

# LES FAUX NÉGATIFS DU VOLET MÉNAGE DE L'ENQUÊTE HANDICAP-SANTÉ 2008

Gérard Bouvier (\*)

(\*) *Département des statistiques, des études et de la documentation, service statistique du secrétariat général à l'immigration et à l'intégration<sup>1</sup>.*

## Introduction

Le volet ménage de l'enquête Handicap Santé (HSM) a été collecté en surreprésentant les personnes présumées en mauvais état de santé ou en situation de handicap. La surreprésentation est rendue possible par une enquête filtre (VQS 2007, Vie quotidienne et santé). On enquête un grand nombre de personnes, pour les classer sommairement en groupes : en bonne santé (groupe I), ..., très handicapés (groupe IV). On s'intéresse alors au problème des « faux négatifs ». Un faux négatif (FN) est une personne qui répond négativement au critère « avez-vous un problème de santé ou de handicap » selon l'enquête filtre, mais que l'enquête elle-même va caractériser comme ayant ces problèmes (de santé ou handicap). Le principe de surreprésentation va attribuer un poids élevé à cette personne (car présumée en bonne santé). Si elle a un problème de santé ou de handicap, cela peut augmenter substantiellement l'effectif estimé des personnes ayant le même problème, en comparaison de ce que cette estimation aurait été si la personne avait été d'emblée classée comme handicapée, donc affectée d'un faible poids.

L'objet de la présentation est de :

- 1) Donner les éléments de contexte nécessaire ;
- 2) Donner une estimation du nombre de faux négatifs ;
- 3) Décrire sommairement ceux-ci dans les termes de l'enquête ;
- 4) Estimer indicativement l'incertitude que cela induit sur des indicateurs synthétiques.

## 1. Eléments de contexte

### 1.1. Vue générale du projet

L'enquête « handicap-santé » (HS) est un projet complexe. Les collectes ont été menées de 2007 à 2009. L'INSEE a assuré la co-maîtrise d'ouvrage du projet, avec la DREES<sup>2</sup>. Le groupe de conception associe des statisticiens à des chercheurs et des médecins spécialisés en épidémiologie, ainsi que des psychiatres. Ces personnes travaillent dans les organismes tels l'INED<sup>3</sup>, l'INSERM<sup>4</sup>, la CNAV<sup>5</sup>, le CTNERHI<sup>6</sup>, l'IRDES<sup>7</sup>, la DREES et la CNSA<sup>8</sup>, outre l'INSEE. Il y a eu des contributions des représentants des associations de personnes en situation de handicap (APF<sup>9</sup>, UNAFAM<sup>10</sup>, UNAPEI<sup>11</sup>...) et des représentants des financeurs de l'enquête. Le pilotage du projet assure la prise

<sup>1</sup> L'étude a été réalisée lorsque l'auteur était à l'INSEE, à la direction des statistiques démographiques et sociales. Elle a bénéficié de nombreux apports de Pascale Breuil et Liliane Lincot. L'annexe 3 est l'œuvre de Loïc Midy.

<sup>2</sup> Soit les bureaux FHD (famille-handicap-dépendance) ESP (état de santé des populations) à la DREES et l'unité UEDS à l'INSEE. Définition des bureaux et unités en vigueur début 2011.

<sup>3</sup> Institut national d'études démographiques.

<sup>4</sup> Institut national de la santé et de la recherche médicale.

<sup>5</sup> Caisse nationale d'assurance vieillesse.

<sup>6</sup> Centre technique national d'études et de recherches sur le handicap et l'inadaptation.

<sup>7</sup> Institut de recherche et de documentation en économie de la santé, anciennement CREDES (centre de recherche, d'études, de documentation en économie de la santé).

<sup>8</sup> Caisse nationale de solidarité pour l'autonomie.

<sup>9</sup> Association des paralysés de France.

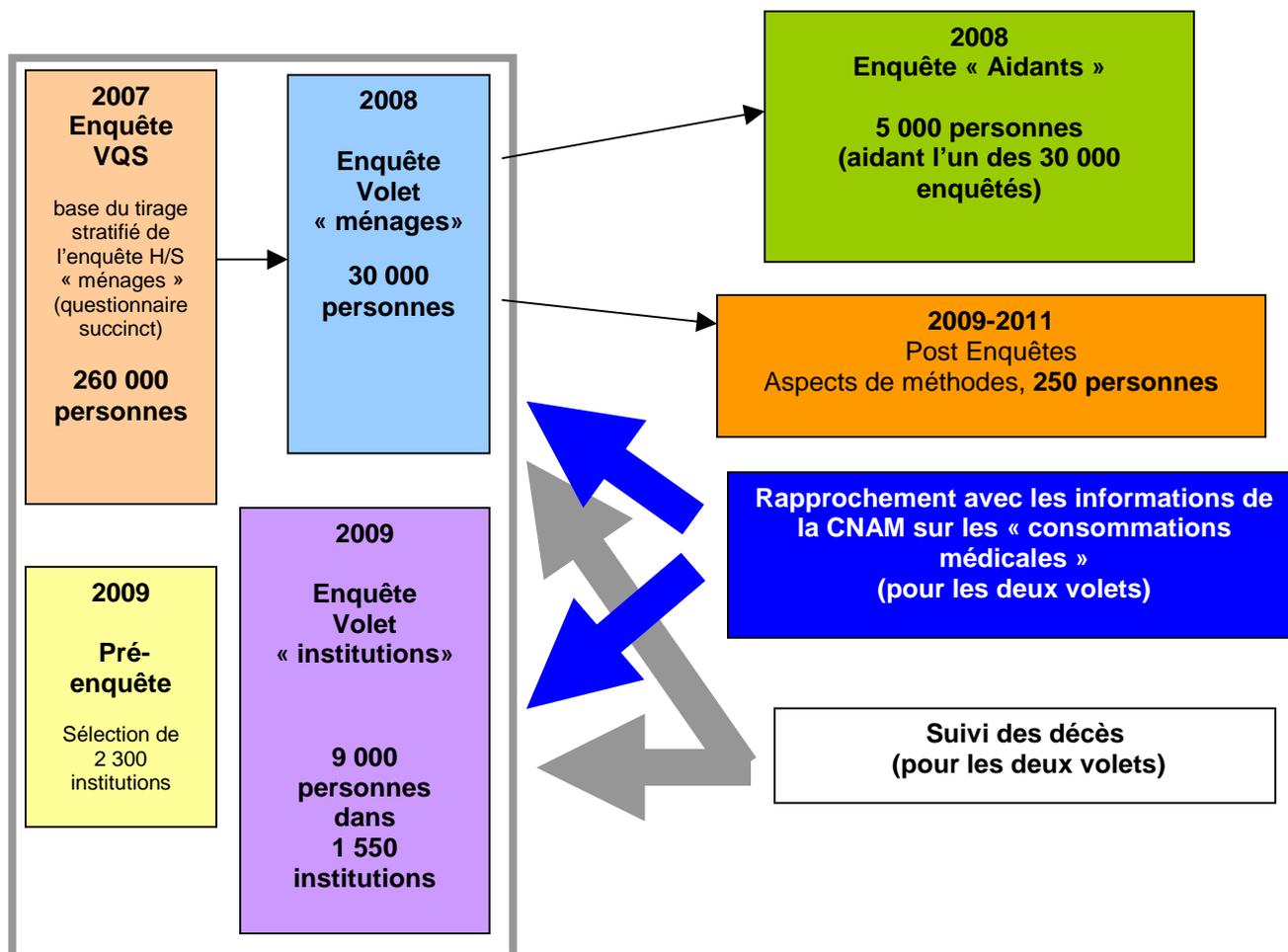
<sup>10</sup> Union nationale des amis et familles de malades psychiques.

<sup>11</sup> Union nationale des associations de parents, de personnes handicapées mentales et de leurs amis.

en compte de l'expérience de l'enquête « Handicap-Incapacités-Dépendance » (de 1998 à 2001, HID), un effet renforcé par la présence dans ce groupe de conception de chercheurs expérimentés ayant déjà participé au projet HID et bons connaisseurs des enquêtes Santé.

Le financement d'HS est assuré d'abord par l'INSEE puis la DREES, avec des participations de la CNSA, des CNAM<sup>12</sup>, CNAV et CNAF<sup>13</sup>, de l'AGEFIPH<sup>14</sup>, la FNMF<sup>15</sup> et l'InVS<sup>16</sup>. Comme pour HID, il y a une enquête auprès des personnes vivant en ménage et auprès des personnes vivant en institutions.

**Figure 1 : Schéma d'ensemble :**



Contrairement à HID, il n'y a pas de « deuxièmes vagues » ni d'enquête en prisons. Il y a par contre une enquête adossée « aidant » et un appariement avec les données de consommations médicales (voir exposé d'Alexis Montaut et de Lucie Gonzalez).

<sup>12</sup> Caisse nationale d'assurance maladie (maternité et accidents du travail) pour les travailleurs salariés.

<sup>13</sup> Caisse nationale des assurances familiales.

<sup>14</sup> Association de Gestion du Fonds pour l'Insertion Professionnelle des Personnes Handicapées.

<sup>15</sup> Fédération nationale de la mutualité française.

<sup>16</sup> Institut de Veille Sanitaire.

## 1.2. L'enquête "Vie quotidienne et santé 2007" (VQS)

A partir d'un questionnaire court, elle sert de filtre pour constituer la base de sondage de l'enquête handicap/santé en ménages ordinaires (HSM), afin d'être sûr de disposer d'un nombre suffisant de personnes concernées par les problèmes de handicap et de dépendance. Elle ne pouvait être adossée au recensement, contrairement à ce qui avait été fait en 1999. Le questionnaire est un peu plus étoffé que dans l'opération de 1999, ce qui permet de viser une stratification plus performante. Bien qu'initialement non conçue comme enquête autonome, elle aborde des questions non reprises dans l'enquête d'une part, et sa taille (260 000 répondants) permet certaines analyses jusqu'à un niveau régional<sup>17</sup>. Elle sera encore utilisée pour permettre des estimations locales via des méthodes de petits domaines (voir présentation de Josiane Le Guennec). Mais il convient d'insister : cette enquête est surtout un outil statistique pour assurer une bonne qualité à l'enquête en ménages.

Le questionnaire est envoyé à un échantillon de ménages, soit une feuille A4 avec 26 questions. Le retour par courrier est attendu. Il y a une phase de relance postale. Puis on recherche les numéros de téléphone des non-répondants, pour une collecte téléphonique. Si le numéro n'est pas retrouvé, un enquêteur (du réseau INSEE) est envoyé à l'adresse connue pour une interview « sous CAPI ». Enfin, un protocole particulier est mis en place dans les départements d'outre-mer. Un impact éventuel de ces différents modes de collecte a été étudié lors de l'exploitation de l'enquête pour la stratification, sans que l'on en relève un.

A partir des informations collectées, des analyses statistiques ont permis de proposer une classification en quatre<sup>18</sup> groupes, le premier (groupe VQS I) constitué de personnes « en bonne santé », suivant cette enquête, le quatrième (groupe VQS IV), *a contrario*, constitué de personnes déclarant de fortes limitations, des handicaps, des situations de dépendance, un mauvais état de santé, ... avec en outre, deux groupes concernant des personnes en situations intermédiaires.

