs socioéconomiques individuels et incidence de maladies coronariennes a été constatée (8). Il a aussi été observé depuis longtemps que la prévalence des cardiopathies ischémiques varie dans l'espace (9). Plusieurs travaux ont suggéré l'existence d'un gradient nord-sud sur le territoire français métropolitain. Les variations géographiques entre différents pays et à l'intérieur même des pays restent en partie inexplicables et posent l'hypothèse d'effets contextuels, opérant à différentes échelles (10). Certains facteurs de risque de ces pathologies ischémiques sont probablement liés au milieu dans lequel l'individu évolue (activité physique, surpoids, accès aux soins) (5, 11). Plusieurs travaux ont déjà mis à jour une association entre caractéristiques socio-économiques du contexte et survenue de maladies coronariennes (12-16). Cependant, aucune étude n'a été publiée sur ce thème à partir de données françaises.

En estimant des modèles de Cox multiniveaux nous avons cherché à quantifier la variabilité du risque coronarien à des échelles géographiques plus ou moins locales (codes postaux, départements, régions) (17, 18). Une étude menée en Suède a montré qu'il existait une plus forte hétérogénéité du risque coronarien au sein des zones urbaines (19) et nous avons examiné si un phénomène identique existait en France.

L'objectif était d'abord de déterminer à quelle échelle géographique se situe cette variabilité. Nous avons ensuite estimé des associations entre facteurs contextuels (définis à

différentes échelles spatiales) et risque coronarien, des processus liés à l'environnement socio-économique pouvant agir au delà des effets au niveau individuel. Nous avons enfin cherché à voir si les effets des caractéristiques du contexte opéraient par l'intermédiaire des comportements ou facteurs de risque liés à la santé : le surpoids, le tabagisme et la consommation d'alcool.

Méthodes

Population d'étude

La base de donnée utilisée dans notre étude provient de la cohorte GAZEL. Celle-ci comptait 20542 participants à l'inclusion, en 1989. Par construction, l'âge des femmes au moment de l'inclusion (27% de la population d'étude) allait de 35 à 50 ans, celui des hommes (73%) de 40 à 50 ans. Les participants sont tous des volontaires travaillant pour les entreprises EDF et GDF. Depuis 1989, ils reçoivent un questionnaire annuel portant notamment sur leur situation socioprofessionnelle, leur état de santé physique et mental et leurs habitudes de vie. Dans la présente étude, nous avons défini une cohorte au 1^{er} janvier 1990, considérant les caractéristiques des individus en 1989 et 1990. Nous nous sommes intéressés à la survenue de cardiopathies ischémiques chez ces individus, du 1^{er} janvier 1990 jusqu'au 31 décembre 2000 (date de point). Les fichiers suivants ont été utilisés pour nos analyses :

- un fichier de population indiquant l'âge, le genre, les dates de retraite et de décès, le code postal du lieu de résidence des individus en 1990.
- le fichier des réponses aux auto-questionnaires de 1989 et 1990. Il fournit des informations sur les ménages des individus de la cohorte (composition, revenu, etc.). Ces derniers y indiquaient aussi leur niveau de diplôme ainsi que leur poids et leur taille. Des questions sur le tabagisme et la consommation d'alcool ont également été posées.
- un fichier concernant la carrière professionnelle des participants. Il a été utilisé pour déterminer la catégorie socioprofessionnelle des individus à l'inclusion.
- un fichier fourni par le service général de médecine de contrôle qui rapporte les cas de cardiopathies ischémiques entre le 1^{er} janvier 1990 et le 31 décembre 2000. Il nous renseigne sur la date de survenue de l'événement coronarien et sur son type (angine de poitrine, infarctus du myocarde, mort subite). Les individus qui ont cessé de répondre aux auto-questionnaires continuent néanmoins à être suivis par le service général de médecine de contrôle. Ceux qui ont quitté l'entreprise à un moment donné (départ, retraite...) ne sont plus suivis par cet organisme.

Pour effectuer nos analyses, nous avons supprimé 737 individus de l'échantillon :

- ceux dont nous ne connaissons pas la catégorie socioprofessionnelle en 1989;
- ceux qui ont connu un évènement coronarien avant le 1^{er} janvier 1990;
- ceux qui habitaient dans les DOM-TOM;
- ceux qui n'étaient pas localisables géographiquement (code postal mal renseigné).

La cohorte étudiée compte finalement 19808 individus.

Afin de déterminer les caractéristiques du contexte de résidence des individus, nous avons utilisé les données du recensement général de la population de 1990 échantillonnées « au quart » (c'est-à-dire ne contenant qu'un habitant sur quatre) ainsi que la base de données exhaustive des logements du même recensement. Les individus de GAZEL sont localisables par leur code postal. L'INSEE fournit des données au niveau des communes. Un tableau de correspondance entre les codes postaux et les codes communes a été établi.

Mesures

Dans la présente étude, nous nous intéressons à la survenue de cardiopathies ischémiques chez les individus de la cohorte (angine de poitrine ou infarctus). Certains individus ont subi plusieurs évènements coronariens (par exemple un individu peut avoir contracté une angine de poitrine puis faire un infarctus quelques années plus tard). Dans nos analyses, nous n'avons retenu que la date du premier incident (seuls 4% des individus ont eu plus d'un incident).

Afin de prédire la survenue d'évènements coronariens chez les individus de la cohorte, nous nous sommes à la fois intéressés à leurs caractéristiques individuelles et aux caractéristiques de leur contexte de résidence. Après avoir vérifié que l'effet de l'âge était linéaire, nous en avons tenu compte à l'aide d'une variable continue, qui a été divisée par cinq afin d'estimer le risque correspondant à une augmentation de 5 ans. Le niveau d'éducation, représenté ici par le diplôme le plus élevé obtenu, a été codé en deux classes : d'une part ceux qui ont obtenu un baccalauréat de la voie générale ou un diplôme de l'enseignement supérieur, d'autre part ceux qui ont un diplôme inférieur au baccalauréat (ou pas de diplôme) ou qui ont suivi une voie technologique et technique. La catégorie socio-professionnelle des individus est définie selon la nomenclature des Professions et Catégories Sociales de l'INSEE (20). Nous avons distingué les cadres, les professions intermédiaires, et les employés et ouvriers regroupés en une seule et même classe. Pour caractériser le niveau de vie des individus, nous ne disposions que d'une fourchette du revenu du ménage et de la composition de celui-ci. Nous avons divisé le revenu du ménage par le nombre d'unités de consommation (un adulte compte pour une part, un enfant pour une demi part). Ce revenu a été discrétisé en

trois classes. Le statut matrimonial des individus a été codé en deux classes, les individus célibataires, divorcés ou veufs d'une part, les individus mariés ou vivant en couple d'autre part.

