

# ESTIMATIONS DÉPARTEMENTALES DU HANDICAP DANS L'ENQUÊTE NATIONALE HANDICAP-SANTÉ 2008

*Josiane Le Guennec, INSEE-Rennes*

L'INSEE réalise, avec une périodicité décennale, une enquête sur la santé destinée à mesurer les prévalences des maladies, le recours aux soins et les comportements ayant une influence sur la morbidité de la population. Depuis 1999, ce dispositif s'est enrichi d'une enquête sur le handicap et les situations d'incapacité résultant de problèmes de santé. Une première enquête avait été réalisée en 1999-2001. Elle a été reconduite, sous une forme proche, en 2007-2008, en collaboration avec la Direction de la Recherche, des Etudes, de l'Evaluation et des Statistiques (DREES) du Ministère de la Santé.

L'ensemble du système d'observation de 2008 comprenait trois volets :

- Le volet principal de l'enquête, dit Enquête Handicap-Santé Ménages (HSM), a pour cible la population vivant en « ménages ordinaires », interrogée par enquêteur au domicile résidence principale. Il s'agit d'une enquête en deux phases.
- Un second volet concernait les personnes hébergées dans des institutions telles que les établissements hospitaliers spécialisés ou les maisons de retraite.
- Une troisième enquête a été réalisée auprès des individus apportant une aide régulière à une personne handicapée ou dépendante, identifiés par le questionnaire de l'enquête HSM.

Seul le premier volet a été utilisé dans cette étude.

L'enquête HSM vise principalement à estimer les prévalences de handicap et de dépendance parmi la population résidant à son domicile, au travers d'un questionnaire détaillé sur les déficiences physiques, les altérations fonctionnelles en résultant, les difficultés rencontrées dans la participation à la vie sociale, les aides reçues ou souhaitées. L'édition de 2008 comportait également un module sur l'état de santé général perçu dans l'ensemble de la population, concernée ou non par le handicap, dont la formulation répondait à une directive d'Eurostat.

Cette enquête, réalisée sur le territoire national en métropole et outre-mer, fournit des résultats France entière avec une précision suffisante, mais les estimateurs Horvitz-Thomson traditionnels sont difficilement utilisables sur des territoires restreints, en raison de coefficients de variation trop élevés à ces niveaux.

La mise en œuvre de la politique sociale nécessite cependant une information concernant des ensembles géographiques plus fins. Les lois de décentralisation ont en effet transféré aux Conseils généraux la gestion des dispositifs d'aide financière aux personnes handicapées et aux personnes âgées dépendantes. Ils ont en charge des budgets départementaux d'aide sociale en partie alimentés par les fonds recueillis par la Caisse Nationale de Soutien à l'Autonomie (CNSA). C'est pourquoi le dénombrement par département des populations souffrant de handicap et des personnes âgées dépendantes, susceptibles d'être éligibles aux mesures compensatrices, est apparu nécessaire : au niveau central comme au niveau local.

L'objet de cette étude a consisté à appliquer une méthode d'estimation sur « petit domaine », au moyen d'une modélisation des variables d'intérêt, pour estimer des prévalences de handicap ou de dépendance, dans la population des ménages, par département et région de France métropolitaine. La méthode utilisée mobilise l'information recueillie dans l'enquête de première phase et celle disponible dans la base de sondage pour réduire la variance de sondage de l'estimateur direct.

## 1. L'enquête Handicap-Santé auprès des Ménages

L'enquête a été réalisée en deux phases. La première phase avait pour but de repérer la population cible de l'étude, tandis que le questionnaire principal était renseigné en deuxième phase.

### 1.1. L'enquête de première phase : « vie quotidienne et santé »

Les personnes handicapées ou dépendantes, objet principal de l'enquête, constituant une population rare, l'enquête de première phase, dite « Vie Quotidienne et Santé » (VQS), a été réalisée auprès d'un très large échantillon : en métropole, environ 218100 personnes dans 94300 ménages ont répondu à un questionnaire auto-administré.

L'échantillon a été sélectionné selon la procédure classique des enquêtes de l'INSEE auprès des ménages, par un plan de sondage qu'on considérera à deux degrés pour simplifier. Le premier degré, aréolaire, était formé par l'intersection entre l'échantillon géographique appelé « échantillon maître » tiré à la suite du recensement de 1999, et l'échantillon de l'enquête annuelle de recensement de 2006. Au second degré étaient tirés les ménages dans la base de sondage formée des résidences principales recensées en 2006. Tous les membres d'un ménage échantillonné répondaient au questionnaire VQS.

Le questionnaire de première phase abordait la situation sanitaire dans l'ensemble de la population au travers de trois questions sur l'état de santé général, la présence d'une maladie chronique, les limitations d'activité ressenties. Les autres questions, destinées à déceler l'existence éventuelle d'un handicap, relevaient le niveau (nul, moyen, important) de difficulté ressentie dans trois domaines: sensoriel, moteur, intellectuel ou psychique, ainsi que le degré de dépendance à l'égard d'une aide extérieure, technique ou humaine. Le questionnaire, adressé à l'ensemble du ménage, se présentait sous forme d'une feuille recto-verso comprenant une colonne à renseigner pour chaque membre du ménage.

La collecte a eu lieu principalement par voie postale, avec relances par téléphone ou par enquêteur. Le protocole de collecte permettait à une personne du ménage de renseigner le questionnaire pour tous ses membres.

La formulation des questions et leur nombre limité font de cette première phase une enquête d'opinion sur la manière dont la personne perçoit son état de santé. Son exploitation directe ne suffit pas à une mesure exacte des prévalences de handicap. Cette enquête « filtre » avait pour objectif la constitution d'une base de sondage pour l'enquête principale de deuxième phase. Elle permettait de hiérarchiser les situations de santé individuelles afin de stratifier le tirage de deuxième phase.

### 1.2. L'enquête de deuxième phase : « Handicap-Santé des ménages »

La deuxième phase, dite « Handicap, santé des ménages » (HSM) est une enquête en face à face. Elle a recueilli les réponses d'environ 29900 personnes, dont 26400 en métropole. L'unité échantillonnée était l'individu. Les répondants à l'enquête VQS ont été sélectionnés selon un plan de sondage destiné à surreprésenter les personnes ayant déclaré les problèmes de santé les plus aigus.

L'ensemble des informations recueillies à l'enquête-filtre ont permis d'attribuer un score, sur une échelle de 1 à 100, à chaque individu répondant, d'autant plus élevé que son état de santé était présumé plus dégradé. Les personnes étaient ensuite regroupées en quatre classes de sévérité croissante de handicap présumé en fonction de leur score. Le niveau requis de score définissant chaque groupe dépend lui-même de l'âge de l'individu, comme indiqué dans le tableau 1.

Le tirage a été réalisé par sondage aléatoire simple dans chaque strate, avec un taux de sondage croissant du groupe 1 ( personnes présumées en bonne santé) au groupe 4 (personnes présumées les plus en difficulté), sous la contrainte de ne pas interroger plus de deux personnes dans un même ménage. L'enquête tendait à l'exhaustivité dans le groupe 4.

Selon la définition retenue par l'Organisation mondiale de la Santé, le handicap résulte de l'interaction entre un problème de santé, des facteurs environnementaux et des facteurs personnels, entraînant pour la personne des limitations dans ses activités et des restrictions dans sa participation à la vie collective.

**Tableau 1. Les strates de tirage de l'échantillon de deuxième phase**

	moins de 19 ans	20-59 ans	60-79 ans	80 ans ou plus	Groupe VQS	Taux de sondage en 2 <sup>ème</sup> phase (%)
<b>Score &lt;</b>	1	4	5	6	<b>I</b>	7,5
	10	12	25	40	<b>II</b>	21,0
	25	30	45	65	<b>III</b>	51,4
<b>Score &gt;=</b>	<b>25</b>	<b>30</b>	<b>45</b>	<b>65</b>	<b>IV</b>	97,0

Note de lecture : un individu répondant âgé de 40 ans et dont le score est compris entre 12 et 29, bornes incluses, est classé dans le groupe III. Un répondant de même âge dont le score atteint ou dépasse 30 est classé en groupe IV<sup>1</sup>. Le score s'accroît avec le nombre et la sévérité des problèmes de santé déclarés.

Le questionnaire, très détaillé, cerne les situations de handicap sous ces trois angles. Il relève l'état de santé fonctionnel de la personne (santé, déficiences de l'organisme, limitations fonctionnelles), décrit son environnement (aides techniques et humaines, logement, accessibilité,...) et retrace ses restrictions d'activité ou de participation sociale. L'ensemble des réponses permet notamment le classement de l'individu dans la nomenclature AGGIR spécifique aux personnes âgées de 60 ans ou plus.

Quatre départements métropolitains et deux départements d'outre-mer<sup>2</sup>, ont réalisé une extension d'échantillon par rapport au plan de sondage national, à chacune des deux phases. Dans le Nord, le Pas-de-Calais, le Rhône et les Hauts de Seine, les estimations directes fournies par l'enquête avaient une précision suffisante pour constituer des valeurs de référence. Ces départements ont ainsi pu servir de zones témoins pour valider les résultats obtenus par modélisation.

## 2. Les variables estimées par région, département

Les estimations départementales ont porté sur les prévalences, dans la population adulte de 20 ans ou plus, des handicaps par type, ainsi que sur celle des personnes âgées dépendantes. Dans le premier cas, les estimations ont été menées successivement dans la population de 20 à 59 ans et dans celle âgée de 60 ans ou plus. La dépendance n'est analysée que dans la population de 60 ans ou plus.

L'échantillon HSM comprenait 12602 personnes âgées de 20 à 59 ans en 2008 et 9165 personnes de 60 ans ou plus.

### 2.1. Les handicaps par type

La définition du handicap adoptée dans cette étude n'en retient que les dimensions physiologiques et fonctionnelles, non sa dimension sociale. Le handicap est ici défini par la présence simultanée d'une déficience d'un organe et d'une limitation fonctionnelle connexe sévère. Une personne est potentiellement handicapée lorsqu'elle souffre d'une altération physique et ne peut réaliser la fonction normalement assurée par l'organe déficient. Les déficiences relevées en clair dans le questionnaire ont été codifiées dans la nomenclature de l'enquête par l'Institut de recherche et de documentation en économie de la santé (IRDES).

Les formes de handicap sont regroupées en trois grands types génériques : sensoriels, moteurs, cognitifs, dont on a estimé les prévalences successivement dans la population de 20 à 59 ans et dans la population de 60 ans ou plus.

<sup>1</sup> Le tableau est repris du document de travail Insee n°F1101 : « Le volet Ménages de l'enquête Handicap-Santé, présentation, calcul des poids », Gérard Bouvier.

<sup>2</sup> Nord, Pas-de-Calais, Rhône, Hauts de Seine, Martinique et Guadeloupe.

**Les handicapés sensoriels** incluent les personnes souffrant de troubles visuels ou auditifs sévères. Une personne est considérée comme malvoyante lorsqu'elle déclare être aveugle totale (déficience 21) ou lorsqu'elle présente au moins l'une des déficiences suivantes :

- 22 - malvoyant (déficience visuelle grave, mais sans être aveugle)
- 23 - un œil ne voit rien
- 24 - difficulté à voir de près ou de loin, mais ni malvoyant, ni aveugle
- 25 - limitation du champ visuel
- 2901 à 2909, 2999 - autres problèmes visuels (perception des couleurs, fatigue visuelle...)

et déclare éprouver une forte limitation (modalité de réponse 3 ou 4 à la question) dans l'une des deux situations :

- lire clairement les caractères d'un journal (avec des lunettes éventuellement)
- voir clairement un visage à 4 mètres (avec des lunettes éventuellement)

Une personne est reconnue malentendante lorsqu'une surdité totale est constatée (déficience 31 de la nomenclature des déficiences) ou lorsqu'elle présente l'une des déficiences suivantes :

- 32 - malentendant
- 33 - surdité d'une oreille
- 34 - autre difficulté à entendre, mais ni malentendant, ni sourd
- 3901 à 3905 - autres déficiences auditives (douleurs, bruits parasite, altération du tympan...)

et déclare ressentir une forte limitation pour entendre une conversation.

**Le handicap moteur** est défini par la présence simultanée d'une déficience d'un organe nécessaire à la motricité et d'une limitation des possibilités de mobilité. On retient ici les personnes présentant au moins l'une des déficiences du groupe 1 de la nomenclature de l'enquête :

- 1100 - paralysie complète d'une ou plusieurs parties du corps
- 1200 - paralysie partielle d'une ou plusieurs parties du corps
- 1300 - amputation d'un membre
- 1400 - gêne importante dans les articulations
- 1500 - limitation de la force musculaire
- 1600 - mouvements incontrôlés ou involontaires
- 1700 - troubles de l'équilibre
- 1901 à 1904 - autres problèmes limitant les mouvements

et éprouvant des difficultés importantes à réaliser l'une des actions suivantes :

- marcher 500 mètres sur terrain plat
- monter et descendre un étage d'escalier
- lever le bras
- se servir des mains et des doigts
- se baisser, s'agenouiller
- porter un sac de 5 kilos.

**Les troubles cognitifs** rassemblent l'ensemble des problèmes de nature psychique, intellectuelle ou mentale. Les personnes concernées sont celles ayant déclaré une déficience relevant du groupe 5 de la nomenclature des déficiences de l'enquête :

- 51-troubles d'orientation dans le temps ou dans l'espace, non liés à des problèmes de vue
- 52-troubles de mémoire importants
- 53-troubles de l'humeur (découragements, démotivations)
- 54-troubles anxieux
- 55-difficultés de relations avec autrui
- 56-difficultés d'apprentissage
- 57-difficultés de compréhension
- 58-retard intellectuel
- 59-autres troubles intellectuels.

et avoir « souvent » des difficultés dans l'un au moins des domaines suivants :

- compréhension des autres
- capacité d'apprentissage
- résoudre les problèmes de la vie quotidienne
- perception du temps
- trous de mémoire importants
- propension à se mettre en danger
- agressivité
- concentration
- difficultés à nouer des relations.

Certains handicaps de communication par la parole s'accompagnent souvent d'un handicap intellectuel ou psychique. Pour cette raison, ils ont également été pris en compte, de façon contrôlée : un individu souffrant d'une déficience d'élocution (groupe 4 de la nomenclature des déficiences) est également inclus dans le groupe des personnes présentant un handicap intellectuel lorsqu'il déclare de grosses difficultés de compréhension des autres.

## **2.2. La dépendance des personnes âgées**

La perte d'autonomie due au vieillissement est évaluée au moyen d'un outil, la grille AGGIR, qui classe les personnes âgées dans six groupes iso-ressources (GIR) selon leurs facultés physiques et mentales (leur définition figure en annexe). La classification de 1 à 6 constitue une échelle d'autonomie croissante.

Dans l'enquête HSM, un algorithme utilisant l'ensemble des réponses au questionnaire a permis d'évaluer à quel groupe GIR pouvait être rattachée chaque personne interrogée âgée de 60 ans ou plus au moment de l'enquête. Le classement individuel obtenu résulte des réponses à un questionnaire statistique. Il pourrait donc s'écarter du diagnostic qu'aurait effectué une équipe de professionnels avec le protocole adapté d'un examen médical. Les effectifs de personnes âgées appartenant à chacun des groupes de la grille AGGIR sont donc ici doublement estimatifs : en raison de l'aléa propre au sondage, mais aussi d'un aléa sur l'observation individuelle.

Les estimations départementales et régionales ont porté sur la prévalence des personnes âgées de 60 ans ou plus appartenant à l'un des GIR 1 à 4, de façon indifférenciée.

### 3. La méthode d'estimation sur « petit domaine »

La méthode retenue repose sur une approche prédictive, avec modélisation explicite des comportements individuels.

#### 3.1. Principe et formules d'estimation

Elle utilise les informations recueillies en première phase VQS de l'enquête Handicap-Santé pour imputer, à l'aide d'un modèle, des valeurs du paramètre aux individus non réinterrogés en deuxième phase. L'extrapolation à la population totale de la région se fait alors à l'aide des pondérations de l'échantillon de première phase. L'estimateur final supporte une variance de sondage réduite par rapport à l'estimateur direct, la taille de l'échantillon de première phase étant beaucoup plus importante, mais la possibilité d'un biais dû au modèle, si celui-ci n'est pas suffisamment pertinent.