Ces groupes constituent la base de sondage de l'enquête HSM : un processus de tirage au sort sélectionne les futurs enquêtés, soit toutes les personnes du groupe « IV », et des proportions d'autant moindres dans les autres groupes que la situation présumée de handicap est présumée moins sévère.

Interroger un nombre assez important de personnes « en bonne santé » répond à un double objectif, chacun essentiel pour assurer une bonne qualité d'enquête :

- 1) La dimension populationnelle du projet signifie qu'on ne saurait se contenter de décrire les problèmes de santé ou de handicap des personnes qui en ont. Il convient de mesurer des prévalences, voire des incidences, mais encore de permettre des comparaisons sur des indicateurs sociodémographiques entre personnes « en bonne santé » et « personnes en situation de handicap. Un questionnement commun des deux types<sup>19</sup> de personnes est donc nécessaire ;
- 2) Le résultat de la classification en groupe VQS est nécessairement grossier et ne peut être parfait. Il est donc certain que l'enquête filtre classera certaines personnes comme en bonne santé alors qu'elles ne le sont pas (« faux négatifs »). Inversement, on s'attend aussi à avoir des « faux positifs ». La présence de « faux négatifs » est plus gênante que celles de faux positifs, car s'il y a peu de « négatifs » dans l'échantillon interrogé, chacun va représenter beaucoup de personnes de l'ensemble de la population, donc une erreur a priori est susceptible de fausser très largement les valeurs des indicateurs d'intérêt<sup>20</sup>.

---

<sup>17</sup> Voir annexe 3.

<sup>18</sup> Soit une classification plus grossière en apparence que ce qui avait été fait pour HID (6 groupes).

<sup>19</sup> Le concept d'une population qui se diviserait en « deux types », avec / sans handicap est par ailleurs totalement irréaliste, non seulement, il convient d'appréhender les situations de handicap plutôt comme un continuum, mais encore faut-il remarquer que ce continuum est multidimensionnel, car la nature des problèmes de santé / handicap / dépendance est loin d'être univoque.

<sup>20</sup> Pour prendre un exemple fictif et extrêmement simplifié : Une enquête filtre détermine deux groupes : A (positifs : en mauvaise santé) et B (négatifs : en bonne santé). On souhaite sur-échantillonner les personnes du groupe A (enquête dédiée à une analyse détaillée des problèmes de santé des personnes). Dans le groupe A de 3 millions de personnes, on en sélectionne une sur mille soit 3000 ; dans le groupe B de 60 millions de personnes, on en sélectionne une sur 100 000 soit 600. Une prévalence est mesurée, qui concerne 10 % des personnes du premier groupe, donc 300 000 personnes dans la population réelle (soit environ 0,5 %). Une erreur dans le groupe B (« faux négatif ») va rajouter 100 000 personnes soit + 33%. Une erreur dans le groupe A (« faux positif ») va enlever 1000 personnes soit -0,3 % .

Un enseignement d'HID est de se prémunir contre une influence excessive des « faux négatifs » en interrogeant suffisamment de personnes en bonne santé.

### **1.3. L'enquête "Handicap Santé Ménages 2008" (HSM)**

Il s'agit de la partie « enquête auprès des personnes vivant dans un logement ordinaire » de l'enquête Handicap-Santé. Avec le volet institution, il s'agit du cœur du dispositif. Le questionnaire est volumineux : en format édité, celui-ci comporte 162 pages. Sa durée de passation prévue est d'une heure. C'est à peu près la durée moyenne relevée, mais cette moyenne cache des écarts importants, le temps pouvant augmenter considérablement si plusieurs maladies, déficiences, ... sont déclarées. L'enquêteur dépose un auto-questionnaire, que l'enquêté remplit et envoie par la poste. Comme cette enquête est longue et peut être délicate, la méthode d'échantillonnage limite à deux personnes maximum par ménage le nombre d'interviewés.

Dans l'ensemble, l'enquête est bien acceptée, même s'il est parfois difficile de convaincre les personnes en bonne santé que leur participation est importante. Pour les personnes ayant des problèmes de santé ou de handicap, on pouvait redouter que certaines questions soient considérées comme indiscrettes, intrusives, ... cela est en fait rarement arrivé. Au contraire, de nombreuses réactions sont positives, une grande confiance s'instaure le plus souvent entre les enquêtés et l'enquêteur. Le sentiment général est que « l'on » n'a pas assez d'informations sur ces sujets et que l'enquête est donc utile et bienvenue, surtout si elle permettra « de faire avancer les choses » et d'améliorer les conditions de vie ou la prise en charge des personnes handicapées. Certaines situations sont très lourdes et les enquêteurs en gardent des souvenirs forts. D'ailleurs, pour certains mêmes, les impressions d'HID, 10 ans après, sont encore vives....

## **2. Estimation du nombre de faux négatifs**

### **2.1. Introduction**

Le volet ménage de l'enquête Handicap Santé (HSM) a donc été collecté en surreprésentant les personnes présumées en mauvais état de santé ou en situation de handicap. Pour l'exploitation de l'enquête, le calcul des poids rétablit la représentativité de chaque individu, la surreprésentation autorisant une meilleure précision sur les indicateurs, voire la détection de phénomènes rares. La surreprésentation est faite à partir de l'enquête filtre Vie quotidienne et santé (VQS). On enquête d'abord un grand nombre de personnes, pour les classer sommairement en groupes : en bonne santé (groupe I) , .... , très handicapés (groupe IV). Il va de soi que l'enquête filtre est peu précise, donc que la classification est très imparfaite. On s'intéresse donc au problème des « faux négatifs ». Un faux négatif est une personne qui répond négativement au critère « avez-vous un problème de santé ou de handicap » selon l'enquête filtre, mais que l'enquête elle-même va caractériser comme ayant ces problèmes (de santé ou handicap). Le principe de surreprésentation va attribuer un poids élevé à cette personne (car présumée en bonne santé). Si elle a un problème de santé ou de handicap, cela peut augmenter substantiellement l'effectif estimé des personnes ayant le même problème, en comparaison avec ce que cette estimation aurait été si la personne avait été d'emblée classée comme handicapée, donc affecté d'un faible poids. Un problème symétrique de « faux positifs » existe mais retient moins l'attention : des personnes en mauvaise santé présumée, mais en réalité en bonne santé, ont un faible poids, donc une erreur dans leur classification suivant tel ou tel indicateur aura moins de conséquences.

L'enseignement de HID est que l'on se prémunit de l'impact des « faux négatifs » en augmentant le nombre de personnes enquêtées et présumées en bonne santé : cette augmentation diminue le poids de chacun des représentants de cette catégorie de population. Cette conclusion a été suivie pour HS. Toutefois, le problème théorique demeure.

### **2.2. Premières estimations**

En première approximation, un « faux négatif » peut traduire :

- 1) un changement réel de la situation de la personne, le filtre et l'enquête ayant eu lieu à neuf mois d'écart (environ) : on espère que ces transitions sont en nombre limité, notamment du fait de la proximité temporelle du filtre et de l'enquête ;
- 2) une incapacité du filtre à détecter la situation réelle de l'enquêté, soit que :
  - a) la situation n'est pas dans le champ du questionnaire du filtre (certaines difficultés psychiques, les limitations plus spécifiques du grand âge en sont des exemples) ;
  - b) la situation est dans le champ mais sa caractérisation échappe à l'examen rapide (filtre) mais va ressortir dans après le questionnement détaillé de l'enquête elle-même ;
  - c) sans exclure des incertitudes liées aux modes de questionnement (réponses par proxy, taille du questionnaire, rôle des enquêteurs, par exemple).

Ces effets sont présents, et on ne peut véritablement les estimer.

On connaît pour chaque enquêté HSM son groupe VQS. Ce groupe est déterminé<sup>21</sup> en utilisant toutes les questions de l'enquête filtre VQS. Il s'agit donc de déterminer « un groupe HSM ».

Le questionnaire VQS n'est pas repris dans le questionnaire HSM, on ne peut donc déterminer un groupe HSM en appliquant l'algorithme VQS aux réponses HSM.

Une première méthode : reconstitution de groupes VQS et HSM avec un algorithme tronqué

Cependant, quelques questions VQS sont reprises en questions HSM de façon très proche ou proche. C'est le cas des trois questions « du mini-module » européen, soit les indicateurs de santé déclarée, état général, état fonctionnel, maladies chroniques.

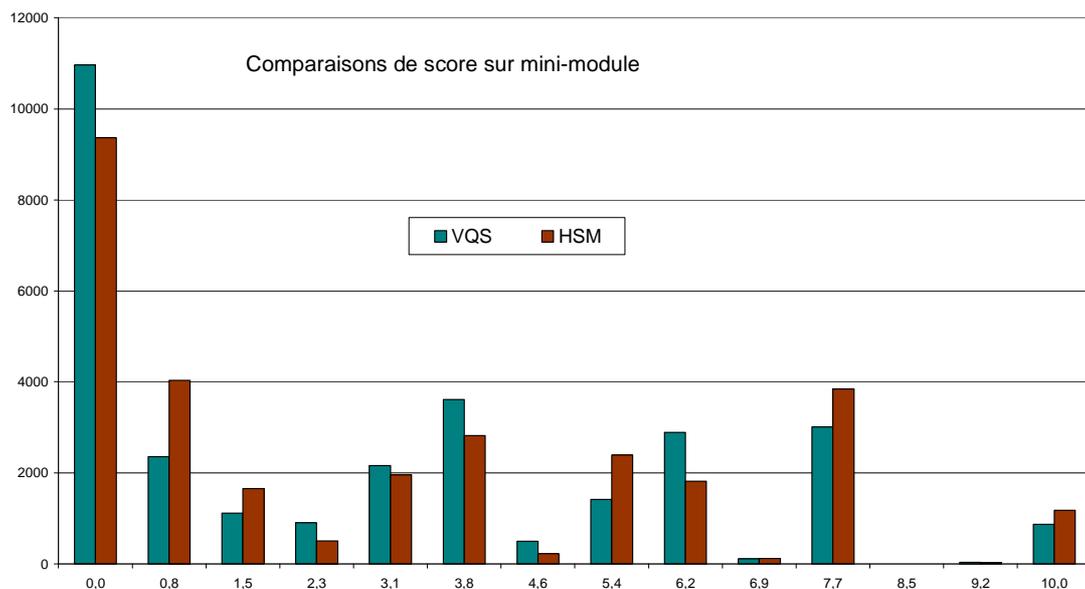
L'algorithme VQS se résume en deux étapes

- 1) Calcul d'un score ou d'une note : on part de 0 et chaque réponse apporte de 0 à 6 points ; on normalise, la note est ramenée entre 0 et 100.
- 2) Répartition en groupes : suivant la note et l'âge, on affecte l'individu dans un groupe VQS.

Pour tronquer, on procède de la même façon, mais le score ne tient compte que des 3 questions. On normalise entre 0 et 10.

La comparaison des notes est alors aisée :

**Figure 2 : Comparaison des notes VQS et HSM, calculées avec les trois questions « de santé déclarée ».**



<sup>21</sup> Voir également le document de travail sur VQS et l'annexe 3.

*Note de lecture : 11 000 enquêtés (HSM) ont une note de 0 (parfait état de santé) selon VQS contre seulement environ 9 400 selon HSM.*

*Source et champ : enquête Handicap-Santé, INSEE 2008-2009, volet ménages (29 931 enquêtés).*

On peut calculer la différence des deux notes. Il y a alors 82,2 % des enquêtés dont les notes ont un écart inférieur à 2,5, mais 2,1 % pour lesquels la note augmente de plus de 5.