Les questionnaires de 1989 et 1990 comportent des questions sur différents comportements de santé ou facteurs de risque. Nous en avons inclus trois dans nos modèles :

- L'activité tabagique a été codée en quatre classes : non fumeurs, ex-fumeurs, petits fumeurs (équivalent 20 cigarettes par jour ou moins), gros fumeurs (équivalent à plus de 20 cigarettes).

- La consommation d'alcool a été codée en trois classes. Outre les petits buveurs et gros buveurs (en prenant comme seuil celui utilisé dans les travaux réalisés à partir de la cohorte GAZEL, équivalent à trois verres de vin ou bière par jour pour les hommes, deux pour les femmes), nous avons regroupé les individus abstinentes dans une classe à part.

- L'indice de masse corporelle (IMC) – défini par le rapport du poids sur le carré de la taille en mètres – a été utilisé comme indicateur du surpoids. Nous l'avons découpé en trois classes selon les valeurs 25 et 30 (un IMC entre 25 et 30 correspond à une situation de surpoids, un IMC supérieur ou égal à 30 correspondant à une obésité).

Nous avons finalement défini et créé trois caractéristiques contextuelles à l'échelle du code postal et du département. Il s'agit ici du pourcentage des individus de la zone géographique âgés de 25 à 60 ans 1) qui sont chefs d'une entreprise de 10 salariés ou plus, 2) qui sont ouvriers non qualifiés, 3) qui ont un diplôme post-bac. Le recensement n'incluant aucune question sur le revenu individuel, nous n'avons pu définir de variable contextuelle du revenu moyen. Les variables contextuelles obtenues ont été discrétisées en quatre classes au moyen des quartiles.

Analyse statistique

Le suivi des individus étant interrompu lors de leur départ en retraite ou de leur décès, nous avons employé des modèles de Cox pour estimer les effets des caractéristiques individuelles sur le risque d'événements coronariens.

Avant d'étudier les effets des variables individuelles et contextuelles, nous avons quantifié la variabilité du phénomène dans l'espace. Pour ce faire, nous avons estimé des modèles de Cox multiniveaux vides (n'incluant aucune variable explicative à part l'âge et le genre) et observé la variance inter-zone de l'effet aléatoire défini au niveau des zones de résidence (21, 22).

Un Modèle de Cox multiniveau est un modèle de Cox qui comprend un terme aléatoire correspondant au niveau des zones de résidence. Le modèle de Cox multiniveau vide s'écrit de la façon suivante :

$$\lambda(t, \mathbf{X}) = \lambda_0(t)\exp(\beta_0 + U_{0j}) \quad \text{où } U_{0j} \sim N(0, \sigma_{u0}^2)$$

Le terme aléatoire U_{0j} permet de tenir compte du niveau de risque spécifique dans chaque zone j . Plutôt que d'estimer directement un terme pour chaque zone j , le modèle suppose que cet effet aléatoire de niveau zone suit une distribution normale et se contente d'estimer la variance σ_{u0}^2 de sa distribution (23). Ces modèles permettent de voir si deux individus d'une même zone sont plus similaires quant au risque coronarien que deux individus issus de zones différentes. Des modèles multiniveaux séparés ont été estimés en considérant successivement au niveau 2 de la hiérarchie les codes postaux, les départements ou les régions de résidence des individus. Nous avons estimé ces modèles vides parmi les individus les plus urbains de la cohorte et parmi les autres individus (moins urbains) pour essayer de quantifier la variabilité géographique dans ces deux groupes. Le caractère urbain du contexte de résidence a été déterminé à partir d'une typologie de commune définie par l'INSEE dans le recensement de 1990. Nous appelons « très urbain » un individu qui habite dans une unité urbaine de plus de 100.000 habitants. Un graphique rapporté en annexe donne les raisons de ce choix.

Les modèles ont ensuite été ajustés sur les caractéristiques des individus afin d'examiner si la variabilité géographique d'incidence des troubles coronariens était due à des effets de compositions (par exemple, à un regroupement d'individus riches ou pauvres dans certaines régions). Toujours à différentes échelles, nous avons donc examiné si une variabilité géographique subsistait après avoir tenu compte des caractéristiques individuelles (état matrimonial, éducation, catégorie socioprofessionnelle et revenu).

Nous avons mené ces analyses pour l'ensemble de la cohorte mais aussi chez les seuls hommes.

L'étape suivante a consisté en l'introduction de différentes variables contextuelles dans les modèles. Nous avons cherché à voir si les 3 caractéristiques contextuelles étudiées étaient associées au risque coronarien. La dernière étape a consisté à ajuster le modèle sur 3 variables de comportements de santé ou facteurs de risque (surpoids, consommation d'alcool et de tabac) dans le but de voir si elles intervenaient comme facteurs médiateurs de l'environnement socio-économique sur la santé coronarienne.

Les statistiques descriptives ont été calculées avec le logiciel SAS. Nous avons estimé les modèles de Cox avec le logiciel R (24).

Résultats

L'échantillon retenu pour l'analyse compte 19808 individus, dont 27,3% de femmes.

Les individus avaient tous entre 36 et 51 ans au 1^{er} janvier 1990. L'âge médian était de 44,7 ans (écart interquartile : 42,3 – 47,5).

Entre le 1^{er} janvier 1990 et le 31 décembre 2000, 325 individus de la cohorte (1,6%) ont eu une cardiopathie ischémique (angine de poitrine ou infarctus), dont 195 (0,9%) un infarctus du myocarde.

Dans notre population, 0,28% des femmes ont eu une cardiopathie ischémique (15/5387 individus). Ce pourcentage est de 2,15% pour les hommes (310/4096).

Si l'on s'intéresse à la population urbaine de la cohorte (définie comme habitant dans une unité urbaine de plus de 100000 habitants), 0,26% des femmes ont eu une cardiopathie ischémique (9/3451) contre 2,25% des hommes (154/6698). Aucune étude ne pouvant être conduite sur la population féminine de la cohorte pour des raisons d'effectifs insuffisants, l'essentiel des analyses qui suivent n'ont été conduites que parmi les hommes.

Le tableau 1 décrit les caractéristiques socio-économiques des individus très urbains et moins urbains de la cohorte. La population très urbaine de la cohorte comprend une proportion de femmes supérieure à la population moins urbaine. Elle a un revenu moyen plus élevé et moins fréquemment un diplôme technique ou inférieur au baccalauréat. Les individus les plus urbains ont connu autant d'incidents coronariens que les moins urbains.

Tableau 1 : Caractérisation des sous échantillons de GAZEL, 1990-2000 (hommes et femmes).