L'effectif total dans la population  $U$  des personnes affectées d'une maladie ou d'un handicap peut se décomposer entre l'échantillon  $S_2$  enquêté en deuxième phase et la population non enquêtée :

$$Y_U = \sum_{i \in U} Y_i = \sum_{k \in S_2} Y_k + \sum_{j \notin S_2} Y_j$$

Si tout l'échantillon de première phase était réenquêté en deuxième phase,  $Y$  pourrait être estimé par l'estimateur Horvitz Thomson muni des pondérations de l'échantillon VQS  $S_1$  :

$$\hat{Y}_{VQS,HT} = \sum_{i \in S_1} w_{1i} y_i = \sum_{k \in S_2} w_{1k} y_k + \sum_{j \in S_1 \cap \bar{S}_2} w_{1j} y_j$$

Le total départemental  $Y_d$  de la variable  $\mathcal{Y}$  est estimé en substituant, dans cette expression, aux valeurs réelles inconnues, les probabilités prédites par un modèle construit dans l'échantillon HSM de deuxième phase. Deux formes alternatives ont été calculées, dont la convergence constitue un test de bonne adéquation du modèle.

La première utilise les prédictions dans tout l'échantillon de première phase :

$$\hat{Y}_{1d} = \frac{1}{\hat{N}_d} \sum_{i \in U_d \cap S_1} w_{1i} \tilde{y}_i \quad (1)$$

La seconde conserve les valeurs observées dans l'échantillon de deuxième phase auxquelles elle ajoute les valeurs imputées par le modèle dans la partie non réenquêtée de l'échantillon de première phase :

$$\hat{Y}_{2d} = \frac{1}{\hat{N}_d} \left( \sum_{k \in U_d \cap S_2} w_{1k} y_k + \sum_{j \in U_d \cap S_1 \cap \bar{S}_2} w_{1j} \tilde{y}_j \right) \quad (2)$$

où :  $U_d$  est la population du département  $d$

$Y_k = 1 \Leftrightarrow$  l'individu  $k$  présente le handicap étudié

$Y_k = 0$  sinon.

Pour tout individu  $i$  répondant à VQS :

$\tilde{y}_i = \text{Pr} \hat{\text{ob}}\{Y_i = 1\}$  est la probabilité d'être handicapé prédite par le modèle (3)

$w_{1i}$  est le poids redressé de l'individu  $i$  dans l'enquête de 1<sup>ère</sup> phase VQS

$\hat{N}_d = \sum_{i \in U_d \cap S_1} w_{1i}$  est la taille estimée de la population dans le département  $d$ .

De nouvelles pondérations de l'échantillon VQS ont été calculées par calage sur les populations départementales au 1<sup>er</sup> janvier 2008, ventilées par sexe et âge croisés en quatre tranches : moins de 19 ans, 20 à 59 ans, 60 à 74 ans, 75 ans et plus.

Le modèle adopté est un modèle de régression logistique mixte, dans lequel le département de résidence est spécifié en effet aléatoire.

Pour tout individu  $i$  résidant dans le département  $d$ , on postule une relation de la forme :

$$\text{Prob}\{Y_{d,i} = 1\} = p_{d,i} = \frac{1}{1 + \exp[-(\mathbf{b}' \mathbf{x}_i + v_d)]}.$$

$\mathbf{x}_i$  est un vecteur de variables disponibles dans les deux échantillons  $S_1$  et  $S_2$ , observées dans l'enquête VQS ou présentes dans la base de sondage

$\mathbf{b}$  est le vecteur des coefficients de régression entre  $\mathcal{Y}$  et les facteurs fixes  $\mathcal{X}$

$v_d$  est un coefficient qui permet de prendre en compte une composante explicative du handicap spécifique au département  $d$  où réside l'individu  $i$ .

Le coefficient  $v_d$  résulte d'une décomposition du résidu  $\varepsilon_i$  du modèle de régression logistique classique :

$$z_{d,i} = \mathbf{b}' \mathbf{x}_i + \varepsilon_i,$$

où 
$$z_{d,i} = \text{Log} \left( \frac{p_{d,i}}{1 - p_{d,i}} \right),$$

en une somme de deux éléments :

$$\varepsilon_{d,i} = v_d + u_i$$

où  $v_d$  est un coefficient inconnu à estimer.

Cette formulation permet d'introduire, dans l'ensemble des effets inobservés non pris en compte par les effets fixes du modèle, l'influence éventuelle des différences de politiques locales, en la traitant comme un effet d'espérance nulle et de variance  $\sigma_v^2$ , conditionnellement aux cofacteurs  $X$ . La variance totale du résidu  $\varepsilon_i$  est dissociée en une variance associée au domaine  $d$  et une variance résiduelle de  $u_i$ .

Le modèle repose sur les hypothèses :

$$E(\varepsilon_i) = 0$$

$$\text{Var}(\varepsilon_i) = \sigma_v^2 + \sigma_u^2$$

$$\text{Cov}(\varepsilon_{d,i}, \varepsilon_{d,j}) = \sigma_v^2$$

$$\text{Cov}(\varepsilon_{d_1,i}, \varepsilon_{d_2,j}) = 0$$

Les coefficients  $\mathbf{b}$  et  $v_d$  du modèle sont estimés dans l'ensemble de l'échantillon national de deuxième phase HSM, hors départements d'outre-mer.

En construisant l'estimateur « petit domaine », on substitue à la phase de sondage HSM, donnant des résultats très instables dans de petites sous populations, l'imputation de valeurs individuelles à variance réduite, grâce à des coefficients de régression robustes, puisque estimés dans un très grand échantillon.

Rappelons par ailleurs que la variance totale de l'estimateur direct Horvitz-Thomson dans l'enquête HSM est celle d'un sondage en deux phases, par conséquent de la forme :

$$\text{Var}(\hat{Y}_{HT}) = \text{Var}\left[E(\hat{Y}_{HT} / \mathbf{s}_1)\right] + E\left[\text{Var}(\hat{Y}_{HT} / \mathbf{s}_1)\right].$$

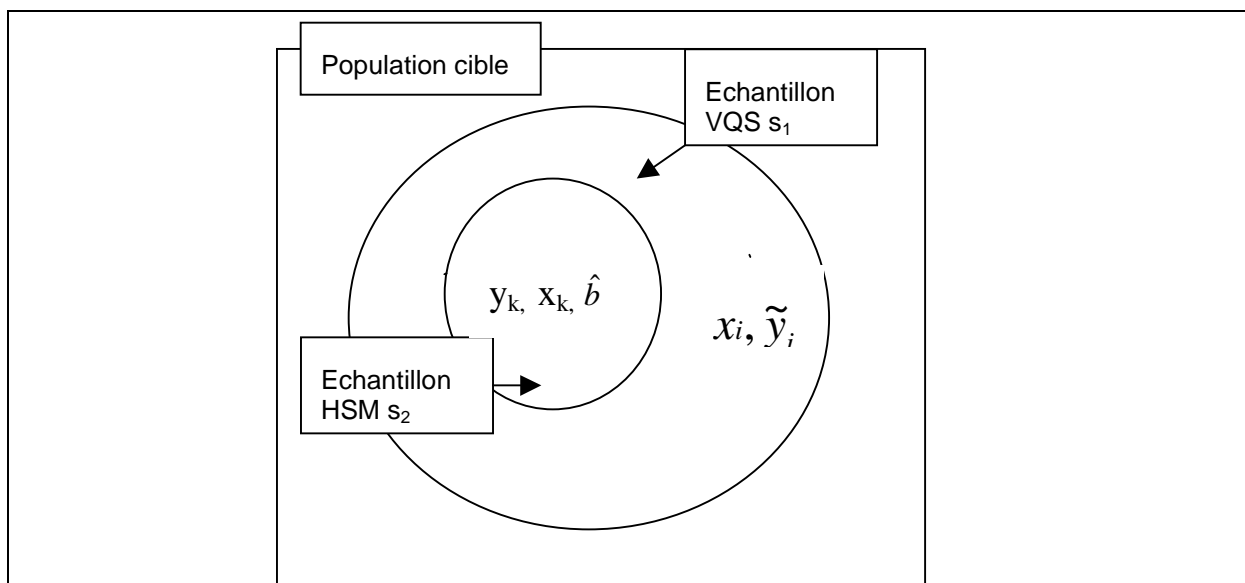
C'est la somme de deux éléments : le premier exprime la variance due à l'échantillonnage de première phase, tandis que le second représente la variance due à l'échantillonnage de deuxième phase, conditionnellement à l'échantillon de phase 1.

En imputant une valeur de la variable d'intérêt à chaque individu de l'échantillon de phase 1, on supprime toute variance de sondage propre à l'échantillonnage HSM conditionnel à l'échantillon VQS. Une grande partie de l'amélioration de la précision d'un estimateur départemental ou régional repose ici sur cette propriété.

Avec ses deux expressions alternatives (1) et (2), l'estimateur modélisé reste néanmoins dépendant de l'échantillonnage de première phase. Outre l'aléa créé par le modèle, il conserve une variance de sondage incompressible, égale à celle de l'estimateur Horvitz-Thomson d'un total ou d'une moyenne dans l'échantillon VQS.

La réduction attendue de la variance, grâce à la modélisation, sera donc d'autant plus importante, que la contribution de la deuxième phase HSM à la variance totale de l'estimateur direct est plus élevée. Celle-ci n'est pas uniforme sur le territoire, et varie selon le paramètre estimé. Le gain de précision ne pourra, par conséquent, être de même ampleur dans toutes les zones pour les sept indicateurs analysés ici.

**Figure 1. Schéma d'estimation**



*Lecture :  $y_k$  est la valeur observée pour l'individu  $k$  enquêté en deuxième phase dans HSM*

*$x_i$  est un vecteur de variables auxiliaires connues pour chaque individu  $i$  de l'échantillon*

*VQS  $S_1$ , par conséquent aussi pour chaque individu  $k$  de l'échantillon HSM  $S_2$*

*$\hat{b}$  est le vecteur des coefficients de régression d'une relation entre  $y_k$  et  $x_k$  estimée dans l'échantillon de deuxième phase  $S_2$*

*$\tilde{y}_i$  est la valeur imputée par le modèle à l'individu  $i$  répondant en première phase*

*VQS mais non enquêté en deuxième phase.*



### 3.2. L'information auxiliaire du modèle

La modélisation consiste à rechercher, au moyen d'une régression logistique, une relation entre la situation de handicap ou de dépendance de l'individu, observée dans l'enquête Handicap Santé, et l'ensemble des réponses de la personne à l'auto questionnaire simplifié de première phase VQS.

A ces premiers facteurs explicatifs, les plus déterminants, ont également été rajoutées les caractéristiques sociodémographiques de l'individu ou du ménage auquel il appartient, présentes dans la base de sondage ou dans les deux échantillons avec la même définition.

Les variables disponibles étaient les suivantes :

- sexe et âge, en 2008 et en année civile, de l'individu
- type d'espace de résidence (espace rural / unités urbaines de moins de 20000 habitants / unités urbaines de 20000 à moins de 100000 habitants / unités urbaines de 100000 habitants ou plus hors Paris / unité urbaine de Paris)
- type de ménage (personne vivant seule, ménage de plusieurs personnes)
- type d'habitat (individuel ou collectif)
- statut d'occupation du logement
- appartenance du logement au secteur HLM
- diplôme de la personne de référence du ménage
- type d'activité du chef de ménage (actif occupé, chômeur, inactif)
- état matrimonial de la personne de référence.

Le questionnaire VQS relevait le sexe et l'année de naissance de chaque personne interrogée en première phase. Ces deux variables peuvent donc être utilisées dans la modélisation.

La taille du ménage a été calculée par dénombrement des personnes ayant répondu en première phase.

L'espace de résidence est codifié dans la typologie des communes de l'INSEE en « zonage en aires urbaines et espace rural » (ZAUER).

Toutes les autres variables proviennent de la base de sondage, c'est-à-dire de l'enquête annuelle de recensement de 2006. Elles souffrent donc d'un décalage de deux ans par rapport à l'enquête principale.

Les coefficients de régression sont estimés dans l'échantillon national métropolitain HSM. La méthode repose sur l'hypothèse suivante :

- il existe une relation entre les réponses au questionnaire de première phase et la situation de l'individu observée dans l'enquête de deuxième phase
- cette relation est homogène sur tout le territoire métropolitain, et n'est pas modifiée lorsqu'on se restreint à la sous population enquêtée en phase HSM.

Les départements d'Outre-mer ont été exclus du champs de l'étude pour deux raisons. La Martinique et la Guadeloupe ayant réalisé des extensions d'échantillon, il leur était possible d'exploiter directement l'enquête HSM sans modélisation. Par ailleurs, la spécificité des problèmes de santé dans les DOM rendrait peu opératoire dans ces territoires un modèle fondé sur des observations réalisées majoritairement en métropole. Les résultats obtenus ici ne sont donc pas susceptibles d'une application aux départements et territoires ultra-marins.

La recherche des facteurs explicatifs a été effectuée avec la procédure GLMSELECT du logiciel SAS (version 9.2), au moyen d'une procédure itérative de type « stepwise », et du critère de sélection bayésien de Schwartz (SBC). Les observations sont pondérées par leur poids de sondage redressé dans l'enquête HSM, la procédure se chargeant de normer automatiquement les pondérations.

Les coefficients de régression sont ensuite estimés par modèle logistique mixte, dans lequel le niveau géographique d'estimation est déclaré en effet aléatoire. Un effet local résiduel, non pris en compte par les facteurs explicatifs retenus, est donc pris en charge par le modèle. La régression est réalisée

au moyen de la procédure GLIMMIX du logiciel SAS, pondérée par le poids de sondage redressé de première phase.

Chaque paramètre à estimer a été modélisé de façon spécifique.

Par souci d'homogénéité, les sous populations champs de l'estimation ont été définies par l'année de naissance déclarée en première phase VQS. L'échantillon HSM comprenait 12602 individus de 20 à 59 ans et 9165 personnes de 60 ans ou plus, âge atteint dans l'année civile 2008, d'après ce critère.

Pour l'estimation du modèle, on a imposé au domaine, spécifié en effet aléatoire, de comprendre au moins 1000 personnes répondantes dans l'échantillon de première phase. Des départements ont dû être regroupés pour satisfaire cette condition, sous la contrainte du respect des limites administratives régionales. La typologie départementale établie par un groupe de travail de la DREES [11] a été utilisée pour réunir les départements les plus proches sur le plan sanitaire, à l'intérieur de chaque région. Après regroupement, on a obtenu 75 entités géographiques.

Hors la Corse et le Limousin, où 654 et 1886 personnes ont respectivement répondu à l'enquête filtre, tous les échantillons régionaux de première phase comprenaient au moins 4000 individus.

Ce modèle a été préféré à l'utilisation de la seule strate de tirage VQS croisée à des critères démographiques, à l'instar du modèle mis en œuvre dans l'enquête HID de 1999-2000 [6], afin de mettre à profit l'information apportée par le questionnaire directement attribuable à chaque type spécifique de handicap. L'usage direct de la strate de deuxième phase en facteur explicatif dans un modèle de régression, ou en critère de post-stratification, est efficace lorsque la stratification VQS laisse peu d'individus handicapés ou dépendants dans les strates 1 et 2. C'était moins le cas avec les limitations sensorielles, motrices ou cognitives.

### 3.3. Les instruments de validation des résultats

On ne dispose pas d'estimateur sans biais de l'erreur quadratique moyenne pour un estimateur modélisé selon les méthodes (1) ou (2). Celle-ci est la somme de la variance et du carré du biais, par construction inconnu :

$$EQM(\hat{Y}) = E[\hat{Y} - E(\hat{Y})]^2 + [E(\hat{Y}) - Y]^2 = V(\hat{Y}) + \text{biais}^2(\hat{Y})$$

La pertinence des résultats est par conséquent recherchée par des moyens indirects.

En premier lieu, les quatre départements métropolitains ayant financé des extensions d'échantillon ont servi de zones-témoins. L'estimation directe y était suffisamment robuste pour être prise en référence et comparée à l'estimation résultant du modèle.

En second lieu, l'adéquation du modèle, donc l'ampleur du biais, est appréciée par la convergence entre la somme des valeurs individuelles prédites et l'estimation directe sur des ensembles agrégés dans lesquels celles-ci est jugée fiable. L'estimation nationale d'une prévalence, telle qu'elle ressort de l'exploitation directe de l'enquête HSM, constitue ici la référence à la somme des prédictions individuelles dans l'échantillon métropolitain. Un écart important entre les deux estimations serait l'indice d'un biais systématique induit par le modèle. Les grandes régions, bénéficiant en phase 2 d'une taille d'échantillon plus importante, constituent également des références utiles.

Cette convergence des estimations directes et modélisées a également été vérifiée en calculant l'intervalle de confiance à 95 % de l'estimateur direct Horvitz-Thomson, au niveau national et par région. Le recouvrement de l'estimateur modélisé par cet intervalle constitue une condition minimale d'absence de biais systématique.