Entre VQS et HSM, la note augmente de 0,15 en moyenne. C'est d'abord la conséquence d'une différence importante des réponses à la question concernant les maladies chroniques<sup>22</sup>, qu'on ne peut attribuer à une transition réelle (développement de maladies chroniques entre l'enquête VQS et l'enquête HSM).

On relève donc deux défauts immédiats de cette approche :

- 1) Dans quelle mesure la troncature fausse-t-elle la répartition par groupes VQS ;
- 2) Dans quelle mesure le changement des réponses entre VQS et HSM est-il dû à un effet réel ou à une imperfection du filtre ?

### **2.3. La méthode retenue**

D'où un essai plus sophistiqué : imputation économétrique de groupes VQS et HSM

L'approche initiale est reprise avec le souci d'utiliser le mieux possible les questions communes à VQS et HSM.

Comme ces questions communes ne peuvent englober tout le questionnaire VQS, on doit donc reconstituer les groupes VQS avec un nombre réduit de questions. Plutôt que de reprendre l'algorithme en deux étapes, on détermine directement le groupe VQS à l'aide d'une régression logistique généralisée : à partir d'une partie du questionnaire VQS, la régression permet d'estimer les paramètres à affecter aux modalités des réponses pour prévoir le mieux possible le groupe VQS.

On dispose donc :

- Du « vrai » groupe VQS, tel qu'il a été calculé pour la stratification ;
- Du groupe VQS « prédit » par régression logistique généralisée ;

Enfin, en effectuant la « prédiction » avec les réponses HSM au lieu des réponses VQS, on obtiendra un groupe HSM. Ce dernier groupe dépend du choix de la partie du questionnaire VQS utilisée.

### **2.4. Questions de questionnaire**

Le questionnaire VQS comprend 26 questions.

Les deux premières concernent le sexe et l'âge, ce dernier intervient dans le calcul des groupes et sera donc introduit dans les modèles. Elles sont reprises dans HSM.

Les trois suivantes déclinent le mini-module de santé. Elles sont reprises dans HSM, de façon très proche ou proche (la question sur la limitation comprend une référence à une période de six mois dans HSM mais pas dans VQS). Son rôle étant de fournir une vue synthétique de l'état de santé, le premier modèle intègre donc l'âge et ces trois questions.

Les quatorze questions suivantes explorent des limitations fonctionnelles. Ces limitations sont également explorées dans le module « E » de HSM, mais également dans le module F (restrictions d'activités). Elles sont complétées par une question de synthèse sur les limitations absolues.

Neuf d'entre elles sont reprises de façon proche (formulations un peu différentes, mention ou non de l'appareil auditif, ...), d'où un second modèle intégrant ces neuf questions aux quatre précédentes. Quatre d'entre elles sont vraisemblablement reconstituables assez directement : par exemple la question sur la capacité à se déplacer et utiliser un escalier dans VQS, correspond à deux questions dans HSM. On peut encore reconstruire la question de synthèse. Par contre une question sur la capacité à prendre des initiatives ne paraît pas reconstituable avec le questionnaire HSM.

Cinq questions appartiennent au champ des aspects environnementaux. Aucune n'a d'homologue dans HSM. Cependant, une reconstruction peut être envisagée pour :

L'aide humaine ;

---

<sup>22</sup> Voir annexe 2

L'aménagement du logement ;  
L'aide technique ;  
La reconnaissance officielle du handicap ;  
La fréquentation (pour les 0-18 ans) d'une classe spécialisée.

Une question porte enfin sur le handicap ressenti. Elle ne paraît pas reconstituable avec le questionnaire HSM.

## 2.5. Choix des variables, construction du modèle

Différents groupes de variables ont été testés à partir de l'analyse du questionnaire VQS présentée. Le choix de base reprend simplement l'âge et le mini-module, les résultats sont détaillés en annexe 1. L'accord entre VQS et HSM est assez bon :

- 1) Les questions sont assez proches dans leurs formulations ;
- 2) Les résultats sont également assez proches pour les deux questions « état général de santé » et « limitation » ;
- 3) Mais on a déjà mentionné qu'il y a beaucoup plus de personnes qui déclarent une maladie chronique dans HSM (39 % contre 25 %).

On a donc déterminé le groupe HSM avec des parties de questionnaires VQS omettant cette dernière question.

Le choix final comprend finalement :

L'âge ;  
Les trois questions du mini-module ;  
Les neuf questions sur limitations « quasi-communes » à VQS et HSM ;  
Une question d'environnement : l'aide au logement.

L'arbitrage est le suivant : pour obtenir le meilleur accord entre la prédiction du groupe VQS et le groupe VQS réel, il faut naturellement prendre toutes les variables VQS, à défaut le plus possible. Mais le choix des variables impacte l'estimation de la transition, c'est-à-dire le changement de groupe (prédit) entre VQS et HSM. On a relevé les effets suivants :

Même question, mêmes réponses : cette situation est en fait rare : les deux questions du mini-module, hors maladies chroniques.

Même question (ou presque), réponses proches : c'est le cas pour sept limitations, toutefois, il y a la plupart du temps davantage de répondants dans la modalité intermédiaire (quelques difficultés) et un peu moins dans la modalité négative (beaucoup de difficultés) selon VQS. On doit relever encore qu'il y a une différence de formulation et dans les modalités de réponse : pouvez faire .... dans HSM + modalité « ne peux pas du tout » / avez vous des difficultés pour faire .... dans VQS.

Mêmes questions, réponses un peu différentes : on a déjà cité la question sur les maladies chroniques. Le fait que plus de personnes selon HSM en déclarent va augmenter les personnes en transitions, dans le sens d'une augmentation du groupe. Inversement, pour deux limitations, « vue près » et « main », il y a plus de personnes en difficultés selon VQS que selon HSM<sup>23</sup>. Enfin la question sur les aménagements du logement peut relever de cette catégorie, sachant que davantage de personnes en déclarent selon HSM que VQS. On peut attribuer cela au fait que le questionnement est en fait un peu différent : question simple dans VQS, présentation d'une liste d'aménagements dans HSM, d'où un effet de liste probable<sup>24</sup>.

Questions VQS à reconstruire pour faire un pendant HSM :

Il y a deux exemples simples, la limitation déjà évoquée<sup>25</sup> et la question sur les aides techniques. Pour cette dernière, on observe une forte augmentation de la proportion des personnes en déclarant selon

<sup>23</sup> Le groupe HSM a donc été prédit avec une partie de questionnaire omettant ces deux variables. Les conclusions (environ 100 faux négatifs) ne sont pas remises en cause.

<sup>24</sup> S'appuyer sur une liste lors de l'enquête permet de s'assurer que l'enquêté pense à tous les items souhaités par le concepteur de l'enquête. Cette technique d'enquête plus coûteuse en temps permet un recueil d'information plus précis. Elle conduit en général à recueillir plus d'items qu'une question formulée sans liste.

<sup>25</sup> La question sur la capacité à se déplacer et utiliser un escalier dans VQS.

HSM. L'effet de liste est très probable puisqu'une carte de 33 aides est présentée, puis quelques questions sont posées sur des aides particulières. De plus, la formulation de la question VQS n'évoque pas comme aides techniques ce qui est présenté dans HSM comme aides techniques pour les soins, la protection et les traitements. On a donc reconstruit la question d'aide avec différents choix dans la liste des aides HSM. Le premier constat est cohérent, le nombre de personnes déclarant une aide technique est croissant avec la taille de la liste. Le second est que le choix des aides est assez vite arbitraire, le dernier que le nombre de faux négatifs change assez peu suivant les différents choix. Tout ceci plaide donc pour ne pas retenir cette question et plus généralement ne pas intégrer d'autres reconstructions de questions VQS, encore plus hasardeuses, sur les autres variables.

En effet, il apparaît clairement aussi dans les modèles testés que l'on a bien deux effets antagonistes sur la transition entre groupe prédit VQS et groupe prédit HSM :

Si l'on intègre davantage de questions, on augmente la probabilité d'avoir quelques changements par questions et donc la probabilité d'une transition d'un groupe. Mais on diminue la probabilité d'avoir des changements sur beaucoup ou toutes les questions, donc la probabilité d'avoir un changement radical de groupe (échange entre groupes I et IV pour le problème qui nous intéresse).

Finalement le modèle retenu réalise un compromis entre :

- 1) Effectuer une bonne prédiction du groupe VQS réel ;
- 2) Se restreindre aux questions VQS ayant un pendant HSM ;
- 3) Mais équilibrer (un peu) le choix du sous-questionnaire entre questions sur l'état fonctionnel (limitations) et l'environnement (introduction de la question sur les aménagements) ;
- 4) Équilibrer les questions avec écarts de réponse entre VQS et HSM (deux dans chaque sens).

Il va de soi que les faux négatifs reconnus par cette étude dépendent du choix de la partie du questionnaire retenue. Il est important de noter que leur nombre est toujours proche de 100.

## 2.6. La prédiction du "groupe VQS réel"

Cette prédiction peut être considérée comme très satisfaisante : en effectifs pondérés, il y a 0,25 % d'erreurs de deux groupes ou plus, et 89,52 % de prédictions parfaites. Il n'y a que 7 enquêtés pour lesquelles l'erreur est maximale (échange des groupes I et IV).

Figure 3 : Comparaison détaillée des groupes VQS "réel" et "prédit".

Prédiction : observations						Prédiction : pondéré					
		Groupe VQS réel						Groupe VQS réel			
		1	2	3	4			1	2	3	4
Groupe VQS modèle	1	10 116	712	123	5	Groupe VQS modèle	1	44 263	1 171	86	1
	2	353	5 096	1 332	92		2	1 502	7 553	867	30
	3	1	1 270	3 698	1 002		3	14	1 872	2 312	325
	4	2	13	1 034	5 082		4	6	19	650	1 786

Note de lecture : Il y a 10 116 enquêtés qui sont à la fois dans le groupe I « réel » et le groupe I « prédiction du modèle ». Ils sont 353 à être dans le groupe I mais pour lesquels le modèle les prédit en groupe II. Les effectifs pondérés correspondants sont en milliers et de 44 263 puis 1 502 respectivement.

Source : Enquête Handicap-santé 2008, volet ménage.

## 2.7. Les changements entre groupe VQS prédit et groupe HSM prédit (transitions)

Si on regarde maintenant les changements entre les groupes VQS prédits et les groupes HSM prédits (à partir des modèles), les transitions peuvent paraître nombreuses, puisque 39,3 % des enquêtés changent de groupe. Cependant ils ne sont que 7,8 % à changer de plus de deux groupes dont 1,2 %

à échanger groupes I et IV. En effectifs pondérés, ces proportions sont respectivement de 29,0 %, 5,5 % et enfin 0,8 % .

Une étude des transitions avec séparations d'effets de transitions véritables et d'effets statistiques n'a pas été tentée. On s'intéresse seulement aux « faux négatifs », soit plus particulièrement aux 140 enquêtés, représentant 420 000 personnes qui apparaissent comme en groupe I selon le modèle VQS et groupe IV selon le modèle HSM.