	Individus très urbains (n = 10303)	Individus moins urbains (n = 9505)	Echantillon total (n = 19808)	p
Age moyen au 1 ^{er} janvier 1990 (en années) *	44,7	44,8	44,7	0,15
Proportion de femmes §	33,5%	20,5%	27,3%	<0,0001
Proportion de diplômes techniques ou inférieurs au baccalauréat §	67,2%	78,1%	72,4%	<0,0001
Proportion d'individus cadres et professions intellectuelles supérieures §	27,9%	17,3%	22,8%	<0,0001
Revenu moyen (en francs) *	6464F	5325F	5916F	<0,0001
Proportion de patrons dans le code postal ‡	0,7%	0,6%	0,7%	<0,001
Proportion d'ouvriers non qualifiés dans le code postal ‡	6,6%	9,4%	7,9%	<0,001
Proportion d'individus ayant présenté au moins un incident coronarien au cours du suivi §	1,6%	1,7%	1,6%	0,5

* Test de Student, § Test du Khi-2, ‡ Test de Wilcoxon

Tableau 2: Modèles vides (ajustés sur l'âge et genre) pour le fait d'avoir un incident coronarien au cours de la période de suivi : variance de l'effet aléatoire. Cohorte GAZEL, 1990-2000 (hommes et femmes).

	Population totale N=19808	Population moins urbaine N=9505	Population très urbaine N=10303
Niveau géographique			
Code Postal	0,235 (p=0,32)	0,083 (p=0,43)	0,262 (p=0,32)
Département	0,048 (p=0,19)	0,001 (p=0,45)	0,132 (p=0,06)
Région	0,038 (p=0,04)	0,001 (p=0,5)	0,135 (p=0,003)

Nous constatons qu'il existe une variabilité géographique du risque coronarien dans la population de GAZEL (Tableau 2). L'estimation ponctuelle de la variance augmente à mesure que la taille de la zone diminue, indiquant que l'on capte des variations géographiques plus importantes quand on travaille à une échelle plus fine. Cependant, le degré de signification associé à cette variance inter-zone augmente à mesure que la taille de la zone diminue, probablement pour des questions de puissance statistique. En effet, au niveau des codes postaux, il existe beaucoup de zones contenant peu d'individus (70% des codes postaux contiennent cinq individus ou moins, 30% un individu seulement). Sur le plan statistique, seule la variance interrégionale est significativement supérieure à zéro (à 5% d'erreur). Par ailleurs, il apparaît que les variations inter-régions sont fortes lorsque l'on considère la population très urbaine mais qu'elles sont pratiquement inexistantes pour la population moins urbaine. En conséquence, la plupart des analyses qui suivent ont été conduites sur la population la plus urbaine de la cohorte.

Tableau 3 : Modèles vides (ajustés sur l'âge) pour le fait d'avoir un incident coronarien au cours de la période de suivi chez les hommes : variance de l'effet aléatoire, cohorte GAZEL 1990-2000.

	Population totale N= 14406	Population moins urbaine N=7554	Population très urbaine N=6852
Niveau géographique			
Code Postal	0,150 (p=0,39)	0,021 (p=0,39)	0,204 (p=0,370)
Département	0,044 (p=0,21)	0,001 (p=0,45)	0,103 (p=0,110)
Région	0,040 (p=0,055)	0,001 (p=0,49)	0,125 (p=0,008)

Nous avons répété ces analyses multiniveaux en ne tenant compte que des hommes (Tableau 3). Dans cet échantillon, la variabilité géographique du risque coronarien est très proche de celle estimée à partir de l'ensemble de la cohorte. De même, on retrouve le fait que la variabilité régionale du risque coronarien observée est avant tout due à la variabilité régionale du risque dans l'échantillon des individus très urbains.

Cette variabilité régionale peut provenir de plusieurs sources. Une partie des variations peut provenir d'effets de composition (par exemple, les cadres de l'entreprise ne sont pas forcément répartis de façon uniforme d'une région à l'autre). Par ailleurs, une partie de la variabilité du risque peut être liée aux effets des caractéristiques du contexte de résidence.

Nous avons ensuite introduit les caractéristiques socio-démographiques des individus dans le modèle. Comme l'indique le tableau 6, un risque d'événements coronariens accru a été identifié chez les individus âgés, chez les non cadres et chez les individus n'ayant pas le baccalauréat ou un diplôme de l'enseignement supérieur. En ajustant sur les caractéristiques socio-économiques des individus, la mesure de la variabilité régionale diminue de 38,4% (de 0,125 à 0,077). Le revenu individuel semble avoir un effet moins important que les autres caractéristiques socio-économiques.

Tableau 4 : Fréquence d'incidents coronariens dans chaque quartile des variables contextuelles : Hommes urbains. Cohorte GAZEL, 1990-2000.

	1er quartile	2e quartile	3e quartile	4e quartile	Z, p *
Pourcentage de patrons					
Niveau code postal	2,92%	2,52%	1,86%	1,70%	Z= -2,70 p=0,003
Niveau département	2,77%	2,51%	1,64%	1,96%	Z= -2,06 p=0,019
Pourcentage d'ouvriers non qualifiés					
Niveau code postal	1,69%	2,34%	2,68%	2,28%	Z= 1,31 p=0,095
Niveau département	1,96%	1,85%	2,34%	2,86%	Z= 1,99 p=0,022
Pourcentage de population qui a un diplôme post bac					
Niveau code postal	2,64%	2,56%	1,93%	1,86%	Z= -1,84 p= 0,032
Niveau département	2,82%	2,13%	1,73%	2,06%	Z= -1,94 p=0,026

* Test unilatéral de tendance

Nous nous sommes ensuite intéressés aux effets des facteurs contextuels. Pour la plupart des variables contextuelles on observe une décroissance de la fréquence d'incidents lorsque l'on passe d'un quartile à un autre, supposé plus favorisé (Tableau 4). Ce gradient est particulièrement fort et régulier pour la proportion de chefs d'entreprise dans le département ou code postal de l'individu. Les valeurs seuil utilisées pour discrétiser les variables contextuelles sont rapportées en annexe.

Tableau 5 : effets des caractéristiques contextuelles issus de modèles multiniveaux ajustés sur les caractéristiques des individus.
Hommes Urbains: Variables définies au niveau du Code postal.