Enfin, la régression des valeurs locales prédites sur les estimations locales directes de l'enquête permettait par ailleurs d'apprécier la proximité de la distribution des départements, avec l'un ou l'autre des estimateurs. Dans le nuage des points formé des départements ayant en coordonnées l'estimation directe de l'enquête  $\hat{Y}_{d,HT}$  et l'estimation « petit domaine »  $\tilde{Y}_d$ , la droite de régression des moindres carrés ordinaires d'équation  $\tilde{Y}_d = a \times \hat{Y}_{d,HT} + b$  est comparée à la bissectrice des axes d'équation  $y = x$ . En l'absence d'un décalage trop marqué entre les deux, on peut conclure à une absence de biais significatif de l'estimateur modélisé.

Comme il a été évoqué plus haut (section 3.1), l'estimateur « petit domaine » conserve une variance de sondage due à l'échantillonnage de première phase, à laquelle s'ajoute également un aléa de

sondage du coefficient de régression, lui-même estimé dans un échantillon. Le gain de précision de ce nouvel estimateur repose notamment sur l'importance de la variance de seconde phase dans la variance totale de l'estimateur direct. La complexité du plan de sondage de l'enquête VQS, lui-même à plusieurs degrés et faisant appel à la seule partie de l'échantillon géographique de l'INSEE<sup>3</sup> recensée en 2006, a pour conséquence des « effets de plan » variables d'un département à l'autre, même à taille d'échantillon équivalente. La structure de l'échantillon-maître rend en particulier plus difficile l'estimation dans les zones moins densément urbanisées. Il s'ensuit, pour cette enquête, une répartition de la variance totale de l'estimateur direct entre les deux phases très variable sur le territoire.

## 4. Les résultats

### 4.1. Les relations économétriques

Les situations de handicap par nature font l'objet de questions spécifiques dans l'enquête de première phase (voir annexe 1). Les questions 6 à 9 de l'enquête VQS sur les problèmes de vision et d'ouïe permettent de cerner les difficultés d'ordre sensoriel. Les questions 10 à 13 du questionnaire VQS relèvent les difficultés de mobilité ressenties. Les questions 14 à 16 et 19 renseignent sur l'état psychique et mental de l'individu.

On a donc une corrélation forte entre la réponse à ces questions en première phase et la situation relevée dans l'enquête HSM dans le domaine correspondant. A ces facteurs peuvent s'ajouter les réponses aux autres questions VQS, moins spécifiques d'un handicap particulier, concernant le recours à une aide technique ou humaine ou les restrictions d'activité ressenties.

Aucune variable du questionnaire VQS ne peut être considérée spécifique de la dépendance des personnes âgées. La probabilité de classement dans l'un des GIR 1 à 4 à l'enquête HSM est sensible à un ensemble de facteurs exprimant le besoin d'une aide auxiliaire dans les actes de la vie quotidienne, les limitations ressenties dans ses activités, une perte d'autonomie intellectuelle.

Dans tous les cas, la probabilité d'un état de santé dégradé s'accroît avec l'âge. Cette progression n'est pas toujours linéaire, et procède par paliers pour certains des indicateurs à estimer. Lorsque c'était le cas, l'âge a été spécifié par des fonctions affines dans chacune des tranches d'âge significatives.

Si l'on observe, dans l'enquête HSM, une corrélation entre le milieu social et l'état de santé ou le risque de handicap, les variables disponibles dans la base de sondage, susceptibles de constituer des « marqueurs » sociaux, apparaissent rarement dans les modèles retenus.

D'abord, l'on doit déplorer l'absence de la catégorie sociale de l'individu dans les facteurs explicatifs à disposition pour cette étude. Elle n'était pas demandée dans l'enquête de première phase, en raison du protocole de collecte. Dans les nouvelles procédures de recensement, la catégorie sociale n'est pas codifiée sur l'intégralité de la population des communes de moins de 10000 habitants. Cette variable n'est donc pas présente dans la base de sondage.

Par ailleurs, la situation de santé ressentie, relevée en première phase et utilisée en facteur explicatif, inclut déjà l'influence du milieu social sur la morbidité.

Quel que soit l'indicateur estimé, la stratification conditionnelle aux réponses de première phase a discriminé efficacement les personnes selon leur état de santé. Les prévalences de handicap ou de dépendance dans chaque groupe VQS sont croissantes de la strate 1 à la strate 4. Néanmoins, la strate 4 a mieux isolé les personnes âgées dépendantes que les individus souffrant d'altérations sensorielles, motrices ou cognitives (tableau 2).

---

<sup>3</sup> Appelé « échantillon-maître », il s'agit d'un échantillon de grappes de territoire sélectionnées pour une durée intercensitaire, dans lesquelles sont tirés les échantillons de logements des enquêtes ménages, afin de limiter les déplacements d'enquêteurs. Dans l'échantillon-maître tiré en 1999, utilisé pour l'enquête VQS, les unités échantillonnées sont des cantons en milieu rural, des unités urbaines, en milieu urbain hors grandes métropoles, des îlots d'agglomération, dans les unités urbaines de plus de 100000 habitants. Les unités de l'échantillon-maître ont été sélectionnées avec des probabilités proportionnelles à leur population.

**Tableau 2. Répartition par strate VQS des personnes concernées par le handicap ou la dépendance dans l'échantillon HSM non pondéré (%)**

Obs	variable	strate1	strate2	strate3	strate4	Total
1	limitations sensorielles - 20-59 ans	8.2	10.0	26.5	55.3	100
2	limitations sensorielles - 60 ans et plus	3.5	26.1	32.1	38.3	100
3	limitations motrices - 20-59 ans	2.7	5.3	25.9	66.1	100
4	limitations motrices - 60 ans et plus	1.7	23.3	35.2	39.9	100
5	troubles cognitifs - 20-59 ans	6.3	6.5	22.1	65.1	100
6	troubles cognitifs - 60 ans et plus	1.9	15.2	29.8	53.1	100
7	GIR 1 à 4	0.3	6.4	26.8	66.5	100

Lecture : 8,2 % des personnes de 20 à 59 ans souffrant de limitations sensorielles appartiennent à la strate VQS n°1, 55,3 % à la strate VQS n°4.

Il en résulte une relation entre réponses à VQS et situation de dépendance d'après HSM différente selon les strates. Pour cette raison, deux modèles ont été construits pour estimer l'appartenance aux GIR 1 à 4. L'un dans le sous échantillon HSM constitué des strates de tirage 1 à 3 d'une part, le second dans le sous échantillon HSM constitué de la strate 4 d'autre part. L'estimation régionale est alors la somme des valeurs prédites par chacun des modèles dans les deux sous échantillons, ce qui a permis de limiter le risque de biais de surestimation.

La structure des modèles retenus, pour chaque indicateur de handicap estimé, est résumée dans le tableau 3.

#### 4.2. Les estimations localisées

Les deux modes de calcul  $\hat{Y}_1$  et  $\hat{Y}_2$  présentés en section 3.1 donnent des résultats presque identiques. Aux niveaux national et régional, comme dans les quatre départements avec extension, les proportions estimées selon les formules (1) ou (2) s'écartent rarement de plus de deux dixièmes. Le nuage de points des départements ayant en coordonnées la prévalence prédite par la méthode 1 et celle prédite par la méthode 2 s'aligne le long de la bissectrice des axes. La droite de régression de l'une sur l'autre a une pente au moins égale à 0,90, quel que soit le paramètre estimé.

La somme des probabilités individuelles prédites fournit, France hors outre-mer, des prévalences qui ne s'écartent pas de plus de 4 % de celles estimées directement par l'enquête HSM avec les pondérations de sondage redressées (tableau 4). Pour les sept indicateurs analysés, l'intervalle de confiance à 95 % de l'estimateur Horvitz-Thomson recouvre l'estimation modélisée, indice d'une absence de biais systématique important.

L'écart relatif entre estimation directe et prédiction du modèle reste compris entre -10 % et +10 % dans la moitié au moins des régions (tableau 5), entre -20 % et +20 % dans 17 des 22 régions. La différence absolue entre la proportion estimée par l'enquête et celle prédite par le modèle reste inférieure à 1 dans au moins la moitié des régions pour tous les indicateurs calculés, à 2 dans 80 % des régions (tableau 6).

Outre la Corse, les régions dans lesquelles le modèle réduit ou accroît l'estimation directe de plus de 20 % sont les petites régions, ayant par conséquent les plus faibles tailles d'échantillon, qui obèrent fortement la précision de l'estimateur direct.

Dans trois des quatre départements témoins ayant réalisé une extension d'échantillon, l'écart relatif entre l'estimation directe et la valeur prédite ne dépasse pas 11 %. L'adéquation du modèle est plus problématique dans le département des Hauts de Seine, où l'on constate les écarts relatifs les plus importants par rapport à l'enquête. On peut imputer cette singularité au plan de sondage de deuxième phase de l'extension des Hauts de Seine, qui s'est différenciée du plan national par un sous échantillonnage dans la strate 1 au profit d'une saturation de la strate 3, conditionnellement à l'échantillon VQS de première phase. Il en résulte une surpondération des individus de la strate 1, source de perturbation lorsque certains d'entre eux s'avèrent être handicapés d'après l'enquête HSM, malgré une probabilité nulle selon leurs réponses en première phase.

La prédiction du modèle diffère de l'estimation directe, en hausse ou en baisse, de moins de 10 % dans la moitié des départements (tableau 7). Dans un quart des départements, cet écart relatif dépasse 20 % en hausse ou en baisse.

Le nuage des départements ayant en coordonnées la prévalence estimée directement dans l'enquête HSM et celle prédite par le modèle s'étire linéairement le long de la bissectrice des axes. La hiérarchie des départements selon l'indicateur étudié est à peu près respectée par l'estimateur modélisé. La pente de la droite de régression de la valeur prédite sur l'estimation directe se situe entre 0,79 et 0,87 selon la variable estimée.

Comme souvent dans les estimations de type « petits domaines », on observe un phénomène de « shrinkage ». La méthode tend à réduire la dispersion entre « domaines », par rapport aux estimateurs Horvitz-Thomson, en rapprochant de la moyenne nationale une partie des régions ou départements situés aux extrêmes de la distribution initiale issue de l'exploitation directe de l'enquête HSM.

## Bibliographie

- [1] P. Ardilly. Panorama des principales méthodes d'estimation sur petits domaines : aspects théoriques.- Polycopié, ENSAI, 2006.
- [2] G. Bouvier. Le volet Ménages de l'enquête Handicap-Santé, présentation, calcul des poids. Document de travail n°F1101, INSEE 2011.
- [3] G. Bouvier. L'enquête Handicap-Santé, présentation générale. Document de travail n°F1109, INSEE 2011.
- [4] G. Bouvier. L'approche du handicap par les limitations fonctionnelles et la restriction globale d'activité chez les adultes de 20 à 59 ans. In : France, portrait social, INSEE 2009.
- [5] A.E.Clark, A.Vicard. Conditions de collecte et santé auto-déclarée : une analyse sur données européennes. Economie et Statistique n°403-404, 2007.
- [6] C. Couet. Estimations locales sur les personnes handicapées vivant en domicile ordinaire, enquête HID 1999, résultats détaillés. INSEE-Résultats Société n°12-2002.
- [7] L. Davezies. Modèles à effets fixes, à effets aléatoires, modèles mixtes ou multi-niveaux : propriétés et mises en œuvre des modélisations de l'hétérogénéité dans le cas de données groupées. Document de travail n°G2011/03, INSEE 2011.
- [8] L. Midy. Présentation de l'enquête « Vie quotidienne et Santé » de 2007. Document de travail n°F1001, INSEE 2010.
- [9] L. Midy. Enquête « Vie quotidienne et Santé » : limitations dans les activités et sentiment de handicap ne vont pas forcément de pair. INSEE-Première n°1254, août 2009.
- [10] K. Moussalam. Approximation de la variance de l'enquête HSM 2008. Note interne, INSEE, mars 2010.
- [11] P.Gauthier, M.Mansuy. Rapport du groupe de travail « indicateurs sociaux départementaux », CNIS, décembre 2009.
- [12] Techniques d'enquête, volume 29, n°1, juin 2003 : numéro spécial de la revue de Statistique Canada consacré à l'estimation sur petits domaines.
- [13] J.Le Guennec. Estimations locales du handicap dans l'enquête Handicap-Santé 2008, Document de travail INSEE à paraître.

**Tableau 3. Les modèles de régression logistique**

Paramètre à estimer	Champ d'observation	Effectif tel que Y=1 dans l'échantillon HSM	Effets fixes explicatifs	Coefficients de régression	t-value
Limitations sensorielles	20 à 59 ans	1006	intercept	-6.8021	-74.09
			v_auditif - difficultés légères	1.3611	461.90
			difficultés fortes	3.2444	872.92
			v_vueloin – difficultés légères	0.3391	65.74
			difficultés fortes	0.8163	107.23
			v_vuepres - difficultés légères	0.0886	25.32
			difficultés fortes	1.2231	242.78
			v_impossible - oui	0.5547	123.86
			ne sait pas	0.0073	2.70
			v_reconnai - oui	0.7334	218.48
v_age	0.0558	459.25			
Limitations sensorielles	60 ans et plus	2003	intercept	-3.5719	-50.65
			v_auditif - difficultés légères	1.0988	497.70
			difficultés fortes	2.7581	972.54
			v_vueloin – difficultés légères	0.3168	97.99
			difficultés fortes	1.4036	252.39
			v_vuepres - difficultés légères	0.4210	173.80
			difficultés fortes	1.0539	273.56
			v_impossible - oui	0.3993	134.90
			ne sait pas	0.1732	69.84
			v_aidetech - oui	0.3563	145.14
			fag1 <sup>4</sup>	0.0266	56.29
			fag2	0.0411	220.49
			hag1	-0.0232	-55.89
hag2	0.0550	250.51			
Limitations motrices	20 à 59 ans	2247	intercept	-4.7670	-83.07
			v_marchram_12 <sup>5</sup>	0.8415	225.29
			v_marchram_13	1.4496	214.60
			v_marchram_21	0.8492	190.31
			v_marchram_22	1.2165	296.01
			v_marchram_23	1.7381	277.61
			v_marchram_31	1.5816	231.46
			v_marchram_32	1.8236	301.21
			v_marchram_33	2.3335	448.69
			v_main_2	0.5872	171.28
			v_main_3	0.5896	114.55
			v_attrape_2	0.4767	146.31
			v_attrape_3	1.0523	230.91
			v_limaid_12 <sup>6</sup>	0.9687	103.37
			v_limaid_21	0.9047	280.25
			v_limaid_22	1.2382	210.18
			v_limaid_31	1.0998	269.14
			v_limaid_32	1.5837	275.81
			v_handrec_12 <sup>7</sup>	1.0850	196.72
			v_handrec_21	0.8696	241.92
			v_handrec_22	1.4212	397.34
			fag2 <sup>8</sup>	0.0614	339.98
			hag2	0.0062	29.70

$${}^4 \text{ fag1} = (\text{age} - 70) \times 1_{60 \leq \text{age} \leq 69} \times 1_{\text{sexe} = \text{femme}} \quad \text{fag2} = (\text{age} - 70) \times 1_{70 \leq \text{age}} \times 1_{\text{sexe} = \text{femme}}$$

$$\text{hag1} = (\text{age} - 70) \times 1_{60 \leq \text{age} \leq 69} \times 1_{\text{sexe} = \text{homme}} \quad \text{hag2} = (\text{age} - 70) \times 1_{70 \leq \text{age}} \times 1_{\text{sexe} = \text{homme}}$$

<sup>5</sup> interaction : marche \* ramasse. 1=pas de difficulté, 2=quelques difficultés, 3=beaucoup de difficultés

<sup>6</sup> Interaction : limitation d'activité\*aide technique (appareillage, prothèse...). Modalités aide technique : 1=non, 2=oui.