**Figure 4 : Comparaison détaillée des groupes VQS "prédit" et HSM "prédit".**

Transitions intra-modèle : observations						Transitions intra-modèle : pondéré					
		Groupe HSM (modèle)						Groupe HSM (modèle)			
		1	2	3	4			1	2	3	4
Groupe VQS modèle	1	8 481	1 869	466	140	Groupe VQS modèle	1	36 432	7 090	1 579	420
	2	2 013	3 114	1 369	377		2	3 422	4 455	1 697	377
	3	733	1 099	2 722	1 417		3	789	943	2 006	786
	4	210	414	1 669	3 838		4	109	190	701	1 461

*Note de lecture : Il y a 1 869 enquêtés qui sont dans le groupe I selon le modèle VQS et dans le groupe II selon le modèle HSM. Ils sont 3 114 à être dans le groupe II selon les deux modèles. Les effectifs pondérés correspondants sont en milliers de 7 090 puis 4 455 respectivement.*

*Source : INSEE, Enquête Handicap-santé 2008, volet ménage.*

## 2.8. Comparaisons finales : écarts entre groupe VQS réel et groupe reconstitué selon l'enquête HSM :

L'écart observable<sup>26</sup> se définit donc par la différence entre le groupe VQS réel et le groupe HSM prédit. Il résulte de l'erreur de prédiction (entre groupes VQS réel et prédit) combinée avec le changement entre groupe prédits. On s'intéresse surtout aux enquêtés qui sont en groupe VQS réel I (négatifs) et en groupe HSM prédit IV (faux négatifs). Pour le choix retenu, il y a 98 « faux négatifs », qui correspondent en effectifs pondérés à 385 000 personnes environ. L'analyse sommaire des transitions observables donne :

60,2 % d'accord entre groupe VQS (réel) et groupe HSM (prédit) en enquêtés, soit 71,1% des personnes du champ en effectifs pondérés ;

32,8 % d'enquêtés (donc un total avec les enquêtés précédents de 93,0 %) pour lesquels le groupe HSM diffère d'une unité, soit 24,0 % (total = 95,1%) en effectifs pondérés ;

6,2 % pour lesquels le groupe HSM diffère de deux unités (4,3 % en effectifs pondérés) ;

Enfin un solde de 0,8 % (0,6 % en effectifs pondérés) d'enquêtés qui échangent de façon extrême leur groupe.

**Figure 5 : Comparaison détaillée des groupes VQS "réel" et HSM "prédit".**

Transitions observables : observations						Transitions observables : pondéré					
		Groupe HSM (modèle)						Groupe HSM (modèle)			
		1	2	3	4			1	2	3	4
Groupe VQS réel	1	8 247	1 704	423	<b>98</b>	Groupe VQS réel	1	36 660	7 115	1 625	<b>385</b>
	2	2 300	3 085	1 398	308		2	3 547	4 593	2 056	419
	3	743	1 332	2 710	1 402		3	495	844	1 732	845
	4	147	375	1 695	3 964		4	50	126	570	1 395

*Note de lecture : Il y a 98 enquêtés qui sont dans le groupe VQS (réel) I et qui sont dans le groupe HSM (modélisé) IV. Cela correspond à 385 000 personnes du champ.*

*Source : INSEE, Enquête Handicap Santé 2008, volet ménage.*

<sup>26</sup> Par opposition à l'écart réel qui existerait entre le groupe VQS réel et le groupe HSM réel.

### 3. Éléments sur les faux négatifs

Il y a donc 98 faux négatifs, définis ici comme des personnes sans problèmes de santé important selon VQS (i.e. groupe I pour VQS) se révélant avoir des problèmes de santé important selon HSM (groupe IV de HSM). Un premier constat est qu'ils ne sont pas d'âge supérieur à 60 ans, ou très peu : 59 d'entre eux ont entre 30 et 59 ans, 33 ont moins de 30 ans. Il y a un peu plus d'hommes que de femmes, 54 contre 44.

#### 3.1. Faux négatifs « généralisés »

Une analyse des déterminants du statut de « faux négatif » a été menée. Dans un premier temps, on définit plus largement le statut de faux négatif, car leur effectif n'est pas assez grand. On s'intéresse donc aux personnes qui changent d'au moins deux groupes (avec dégradation de l'état de santé).

Soit 829 personnes :

423 personnes qui passent du groupe VQS = 1 au groupe HSM = 3 ;

308 personnes qui passent du groupe VQS = 2 au groupe HSM = 4 ;

Les 98 personnes véritablement « faux négatifs », qui passent du groupe VQS = 1 au groupe HSM = 4.

Dans un premier temps, on a cherché (statistiques descriptives, régressions logistiques) à distinguer ces personnes par des caractéristiques sociodémographiques : sexe, âge, présence d'un conjoint, catégorie sociale, niveau de diplôme, lieu d'habitation. Il y a un effet d'âge, notamment parce qu'il y a peu de faux négatifs âgés (de plus de 60 ans). L'effet le plus sensible est l'habitat : DOM (influe positivement) contre métropole. Cependant, les modèles construits sont peu robustes et n'expliquent pas très bien le phénomène étudié.

Dans un second temps, on a introduit des indicateurs (assez synthétique, de santé, altérations fonctionnelles, besoin d'aides techniques) dans les modèles explicatifs. D'une part, ces modèles sont le plus souvent de bien meilleure qualité. D'autre part, l'effet de l'indicateur de santé fonctionnelle apparaît très fortement, dans une moindre mesure l'indicateur de santé déclarée, enfin assez significativement et de façon comparable à l'effet DOM, l'indicateur de maladie chronique et de recours aux aides à la mobilité.

Dans un dernier temps, on insère dans les modèles une variable indicatrice de la transition de la modalité la plus favorable selon VQS à la modalité la plus défavorable selon HSM. Les modèles obtenus en sont encore améliorés et les faux négatifs sont fortement déterminés par les transitions « extrêmes » pour les variables du mini-module, et en premier lieu la variable d'état de santé fonctionnelle.

Ces résultats semblent paradoxaux : c'est davantage sur les limitations physiques qu'il y a des écarts entre VQS et HSM, bien que l'on ait initialement considéré que les « problèmes » cognitifs seraient plus difficiles à détecter avec le filtre. La première piste d'explication part du constat qu'HSM comprend un questionnement très détaillé sur les déficiences, limitations, ... , notamment physiques qui a permis de repérer finement un grand nombre de situations. C'est ainsi d'abord probablement un défaut de qualité dans les réponses à l'enquête filtre. Mais, il y a un peu moins de 3 % de personnes avec fortes transitions sur des questions assez générales. Ce qui est donc assez peu.

### 4. Quelques analyses à posteriori

Il s'agit ici d'apprécier l'impact possible des faux négatifs.

#### 4.1. Première approche : en exploitation d'enquête

En termes d'impacts (exploitation de l'enquête), on donne pour neufs indicateurs binaires les trois valeurs (en effectifs pondérés) :

Part des personnes pour lesquelles l'indicateur signale un mauvais état de santé ou un handicap (« Tous ») ;

Part (non pondérée mais ces personnes sont toutes dans le groupe VQS I, donc de poids a priori proches) des enquêtés pour lesquels l'indicateur signale un mauvais état de santé ou un handicap (« FN » comme faux négatifs) ;

Part des personnes pour lesquelles l'indicateur signale un mauvais état de santé ou un handicap (« Tous ») après correction grossière du poids : le poids est divisé par 13,5, qui est le ratio des probabilités de tirage entre groupe VQS IV et groupe VQS I, hors tous ajustements<sup>27</sup> («Corrigé»).

Les indicateurs retenus sont encore les indicateurs synthétiques pour une première exploitation de l'enquête, soit :

Santé (SD\_EG, SD\_LI, SD\_CH) soit les indicateurs de santé du mini-module, binarisés ;  
 Handicap déclaré, altérations de l'état fonctionnel, de nature motrices, sensorielles ou cognitives (AEF\_M, AEF\_C, AEF\_C) ;  
 Handicap déclaré, usage ou besoin d'aides techniques, soit aides de types prothèses ou appareillages (AT\_PA), aides aux fonctions motrices ou sensorielles (AT\_MS), aides pour soins, protection et traitements (AT\_ST).

Un bilan rapide peut être le suivant : la présence de faux négatifs conduit à surestimer la part des personnes en situation de handicap, mais cette surestimation est faible, moins de 5 % en écart relatif. Au regard de la précision de l'enquête, cette erreur est donc au plus peu importante.

**Figure 6 : Une estimation grossière de l'impact des faux négatifs.**

	Taux		Ecart (relatif)	Effectif FN
	Officiel	Corrigé		
SD_EG	24,70	24,42	-1,1	74
SD_LI	21,48	21,10	-1,8	89
SD_CH	38,55	38,29	-0,7	85
AEF_M	7,97	7,79	-2,3	39
AEF_S	4,49	4,40	-2,0	21
AEF_C	5,33	5,14	-3,6	37
AT_PA	6,00	5,88	-2,0	27
AT_MS	7,00	6,99	-0,1	17
AT_ST	4,81	4,81	0,0	5

*Note de lecture : 24,70 % des personnes du champ de l'enquête (personnes vivant en France, métropole ou DOM et en ménages ordinaires) déclarent un mauvais état de santé. C'est le cas de 74 % des faux négatifs. Une correction de la pondération ramène cette part (de la population totale) à 24,42 %, soit une diminution relative de 1,1 %.*

*Source : INSEE, Enquête Handicap Santé 2008, volet ménage.*

## 4.2. Deuxième approche : comparaison des poids et des indicateurs

On a examiné la répartition des répondants suivant la modalité de réponse à certaines questions et suivant leur poids. En principe, il doit y avoir correspondance entre une modalité signalant un mauvais état de santé et des poids faibles. On donne quelques exemples, puis on indique pourquoi l'analyse proposée ne doit pas être considérée comme signalant une mauvaise pondération. En effet, les caractéristiques des personnes « en mauvaise santé » et « de poids (assez) élevé » font qu'elles sont probablement assez représentatives d'un grand nombre de personnes. De plus, la dispersion des poids ne résulte pas seulement de la stratification issue de l'enquête filtre.

Exemple 1 : les questions du mini-module

Le poids médian est de l'ordre de 2 000, car il y a environ 60 000 000 personnes dans le champ et 30 000 personnes interrogées. Le tableau est fait en ne tenant compte que de l'échantillon national pour ne pas avoir à tenir compte de la dispersion induit par les extensions départementales de l'enquête. Le poids médian s'élève alors à un peu plus de 3 000.

Si l'on considère le premier indicateur (état de santé général, en cinq modalités), il est satisfaisant de constater l'étagement des poids médians. Néanmoins, la dispersion est importante, d'autant plus que

<sup>27</sup> Voir document de travail sur les pondérations d'HSM, pour plus de détail.

la modalité est proche de la modalité « intermédiaire ». Enfin, l'analyse des queues de distributions montre par exemple qu'environ 5 % des personnes déclarant un très mauvais état de santé ont un poids supérieur ou égal au poids médian des personnes déclarant la modalité « moyenne ». L'analyse faite à partir des deux autres questions conduit à des conclusions similaires.