	Modèle ajusté sur l'âge		Modèle ajusté sur l'âge, l'état matrimonial, le niveau d'éducation, la profession, le revenu.	
Pourcentage de patrons				
4 ^e quartile	1,00		1,00	
3 ^e quartile	1,27	(0,76 ; 2,11)	1,10	(0,66 ; 1,83)
2 ^e quartile	1,70	(1,05 ; 2,77)	1,39	(0,86 ; 2,27)
1 ^{er} quartile	2,22	(1,38 ; 3,56)	1,74	(1,08 ; 2,80)
Pourcentage d'ouvriers non qualifiés				
1 ^{er} quartile	1,00		1,00	
2 ^e quartile	1,50	(0,92 ; 2,45)	1,35	(0,83 ; 2,20)
3 ^e quartile	1,57	(0,95 ; 2,59)	1,42	(0,87 ; 2,30)
4 ^e quartile	1,27	(0,71 ; 2,26)	1,21	(0,73 ; 2,03)
Pourcentage de population qui a un diplôme post bac				
4 ^e quartile	1,00		1,00	
3 ^e quartile	1,26	(0,76 ; 2,08)	1,02	(0,61 ; 1,69)
2 ^e quartile	1,93	(1,21 ; 3,08)	1,42	(0,88 ; 2,29)
1 ^{er} quartile	1,96	(1,20 ; 3,18)	1,34	(0,82 ; 2,22)

Dans des modèles qui n'étaient ajustés que sur l'âge, on observe des effets de type dose-réponse pour certaines variables contextuelles (pourcentage de chefs d'entreprise) mais pas pour d'autres (tableau 5). Une partie des associations identifiées résistent à un ajustement sur les caractéristiques socio-économiques des individus (pourcentage de chefs d'entreprise), indiquant un risque coronarien accru dans les zones plus défavorisées. Nous avons ensuite en plus réajusté les modèles sur trois variables de comportements de santé ou facteurs de risque (tableau 6). Même si le fait de fumer et d'être obèse était associé à un risque coronarien accru, ces trois variables ne semblent expliquer qu'une toute petite partie des effets du contexte socio-économique de résidence. Ces comportements ne semblent pas expliquer la variabilité géographique résiduelle.

Tableau 6 : Effets des caractéristiques individuelles et contextuelles sur le risque d'événements coronariens.

Hazard ratios et IC à 95%

	Modèle vide	Modèle incluant des caractéristiques socio-économiques	Modèle incluant des variables contextuelles	Modèle incluant des variables liées à des comportements de santé.
Age (Unité=5 ans)	3,15 (2,30 ; 4,31)	4,42 (3,15 ; 6,20)	4,48 (3,19 ; 6,28)	4,47 (3,19 ; 6,29)
Profession (réf : cadres)				
Professions intermédiaires		1,68 (1,04 ; 2,70)	1,62 (1,01 ; 2,61)	1,64 (1,00 ; 2,59)
Employés et Ouvriers		1,82 (0,97 ; 3,43)	1,79 (0,95 ; 3,35)	1,76 (0,93 ; 3,31)
Statut matrimonial (réf : mariés ou en couple)				
Célibataires		1,20 (0,68 ; 2,13)	1,20 (0,68 ; 2,13)	1,00 (0,56 ; 1,79)
Niveau d'éducation (réf : baccalauréat et supérieur)				
Sans diplôme ou diplôme technique		1,80 (1,13 ; 2,85)	1,71 (1,08 ; 2,72)	1,64 (1,02 ; 2,55)
Revenu (réf : revenu supérieur à 6600 francs)				
Revenu entre 4600 et 6600 francs		1,09 (0,69 ; 1,73)	1,09 (0,69 ; 1,72)	1,03 (0,65 ; 1,63)
Revenu inférieur à 4600 francs		1,34 (0,85 ; 2,12)	1,30 (0,82 ; 2,06)	1,19 (0,75 ; 1,89)
Variable contextuelle (réf : Pourcentage de patrons 4 ^e quartile)				
3 ^e quartile			1,10 (0,66 ; 1,83)	1,10 (0,66 ; 1,84)
2 nd quartile			1,39 (0,86 ; 2,27)	1,37 (0,84 ; 2,22)
1 ^{er} quartile			1,74 (1,08 ; 2,79)	1,68 (1,04 ; 2,70)
IMC (réf : IMC inférieur à 25)				
IMC entre 25 et 30				1,94 (1,18 ; 3,17)
IMC supérieur à 30				2,42 (1,36 ; 4,33)
Tabagisme (réf : non-fumeurs)				
Ex-fumeurs				1,24 (0,81 ; 1,89)
Petits fumeurs				1,94 (1,18 ; 3,17)
Gros fumeurs				3,12 (2,00 ; 4,87)
Consommation d'alcool (réf : petits buveurs)				
Gros buveurs				0,83 (0,57 ; 1,22)
Abstinentes				0,93 (0,29 ; 2,99)
Variance de l'effet aléatoire et p-value	0,125 (p=0,008)	0,077 (p=0,10)	0,064 (p=0,15)	0,066 (p=0,13)

Discussion

Dans cette étude, nous nous sommes intéressés aux variations géographiques du risque de cardiopathie ischémique sur le territoire métropolitain. Ce travail pourrait corroborer l'existence d'une association entre le contexte socio-économique de résidence des individus et la survenue de cardiopathies ischémiques, au-delà des effets liés à certaines caractéristiques socio-économiques individuelles (5, 25, 26). Les comportements de santé et facteurs de risque pris en compte (consommation de tabac et d'alcool, surcharge pondérale) ne semblent pas intervenir comme facteurs médiateurs des effets du contexte socio-économique sur la santé coronarienne.

Toutefois, il faut être très prudent lors de l'interprétation de ces résultats car notre travail s'est heurté à de nombreuses limites. La principale faiblesse de cette étude provient du fait que le contexte est ici défini à une échelle très large (au plus fin à l'échelle du code postal). Les précédentes études ont montré un rôle du contexte à l'échelle du quartier (une échelle géographique inférieure à la commune) (7, 12). Deuxièmement, la base de données utilisée ne comportant pas de données biomédicales, il a été impossible d'ajuster nos résultats sur d'importants facteurs de risque comme l'hypertension artérielle, le cholestérol ou le diabète. Des travaux suggèrent que ces facteurs pourraient être associés aux caractéristiques du lieu de vie (9). Troisièmement, les données concernant les comportements de santé ayant été rapportées par les individus eux-mêmes, leur exactitude peut être mise en question.

Ce travail a bénéficié de la qualité des données de GAZEL (recueillies par l'unité 687 de l'INSERM) et notamment de l'importance de la durée de suivi des individus. Nous pouvons penser que les incidents coronariens ont été quasi exhaustivement enregistrés (27). L'originalité de notre analyse provient du fait, qu'à notre connaissance, aucune étude contextuelle des pathologies coronariennes n'avait été conduite sur des données françaises.