Paramètre à estimer	Champ d'observation	Effectif tel que Y=1 dans l'échantillon HSM	Effets fixes explicatifs	Coefficients de régression	t-value
Limitations motrices	60 ans et plus	4158	Intercept	-6.0176	-93.59
			v_marchram_12	0.6436	214.80
			v_marchram_13	1.8700	285.91
			v_marchram_21	0.8924	287.05
			v_marchram_22	1.1816	431.92
			v_marchram_23	1.9124	432.94
			v_marchram_31	1.6555	313.43
			v_marchram_32	1.9212	498.89
			v_marchram_33	2.2657	576.61
			v_attrape - difficultés légères	0.5587	273.74
			v_attrape - difficultés fortes	0.6954	212.88
			v_limaid_12 <sup>9</sup>	1.5382	241.30
			v_limaid_21	0.8800	375.74
			v_limaid_22	1.1043	313.81
			v_limaid_23	1.5267	182.46
			v_limaid_31	1.0930	312.34
			v_limaid_32	1.2004	285.79
			v_limaid_33	1.5285	310.33
			v_handicap - oui	0.5932	284.36
			fage <sup>10</sup>	0.0421	402.91
hage	0.0360	340.78			
Troubles cognitifs	20 à 59 ans	1717	Intercept	-3.9731	-51.41
			v_initiative - difficultés légères	0.5817	162.90
			v_initiative - difficultés fortes	0.8085	134.76
			v_reconnai - oui	0.9102	299.99
			v_concentra - difficultés légères	0.6094	168.42
			v_concentra - difficultés fortes	0.9805	168.59
			v_aidepers- un peu	0.3081	70.29
			v_aidepers - difficultés fortes	0.6570	117.94
			v_memoire- difficultés légères	0.6054	214.90
			v_memoire - difficultés fortes	0.8759	163.61
			v_limitat- difficultés légères	0.7993	303.76
			v_limitat - difficultés fortes	1.0263	308.92
			v_prob_quot- difficultés légères	0.2515	64.38
			v_prob_quot - difficultés fortes	0.6986	117.28
			v_comprendre- difficultés légères	0.4848	126.46
v_comprendre - difficultés fortes	0.6018	88.63			
Troubles cognitifs	60 ans et plus	1653	Intercept	-3.5933	-49.37
			v_concentra - difficultés légères	0.6084	201.65
			v_concentra - difficultés fortes	1.1509	232.21
			v_impossible - oui	0.8342	262.74
			v_initiative - difficultés légères	0.5379	161.68
			v_initiative - difficultés fortes	0.7679	146.13
			v_memoire - difficultés légères	0.2693	97.04
			v_memoire - difficultés fortes	0.8756	195.95
			v_prob_quot - difficultés légères	0.4251	125.02
			v_prob_quot - difficultés fortes	0.6671	127.90
			v_aidepers - un peu	0.4622	147.51
			v_aidepers - beaucoup	0.7197	179.76
			fag1 <sup>11</sup>	-0.0119	-24.32
			fag2	0.0051	22.72
			hag1	0.0248	46.21
hag2	0.0399	158.94			

<sup>7</sup> Interaction : se sentir handicapé \* avoir une reconnaissance administrative de handicap

<sup>8</sup>  $fag2 = (age-40) \times 1_{40 \leq \text{âge} \leq 59} \times 1_{\text{sexe}=femme}$      $hag2 = (age-40) \times 1_{40 \leq \text{âge} \leq 59} \times 1_{\text{sexe}=homme}$

<sup>9</sup> interaction : limitation d'activité \* aide d'un tiers dans la vie quotidienne

<sup>10</sup>  $Fage = (\text{âge} - 20) * (\text{sexe}=femme)$      $Hage = (\text{âge} - 20) * (\text{sexe}=homme)$



Paramètre à estimer	Champ d'observation	Effectif tel que Y=1 dans l'échantillon HSM	Effets fixes explicatifs	Coefficient de régression	t-value
Appartenance aux GIR 1 à 4	60 ans et plus Strates VQS 1 à 3	519	Intercept	-5.5832	-32.29
			v_sorties - difficultés légères	0.8031	170.49
			v_sorties - difficultés fortes	1.0758	189.43
			v_aidepers - un peu	0.7769	171.88
			v_aidepers - beaucoup	1.3982	226.68
			v_limitat - difficultés légères	0.6874	119.22
			v_limitat - difficultés fortes	1.4452	227.68
			v_initiative - difficultés légères	0.0818	17.94
			v_initiative - difficultés fortes	1.0126	162.31
			v_amenalog - difficultés légères	0.4930	88.94
			v_aidetech - oui	0.5078	123.71
			v_main - difficultés légères	0.1957	45.70
			v_main - difficultés fortes	0.8182	143.48
			vag1 <sup>12</sup>	0.0696	57.82
vag2	0.0495	67.51			
vag3	0.0522	98.53			
Appartenance aux GIR 1 à 4	60 ans et plus Strate VQS 4	1029	Intercept	-2.5920	-32.54
			v_sorties - difficultés légères	0.1648	17.07
			v_sorties - difficultés fortes	1.0329	117.03
			v_aidepers - un peu	0.2162	23.91
			v_aidepers - beaucoup	1.1853	143.45
			v_main - difficultés légères	0.5933	76.04
			v_main - difficultés fortes	1.0202	135.17
			v_concentra - difficultés légères	0.3526	48.42
			v_concentra - difficultés fortes	0.7320	102.28
			v_reconnai - oui	0.6263	102.25
			vag1 <sup>13</sup>	0.0106	22.82
			vag2	0.0942	67.48

<sup>11</sup> Voir note 5.

<sup>12</sup>  $Vag1 = (age-70) * (60 \leq age \leq 69)$      $vag2 = (age-70) * (70 \leq age \leq 79) + (80-70) * (age \geq 80)$      $vag3 = (age-80) * (age \geq 80)$

<sup>13</sup>  $Vag1 = (age-85) * (60 \leq age \leq 84)$      $vag2 = (age-85) * (age \geq 85)$

**Liste des variables**

Limitat	limitations d'activité ressenties
Aidepers	aide régulière d'un tiers pour les actes de la vie quotidienne
Aidetech	utilisation d'un appareillage, de prothèses ou autre aide technique
Amenalog	Aménagements du logement pour des raisons de santé
Marche	difficultés à marcher plus de 500 mètres ou à monter un escalier
Ramasse	difficulté à ramasser un objet par terre
Attrape	difficulté à attraper un objet en hauteur
Main	difficulté à se servir de ses mains
Auditif	difficulté à entendre ce qui se dit dans une conversation entre plusieurs personnes
Vueloin	difficulté à voir de loin
Vuepres	difficulté à voir de près
Comprendre	difficulté à comprendre les autres ou à se faire comprendre des autres
Initiative	difficulté à prendre des initiatives
Concentra	difficulté à se concentrer
Prob_quot	difficulté à résoudre les problèmes de la vie quotidienne
Memoire	difficulté à se souvenir
Impossible	impossibilité totale à réaliser l'une des activités essentielles citées (voir, entendre, parler, bouger)
Sorties	difficulté à sortir seul de chez soi
Handicap	handicap ressenti (non/oui)
Reconnai	reconnaissance d'un handicap par l'administration

Tableau 4. Prévalences en France de métropole : estimations directes et valeurs prédites

variable	borne inf. interv. conf. à 95 %	estimateur direct (0)	borne sup. interv. conf. à 95 %	Prévalence prédite par le modèle méthode 1 (1)	Prévalence prédite par le modèle méthode 2 (2)	Ecart absolu (2)-(0)	Ecart relatif (% (2)/(0)
<b>Population de 20 à 59 ans</b>							
limitations motrices	4.1	4.4	4.8	4.3	4.3	-0.1	-3.2
limitations sensorielles	2.7	3.1	3.5	3.0	3.1	0.0	-0.2
troubles cognitifs	4.0	4.5	5.0	4.4	4.5	0.0	0.2
<b>Population de 60 ans ou plus</b>							
limitations motrices	23.2	24.4	25.6	24.7	24.8	0.4	1.6
limitations sensorielles	12.5	13.5	14.4	13.3	13.3	-0.1	-1.1
troubles cognitifs	8.3	9.1	9.9	9.2	9.4	0.3	3.1
GIR 1 à 4	6.0	6.5	7.1	6.8	6.8	0.3	4.2

Tableau 5. Quantiles des écarts relatifs entre prévalences régionales estimées par l'enquête et prédites par le modèle (méthode 2)

Variable	minimum	centile 10	1er quartile	médiane	3ème quartile	centile 90	maximum
<b>Population de 20 à 59 ans</b>							
limitations motrices	-34.8	-15.0	-10.1	-3.3	3.6	9.4	32.1
limitations sensorielles	-47.2	-16.2	-4.5	0.4	5.2	20.1	41.0
limitations cognitives	-28.8	-10.1	-5.5	-1.6	3.9	25.8	38.8
<b>Population de 60 ans et plus</b>							
limitations motrices	-18.7	-7.4	-2.3	1.4	5.5	14.4	23.6
limitations sensorielles	-31.8	-9.8	-3.1	1.4	7.5	18.6	33.8
limitations cognitives	-15.4	-11.8	-0.9	5.3	9.3	13.2	58.4
GIR 1 à 4	-25.7	-13.1	-3.9	2.8	16.4	27.9	34.4

Tableau 6. Quantiles des écarts absolus entre prévalences régionales estimées par l'enquête et prédites par le modèle (méthode 2)

Variable	minimum	centile 10	1er quartile	médiane	3ème quartile	centile 90	maximum
<b>Population de 20 à 59 ans</b>							
limitations motrices	-2.6	-0.8	-0.5	-0.1	0.2	0.3	0.9
limitations sensorielles	-2.4	-0.6	-0.1	0.0	0.2	0.4	1.4
limitations cognitives	-1.4	-0.5	-0.2	-0.1	0.2	1.5	2.4
<b>Population de 60 ans et plus</b>							
limitations motrices	-4.2	-2.1	-0.7	0.3	1.3	3.0	4.9
limitations sensorielles	-3.3	-2.2	-0.5	0.2	0.9	1.4	4.4
limitations cognitives	-1.8	-1.4	-0.1	0.3	0.8	1.2	2.3
GIR 1 à 4	-2.7	-0.9	-0.3	0.2	1.1	1.2	1.8

**Tableau 7. Quantiles des écarts relatifs entre prévalences départementales estimées par l'enquête et prédites par le modèle (méthode 2)**

Variable	minimum	centile 10	1er quartile	médiane	3ème quartile	centile 90	maximum
<b>Population de 20 à 59 ans</b>							
limitations motrices	-34.8	-18.1	-12.7	-1.9	8.2	23.0	53.9
limitations sensorielles	-47.2	-15.9	-5.7	5.4	16.1	37.7	107.7
limitations cognitives	-28.9	-15.5	-6.5	1.2	12.0	27.1	83.6
<b>Population de 60 ans et plus</b>							
limitations motrices	-22.0	-8.9	-4.2	1.8	14.2	24.3	37.7
limitations sensorielles	-21.9	-13.9	-6.7	2.5	12.0	25.5	92.7
limitations cognitives	-20.0	-7.9	-1.6	6.4	19.4	34.2	63.4
<b>GIR14</b>	-46.0	-18.2	-7.7	5.2	16.2	37.0	71.4

**Tableau 8. Quantiles des écarts absolus entre prévalences départementales estimées par l'enquête et prédites par le modèle (méthode 2)**

Variable	minimum	centile 10	1er quartile	médiane	3ème quartile	centile 90	maximum
<b>Population de 20 à 59 ans</b>							
limitations motrices	-2.9	-1.2	-0.5	-0.1	0.3	0.7	1.6
limitations sensorielles	-3.2	-0.8	-0.3	0.1	0.4	0.6	1.5
limitations cognitives	-2.9	-0.9	-0.3	0.1	0.4	0.7	1.2
<b>Population de 60 ans et plus</b>							
limitations motrices	-7.0	-2.9	-1.2	0.4	3.0	4.3	8.1
limitations sensorielles	-4.5	-2.3	-1.1	0.4	1.1	2.2	6.7
limitations cognitives	-3.7	-0.8	-0.2	0.5	1.1	2.0	2.7
<b>GIR14</b>	-3.5	-1.5	-0.5	0.4	1.0	1.5	3.1

### Les limitations sensorielles dans la population de 20 à 59 ans

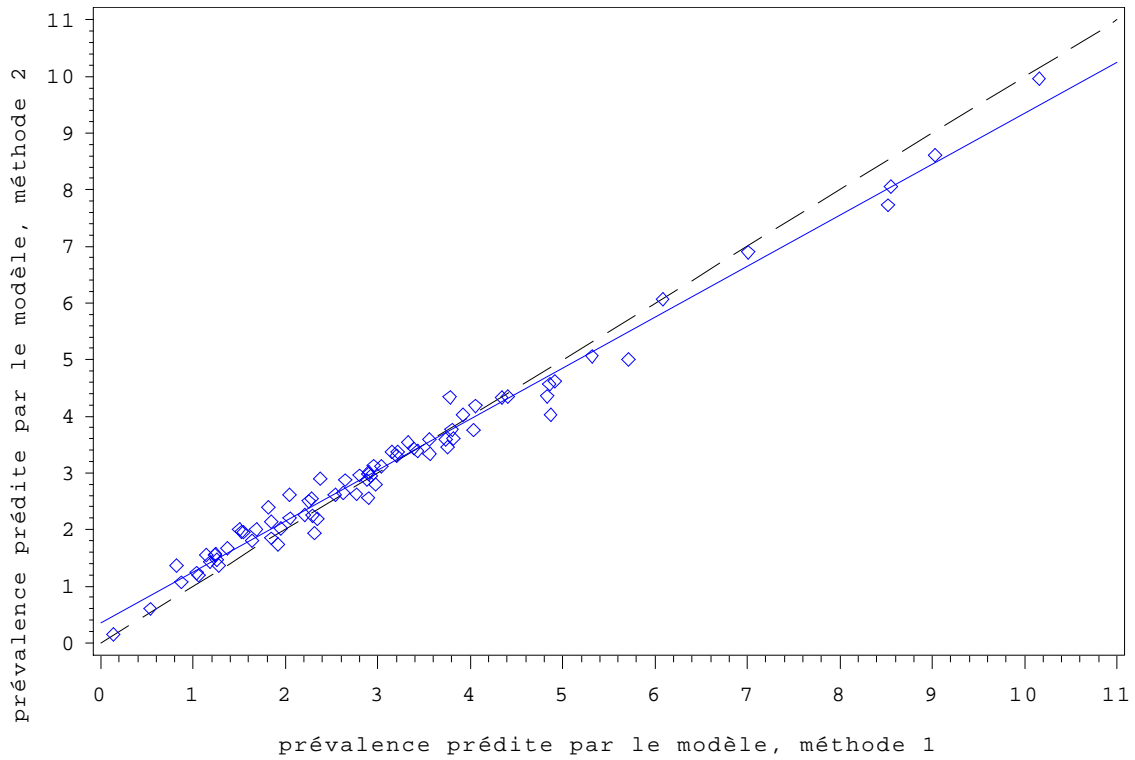
Région	Borne inf. de l'IC	Estimation directe (0)	Borne sup. de l'IC	Part de la phase HSM dans la variance totale (%)	Prévalence prédite par le modèle méthode 1 (1)	Prévalence prédite par le modèle méthode 2 (2)	Ecart relatif (%) (2)/(0)	Ecart absolu (2)-(0)
Ile de France	1.6	2.5	3.3	90.3	2.5	2.5	0.5	0.0
Champagne-Ardenne	0.8	3.8	6.8	47.4	4.0	3.7	-2.3	-0.1
Picardie	0.5	2.7	5.0	28.9	2.7	2.9	5.0	0.1
Haute-Normandie	0.2	2.9	5.5	73.2	3.2	3.1	7.1	0.2
Centre	0.8	3.0	5.2	53.7	2.6	2.6	-11.9	-0.4
Basse-Normandie	-0.2	1.9	4.0	80.0	2.1	2.3	20.2	0.4
Bourgogne	0.6	3.4	6.2	72.6	4.7	4.7	41.0	1.4
Nord	2.5	3.3	4.1	76.0	3.2	3.3	-2.2	-0.1
Lorraine	0.2	2.1	3.9	50.0	2.2	2.5	20.1	0.4
Alsace	0.2	2.1	4.0	84.8	1.9	2.2	2.6	0.1
Franche-Comté	0.2	2.1	4.0	52.4	1.9	2.1	-1.8	-0.0
Pays de la Loire	1.6	3.6	5.5	61.8	3.7	3.7	5.2	0.2
Bretagne	0.8	2.2	3.5	55.0	2.1	2.3	7.2	0.2
Poitou-Charentes	0.7	2.7	4.8	66.8	1.9	2.1	-23.3	-0.6
Aquitaine	1.1	2.8	4.6	79.5	2.8	2.9	3.9	0.1
Midi-Pyrénées	1.3	3.3	5.4	62.3	3.1	3.2	-4.5	-0.1
Limousin	-1.0	3.5	7.9	63.7	3.6	3.5	1.4	0.0
Rhône-Alpes	1.4	3.6	5.8	42.1	3.6	3.6	0.2	0.0
Auvergne	-0.6	3.1	6.9	25.8	2.2	2.6	-16.2	-0.5
Languedoc-Roussillon	-0.0	4.7	9.4	23.4	4.0	4.1	-12.6	-0.6
Provence Côte d'Azur	2.5	4.5	6.5	71.1	4.7	4.4	-2.5	-0.1
Corse	-3.9	5.0	13.9	50.8	2.6	2.6	-47.2	-2.4
France	2.7	3.1	3.5	73.0	3.0	3.1	-0.2	-0.0