**Figure 7 : Correspondance entre poids et modalités de réponse aux questions du mini-module.**

Etat de Santé		Effectifs	Poids m	Q3/Q1	D9/D1	D9	V19	Max
	T.Bon	3 947	5 376	1,96	4,71	8 737	9 924	17 388
	Bon	4 924	3 872	3,55	10,95	7 531	8 732	14 721
	Moyen	4 317	2 121	3,86	12,58	5 144	6 443	13 784
	Mauvais	3 170	1 089	2,76	6,86	2 204	3 075	14 039
	T.Mauvais	712	816	2,01	4,49	1 484	2 142	7 651
Limitation								
	Beaucoup	4 341	986	2,56	6,19	1 966	2 822	12 563
	Un peu	3 718	1 944	3,49	11,34	4 694	6 099	14 039
	Non	9 028	4 650	2,66	6,55	8 177	9 393	17 388
Maladie chronique								
	Oui	10 334	2 018	4,59	14,27	5 289	6 644	14 512
	Non	6 754	4 831	2,50	7,06	8 436	9 652	17 388
<b>Ensemble</b>		<b>17 070</b>	<b>3 133</b>	6,39	16,19	7 107	8 415	17 388

*Note de lecture : 3 947 personnes se sont déclarées en très bonne santé, leur poids médian étant de 5 376. Le seuil du dernier vingtile est (pour ces 3 947 personnes) de 9 924.*

Exemple 2 : limitations fonctionnelles physiques ou "cognitives".

On s'intéresse toujours aux personnes ayant un poids élevé. On explore ici les limitations fonctionnelles, d'abord "physique", puis "cognitives". On construit un indicateur grossier : d'abord un score, on somme le nombre de modalités « quelques » et trois fois le nombre de modalités « beaucoup ». Cela permet de constituer 5 groupes : du groupe I, peu de limitations, et pas très fortes, au groupe V, beaucoup de limitations, voire d'impossibilités.

Les cinq groupes sont d'effectifs à peu près égaux dans l'enquête.

On compare la répartition des personnes avec l'altération de l'état fonctionnel correspondant, sans cette altération et enfin les personnes dans le vingtile (pour la répartition des poids) suivant les groupes ainsi constitués

Pour la correspondance entre les altérations et les limitations motrices, la cohérence est bonne : 1 % seulement des personnes avec altération ne sont pas dans le groupe des personnes ayant peu de limitations. Mais il y a des personnes avec poids élevés dans les groupes 1 et 2, davantage que dans le cas des personnes déclarant une altération (71 % contre 55 %).

Ces correspondances sont un peu moins bonnes entre altérations et limitations cognitives.

On complète cette analyse indicative en examinant les déficiences déclarées par les personnes avec poids élevé et altérations. Il ressort alors que ces personnes déclarent souvent plusieurs déficiences. De plus ces déficiences déclarées sont peut-être des déficiences pas forcément très lourdes<sup>28</sup>. Ces personnes sont âgées (75 % ont plus de 60 ans, 25 % ont plus de 80 ans) dans le cas des altérations / déficiences motrices.

L'âge est mieux réparti dans le cas des altérations / déficiences cognitives.

On semble ainsi confirmer que le filtre n'a pas toujours pu classer correctement des personnes dont l'état fonctionnel moteur est dégradé car on ne peut bien faire la part de ce qui relève d'un aspect

<sup>28</sup> Gènes articulaires, force limitée, troubles de l'équilibre, difficultés de vision (mais pas aveugle, ...), troubles de mémoire, troubles de l'humeur, troubles anxieux. On suppose donc que ces déficiences sont moins "graves" que d'autres (exemple : paralysie). Il faut toutefois noter qu'un libellé comme « gêne dans les articulations » peut recouvrir des situations diverses, graves comme plus bénignes.

santé, et de ce qui relève de l'âge. Et il n'a pu classer certaines personnes ayant des "problèmes" cognitifs.

D'un point de vue de l'exploitation, le calage sur la pyramide des âges assure en principe une représentativité correcte.

**Figure 8 : Correspondance entre altérations et limitations.**

	Limitations motrices, par groupes				
	1	2	3	4	5
Avec	29	26	18	16	11
Sans	99	1	0	0	0
Vingtile	41	30	15	9	5
	Limitations cognitives, par groupes				
	1	2	3	4	5
Avec	25	18	20	17	20
Sans	97	2	1	0	0
Vingtile	57	20	15	6	1

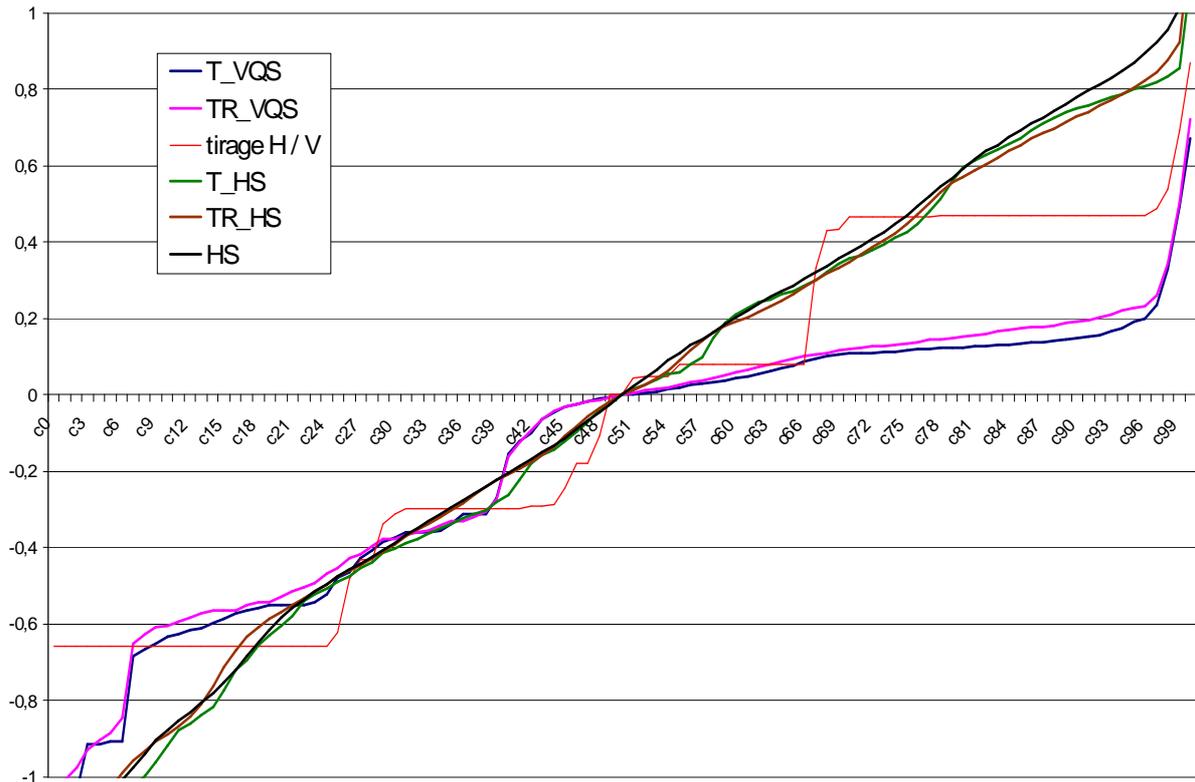
*Note de lecture : 29 % des personnes ayant une altération fonctionnelle motrice sont dans le groupe 1 pour les limitations motrices (99 % sont dans ce groupe pour les personnes sans cette altération).*

### 4.3. Remarques sur la dispersion des poids

Il faut en effet relever que la dispersion des poids n'est pas seulement due à la stratification en groupe "VQS".

Une dispersion est également induite du fait des extensions départementales, mais aussi parce que l'échantillon du filtre est fait à partir de l'enquête annuelle de recensement (2006).

**Figure 9 : Dispersion des poids.**



*Note de lecture : On représente le logarithme du ratio  $p / p50$ ,  $p$  étant le seuil du nième centile et  $p50$  le poids médian.*

*Courbe bleue : tirage pour pré-enquête : essentiellement les écarts entre départements à extensions et les autres.*

*Courbe rose : effet du redressement (VQS) : lissage*

*Courbe rouge : tirage pour enquête dans les répondants de la pré-enquête : essentiellement les 4 groupes, quelques corrections.*

*Courbe verte : combinaison des précédentes : représentation du poids « initial » du volet ménage.*

*Courbe brune : effet du redressement, lissage.*

*Courbe noire : poids redressés et calés, lissage accentué.*

## 5. Annexes

### 5.1. Annexe 1 : Détails du modèle « réduit » pour la détermination des faux négatifs

Modèle réduit : modélise le groupe VQS avec les trois variables du mini module

**Prédiction : observations**

		Groupe VQS réel			
		1	2	3	4
Groupe VQS modèle 1	1	10 080	1 949	545	79
	2	253	3 510	1 254	336
	3	137	1 407	2 757	1 263
	4	2	225	1 631	4 503

Modèle =réel	20 850
Modèle =réel +/- 1	7 757
Modèle =réel +/- 2	1 243
Reste	81
<b>TOTAL</b>	<b>29 931</b>

		Groupe VQS réel			
		1	2	3	4
Groupe VQS modèle 1	1	33,7	6,5	1,8	0,3
	2	0,8	11,7	4,2	1,1
	3	0,5	4,7	9,2	4,2
	4	0,0	0,8	5,4	15,0

Modèle =réel	69,7
Modèle =réel +/- 1	25,9
Modèle =réel +/- 2	4,2
Reste	0,3
<b>TOTAL</b>	<b>100</b>

**Prédiction : pondéré**

		Groupe VQS réel			
		1	2	3	4
Groupe VQS modèle 1	1	44 344	3 168	386	23
	2	939	5 028	764	109
	3	496	2 119	1 746	420
	4	6	299	1 019	1 589

Modèle =réel	52 707
Modèle =réel +/- 1	8 429
Modèle =réel +/- 2	1 291
Reste	29
<b>TOTAL</b>	<b>62 457</b>

		Groupe VQS réel			
		1	2	3	4
Groupe VQS modèle 1	1	71,0	5,1	0,6	0,0
	2	1,5	8,1	1,2	0,2
	3	0,8	3,4	2,8	0,7
	4	0,0	0,5	1,6	2,5

Modèle =réel	84,4
Modèle =réel +/- 1	13,5
Modèle =réel +/- 2	2,1
Reste	0,0
<b>TOTAL</b>	<b>100</b>

**Transitions intra-modèle : observations**

		Groupe HSM (modèle)			
		1	2	3	4
Groupe VQS modèle 1	1	9 948	1 459	890	356
	2	1 141	2 588	899	725
	3	905	579	2 167	1 913
	4	325	412	1 281	4 343

VQS=HSM	19 046
VQS=HSM +/- 1	7 272
VQS=HSM +/- 2	2 932
Reste	681
<b>TOTAL</b>	<b>29 931</b>

		Groupe HSM (modèle)			
		1	2	3	4
Groupe VQS modèle 1	1	33,2	4,9	3,0	1,2
	2	3,8	8,6	3,0	2,4
	3	3,0	1,9	7,2	6,4
	4	1,1	1,4	4,3	14,5

VQS=HSM	63,6
VQS=HSM +/- 1	24,3
VQS=HSM +/- 2	9,8
Reste	2,3
<b>TOTAL</b>	<b>100</b>

**Transitions intra-modèle : pondéré**

		Groupe HSM (modèle)			
		1	2	3	4
Groupe VQS modèle 1	1	39 681	4 628	2 756	856
	2	1 776	3 412	997	656
	3	1 139	456	1 948	1 239
	4	230	211	639	1 835

VQS=HSM	46 876
VQS=HSM +/- 1	9 733
VQS=HSM +/- 2	4 762
Reste	1 086
<b>TOTAL</b>	<b>62 457</b>