Nous avons constaté que la variabilité du risque coronarien était plus importante entre les codes postaux qu'entre les régions. Toutefois, pour des raisons de puissance statistique, ce résultat n'a pu être établi. En effet, la variabilité géographique au niveau des codes postaux n'a pu être précisément étudiée (notre population étant dispersée sur tout le territoire métropolitain, il y avait souvent très peu d'individus dans chaque code postal).

Même si cela n'apporte pas de connaissance directe sur les pathologies coronariennes, le fait que la variabilité régionale soit uniquement due aux individus vivant dans les zones les plus urbaines est un résultat qui mérite d'être mieux compris dans des études futures. Une étude menée en Suède avait relevé un phénomène comparable, suggérant que seules les

disparités socio-spatiales existant en milieu urbain donnent lieu à des variations géographiques du risque coronarien (19).

Dans notre étude, l'impact des caractéristiques socio-économiques individuelles sur la santé coronarienne semble capté en grande partie par la catégorie socioprofessionnelle des individus. Une des raisons à cela peut provenir du fait que la variable caractérisant le revenu individuel a été estimée à partir d'une fourchette de revenu du ménage, rapportée par les individus, donc de façon très imprécise. Le fait que les individus de la cohorte soient tous des salariés des entreprises EDF et GDF implique que les gradients socio-économiques individuels et contextuels peuvent avoir été sous-estimés par rapport à ceux qu'on aurait observé dans un échantillon représentatif de la population française (la population de la cohorte étant très homogène du point de vue du statut d'emploi). Par ailleurs, nous n'avons pas observé d'effet du statut matrimonial alors que la littérature passée suggère un effet protecteur du fait de vivre en couple (10).

Pour caractériser le contexte socio-économique de résidence, nous avons notamment tenu compte du pourcentage de chefs d'entreprise de plus de dix salariés qui habitent dans le même code postal que l'individu. Nous pensons que les individus à très hauts salaires ont une plus grande capacité à choisir leur lieu d'habitation; leur présence en nombre important dans une zone géographique est le signe d'une zone socialement très favorisée. Cette variable donne une indication du niveau de richesse « agrégé » du lieu d'habitation des individus, constituant un marqueur d'un certain nombre de caractéristiques du contexte de résidence (accès aux services et infrastructures, interrelations sociales etc.). Après ajustement sur quelques caractéristiques socio-économiques des individus, nous avons trouvé que le risque coronarien augmentait avec le niveau de pauvreté de la zone de résidence. Toutefois, une limite de ce résultat est qu'il n'a pas été trouvé pour l'ensemble des variables contextuelles utilisées (pourcentage d'ouvriers non qualifiés, pourcentage de population qui a un diplôme post-bac). S'appuyant sur la littérature (7, 28), nous pouvons supposer que des effets contextuels plus forts auraient été trouvés si nous avions pu définir les variables contextuelles à une échelle plus locale (du « quartier » par exemple), échelle à laquelle nous n'avons pas de données.

Nous sommes cependant bien incapables de déterminer précisément les phénomènes sous-jacents à cet effet du contexte socio-économique de résidence. Le fait de s'intéresser à la survenue d'événements coronariens plutôt qu'à la mortalité nous autorise toutefois à penser que la variabilité géographique constatée ne provient pas de disparités d'accès aux soins d'urgence (temps pour aller à l'hôpital ou pour être secouru).

Une hypothèse avancée dans la littérature est que certains comportements de santé et facteurs de risque (alcool, tabac, surpoids) interviennent comme facteurs médiateurs de l'effet du contexte socio-économique sur la santé coronarienne. Il a parfois été montré que ces comportements expliquaient une grande partie des gradients sociaux du risque coronarien (29). Nous n'avons pas retrouvé ce résultat. De plus, ces comportements ne semblent que très partiellement expliquer la variabilité géographique du risque de cardiopathie ischémique.

Au final, d'un point de vue de santé publique, il apparaît important de mener des analyses à une échelle géographique plus fine (si possible à l'échelle du « quartier ») afin de décrire et expliquer les éventuelles associations entre caractéristiques socio-économiques du lieu de résidence et santé coronarienne en France.