### Départements avec extension d'échantillon

Département	Borne inf. de l'IC	Estimation directe (0)	Borne sup. de l'IC	Prévalence prédite par le modèle méthode 1 (1)	Prévalence prédite par le modèle méthode 2 (2)	Ecart relatif (%) (2)/(0)	Ecart absolu (2)-(0)
Nord	2.5	3.6	4.7	3.4	3.4	-5.1	-0.2
Pas-de-Calais	1.8	2.8	3.8	2.8	3.0	4.9	0.1
Rhône	1.3	2.7	4.2	2.9	3.0	8.6	0.2
Hauts-de-Seine	0.5	1.6	2.7	1.6	1.8	13.3	0.2

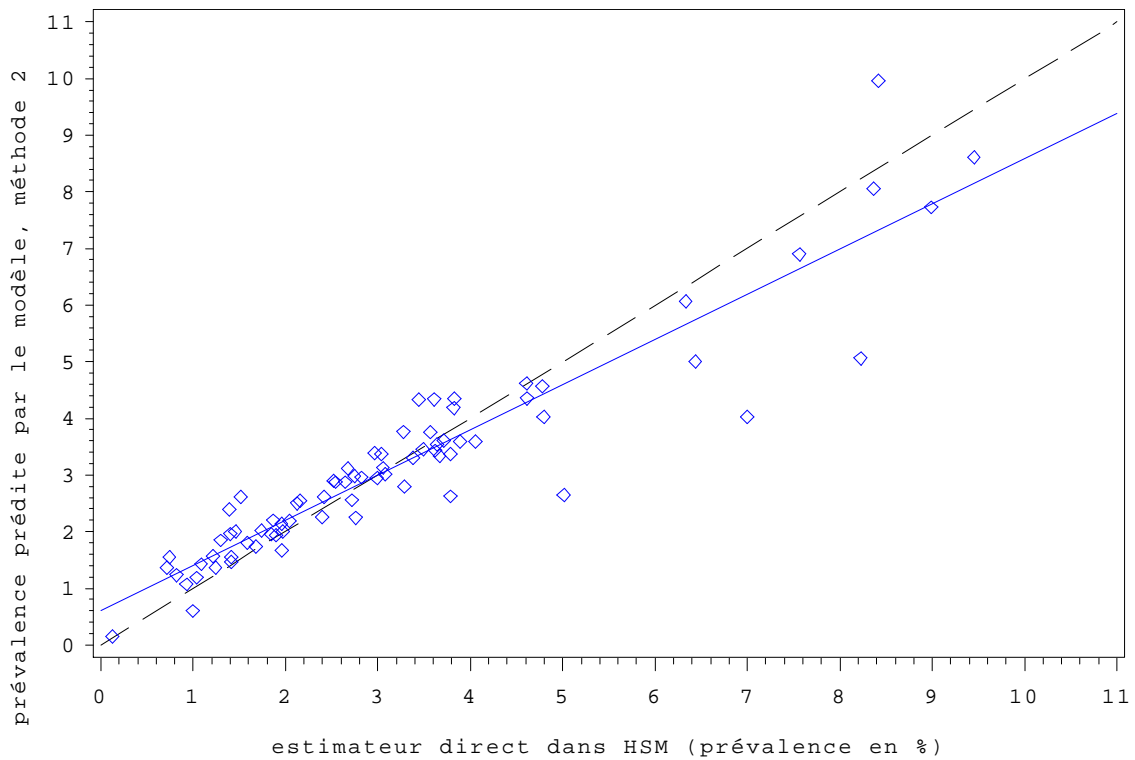
Note : IC = intervalle de confiance à 95 %

**Les limitations sensorielles dans la population de 20 à 59 ans :  
prévalences départementales prédites selon les méthodes 1 et 2**



*Equation de régression :* prédiction méthode 1 =  $0.34 + 0.90 * \text{prédiction méthode 2}$

**Prévalences départementales prédites (méthode 2) et directes**



*Equation de régression :* prevalence prédite =  $0.60 + 0.80 * \text{estimation directe}$

### Les limitations sensorielles dans la population de 60 ans et plus, par région

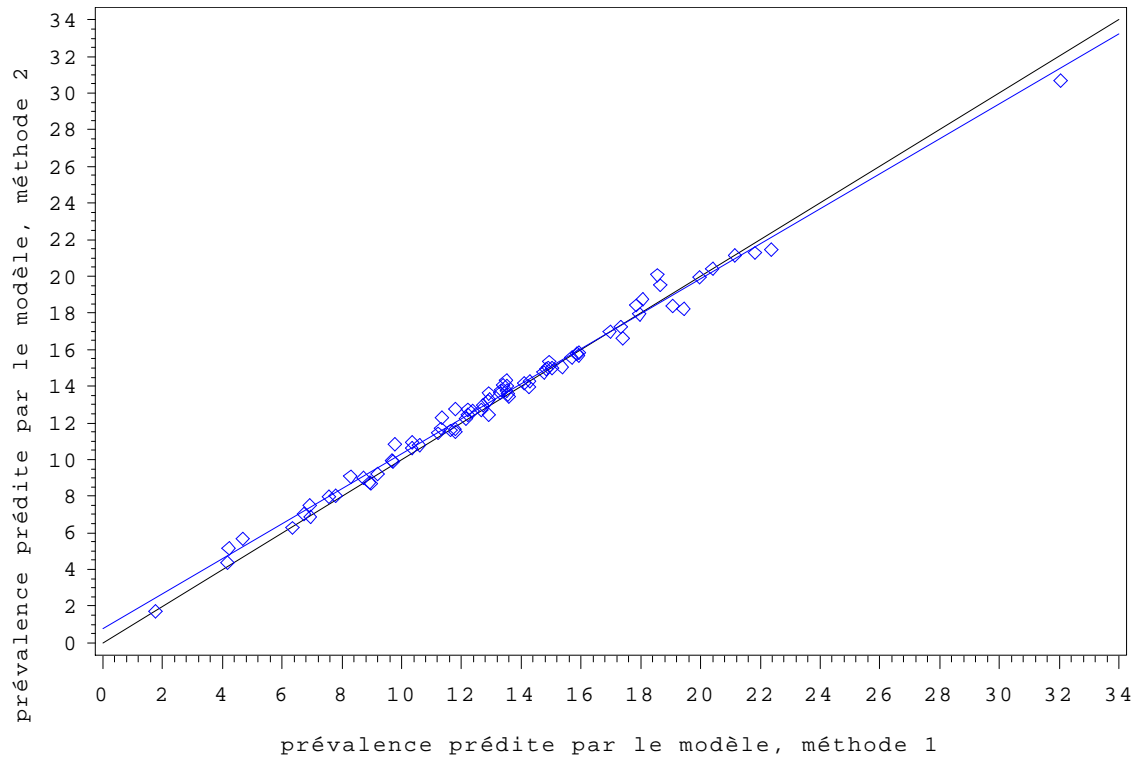
Région	Borne inf. de l'IC	Estimation directe (0)	Borne sup. de l'IC	Part de la phase HSM dans la variance totale (%)	Prévalence prédite par le modèle méthode 1 (1)	Prévalence prédite par le modèle méthode 2 (2)	Ecart absolu (2)-(0)	Ecart relatif (%) (2)/(0)
Ile de France	8.2	10.4	12.7	86.0	10.1	10.2	-0.2	-2.1
Champagne-Ardenne	7.0	15.1	23.2	52.1	15.8	15.8	0.7	4.4
Picardie	2.8	14.5	26.2	18.6	11.7	12.3	-2.2	-15.2
Haute-Normandie	2.9	8.7	14.5	63.2	9.1	9.2	0.5	5.3
Centre	3.8	12.3	20.8	27.4	13.3	13.1	0.8	6.7
Basse-Normandie	2.9	11.6	20.3	28.1	12.6	12.7	1.1	9.3
Bourgogne	4.4	11.1	17.8	45.3	12.1	12.5	1.4	12.3
Nord	13.1	15.1	17.0	91.0	15.5	15.4	0.3	2.2
Lorraine	6.3	14.9	23.5	32.2	14.5	14.4	-0.5	-3.1
Alsace	3.2	9.9	16.7	62.1	11.6	11.9	2.0	19.8
Franche-Comté	1.1	10.4	19.8	21.7	6.9	7.1	-3.3	-31.8
Pays de la Loire	7.2	13.7	20.3	29.7	14.4	14.8	1.0	7.5
Bretagne	5.4	10.5	15.5	66.0	9.3	9.4	-1.0	-9.8
Poitou-Charentes	3.0	10.3	17.7	21.7	9.7	10.1	-0.3	-2.5
Aquitaine	10.0	17.3	24.7	43.0	16.0	16.0	-1.3	-7.5
Midi-Pyrénées	6.9	12.4	17.8	38.0	11.9	12.2	-0.2	-1.4
Limousin	3.7	13.2	22.6	30.9	16.3	17.6	4.4	33.8
Rhône-Alpes	9.6	14.1	18.6	38.0	13.8	13.9	-0.2	-1.3
Auvergne	8.4	22.6	36.8	35.1	21.4	20.4	-2.2	-9.7
Languedoc-Roussillon	4.7	16.0	27.3	17.6	16.4	16.3	0.3	2.2
Provence Côte d'Azur	11.8	17.0	22.3	59.9	17.6	17.2	0.1	0.7
Corse	-0.8	4.8	10.3	67.2	4.7	5.7	0.9	18.6
France	12.5	13.5	14.4	81.3	13.3	13.3	-0.1	-1.1

### Départements avec extension d'échantillon

Département	Borne inf. de l'IC	Estimation directe (0)	Borne sup. de l'IC	Prévalence prédite par le modèle méthode 1 (1)	Prévalence prédite par le modèle méthode 2 (2)	Ecart relatif (%) (2)/(0)	Ecart absolu (2)-(0)
Nord	12.7	15.5	18.2	15.9	15.7	1.2	0.2
Pas-de-Calais	11.8	14.4	17.0	14.8	15.0	3.9	0.6
Rhône	11.7	14.2	16.8	13.6	13.6	-4.7	-0.7
Hauts-de-Seine	7.3	10.2	13.1	10.6	10.8	5.8	0.6

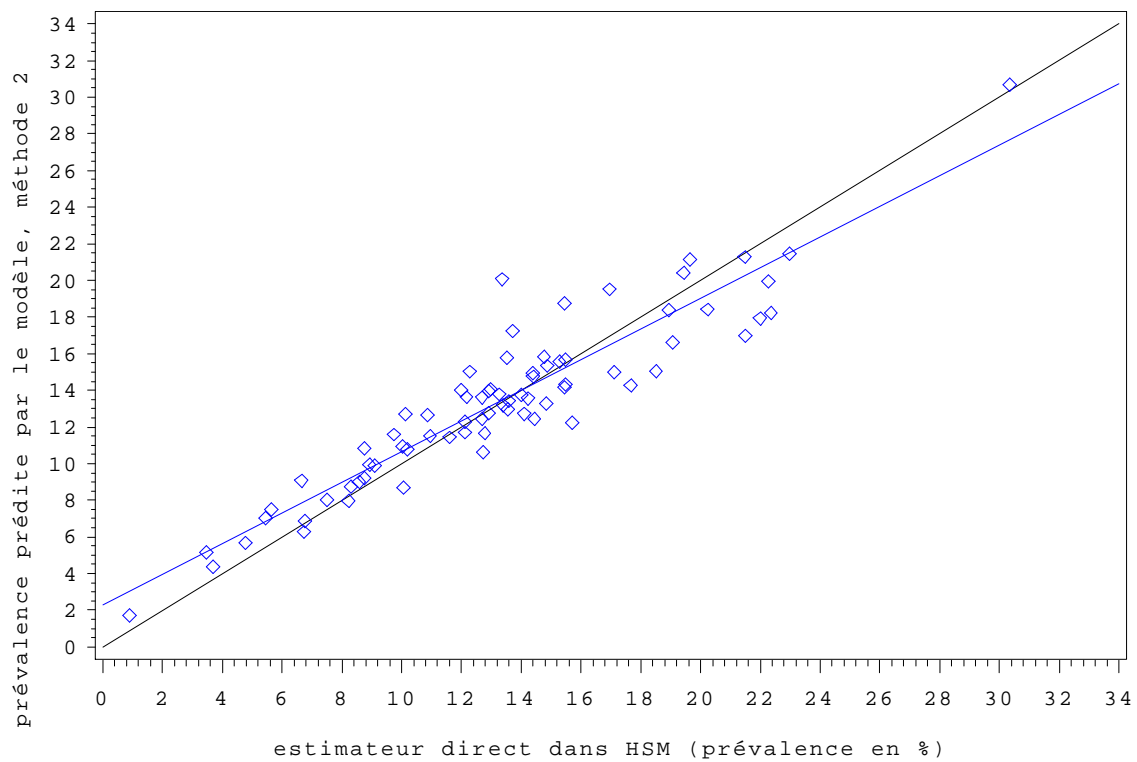
Note : IC = intervalle de confiance à 95 %

**Les limitations sensorielles dans la population de 60 ans et plus :  
prévalences départementales prédites selon les méthodes 1 et 2**



*Equation de régression :* prédition méthode 1 =  $0.75 + 0.95 * \text{prédiction méthode 2}$

**Prévalences départementales prédites et directes**



*Equation de régression :* prevalence prédite =  $2.29 + 0.84 * \text{estimation directe}$



### Les limitations de mobilité dans la population de 20 à 59 ans, par région

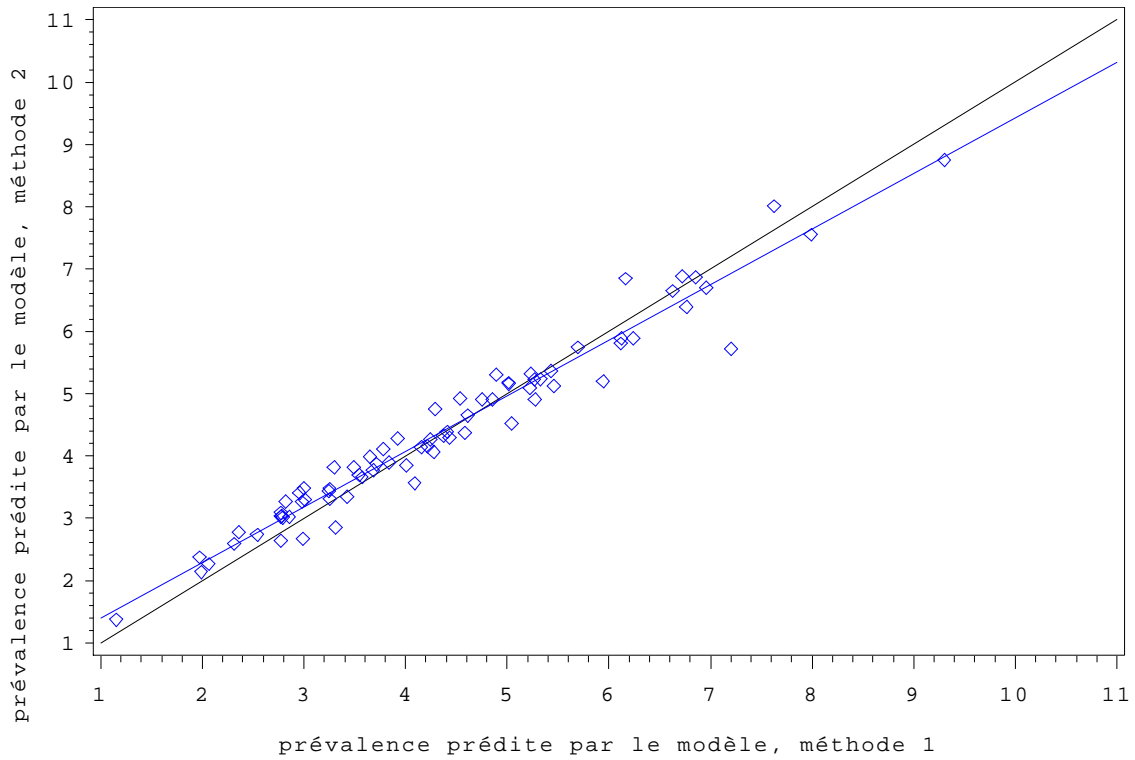
Région	borne inf. de l'IC	Estimation directe (0)	borne sup. de l'IC	Part de la phase HSM dans la variance totale (%)	Prévalence prédite par le modèle méthode 1 (1)	Prévalence prédite par le modèle méthode 2 (2)	Ecart relatif (%) (2)/(0)	Ecart (2)-(0)
Ile de France	2.7	3.5	4.4	90.2	3.4	3.4	-3.6	-0.1
Champagne-Ardenne	1.1	3.6	6.1	27.7	3.6	3.8	4.2	0.2
Picardie	1.4	5.6	9.8	23.0	5.8	5.6	0.2	0.0
Haute-Normandie	0.8	5.6	10.5	35.4	5.2	5.1	-9.2	-0.5
Centre	1.0	4.8	8.6	17.4	4.1	4.1	-15.3	-0.7
Basse-Normandie	1.3	3.5	5.7	28.8	3.6	3.9	11.7	0.4
Bourgogne	2.1	5.8	9.5	58.8	6.3	6.0	3.6	0.2
Nord	4.6	5.3	6.1	72.2	5.3	5.3	-1.1	-0.1
Lorraine	2.4	4.7	6.9	47.6	4.9	4.8	3.2	0.1
Alsace	2.0	4.0	6.0	62.8	3.7	3.9	-3.0	-0.1
Franche-Comté	1.0	5.6	10.3	24.7	4.8	5.0	-11.0	-0.6
Pays de la Loire	2.1	4.3	6.5	41.7	4.3	4.3	-0.3	-0.0
Bretagne	1.7	4.1	6.6	52.5	3.7	3.7	-9.6	-0.4
Poitou-Charentes	2.2	5.5	8.7	28.1	4.5	4.7	-15.0	-0.8
Aquitaine	1.8	3.7	5.7	55.4	3.9	4.0	8.2	0.3
Midi-Pyrénées	2.0	3.6	5.1	47.0	3.7	3.9	9.4	0.3
Limousin	0.6	7.4	14.2	83.7	6.8	6.4	-14.0	-1.0
Rhône-Alpes	2.7	4.4	6.1	46.4	4.1	3.9	-10.1	-0.4
Auvergne	1.1	2.9	4.7	45.3	3.5	3.8	32.1	0.9
Languedoc-Roussillon	1.5	4.8	8.2	28.4	4.3	4.5	-7.3	-0.4
Provence Côte d'Azur	3.7	5.5	7.4	58.7	5.3	5.3	-5.1	-0.3
Corse	-0.8	7.5	15.9	57.2	4.5	4.9	-34.8	-2.6
France	4.1	4.4	4.8	77.1	4.3	4.3	-3.2	-0.1