		Groupe HSM (modèle)			
		1	2	3	4
Groupe VQS modèle 1	1	63,5	7,4	4,4	1,4
	2	2,8	5,5	1,6	1,1
	3	1,8	0,7	3,1	2,0
	4	0,4	0,3	1,0	2,9

VQS=HSM	75,1
VQS=HSM +/- 1	15,6
VQS=HSM +/- 2	7,6
Reste	1,7
<b>TOTAL</b>	<b>100</b>

**Transitions observable : observations**

		Groupe HSM (modèle)			
		1	2	3	4
Groupe VQS réel	1	8 563	1 066	657	<b>186</b>
	2	2 541	2 503	1 292	755
	3	887	994	2 027	2 279
	4	328	475	1 261	4 117

Modèle =réel	17 210
Modèle =réel +/- 1	9 433
Modèle =réel +/- 2	2 774
Reste	514
<b>TOTAL</b>	<b>29 931</b>

		Groupe VQS réel			
		1	2	3	4
Groupe VQS réel	1	28,6	3,6	2,2	0,6
	2	8,5	8,4	4,3	2,5
	3	3,0	3,3	6,8	7,6
	4	1,1	1,6	4,2	13,8

Modèle =réel	57,5
Modèle =réel +/- 1	31,5
Modèle =réel +/- 2	9,3
Reste	1,7
<b>TOTAL</b>	<b>100</b>

**Transitions observable : pondéré**

		Groupe VQS réel			
		1	2	3	4
Groupe VQS réel	1	38 135	4 249	2 678	<b>723</b>
	2	3 951	3 713	1 897	1 054
	3	622	588	1 335	1 371
	4	119	155	430	1 438

Modèle =réel	44 620
Modèle =réel +/- 1	12 487
Modèle =réel +/- 2	4 509
Reste	842
<b>TOTAL</b>	<b>62 457</b>

		Groupe VQS réel			
		1	2	3	4
Groupe VQS réel	1	61,1	6,8	4,3	1,2
	2	6,3	5,9	3,0	1,7
	3	1,0	0,9	2,1	2,2
	4	0,2	0,2	0,7	2,3

Modèle =réel	71,4
Modèle =réel +/- 1	20,0
Modèle =réel +/- 2	7,2
Reste	1,3
<b>TOTAL</b>	<b>100</b>

Ce modèle détecte 186 faux négatifs, qui correspondent à un effectif pondéré de 723 000. La transition observable (erreur de prédiction combinée avec erreur intra-modèle est assez comparable à l'erreur intra-modèle.

**5.2. Annexe 2 : Détails du modèle retenu pour la détermination des faux négatifs.**

**Annexe 2-A : Tableaux pour prédictions et transitions**

La prédiction : retrouver le groupe VQS avec les 14 variables retenues seulement

**Prédiction : observations**

		Groupe VQS réel			
		1	2	3	4
Groupe VQS modèle 1	1	10 116	712	123	5
	2	353	5 096	1 332	92
	3	1	1 270	3 698	1 002
	4	2	13	1 034	5 082

Modèle =réel	23 992
Modèle =réel +/- 1	5 703
Modèle =réel +/- 2	229
Reste	7
<b>TOTAL</b>	<b>29 931</b>

		Groupe VQS réel			
		1	2	3	4
Groupe VQS modèle 1	1	33,8	2,4	0,4	0,0
	2	1,2	17,0	4,5	0,3
	3	0,0	4,2	12,4	3,3
	4	0,0	0,0	3,5	17,0

Modèle =réel	80,2
Modèle =réel +/- 1	19,1
Modèle =réel +/- 2	0,8
Reste	0,0
<b>TOTAL</b>	<b>100</b>

**Prédiction : pondéré**

		Groupe VQS réel			
		1	2	3	4
Groupe VQS modèle 1	1	44 263	1 171	86	1
	2	1 502	7 553	867	30
	3	14	1 872	2 312	325
	4	6	19	650	1 786

Modèle =réel	55 915
Modèle =réel +/- 1	6 387
Modèle =réel +/- 2	149
Reste	7
<b>TOTAL</b>	<b>62 457</b>

		Groupe VQS réel			
		1	2	3	4
Groupe VQS modèle 1	1	70,9	1,9	0,1	0,0
	2	2,4	12,1	1,4	0,0
	3	0,0	3,0	3,7	0,5
	4	0,0	0,0	1,0	2,9

Modèle =réel	89,52
Modèle =réel +/- 1	10,2
Modèle =réel +/- 2	0,24
Reste	0,01
<b>TOTAL</b>	<b>100</b>

La transition : changement de modalités dans les 14 variables suivant les deux enquêtes

**Transitions intra-modèle : observations**

		Groupe HSM (modèle)			
		1	2	3	4
Groupe VQS modèle 1	1	8 481	1 869	466	140
	2	2 013	3 114	1 369	377
	3	733	1 099	2 722	1 417
	4	210	414	1 669	3 838

VQS=HSM	18 155
VQS=HSM +/- 1	9 436
VQS=HSM +/- 2	1 990
Reste	350
<b>TOTAL</b>	<b>29 931</b>

		Groupe HSM (modèle)			
		1	2	3	4
Groupe VQS modèle 1	1	28,3	6,2	1,6	0,5
	2	6,7	10,4	4,6	1,3
	3	2,4	3,7	9,1	4,7
	4	0,7	1,4	5,6	12,8

VQS=HSM	60,7
VQS=HSM +/- 1	31,5
VQS=HSM +/- 2	6,6
Reste	1,2
<b>TOTAL</b>	<b>100</b>

**Transitions intra-modèle : pondéré**

		Groupe HSM (modèle)			
		1	2	3	4
Groupe VQS modèle 1	1	36 432	7 090	1 579	420
	2	3 422	4 455	1 697	377
	3	789	943	2 006	786
	4	109	190	701	1 461

VQS=HSM	44 354
VQS=HSM +/- 1	14 639
VQS=HSM +/- 2	2 935
Reste	529
<b>TOTAL</b>	<b>62 457</b>

		Groupe HSM (modèle)			
		1	2	3	4
Groupe VQS modèle 1	1	58,3	11,4	2,5	0,7
	2	5,5	7,1	2,7	0,6
	3	1,3	1,5	3,2	1,3
	4	0,2	0,3	1,1	2,3

VQS=HSM	71,0
VQS=HSM +/- 1	23,4
VQS=HSM +/- 2	4,7
Reste	0,8
<b>TOTAL</b>	<b>100</b>

D'où les groupes VQS et HSM résultants

**Transitions observables : observations**

		Groupe HSM (modèle)			
		1	2	3	4
Groupe VQS réel	1	8 247	1 704	423	<b>98</b>
	2	2 300	3 085	1 398	308
	3	743	1 332	2 710	1 402
	4	147	375	1 695	3 964

Modèle =réel	18 006
Modèle =réel +/- 1	9 831
Modèle =réel +/- 2	1 849
Reste	245
<b>TOTAL</b>	<b>29 931</b>

		Groupe HSM (modèle)			
		1	2	3	4
Groupe VQS réel	1	27,6	5,7	1,4	0,3
	2	7,7	10,3	4,7	1,0
	3	2,5	4,5	9,1	4,7
	4	0,5	1,3	5,7	13,2

Modèle =réel	60,2
Modèle =réel +/- 1	32,8
Modèle =réel +/- 2	6,2
Reste	0,8
<b>TOTAL</b>	<b>100</b>

**Transitions observables : pondéré**

		Groupe HSM (modèle)			
		1	2	3	4
Groupe VQS réel	1	36 660	7 115	1 625	<b>385</b>
	2	3 547	4 593	2 056	419
	3	495	844	1 732	845
	4	50	126	570	1 395

Modèle =réel	44 381
Modèle =réel +/- 1	14 977
Modèle =réel +/- 2	2 664
Reste	436
<b>TOTAL</b>	<b>62 457</b>

		Groupe HSM (modèle)			
		1	2	3	4
Groupe VQS réel	1	58,7	11,4	2,6	0,6
	2	5,7	7,4	3,3	0,7
	3	0,8	1,4	2,8	1,4
	4	0,1	0,2	0,9	2,2

Modèle =réel	71,1
Modèle =réel +/- 1	24,0
Modèle =réel +/- 2	4,3
Reste	0,7
<b>TOTAL</b>	<b>100</b>

## Annexe 2-B : Comparaison détaillée des variables VQS et HSM retenues pour le modèle.

Pour chaque variable (sauf l'âge) on donne quatre tableaux :

- répartition suivant les groupes, VQS et HSM, tous enquêtés, en effectifs pondérés
- répartition suivant les groupes, VQS et HSM, faux négatifs
- pour apprécier l'importance de la variable dans la distinction des faux négatifs : effectifs et pourcentages pour :
  - VQS = HSM : mêmes modalités suivant les deux enquêtes
  - VQS=HSM +/- 1 : modalités différant d'une unité
  - VQS=HSM +/- 2 : modalités différant de deux unités
  - Reste = VQS= HSM + /- 3 : modalités différant de trois unités
- Comparaison en effectifs pondérés des variables VQS et HSM par modalité. Une « note » permet de se faire une idée de la correspondance (0 : parfait, 20 aussi opposé que possible).

Ainsi, les variables du mini-module ont un poids important pour distinguer les faux négatifs (ces personnes ont donc répondu très différemment sur le mini-module). Particulièrement pour la question sur les limitations, alors que la répartition par modalités diffère peu entre les deux enquêtes. Les limitations qui pèsent le plus sont « ramasse » et « concentra ». La variable d'environnement joue assez peu.

Santé		Données sur tous, pondéré			
		Variable HSM			
		1	2	3	4
Variable VQS	1	16 955	7 531	1 199	201
	2	7 159	12 674	4 471	670
	3	360	2 162	4 489	2 211
	4	35	154	533	1 651

Santé		Faux négatifs, non pondérés			
		Variable HSM			
		1	2	3	4
Variable VQS	1	5	6	7	22
	2	3	10	8	27
	3	0	0	1	9
	4	0	0	0	0

	Tous, pondéré	FN	
VQS=HSM	35 769	57,3	16
VQS=HSM +/- 1	24 069	38,5	26
VQS=HSM +/- 2	2 383	3,8	34
Reste	236	0,4	22
<b>TOTAL</b>	<b>62 457</b>	<b>100</b>	<b>98</b>

Santé	VQS	HSM
1	25 887	24 510
2	24 975	22 521
3	9 223	10 693
4	2 373	4 734

**Ecart tous / faux négatifs**

**10,56**

**Ecart VQS / HSM**

**2,45**

Limitations		Données sur tous, pondéré			
		Variable HSM			
		1	2	3	
Variable VQS	1	45 574	4 281	1 091	
	2	2 825	3 081	1 540	
	3	668	1 044	2 354	

Limitations		Faux négatifs, non pondérés			
		Variable HSM			
		1	2	3	
Variable VQS	1	13	22	61	
	2	0	0	2	
	3	0	0	0	

	Tous, pondéré		FN
VQS=HSM	51 008	81,7	13
VQS=HSM +/- 1	9 689	15,5	24
VQS=HSM +/- 2	1 760	2,8	61
Reste	0	0,0	0
<b>TOTAL</b>	<b>62 457</b>	<b>100</b>	<b>98</b>