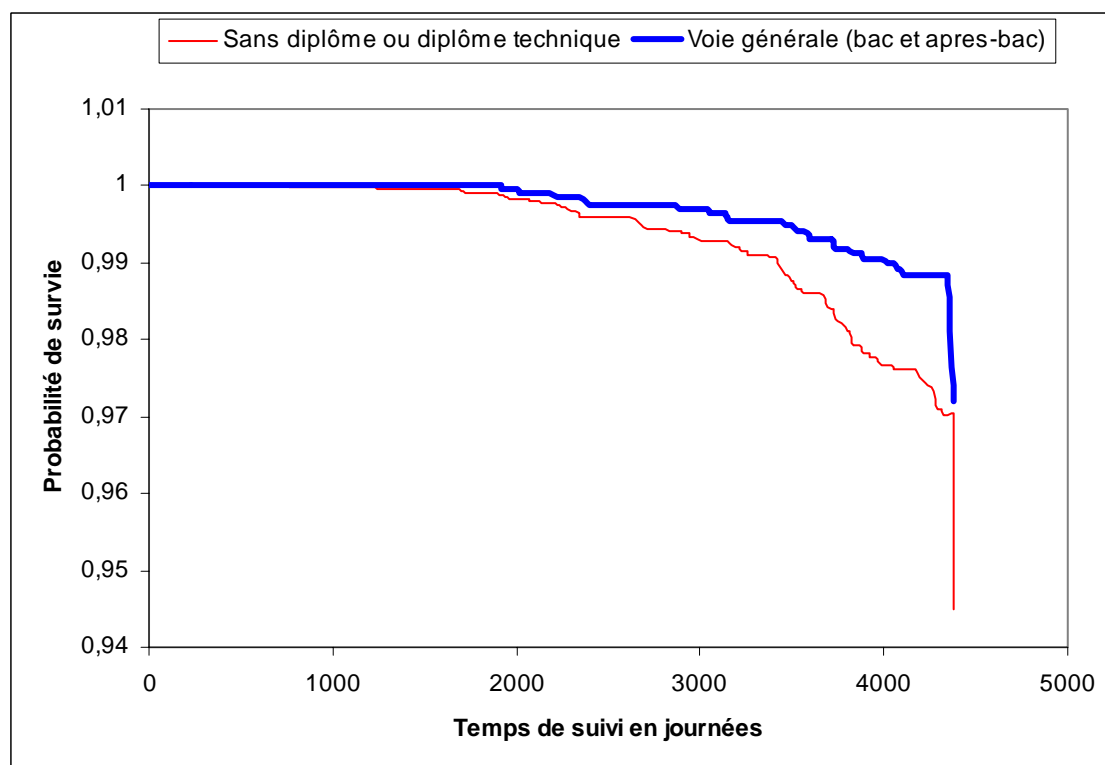
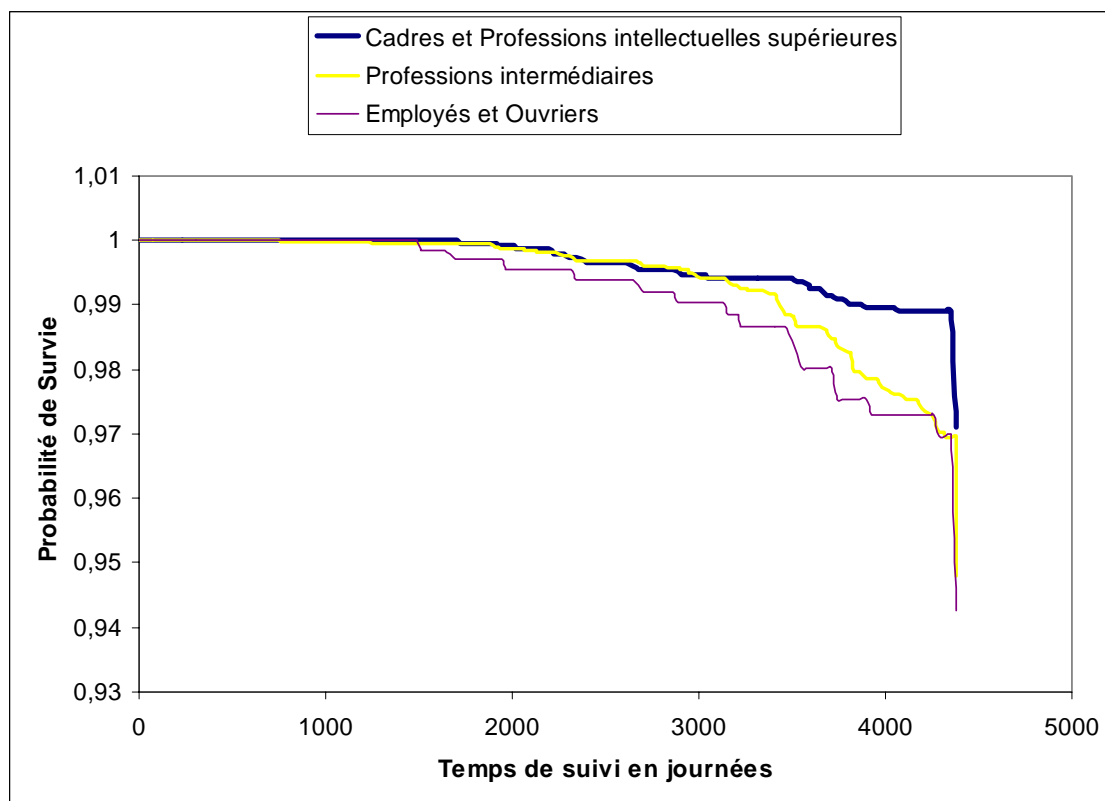
Références

1. Kawachi IB, L. Neighborhoods and health. New York: Oxford University Press, 2003.
2. O'Campo P, Xue X, Wang MC, Caughy M. Neighborhood risk factors for low birthweight in Baltimore: a multilevel analysis. *Am J Public Health* 1997;87:1113-8.
3. Duncan C, Jones K, Moon G. Smoking and deprivation: are there neighbourhood effects? *Soc Sci Med* 1999;48:497-505.
4. 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.
5. Sundquist J, Malmstrom M, Johansson SE. Cardiovascular risk factors and the neighbourhood environment: a multilevel analysis. *Int J Epidemiol* 1999;28:841-845.
6. 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.
7. Waitzman NJ, Smith KR. Phantom of the area: poverty-area residence and mortality in the United States. *Am J Public Health* 1998;88:973-6.
8. Kaplan GA, Keil JE. Socioeconomic factors and cardiovascular disease: a review of the literature. *Circulation* 1993;88:1973-1998.
9. Lawlor DA, Bedford C, Taylor M, Ebrahim S. Geographical variation in cardiovascular disease, risk factors, and their control in older women: British Women's Heart and Health Study. *J Epidemiol Community Health* 2003;57:134-40.
10. Morris RW, Whincup PH, Lampe FC, Walker M, Wannamethee SG, Shaper AG. Geographic variation in incidence of coronary heart disease in Britain: the contribution of established risk factors. *Heart* 2001;86:277-83.
11. Chaix B, Chauvin P. Tobacco and alcohol consumption, sedentary lifestyle and overweightness in France: a multilevel analysis of individual and area-level determinants. *Eur J Epidemiol* 2003;18:531-538.
12. Chaix B, Rosvall M, Merlo J. Increase of neighborhood socioeconomic effects on ischemic heart disease mortality over the recent period: a multilevel survival analysis of two large Swedish cohorts. *Am J Epidemiol* 2006;in press.
13. Diez-Roux AV, Nieto FJ, Muntaner C, et al. Neighborhood environments and coronary heart disease: a multilevel analysis. *Am J Epidemiol* 1997;146:48-63.