### Départements avec extension d'échantillon

Département	Borne inf. de l'IC	Estimation directe (0)	Borne sup. de l'IC	Prévalence prédite par le modèle méthode 1 (1)	Prévalence prédite par le modèle méthode 2 (2)	Ecart relatif (%) (2)/(0)	Ecart absolu (2)-(0)
Nord	3.8	4.8	5.7	4.9	4.9	3.1	0.1
Pas de Calais	5.1	6.3	7.6	6.1	5.9	-7.0	-0.4
Rhône	3.2	4.3	5.4	4.4	4.3	0.8	0.0
Hauts de Seine	1.7	2.3	2.9	2.3	2.6	13.0	0.3

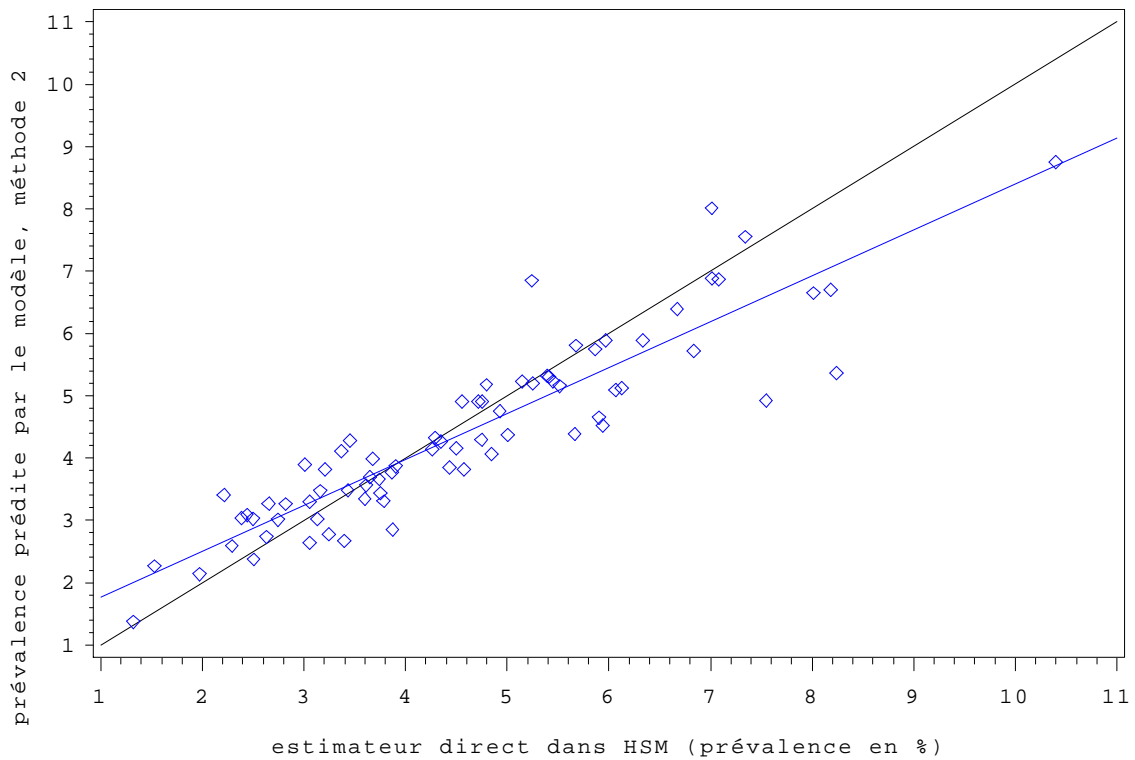
Note : IC = intervalle de confiance à 95 %

**Les limitations de mobilité dans la population de 20 à 59 ans :  
Prévalences départementales prédites selon les méthodes 1 et 2**



*Equation de régression* : prédiction méthode 1 =  $0.508 + 0.891 * \text{prédiction méthode 2}$

**Prévalences départementales prédites et directes**



*Equation de régression* : prévalence prédite =  $1.019 + 0.738 * \text{estimation directe}$   
*Sans la Corse et la Loire* : prévalence prédite =  $0.820 + 0.793 * \text{estimation directe}$

### Les limitations de mobilité dans la population de 60 ans et plus, par région

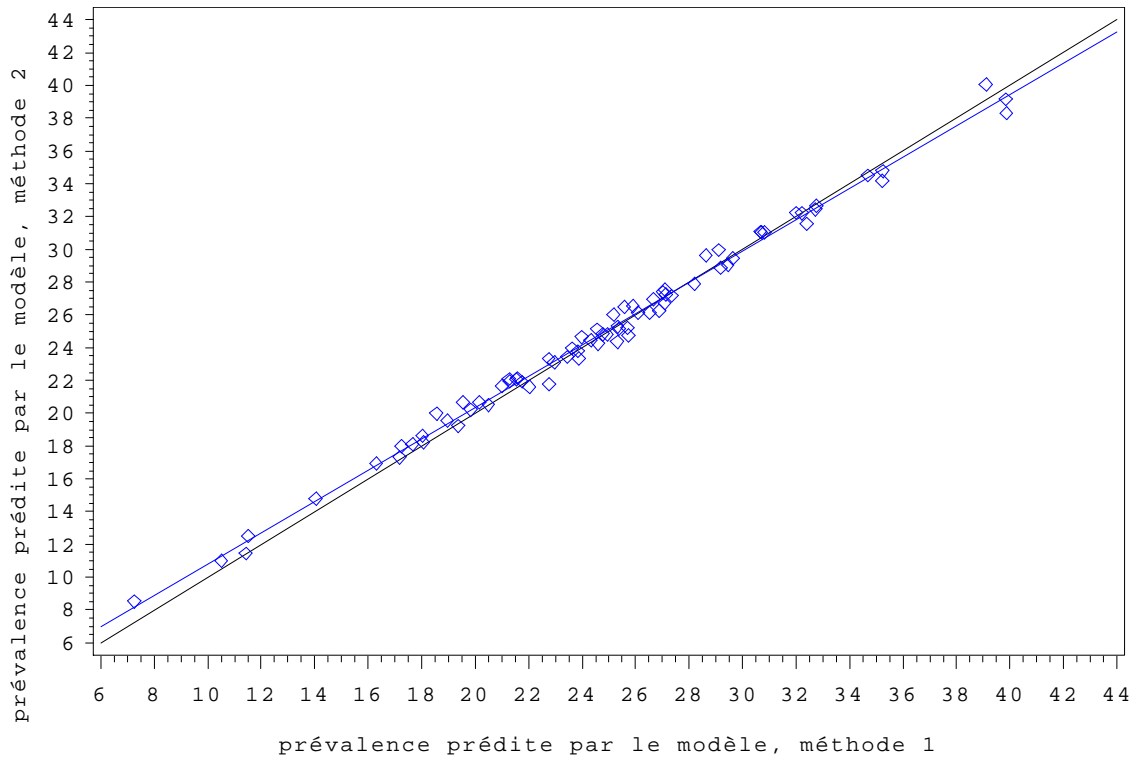
Région	borne inf. de l'IC	Estimation directe (0)	borne sup. de l'IC	Part de la phase HSM dans la variance totale (%)	Prévalence prédite par le modèle méthode 1 (1)	Prévalence prédite par le modèle méthode 2 (2)	Ecart relatif (%) (2)/(0)	Ecart (2)-(0)
Ile de France	19.3	22.2	25.1	89.6	22.9	22.6	2.2	0.5
Champagne-Ardenne	10.4	20.6	30.7	29.4	25.4	25.4	23.6	4.9
Picardie	7.8	33.8	59.8	9.4	30.1	30.3	-10.3	-3.5
Haute-Normandie	7.3	28.1	49.0	20.3	26.2	26.1	-7.4	-2.1
Centre	8.9	25.3	41.7	13.5	27.4	27.8	10.0	2.5
Basse-Normandie	7.1	22.3	37.5	14.6	22.7	23.2	3.7	0.8
Bourgogne	13.8	27.7	41.6	27.3	27.7	27.9	0.7	0.2
Nord	28.3	31.4	34.4	66.1	32.0	32.1	2.2	0.7
Lorraine	11.5	21.0	30.5	24.3	23.8	24.0	14.4	3.0
Alsace	11.4	25.5	39.5	29.6	29.9	29.5	15.9	4.1
Franche-Comté	6.0	22.2	38.5	14.5	17.8	18.1	-18.7	-4.2
Pays de la Loire	13.1	23.1	33.0	19.6	22.8	23.0	-0.4	-0.1
Bretagne	11.5	19.5	27.4	34.5	19.5	19.9	2.1	0.4
Poitou-Charentes	10.9	25.0	39.1	19.0	23.8	24.0	-4.0	-1.0
Aquitaine	16.5	25.0	33.4	37.0	23.9	23.9	-4.4	-1.1
Midi-Pyrénées	15.8	25.2	34.6	21.4	25.0	25.4	0.6	0.1
Limousin	10.8	31.2	51.5	17.0	32.2	33.2	6.5	2.0
Rhône-Alpes	16.9	22.5	28.0	37.9	22.6	22.6	0.6	0.1
Auvergne	13.4	29.4	45.4	23.4	28.9	28.8	-2.3	-0.7
Languedoc-Roussillon	9.0	23.9	38.8	12.3	25.5	25.2	5.5	1.3
Provence Côte d'Azur	18.0	23.7	29.5	37.6	24.0	24.4	2.6	0.6
Corse	0.3	24.7	49.2	40.9	25.3	24.4	-1.5	-0.4
France	23.2	24.4	25.6	73.2	24.7	24.8	1.6	0.4

### Départements avec extension d'échantillon

Département	Borne inf. de l'IC	Estimation directe (0)	Borne sup. de l'IC	Prévalence prédite par le modèle méthode 1 (1)	Prévalence prédite par le modèle méthode 2 (2)	Ecart relatif (%) (2)/(0)	Ecart absolu (2)-(0)
Nord	28.0	32.2	36.4	32.8	32.7	1.4	0.4
Pas de Calais	25.9	29.9	33.9	30.7	31.1	3.9	1.2
Rhône	20.7	23.8	26.9	23.0	23.1	-2.8	-0.7
Hauts de Seine	17.0	21.0	25.0	22.8	21.8	3.7	0.8

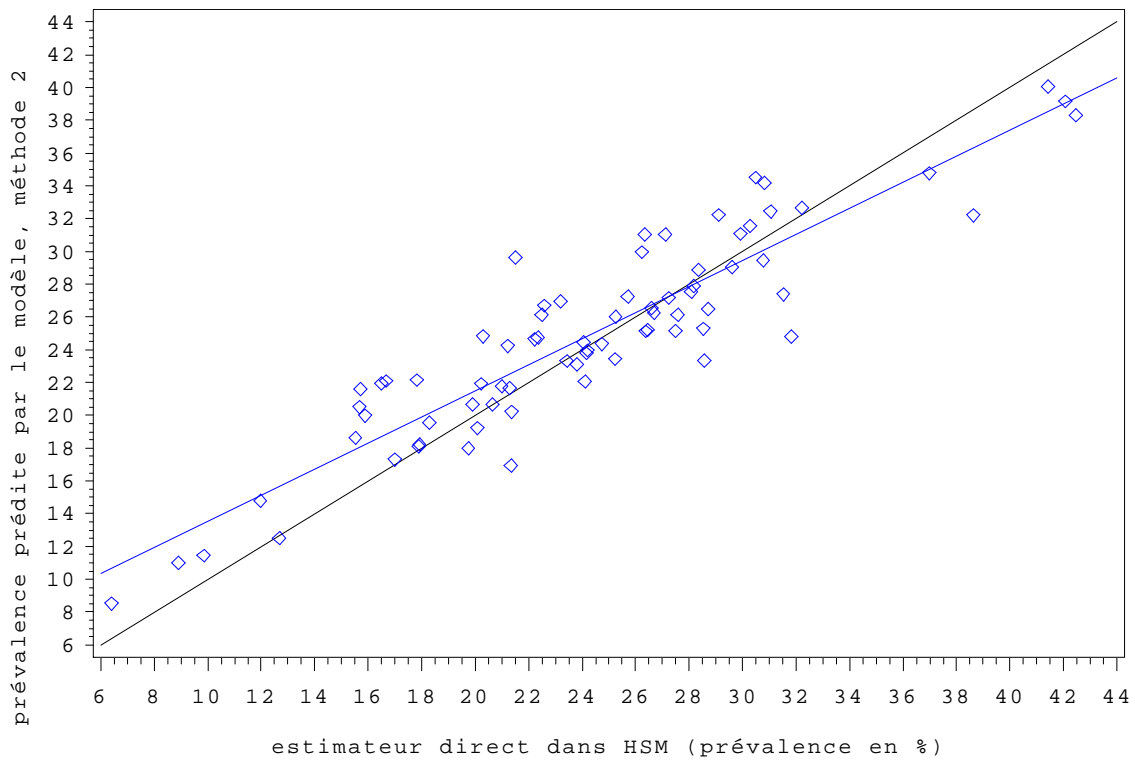
Note : IC = intervalle de confiance à 95 %

**Limitations de mobilité dans la population de 60 ans ou plus :  
Prévalences départementales selon les méthodes 1 et 2**



*Equation de régression* : prédiction méthode 1 =  $1.265 + 0.955 * \text{prédiction méthode 2}$

**Prévalences départementales prédites et directes**



*Equation de régression* : prévalence prédite =  $5.613 + 0.795 * \text{estimation directe}$