**Ecart tous / faux négatifs**

**13,53**

Limitations	VQS	HSM
1	50 946	49 068
2	7 446	8 405
3	4 065	4 985
4	0	0

**Ecart VQS / HSM**

**1,20**

Mal chroniques		Données sur tous, pondéré			
		Variable HSM			
		1	2		
Variable VQS réel	1	36 070	11 399		
	2	2 342	12 647		

Mal chroniques		Faux négatifs, non pondérés			
		Variable HSM			
		1	2		
Variable VQS réel	1	15	74		
	2	1	8		

	Tous, pondéré		FN
VQS=HSM	48 717	78,0	23
VQS=HSM +/- 1	13 740	22,0	75
VQS=HSM +/- 2	0	0	0
Reste	0	0	0
<b>TOTAL</b>	<b>62 457</b>	<b>100</b>	<b>98</b>

**Ecart tous / faux négatifs**

**10,80**

Mal chroniques	VQS	HSM
1	47 469	38 411
2	14 989	24 046
3	0	0
4	0	0

**Ecart VQS / HSM**

**5,80**

Vue près		Données sur tous, pondéré			
		Variable HSM			
		1	2	3	
Variable VQS	1	53 839	1 551	541	
	2	4 151	672	202	
	3	794	261	447	

Vue près		Faux négatifs, non pondérés			
		Variable HSM			
		1	2	3	
Variable VQS	1	72	11	13	
	2	2	0	0	
	3	0	0	0	

	Tous, pondéré		FN
VQS=HSM	54 958	88,0	72
VQS=HSM +/- 1	6 165	9,9	13
VQS=HSM +/- 2	1 334	2,1	13
Reste	0	0,0	0
<b>TOTAL</b>	<b>62 457</b>	<b>100</b>	<b>98</b>

**Ecart tous / faux négatifs**

**3,00**

Vue près	VQS	HSM
1	55 931	58 784
2	5 026	2 484
3	1 501	1 189
4	0	0

**Ecart VQS / HSM**

**0,20**

Vue loin		Données sur tous, pondéré			
		Variable HSM			
		1	2	3	
Variable VQS	1	58 554	1 066	408	
	2	1 516	190	139	
	3	269	68	248	

Vue loin		Faux négatifs, non pondérés			
		Variable HSM			
		1	2	3	
Variable VQS	1	85	4	9	
	2	2	0	0	
	3	0	0	0	

	Tous, pondéré		FN
VQS=HSM	58 992	94,5	85
VQS=HSM +/- 1	2 789	4,5	6
VQS=HSM +/- 2	677	1,1	9
Reste	0	0,0	0
<b>TOTAL</b>	<b>62 457</b>	<b>100</b>	<b>100</b>

**Ecart tous / faux négatifs**

**1,89**

Vue loin	VQS	HSM
1	60 028	60 339
2	1 845	1 323
3	584	795
4	0	0

**Ecart VQS / HSM**

**0,14**

Auditif		Données sur tous, pondéré			
		Variable HSM			
		1	2	3	
Variable VQS	1	51 288	2 601	580	
	2	2 959	2 414	895	
	3	348	452	919	

Auditif		Faux négatifs, non pondérés			
		Variable HSM			
		1	2	3	
Variable VQS	1	74	7	11	
	2	1	1	4	
	3	0	0	0	

	Tous, pondéré		FN
VQS=HSM	54 622	87,5	75
VQS=HSM +/- 1	6 907	11,1	12
VQS=HSM +/- 2	928	1,5	11
Reste	0	0,0	0
<b>TOTAL</b>	<b>62 457</b>	<b>100</b>	<b>98</b>

**Ecart tous / faux négatifs**

**2,29**

Auditif	VQS	HSM
1	54 469	54 596
2	6 269	5 467
3	1 720	2 395
4	0	0

**Ecart VQS / HSM**

**0,43**

Attrape		Données sur tous, pondéré			
		Variable HSM			
		1	2	3	
Variable VQS	1	56 002	1 090	321	
	2	2 244	900	424	
	3	504	391	580	

Attrape		Faux négatifs, non pondérés			
		Variable HSM			
		1	2	3	
Variable VQS	1	65	16	14	
	2	0	1	2	
	3	0	0	0	

**Ecart tous / faux négatifs**

**5,01**

	Tous, pondéré		FN
VQS=HSM	57 483	92,0	66
VQS=HSM +/- 1	4 149	6,6	18
VQS=HSM +/- 2	825	1,3	14
Reste	0	0,0	0
<b>TOTAL</b>	<b>62 457</b>	<b>100</b>	<b>98</b>

Attrape	VQS	HSM
1	57 414	58 751
2	3 568	2 381
3	1 475	1 325
4	0	0

**Ecart VQS / HSM**

**0,10**

Main		Données sur tous, pondéré			
		Variable HSM			
		1	2	3	
Variable VQS réel	1	57 065	760	192	
	2	2 415	618	194	
	3	544	348	321	

Main		Faux négatifs, non pondérés			
		Variable HSM			
		1	2	3	
Variable VQS réel	1	72	13	11	
	2	0	0	2	
	3	0	0	0	

	Tous, pondéré		FN
VQS=HSM	58 005	92,9	72
VQS=HSM +/- 1	3 717	6,0	15
VQS=HSM +/- 2	736	1,2	11
Reste	0	0	0
<b>TOTAL</b>	<b>62 457</b>	<b>100</b>	<b>98</b>

**Ecart tous / faux négatifs 3,97**

Main	VQS	HSM
1	58 017	60 024
2	3 228	1 725
3	1 213	708
4	0	0

**Ecart VQS / HSM 0,32**

	Tous, pondéré		FN
VQS=HSM	58 005	92,9	72
VQS=HSM +/- 1	3 717	6,0	15
VQS=HSM +/- 2	736	1,2	11
Reste	0	0	0
<b>TOTAL</b>	<b>62 457</b>	<b>100</b>	<b>98</b>

**Ecart tous / faux négatifs 3,97**

Main	VQS	HSM
1	58 017	60 024
2	3 228	1 725
3	1 213	708
4	0	0

**Ecart VQS / HSM 0,32**

	Tous, pondéré		FN
VQS=HSM	53 978	86,4	55
VQS=HSM +/- 1	6 961	11,1	14
VQS=HSM +/- 2	1 519	2,4	29
Reste	0	0	0
<b>TOTAL</b>	<b>62 457</b>	<b>100</b>	<b>98</b>

**Ecart tous / faux négatifs 6,08**

Ramasse	VQS	HSM
1	54 195	53 872
2	5 848	4 443
3	2 415	4 143
4	0	0

**Ecart VQS / HSM 1,11**

Concentra		Données sur tous, pondéré			
		Variable HSM			
		1	2		
Variable	1	55 126	3 128		
	2	2 601	1 602		
VQS					
réel					

Concentra		Faux négatifs, non pondérés			
		Variable HSM			
		1	2		
Variable	1	62	36		
	2	0	0		
VQS					
réel					

	Tous, pondéré		FN
VQS=HSM	56 728	90,8	62
VQS=HSM +/- 1	5 729	9,2	36
VQS=HSM +/- 2	0	0	0
Reste	0	0	0
<b>TOTAL</b>	<b>62 457</b>	<b>100</b>	<b>98</b>

**Ecart tous / faux négatifs 5,57**

Concentra	VQS	HSM
1	58 255	57 727
2	4 203	4 730
3	0	0
4	0	0

**Ecart VQS / HSM 0,34**

Vieq		Données sur tous, pondéré			
		Variable HSM			
		1	2	3	
Variable	1	57 097	1 168	340	
	2	2 134	401	182	
	3	437	193	504	
VQS					
réel					

Vieq		Faux négatifs, non pondérés			
		Variable HSM			
		1	2	3	
Variable	1	66	16	15	
	2	0	1	0	
	3	0	0	0	
VQS					
réel					

	Tous, pondéré		FN
VQS=HSM	58 002	92,9	67
VQS=HSM +/- 1	3 679	5,9	16
VQS=HSM +/- 2	777	1,2	15
Reste	0	0	0
<b>TOTAL</b>	<b>62 457</b>	<b>100</b>	<b>98</b>

**Ecart tous / faux négatifs 4,97**

Vieq	VQS	HSM
1	58 605	59 668
2	2 718	1 763
3	1 134	1 027
4	0	0

**Ecart VQS / HSM 0,07**

Comprendre		Données sur tous, pondéré			
		Variable HSM			
		1	2		
Variable	1	57 606	1 896		
	2	1 883	1 072		
VQS					
réel					

Comprendre		Faux négatifs, non pondérés			
		Variable HSM			
		1	2		
Variable	1	72	25		
	2	0	1		
VQS					
réel					

	Tous, pondéré	FN
VQS=HSM	58 678	93,9
VQS=HSM +/- 1	3 779	6,1
VQS=HSM +/- 2	0	0
Reste	0	0
<b>TOTAL</b>	<b>62 457</b>	<b>100</b>

Comprendre	VQS	HSM
1	59 502	59 489
2	2 955	2 968
3	0	0
4	0	0

Ecart tous / faux négatifs

3,99

Ecart VQS / HSM

0,01

Amenalog		Données sur tous, pondéré			
		Variable HSM			
		1	2		
Variable	1	59 697	1 788		
	2	508	464		
VQS					
réel					

Amenalog		Faux négatifs, non pondérés			
		Variable HSM			
		1	2		
Variable	1	91	7		
	2	0	0		
VQS					
réel					

	Tous, pondéré	FN
VQS=HSM	60 161	96,3
VQS=HSM +/- 1	2 296	3,7
VQS=HSM +/- 2	0	0
Reste	0	0
<b>TOTAL</b>	<b>62 457</b>	<b>100</b>

Amenalog	VQS	HSM
1	61 485	60 206
2	972	2 252
3	0	0
4	0	0

Ecart tous / faux négatifs

0,86

Ecart VQS / HSM

0,82

## Annexe 2-C : Quelques éléments pour « expliquer » l'écart VQS - HSM

Modèle avec variables sociodémographiques

Odds Ratio Estimates

Effect	Estimate	Confidence Limits	
age20 00-19 vs 40-59	0.456	0.343	0.607
age20 20-39 vs 40-59	0.798	0.652	0.975
age20 60-79 vs 40-59	0.515	0.425	0.624
age20 80et+ vs 40-59	0.265	0.184	0.382
trizone DOM vs Ile-France	2.529	1.920	3.332
trizone province_métro vs Ile-France	1.283	1.009	1.630
dip7 1 vs 7	0.798	0.495	1.287
dip7 2 vs 7	0.659	0.411	1.057
dip7 3 vs 7	0.753	0.525	1.079
dip7 4 vs 7	0.728	0.549	0.965
dip7 5 vs 7	0.947	0.784	1.144
dip7 6 vs 7	0.636	0.462	0.874

cs	1 vs 8	1.191	0.865	1.640
cs	3 vs 8	1.060	0.701	1.603
cs	4 vs 8	1.293	0.965	1.733
cs	5 vs 8	1.433	1.138	1.806
cs	6 vs 8	1.398	1.106	1.767

Modèle avec variables sociodémographiques et indicateurs santé-handicap :

Odds Ratio Estimates

Effect	Estimate	Confidence Limits	
age20 00-19 vs 40-59	1.545	1.174	2.032
age20 20-39 vs 40-59	1.394	1.138	1.707
age20 60-79 vs 40-59	0.465	0.385	0.562
age20 80et+ vs 40-59	0.238	0.164	0.346
copain 0 vs 1	0.694	0.589	0.819
trizone DOM vs Ile-France	2.576	1.956	3.393
trizone province_métro vs Ile-France	1.264	0.994	1.606
igs_D 0 vs 1	0.435	0.348	0.543
igs_L 0 vs 1	0.160	0.123	0.207
igs_K 0 vs 1	0.742	0.582	0.946
AEF_S 0 vs 1	0.612	0.502	0.747
AT_MS 0 vs 1	1.883	1.539	2.302
AT_ST 0 vs 1	1.406	1.117	1.771

à noter les effets d'âge, différents du modèle précédent ....