14. Diez Roux AV. Residential environments and cardiovascular risk. *J Urban Health* 2003;80:569-89.
15. Davey Smith G, Hart C, Watt G, Hole D, Hawthorne V. Individual social class, area-based deprivation, cardiovascular disease risk factors, and mortality: the Renfrew and Paisley Study. *J Epidemiol Community Health* 1998;52:399-405.
16. Diez Roux AV, Merkin SS, Arnett D, et al. Neighborhood of residence and incidence of coronary heart disease. *N Engl J Med* 2001;345:99-106.
17. Diez-Roux AV. Multilevel analysis in public health research. *Annu Rev Public Health* 2000;21:171-192.
18. Leyland AH, Goldstein H. *Multilevel modelling of health statistics*. Chichester, England: Wiley, 2001.
19. Chaix B, Rosvall M, Merlo J. Assessment of the magnitude of geographic variations and socioeconomic contextual effects on ischaemic heart disease mortality: a multilevel survival analysis of a large Swedish cohort. *J Epidemiol Community Health* 2006;in press.
20. INSEE. *Nomenclature des professions et catégories socioprofessionnelles [List of professions and social categories]* (in French). Paris, France: Institut National de la Statistique et des Etudes Economiques, 1994.
21. Merlo J, Chaix B, Yang M, Lynch JW, Rastam L. A brief conceptual tutorial of multilevel analysis in social epidemiology - interpreting neighbourhood differences and the effects of neighbourhood characteristics on individual health. *J Epidemiol Community Health* 2005;59:1022-1029.
22. Merlo J, Chaix B, Yang M, Lynch JW, Rastam L. A brief conceptual tutorial of multilevel analysis in social epidemiology - linking the statistical concept of clustering to the idea of contextual phenomenon. *J Epidemiol Community Health* 2005;59:443-449.
23. Snijders T, Bosker R. *Multilevel Analysis. An introduction to basic and advanced multilevel modelling*. London, UK: Sage Publications, 1999.
24. Therneau T. *Survival R package, version 2.20*: (Accessed November 15, 2005, at <http://www.r-project.org/>):(Available at <http://www.r-project.org/>).
25. Sundquist K, Winkleby M, Ahlen H, Johansson SE. Neighborhood socioeconomic environment and incidence of coronary heart disease: a follow-up study of 25,319 women and men in Sweden. *Am J Epidemiol* 2004;159:655-662.
26. Sundquist K, Malmstrom M, Johansson SE. Neighbourhood deprivation and incidence of coronary heart disease: a multilevel study of 2.6 million women and men in Sweden. *J Epidemiol Community Health* 2004;58:71-77.
27. Chevalier A, Zins M, Godard C, et al. [A registry of ischaemic cardiopathies among active workers at Electricite de France-Gaz de France. Program development and first results]. *Rev Epidemiol Sante Publique* 2001;49:51-60.
28. Boyle MH, Willms JD. Place effects for areas defined by administrative boundaries. *Am J Epidemiol* 1999;149:577-585.
29. Lynch J, Davey Smith G, Harper S, Bainbridge K. Explaining the social gradient in coronary heart disease: comparing relative and absolute risk approaches. *J Epidemiol Community Health* 2006;60:436-41.

ANNEXE

Figures 1-2 : Catégorie Socioprofessionnelle et niveau d'éducation chez les hommes très urbains de la cohorte GAZEL: Courbe De Kaplan Meier (1990-2000)



Figures 3-4 : Revenu individuel et pourcentage de chefs d'entreprise dans le code postal des hommes très urbains de GAZEL: Courbe De Kaplan Meier (1990-2000)

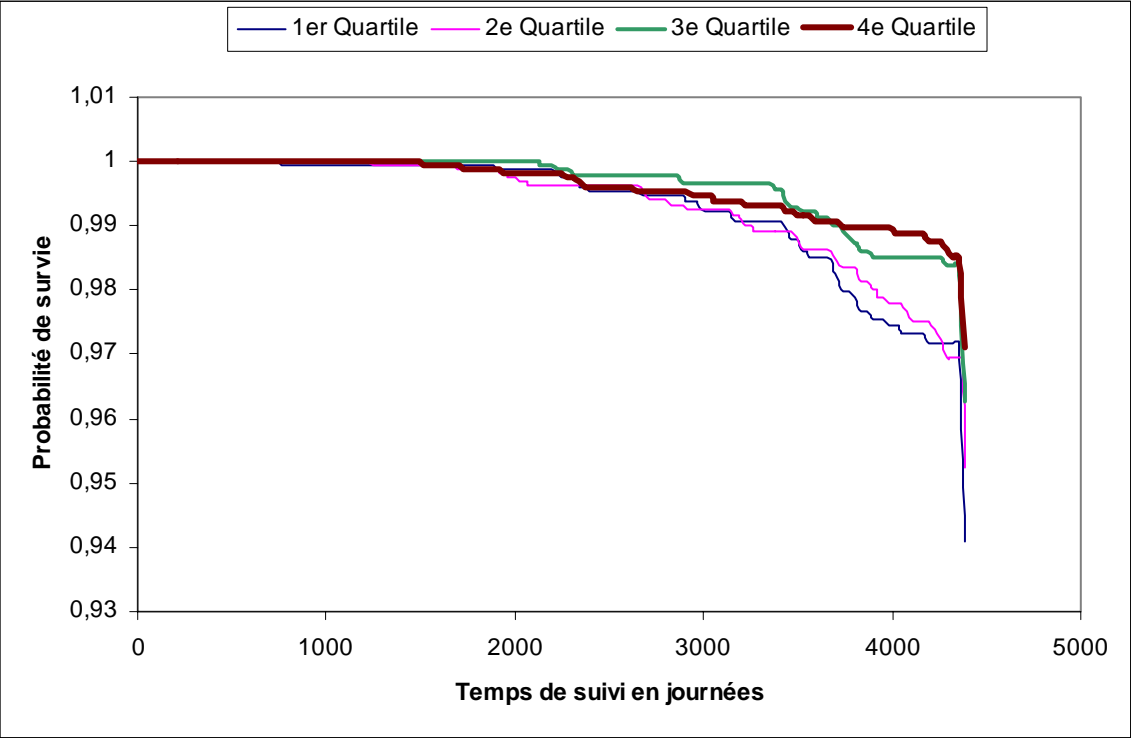
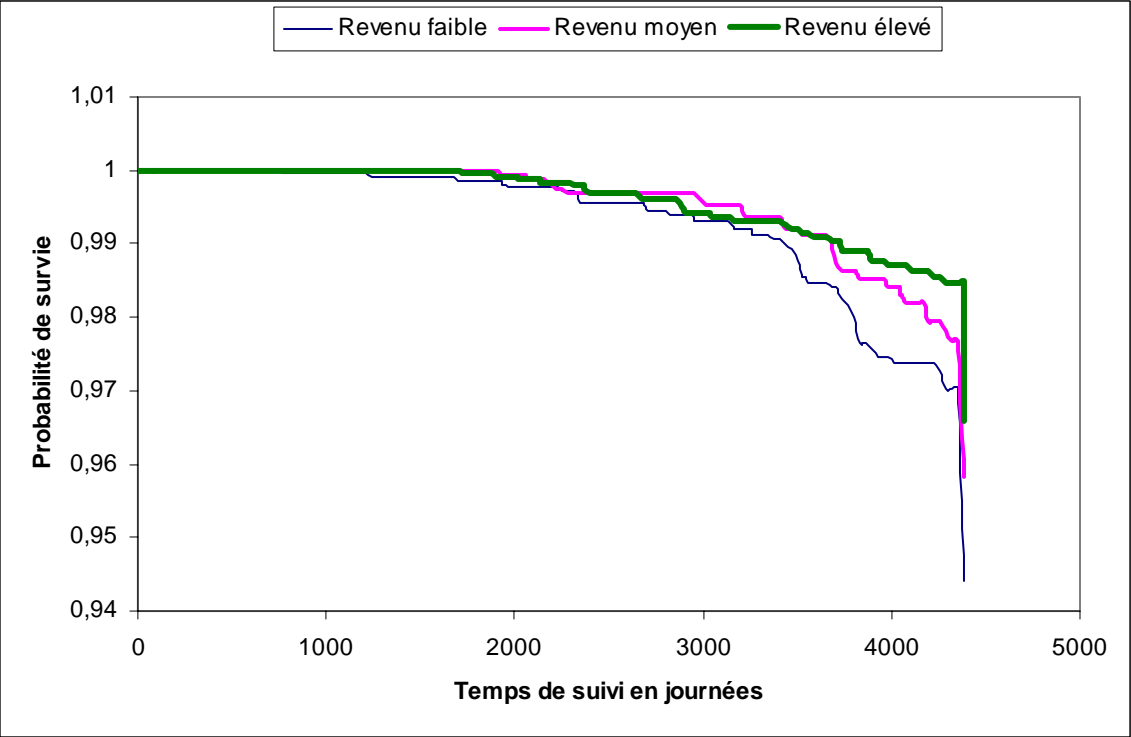