### Troubles cognitifs dans la population de 20 à 59 ans

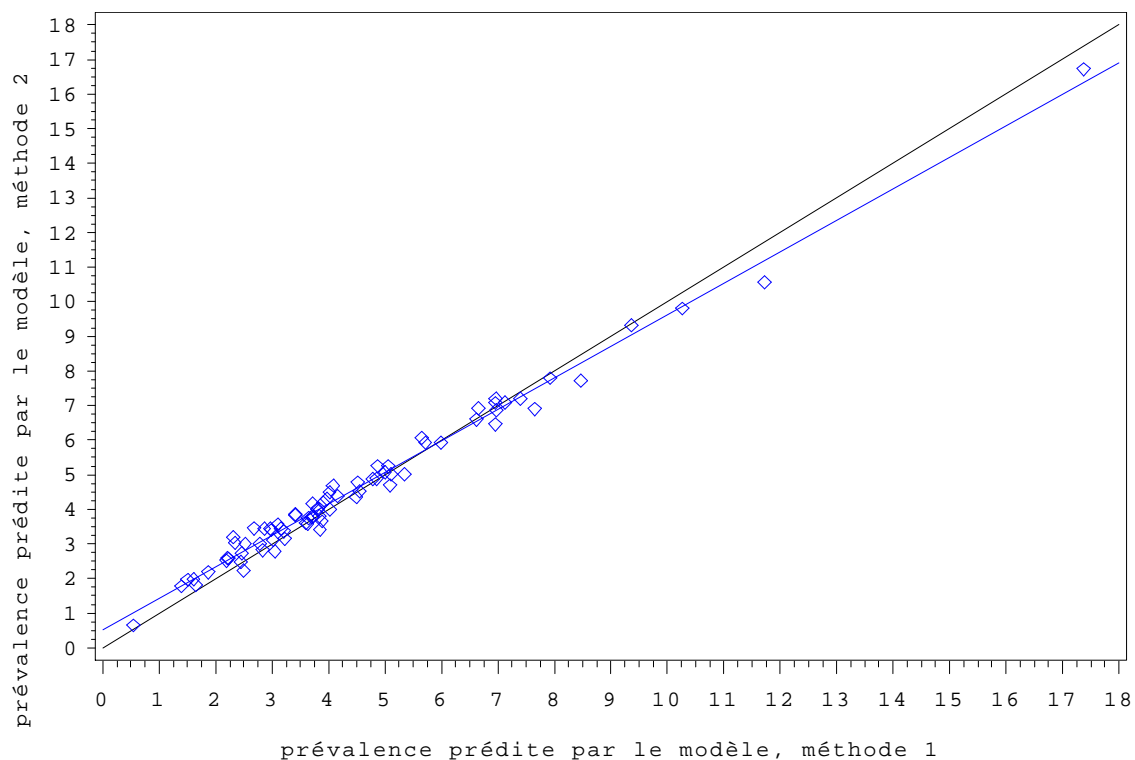
Région	borne inf. de l'IC	Estimation directe (0)	borne sup. de l'IC	Part de la phase HSM dans la variance totale (%)	Prévalence prédite par le modèle méthode 1 (1)	Prévalence prédite par le modèle méthode 2 (2)	Ecart relatif (%) (2)/(0)	Ecart (2)-(0)
Ile de France	2.7	3.6	4.5	90.7	3.3	3.4	-4.4	-0.2
Champagne-Ardenne	1.5	5.3	9.0	35.6	5.2	5.1	-3.3	-0.2
Picardie	1.1	5.4	9.8	24.3	4.8	4.9	-10.1	-0.5
Haute-Normandie	1.1	3.9	6.7	57.1	3.6	3.8	-3.1	-0.1
Centre	0.9	5.1	9.3	24.1	5.0	5.1	0.0	0.0
Basse-Normandie	0.8	4.9	9.1	48.5	5.0	4.8	-1.9	-0.1
Bourgogne	0.8	6.4	11.9	52.8	8.9	8.7	37.2	2.4
Nord	5.0	6.0	7.0	85.4	6.0	6.0	0.1	0.0
Lorraine	1.2	3.4	5.6	58.9	3.3	3.4	-0.3	-0.0
Alsace	1.9	4.6	7.3	55.9	4.4	4.6	0.6	0.0
Franche-Comté	1.2	4.6	8.0	37.3	4.1	4.2	-8.1	-0.4
Pays de la Loire	2.0	4.3	6.6	49.7	4.1	4.3	-1.3	-0.1
Bretagne	2.2	4.5	6.8	67.5	4.3	4.8	5.2	0.2
Poitou-Charentes	0.9	4.4	7.9	41.2	3.7	4.0	-9.7	-0.4
Aquitaine	1.5	4.4	7.4	45.9	4.5	4.6	3.9	0.2
Midi-Pyrénées	1.4	2.6	3.8	55.9	3.1	3.6	38.8	1.0
Limousin	-2.3	6.7	15.6	33.0	8.7	8.2	23.4	1.6
Rhône-Alpes	2.4	3.8	5.2	72.8	3.8	3.6	-5.5	-0.2
Auvergne	1.9	5.7	9.4	73.4	7.2	7.1	25.8	1.5
Languedoc-Roussillon	-0.2	5.9	12.0	14.4	4.8	5.0	-16.6	-1.0
Provence Côte d'Azur	3.6	5.7	7.7	77.8	5.5	5.5	-3.6	-0.2
Corse	-1.1	4.7	10.6	64.2	3.2	3.4	-28.8	-1.4
France	4.0	4.5	5.0	70.6	4.4	4.5	0.2	0.0

### Départements avec extension d'échantillon

Département	Borne inf. de l'IC	Estimation directe (0)	Borne sup. de l'IC	Prévalence prédite par le modèle méthode 1 (1)	Prévalence prédite par le modèle méthode 2 (2)	Ecart relatif (%) (2)/(0)	Ecart absolu (2)-(0)
Nord	5.2	6.6	8.0	6.6	6.6	-0.2	-0.0
Pas-de-Calais	3.5	4.8	6.2	4.8	4.9	1.2	0.1
Rhône	3.1	4.8	6.4	4.5	4.5	-5.3	-0.3
Hauts-de-Seine	1.4	2.9	4.4	2.8	2.8	-2.1	-0.1

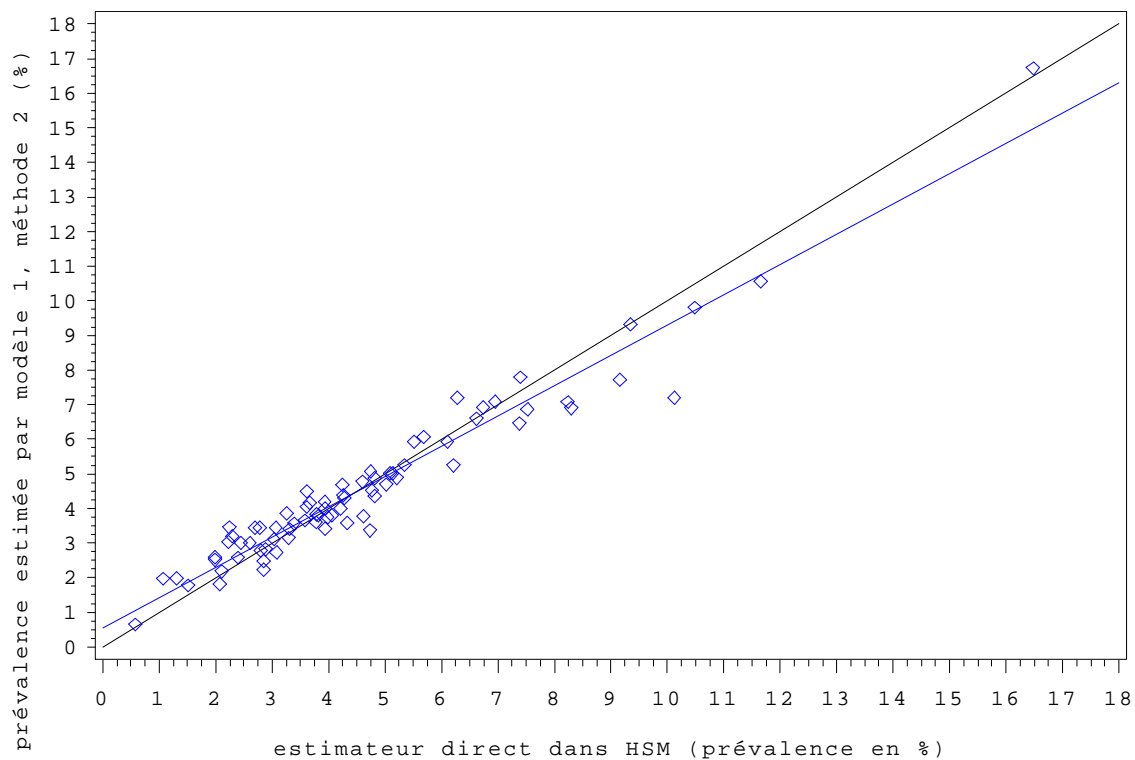
Note : IC = intervalle de confiance à 95 %

**Troubles cognitifs dans la population de 20 à 59 ans :  
Prévalences départementales prédites selon les méthodes 1 et 2**



*Equation de régression* : prédiction méthode 1 =  $0.51 + 0.91 * \text{prédiction méthode 2}$

**Prévalences départementales directes et prédites par le modèle**



*Equation de régression* : prévalence prédite =  $0.54 + 0.87 * \text{estimation directe}$

### Troubles cognitifs dans la population de 60 ans et plus

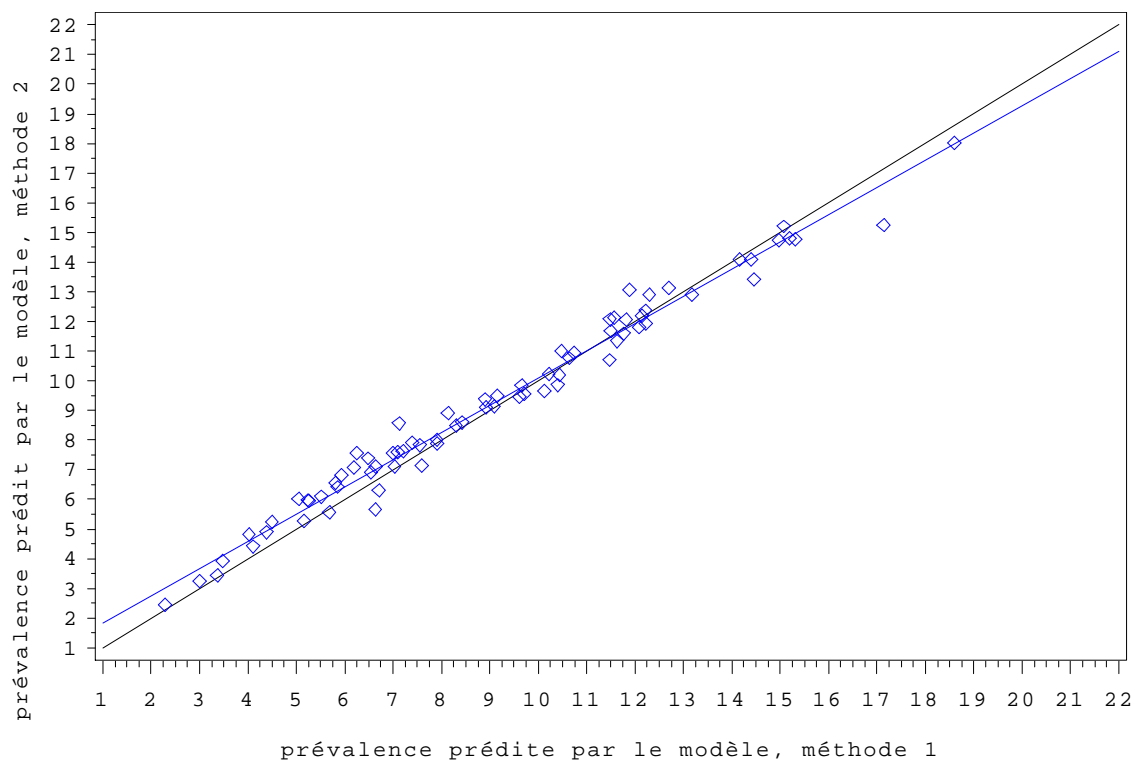
Région	Borne inf. de l'IC	Estimation directe (0)	Borne sup. de l'IC	Part de la phase HSM dans la variance totale (%)	Prévalence prédite par le modèle méthode 1 (1)	Prévalence prédite par le modèle méthode 2 (2)	Ecart relatif (%) (2)/(0)	Ecart (2)-(0)
Ile de France	6.4	8.3	10.1	93.8	8.2	8.4	1.7	0.1
Champagne-Ardenne	3.2	9.7	16.1	39.3	10.8	10.9	12.9	1.2
Picardie	3.1	12.8	22.5	19.4	11.0	11.3	-11.8	-1.5
Haute-Normandie	-0.5	12.8	26.1	23.6	11.2	11.0	-14.3	-1.8
Centre	-0.1	8.8	17.6	17.2	7.6	8.7	-0.9	-0.1
Basse-Normandie	0.3	7.3	14.3	20.4	7.9	7.9	8.5	0.6
Bourgogne	5.0	12.3	19.6	71.0	12.6	12.0	-2.0	-0.2
Nord	7.7	9.2	10.7	71.3	9.7	10.1	9.3	0.9
Lorraine	2.3	5.9	9.5	48.0	6.2	6.5	10.2	0.6
Alsace	1.2	3.9	6.5	43.6	5.3	6.2	58.4	2.3
Franche-Comté	1.0	9.2	17.3	27.3	7.6	7.8	-15.4	-1.4
Pays de la Loire	3.9	8.3	12.7	35.3	8.4	8.7	5.6	0.5
Bretagne	2.7	5.6	8.6	64.0	6.3	6.4	13.2	0.7
Poitou-Charentes	1.6	4.7	7.8	36.3	4.5	4.9	4.9	0.2
Aquitaine	6.6	10.7	14.7	63.0	10.1	10.4	-2.9	-0.3
Midi-Pyrénées	4.7	8.8	12.9	32.5	8.7	9.3	6.4	0.6
Limousin	1.4	7.8	14.3	30.6	7.4	8.1	3.0	0.2
Rhône-Alpes	7.3	10.2	13.1	72.4	10.3	10.2	-0.1	-0.0
Auvergne	3.1	10.3	17.4	33.4	11.5	11.7	13.4	1.4
Languedoc-Roussillon	5.3	13.4	21.6	36.1	15.2	14.2	5.8	0.8
Provence Côte d'Azur	6.8	10.7	14.6	51.0	11.5	11.5	7.3	0.8
Corse	-4.0	5.6	15.2	84.9	6.6	5.7	1.4	0.1
France	8.3	9.1	9.9	70.4	9.2	9.4	3.1	0.3

### Départements avec extension d'échantillon

Département	Borne inf. de l'IC	Estimation directe (0)	Borne sup. de l'IC	Prévalence prédite par le modèle méthode 1 (1)	Prévalence prédite par le modèle méthode 2 (2)	Ecart relatif (%) (2)/(0)	Ecart absolu (2)-(0)
Nord	7.9	9.9	11.9	10.5	11.0	11.0	1.1
Pas-de-Calais	6.1	8.1	10.1	8.4	8.6	6.4	0.5
Rhône	6.8	9.2	11.6	9.1	9.1	-0.8	-0.1
Hauts-de-Seine	4.7	7.0	9.2	7.6	7.8	12.4	0.9

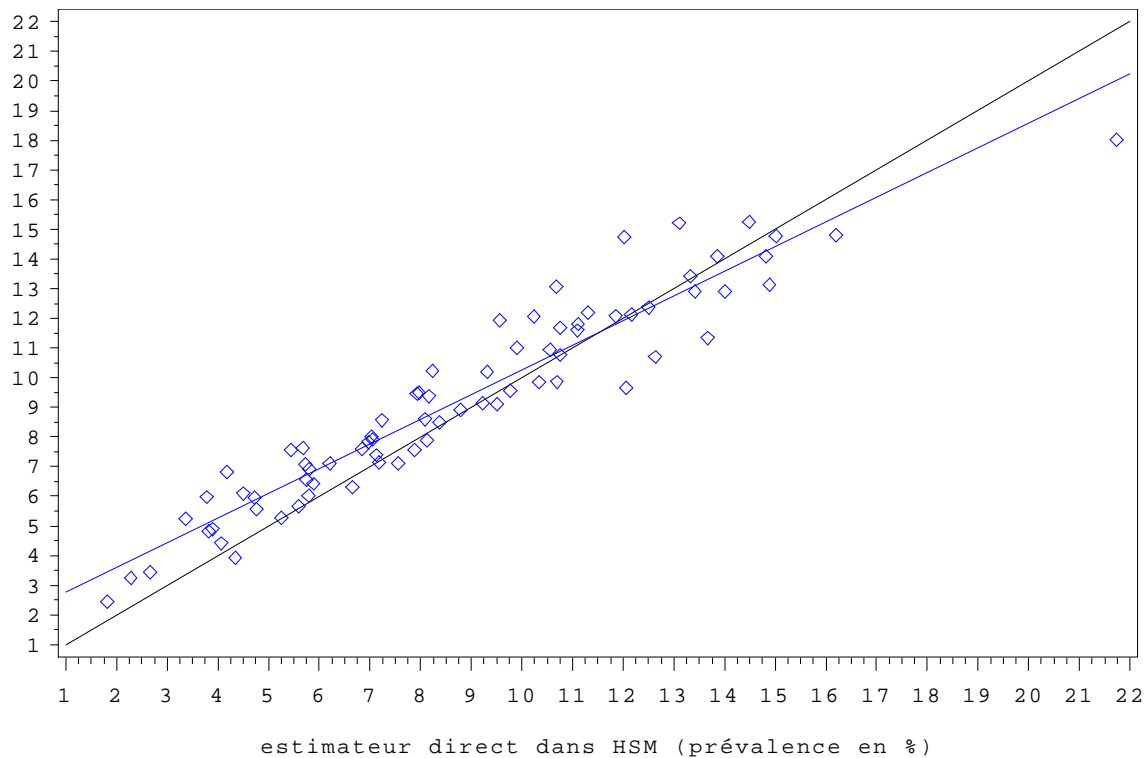
Note : IC = intervalle de confiance à 95 %