Modèle avec variables sociodémographiques, indicateurs santé-handicap dont transitions VQS -HSM :

Odds Ratio Estimates

Effect	Estimate	Confidence Limits	
sexe 1 vs 2	0.854	0.731	0.998
age20 00-19 vs 40-59	0.807	0.607	1.072
age20 20-39 vs 40-59	1.017	0.819	1.263
age20 60-79 vs 40-59	0.419	0.340	0.516
age20 80et+ vs 40-59	0.212	0.142	0.317
copain 0 vs 1	0.743	0.620	0.890
trizone DOM vs Ile-France	2.674	1.991	3.590
trizone province_métro vs Ile-France	1.268	0.983	1.637
D_san 1 vs 0	7.280	6.210	8.535
D_lim 1 vs 0	19.975	16.104	24.776
igs_K 0 vs 1	0.346	0.278	0.430
AEF_M 0 vs 1	0.549	0.454	0.664
AEF_S 0 vs 1	0.528	0.425	0.657
AT_MS 0 vs 1	1.471	1.171	1.849

### 5.3. Annexe 3 : Détermination des groupes VQS pour la stratification d'HSM

**Méthode basée sur la constitution d'un « score » à la main**

A partir des réponses à VQS on calcule un « score » à chaque répondant à l'enquête VQS. Le « score » initial vaut 0 puis il est incrémenté en fonction des réponses à VQS. Par exemple, le « score » d'une personne déclarant un très mauvais état de santé sera augmenté de 6 alors que celui d'une personne déclarant un état de santé moyen sera augmenté de 1. Le « score » final obtenu est ensuite transformé suivant la formule :

$$\text{« score » final} = \frac{\text{"score\_initial"}}{\text{score\_maximal\_possible}} * 100$$

Le score\_maximal\_possible correspond au score le plus élevé que la personne pourrait avoir compte tenu des questions auxquelles elle a répondu. Cette transformation permet donc de tenir compte dans le mode de calcul du « score » de la non réponse partielle qui est fort heureusement assez faible<sup>29</sup>.

Pour créer ce « score » manuel il faut déterminer quelles sont les variables de VQS les plus importantes pour détecter les personnes handicapées. Les enquêtes VQS de 1999 et de 2007 ont un grand nombre de questions dont la formulation est soit identique soit assez proche. Pour ces questions, on peut réaliser une analyse factorielle discriminante ou une régression logistique en utilisant comme variable discriminante une variable synthétique issue de HID.

### Recherche des variables de VQS les plus importantes pour détecter les personnes handicapées

On détermine d'abord une variable discriminante issue de HID qui sera utilisée pour réaliser l'analyse discriminante et la régression logistique. On pourrait utiliser une des grilles de handicap (ex : Katz) reconstituées dans HID. Cependant, ces grilles ne prennent en compte qu'un petit nombre de variables de HID. On décide donc de créer une variable prenant en compte 6 aspects :

- Handicaps sensoriels (questions sur limitations et déficiences) ;
- Handicaps moteurs (questions sur limitations) ;
- Handicaps cognitifs (questions sur limitations, une question de recours aux soins) ;
- Limitations dans les activités de la vie quotidienne (« ADL ») ;
- Limitations dans les activités instrumentales de la vie quotidienne (« IADL ») ;
- Restrictions de participation sociale (questions sur formation et emploi).

Le score de handicap s'obtient en sommant les 6 scores partiels, après normalisation<sup>30</sup>, puis on ramène le résultat entre 0 et 100. On réalise ensuite une classification ascendante hiérarchique, d'où une classification des répondants à HID en 3 groupes de sévérité de handicap<sup>31</sup>. Les résultats sont résumés dans le tableau suivant :

Groupe de sévérité de handicap	Effectifs au sein du groupe	Moyenne du score au sein du groupe
1	10 645	4
2	4 740	22
3	1531	48

On détermine ensuite par régression logistique de la variable Groupe de sévérité de handicap sur les variables de l'enquête VQS.

Variable (VQS 2007)	Odd ratio
Avoir de l'aide humaine	3,2
Avoir des problèmes pour remplir des formulaires simples	2,4
Avoir fait une demande de reconnaissance officielle	2,2
Considérer avoir un handicap	2,0
Avoir des problèmes pour s'habiller	1,9
Être limité	1,8
Avoir des problèmes pour voir de loin	1,8
Avoir des problèmes pour ramasser un objet au sol	1,6
Avoir des problèmes pour parler et se faire comprendre	1,5
Avoir de(s) aide(s) technique(s)	1,5
Avoir des problèmes pour entendre	1,3

<sup>29</sup> Seulement 0,5% des répondants n'ont pas répondu à 6 questions ou plus de VQS Par ailleurs, en raison d'un filtre, (dans la version CAPI), un grand nombre de personnes n'ont eu à répondre qu'à 24 questions et non 26.

<sup>30</sup> La moyenne de chacun des 6 scores a été ramenée à 0 et l'écart type a été ramené à 1.

<sup>31</sup> Il s'agit bien d'une construction différente (et préalable) de la construction des groupes de sévérité présumée, utilisée ensuite pour la stratification. En particulier la sévérité n'est pas ici présumée, elle résulte de l'exploitation de l'enquête HID (1999).

Avoir des problèmes pour voir de près	1,2
Avoir de(s) aménagement(s) du logement	1,2

Cette analyse est confrontée à celle résultant d'une analyse des correspondances multiples (ACM) sur les variables de VQS complétée par une analyse discriminante sur les axes provenant de l'ACM. On ne retient qu'un axe dans l'ACM<sup>32</sup>. Cet axe oppose les personnes handicapées ou en mauvaise santé aux autres personnes : les modalités OUI se projettent toutes positivement sur l'axe et les modalités NON se projettent toutes négativement sur l'axe. Enfin, le premier axe de l'analyse discriminante est déterminé essentiellement par l'axe 1 de l'ACM : ce qui est cohérent avec l'analyse de cet axe dans l'ACM.

Voici le poids de chacune des variables de VQS dans la détermination de l'axe 1 de l'ACM

Variable	Contribution à l'axe 1 de l'ACM
Avoir de l'aide humaine	12,9%
Avoir des problèmes pour s'habiller	12,2%
Considérer avoir un handicap	11,7%
Être limité	11,4%
Avoir des problèmes pour ramasser un objet au sol	10,7%
Avoir de(s) aide(s) technique(s)	8,3%
Avoir des problèmes pour remplir des formulaires simples	6,5%
Avoir fait une demande de reconnaissance officielle	6,4%
Avoir de(s) aménagement(s) du logement	5,5%
Avoir des problèmes pour voir de loin	4,1%
Avoir des problèmes pour voir de près	4%
Avoir des problèmes pour parler et se faire comprendre	3,7%
Avoir des problèmes pour entendre	2,7%

Ce qui permet de conclure : l'analyse discriminante comme l'analyse par régression logistique font ressortir les variables « Avoir de l'aide humaine » et « Considérer avoir un handicap ». L'analyse discriminante fait également ressortir les variables « Avoir des problèmes pour s'habiller » et « Être limité ». L'analyse par régression logistique fait encore ressortir les variables « Avoir des problèmes pour remplir des formulaires simples » et « Avoir fait une demande de reconnaissance officielle ». En outre, on accordera un poids important à la modalité de réponse « très mauvaise » de la question sur l'état de santé et à la question *La personne est-elle dans l'impossibilité totale d'accomplir une ou plusieurs des activités citées précédemment ?*

Ce qui aboutit au programme sas :

```

if Etat_sant='5' then score=score+6;
if Etat_sant='4' then score=score+3;
if Etat_sant='3' then score=score+1;
if Mal_chro='2' then score=score+1;
if Limitat='3' then score=score+6;
if Limitat='2' then score=score+3;
if Vuepres='3' then score=score+3;
if Vuepres='2' then score=score+1;
if Vueloin='3' then score=score+3;
if Vueloin='2' then score=score+1;
if Parle='3' then score=score+3;
if Parle='2' then score=score+1;
if Auditif='3' then score=score+3;
if Auditif='2' then score=score+1;
if Marche='3' then score=score+3;
if Marche='2' then score=score+1;
if Attrape='3' then score=score+3;

```

<sup>32</sup> L'axe 1 explique 27% de l'inertie, l'axe 2 13% et l'axe 3 9%.

```

if Attrape='2' then score=score+1;
if Main='3' then score=score+3;
if Main='2' then score=score+1;
if Ramasse='3' then score=score+3;
if Ramasse='2' then score=score+1;
if Concentra='3' then score=score+3;
if Concentra='2' then score=score+1;
if Memoire='3' then score=score+3;
if Memoire='2' then score=score+1;
if Initiative='3' then score=score+3;
if Initiative='2' then score=score+1;
if Prob_quot='3' then score=score+3;
if Prob_quot='2' then score=score+1;
if Sorties='3' then score=score+3;
if Sorties='2' then score=score+1;
if Comprendre='3' then score=score+3;
if Comprendre='2' then score=score+1;
if Impossible='2' then score=score+6;
if Aidepers='3' then score=score+6;
if Aidepers='2' then score=score+3;
if Amenalog='2' then score=score+2;
if Aidetech='2' then score=score+1;
if Handicap='2' then score=score+6;
if Reconnai='2' then score=score+6;

```

Après la normalisation qui permet ramène ce score entre 0 (très bonne santé présumée) et 100, le groupe se détermine grâce au tableau :

	âge <= à 19	19 < âge <= 59	59 < âge <= 79	79 < âge	Groupe
Score <	1	4	5	6	I
	10	12	25	40	II
	25	30	45	65	III
					IV

*Note de lecture : une personne d'âge compris entre 19 et 59 ans sera affectée dans le groupe I si son score est inférieur à 4, le groupe II si le score est compris entre 4 et 12, le groupe III si le score est compris entre 12 et 30 et le groupe IV si le score dépasse 40.*

## Bibliographie

[1] L'essentiel de ce document provient du document de travail : Présentation générale de l'enquête Handicap-Santé ([http://www.insee.fr/fr/themes/document.asp?reg\\_id=0&ref\\_id=F1109](http://www.insee.fr/fr/themes/document.asp?reg_id=0&ref_id=F1109)).

Des compléments figurent dans les autres documents de travail relatifs à cette enquête :

[2] [http://www.insee.fr/fr/themes/document.asp?reg\\_id=0&ref\\_id=F1101](http://www.insee.fr/fr/themes/document.asp?reg_id=0&ref_id=F1101)

[3] [http://www.insee.fr/fr/themes/document.asp?reg\\_id=0&ref\\_id=F1102](http://www.insee.fr/fr/themes/document.asp?reg_id=0&ref_id=F1102)

[4] [http://www.insee.fr/fr/themes/document.asp?reg\\_id=0&id=2769](http://www.insee.fr/fr/themes/document.asp?reg_id=0&id=2769)