Tableau 1 : Caractéristiques des différentes sous populations masculines « très urbaines » de la cohorte GAZEL en 1990.

	Variance de l'effet aléatoire au niveau régional.	p-value de l'effet aléatoire	Nombre de régions dans l'échantillon	Nombre de régions qui contiennent moins de 50 individus
Unités urbaines de plus de 50.000 hab. N=7900	0.105	0.014	22	1
Unités urbaines de plus de 70.000 hab. N=7293	0.123	0.0097	21	1
Unités urbaines de plus de 100.000 hab. N=6852	0.125	0.008	21	4
Unités urbaines de plus de 150.000 hab. N=6146	0.141	0.0054	19	4
Unités urbaines de plus de 200.000 hab. N=5646	0.096	0.042	16	2
Unités urbaines de plus de 300.000 hab. N=4869	0.035	0.24	10	1

La variabilité du risque coronarien entre régions semble augmenter au fur et à mesure que l'effectif de la population des « plus urbains » diminue, jusqu'à une certaine limite.

Cependant, le nombre de régions d'étude et la significativité de cette variance diminuent avec la taille de l'échantillon. La valeur de 100000 habitants nous a semblé être le meilleur compromis possible pour pouvoir caractériser une unité très urbaine (Tableau 1 et Figure 5).

Figure 5: Evolution de la variance de l'effet aléatoire selon la définition d'un individu très urbain :

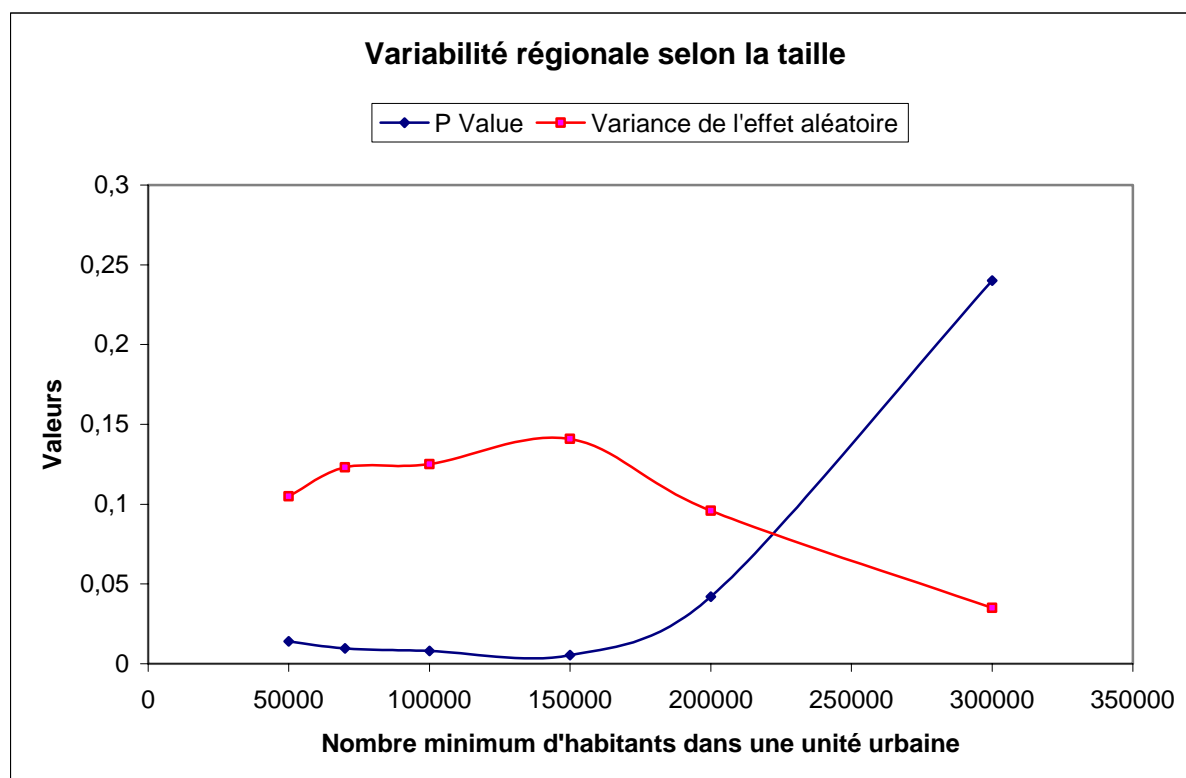


Tableau 2 : Valeurs seuils des variables contextuelles (Tableau 4). Hommes urbains de GAZEL.

	1er Quartile	2e Quartile	3e Quartile	4e Quartile
Pourcentage de patrons				
Niveau Code postal	Inférieur à 0,39%	Entre 0,39% et 0,59%	Entre 0,59% et 0,82%	Supérieur à 0,82%
Niveau Département	Inférieur à 0,53%	Entre 0,53% et 0,62%	Entre 0,62% et 0,77%	Supérieur à 0,77%
Pourcentage d'ouvriers non qualifiés				
Niveau Code postal	Inférieur à 4,33%	Entre 4,33% et 6,22%	Entre 6,22% et 8,55%	Supérieur à 8,55%
Niveau Département	Inférieur à 6,05%	Entre 6,05% et 7,00%	Entre 7,00% et 10,02%	Supérieur à 10,02%
Pourcentage de population qui a un diplôme post bac				
Niveau Code postal	Inférieur à 13,05%	Entre 13,05% et 17,30%	Entre 17,30% et 23,80%	Supérieur à 23,80%
Niveau Département	Inférieur à 13,31%	Entre 13,31% et 14,70%	Entre 14,70% et 19,00%	Supérieur à 19,00%