**Troubles cognitifs dans la population de 60 ans et plus**  
**Prévalences départementales prédites selon les méthodes 1 et 2**



*Equation de régression* : prédiction méthode 1 =  $0.93 + 0.92 * \text{prédiction méthode 2}$

**Prévalences départementales directes et prédites par le modèle**



*Equation de régression* : prévalence prédite =  $1.94 + 0.83 * \text{estimation directe}$



## GIR 1 à 4 : prévalences régionales dans la population de 60 ans ou plus

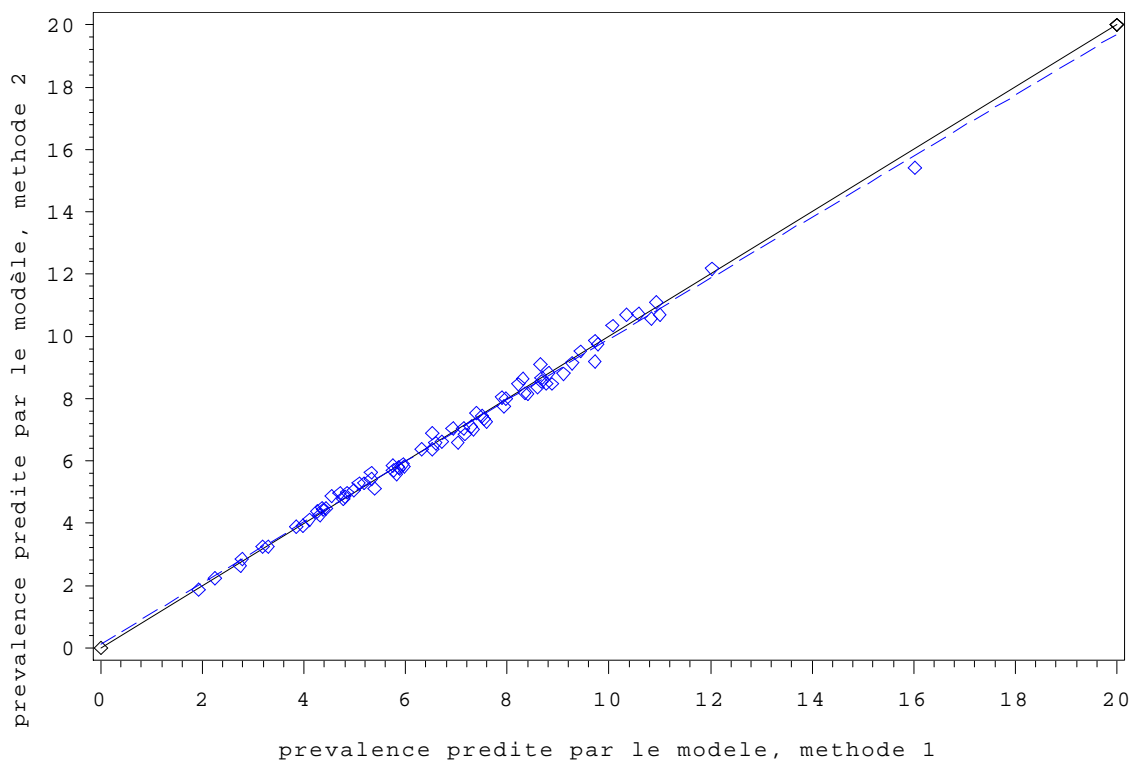
Région	Borne inf. de l'IC	Estimation directe (0)	Borne sup. de l'IC	Part de la phase HSM dans la variance totale (%)	Prévalence prédite par le modèle méthode 1 (1)	Prévalence prédite par le modèle méthode 2 (2)	Ecart relatif (%) (2)/(0)	Ecart (2)-(0)
Ile de France	3.7	5.0	6.2	76.9	4.9	4.9	-0.8	-0.0
Champagne-Ardenne	1.9	4.3	6.7	50.1	5.5	5.4	26.9	1.1
Picardie	1.8	9.7	17.5	17.5	7.3	7.2	-25.7	-2.5
Haute-Normandie	2.0	6.7	11.5	63.7	5.9	5.9	-13.1	-0.9
Centre	2.4	5.2	8.1	48.1	6.9	7.0	34.4	1.8
Basse-Normandie	1.8	8.6	15.4	30.8	9.4	9.3	7.8	0.7
Bourgogne	2.1	6.1	10.0	39.7	5.8	5.7	-6.2	-0.4
Nord	6.3	7.5	8.7	69.1	7.9	8.0	7.8	0.6
Lorraine	3.4	6.6	9.8	54.1	7.0	6.9	3.3	0.2
Alsace	1.0	4.0	7.1	37.8	5.3	5.2	29.3	1.2
Franche-Comté	2.1	6.8	11.4	38.6	7.9	7.9	16.4	1.1
Pays de la Loire	3.9	7.1	10.3	53.2	7.0	6.8	-3.9	-0.3
Bretagne	2.9	5.9	9.0	41.4	7.5	7.6	27.9	1.7
Poitou-Charentes	2.1	6.3	10.4	31.1	6.4	6.3	0.4	0.0
Aquitaine	3.2	5.9	8.6	49.3	5.9	6.0	0.6	0.0
Midi-Pyrénées	3.6	6.7	9.8	28.7	7.5	7.8	16.4	1.1
Limousin	2.9	9.5	16.1	33.7	10.0	9.7	2.3	0.2
Rhône-Alpes	3.9	6.4	8.9	65.5	6.4	6.4	0.5	0.0
Auvergne	2.1	6.6	11.2	26.6	5.7	5.8	-12.2	-0.8
Languedoc-Roussillon	1.7	7.4	13.1	25.9	7.8	7.7	4.5	0.3
Provence Côte d'Azur	4.8	6.8	8.9	53.8	7.9	7.9	15.6	1.1
Corse	-8.1	18.1	44.3	90.8	16.0	15.4	-14.7	-2.7
France	6.0	6.5	7.1	77.5	6.8	6.8	4.2	0.3

## Départements avec extension d'échantillon

Département	Borne inf. de l'IC	Estimation directe (0)	Borne sup. de l'IC	Prévalence prédite par le modèle méthode 1 (1)	Prévalence prédite par le modèle méthode 2 (2)	Ecart relatif (%) (2)/(0)	Ecart absolu (2)-(0)
Nord	5.3	6.8	8.3	7.4	7.5	10.9	0.7
Pas-de-Calais	6.5	8.5	10.6	8.8	8.8	3.8	0.3
Rhône	4.2	5.6	6.9	5.3	5.6	0.9	0.0
Hauts-de-Seine	3.1	4.4	5.7	5.1	5.3	19.2	0.9

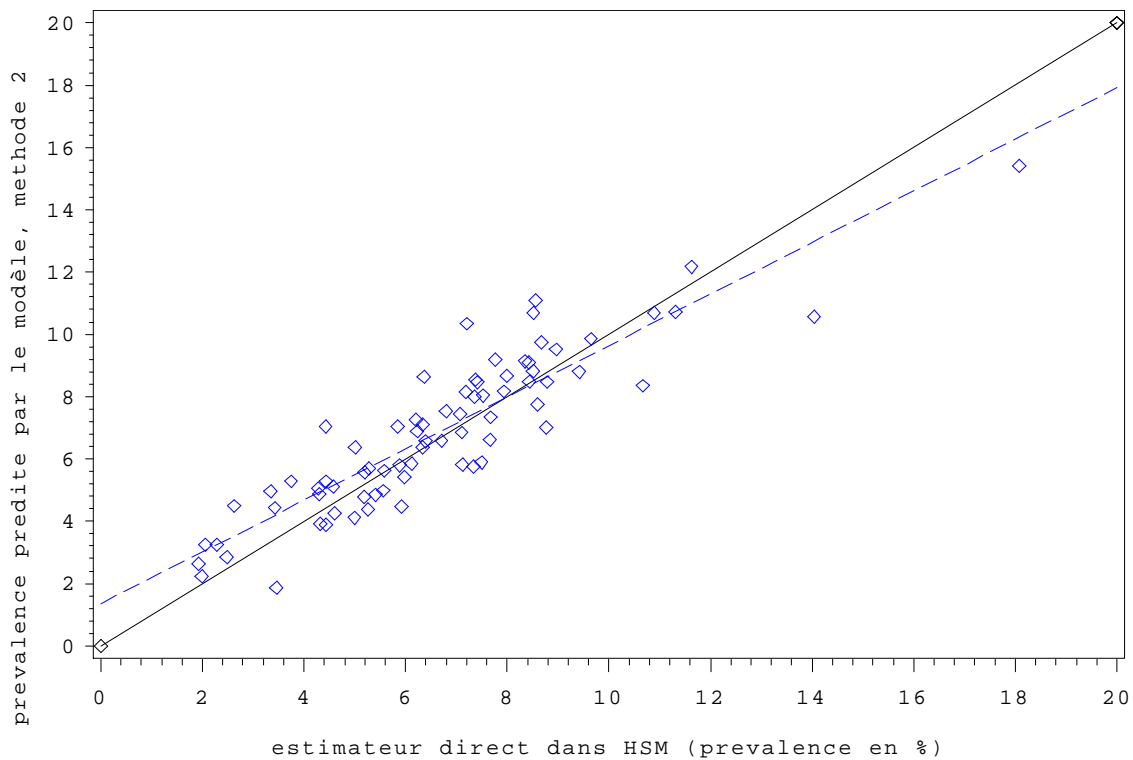
Note : IC = intervalle de confiance à 95 %

**Population de 60 ans ou plus en GIR 1 à 4 : estimations départementales prédites selon les méthodes 1 et 2**



*Equation de régression* : prédiction méthode 1 =  $0.12 + 0.98 * \text{prédiction méthode 2}$

**Estimations départementales directes et prédites par le modèle**



*Equation de régression* : prévalence prédite =  $1.36 + 0.83 * \text{estimation directe}$

## ANNEXE 1 : QUESTIONNAIRE DE L'ENQUÊTE VQS

### Consignes en haut du questionnaire

A remplir pour toutes les personnes du foyer : une colonne par personne, y compris les enfants.

Si une question ne s'applique pas pour un enfant trop jeune, ne répondez pas à cette question.

Si une personne ne peut absolument pas réaliser une des activités citées cochez la case

Oui, beaucoup pour cette activité.

Pour chaque personne du foyer on pose la série de questions suivante

Prénom

1 : sexe de la personne

2 : année de naissance

3 : *Comment est l'état de santé général de la personne?*

Très bon / bon / moyen / mauvais / très mauvais

4 : *La personne a-t-elle actuellement une ou plusieurs maladie(s) chronique(s) ou problème(s) de santé durable(s) ?*

Non / Oui

5 : *La personne est-elle limitée dans les activités qu'elle peut faire en raison d'un problème de santé ou d'un handicap(à la maison, au travail, à l'école...)?*

Non / Oui, un peu / Oui, beaucoup

6 : *La personne a-t-elle des difficultés pour voir les caractères ordinaires d'un article de journal (avec ses lunettes ou ses lentilles si elle en porte habituellement) ?*

Non / Oui, un peu / Oui, beaucoup

7 : *A-t-elle des difficultés pour voir nettement le visage de quelqu'un à l'autre bout de la pièce (avec ses lunettes ou ses lentilles si elle en porte habituellement) ?*

Non / Oui, un peu / Oui, beaucoup

8 : *A-t-elle des difficultés pour parler ?*

Non / Oui, un peu / Oui, beaucoup

9 : *A-t-elle des difficultés pour entendre ce qui se dit au cours d'une conversation avec plusieurs personnes ?*

Non / Oui, un peu / Oui, beaucoup

10 : *A-t-elle des difficultés pour monter un étage d'escalier ou marcher 500 mètres ?*

Non / Oui, un peu / Oui, beaucoup

11 : *A-t-elle des difficultés pour lever le bras (par exemple pour attraper un objet en hauteur)?*

Non / Oui, un peu / Oui, beaucoup

12 : *A-t-elle des difficultés pour se servir de ses mains ou de ses doigts (par exemple pour ouvrir une bouteille d'eau) ?*

Non / Oui, un peu / Oui, beaucoup

13 : *Lorsqu'elle est debout, a-t-elle des difficultés pour se pencher et ramasser un objet?*

Non / Oui, un peu / Oui, beaucoup

14 : *A-t-elle des difficultés pour se concentrer plus de 10 minutes ?*

Non / Oui, un peu / Oui, beaucoup

15 : *A-t-elle des difficultés pour se souvenir de choses importantes ?*

Non / Oui, un peu / Oui, beaucoup

16 : *A-t-elle des difficultés pour prendre des initiatives dans la vie quotidienne?*

Non / Oui, un peu / Oui, beaucoup

17 : *A-t-elle des difficultés pour résoudre les problèmes de la vie quotidienne (comme se repérer sur un itinéraire ou compter l'argent) ?*

Non / Oui, un peu / Oui, beaucoup

18 : *A-t-elle des difficultés pour sortir du domicile ?*

Non / Oui, un peu / Oui, beaucoup

19 : *A-t-elle des difficultés pour comprendre les autres ou se faire comprendre des autres (en dehors des difficultés liées aux différences de langue)?*

Non / Oui, un peu / Oui, beaucoup

20 : *La personne est-elle dans l'impossibilité totale d'accomplir une ou plusieurs des activités citées précédemment (questions 6 à 19). Par exemple une personne aveugle ne peut absolument pas voir les caractères ordinaires d'un article de journal.*

Non / Oui

*21 : En raison d'un problème de santé ou d'un handicap, reçoit-elle de l'aide d'une autre personne dans la vie quotidienne ?*

Non / Oui, un peu / Oui, beaucoup

*22 : En raison d'un problème de santé ou d'un handicap, des aménagements du logement ont-ils été réalisés pour cette personne?*

Non / Oui

*23 : En raison d'un problème de santé ou d'un handicap, utilise-t-elle régulièrement une prothèse, un appareillage ou une aide technique ?*

Non / Oui

*24 : La personne considère-t-elle qu'elle a un handicap ?*

Non / Oui

*25 : La personne a-t-elle une reconnaissance officielle d'un handicap ou d'une perte d'autonomie (allocation, pension ou carte d'invalidité, admission dans un établissement spécialisé...)?*

Non / Oui

*26 : Si c'est un enfant d'âge scolaire: Est-il inscrit dans une classe ou un établissement spécialisé en raison de problèmes de santé ou de difficultés d'apprentissage ?*

Non / Oui

## ANNEXE 2 : LA GRILLE AGGIR

Le principe<sup>14</sup> de la grille AGGIR consiste à repérer ce que la personne âgée fait ou ne fait pas seule, en excluant ce que font les aidants et les soignants. C'est la capacité à accomplir seul les actes essentiels de la vie quotidienne qui est évaluée, et le besoin de l'aide d'un tiers.

- Le GIR 1 comprend les personnes âgées confinées au lit ou au fauteuil, dont les fonctions mentales sont gravement altérées, qui nécessitent la présence indispensable et continue d'aidants.
- Le GIR 2 comprend deux profils de personnes âgées : celles confinées au lit ou au fauteuil, dont les fonctions mentales ne sont pas totalement altérées et qui nécessitent une aide pour la plupart des activités de la vie courante, et celles dont les fonctions mentales sont altérées mais qui ont conservé leurs capacités à se déplacer.
- Le GIR 3 correspond pour l'essentiel, aux personnes âgées ayant conservé leurs fonctions mentales, partiellement leur capacité à se déplacer mais qui nécessitent plusieurs fois par jour des aides pour leur autonomie corporelle. La majorité d'entre elles n'assument pas seules l'hygiène de l'élimination anale et urinaire.
- Le GIR 4 comprend essentiellement deux profils de personnes. D'une part celles qui n'assument pas seules leurs transferts mais qui, une fois levées, peuvent se déplacer à l'intérieur du logement. Elles doivent parfois être aidées pour la toilette et l'habillage. La grande majorité d'entre elles s'alimente seule. D'autre part celles qui n'ont pas de problèmes pour se déplacer mais qu'il faut aider pour les activités corporelles ainsi que les repas.
- Le GIR 5 correspond aux personnes qui assurent seules leurs déplacements à l'intérieur de leur logement, s'alimentent et s'habillent seules. Elles nécessitent une aide ponctuelle pour la toilette, la préparation des repas et le ménage.
- Le GIR 6 regroupe toutes les personnes qui n'ont pas perdu leur autonomie pour les actes discriminants de la vie courante.

---

<sup>14</sup> La définition des GIR présentée ici est reprise du dictionnaire des codes de l'enquête HSM.