



DOSSIERS DE LA DREES

Méthodologie

n° 43 • octobre 2019

Échantillonnage des enquêtes VQS 2014 et CARE-ménages-2015

Amélie Carrere (DREES), Olivier Haag (INSEE) et Noémie Soullier (DREES)

Échantillonnage des enquêtes VQS et CARE-ménages

Amélie Carrere (DREES), Olivier Haag (INSEE) et Noémie Soullier (DREES)

Retrouvez toutes nos publications sur : drees.solidarités-sante.gouv.fr

Retrouvez toutes nos données sur : data.drees.sante.gouv.fr

Sommaire

SYNTHÈSE	5
■ INTRODUCTION	6
■ MOTIVATIONS D'UNE ENQUÊTE FILTRE	8
Échantillonnage de l'enquête CARE	8
Une option écartée : Enquête CARE sans enquête filtre	9
Autre option écartée : un échantillon complémentaire de bénéficiaires de l'APA dans l'enquête CARE	10
L'expérience de l'enquête filtre VQS 2007	11
Enquête filtre VQS 2014	12
Représentativité départementale	16
■ PLAN DE SONDAGE DE VQS	19
Constitution de la base de sondage de VQS	19
Restriction aux personnes de 60 ans ou plus	19
Conservation des logements-foyers	19
Retrait des personnes vivant dans des communautés hors champ de l'enquête : constitution d'une base de « logements-ordinaires »	20
Mise en commun de la base des logements-foyers et des « logements-ordinaires »	20
Caractéristiques de la base finale de sondage	20
Tirage de l'échantillon	21
Calcul des allocations de l'échantillon principal de 140 000 logements	21
Échantillon complémentaire de 60 000 logements	22
Calcul des poids	23
Poids à utiliser pour CARE	23
Poids à utiliser pour VQS	23
■ PRINCIPE DU CALCUL ANTICIPÉ DES ALLOCATIONS POUR CHESTER	24
Les allocations à livrer à Chester	24
Choix de la base de sondage pour les calculs	24
Les hypothèses mises en œuvre	25
Le calcul anticipé des quotités par ZAE pour la planification de l'activité 2015	25
■ TIRAGE DE L'ÉCHANTILLON CARE	27
Le principe général	27
L'estimation du poids des logements répondants à VQS	27
Le poids des individus de la base de sondage CARE	29
Les modifications à la marge des allocations théoriques	29
Les ZAE « cas simples »	30
Les ZAE « cas complexes »	30
Le tirage proprement dit	31
Les résultats en termes de dispersion des poids	32
■ BIBLIOGRAPHIE	33
Annexe 1. Importance des variables dans VQS 2007-HSM 2008 : rang selon la modélisation	34
Annexe 2. Incrémentation de l'ancien et du nouveau score par variable du questionnaire	35
Annexe 3. Structure des réponses à VQS 2007 avec et sans face-à-face	36
Annexe 4. Lexique	39

SYNTHÈSE

L'enquête Capacités, Aides et REssources des seniors (CARE) a pour but d'éclairer le débat public sur les conditions de vie des seniors et en particulier sur la façon dont ils font face à la perte d'autonomie. La perte d'autonomie concerne essentiellement les très grands âges, dans des proportions qui restent néanmoins faibles. Selon la DREES (Enquête Aide sociale) et l'INSEE (estimations provisoires de population au 01/01/2016), tous lieux de vie confondus, 20,8 % des personnes âgées de 75 ans ou plus sont bénéficiaires de l'allocation personnalisée d'autonomie (APA) et 12,3 % le sont à domicile.

Pour interroger un nombre suffisant de personnes dépendantes vivant en logement ordinaire, il est donc nécessaire de les surreprésenter lors du tirage de l'échantillon. Or, aucune source administrative ne permet de repérer exactement cette population cible. Une première enquête permettant de disposer d'une première estimation du niveau de dépendance des personnes de 60 ans ou plus est donc nécessaire : c'est le rôle principal de l'enquête Vie Quotidienne et Santé (VQS). Les répondants à cette enquête serviront alors de base de sondage pour l'enquête CARE, volet « ménages ». En outre, l'enquête VQS devant par définition couvrir un grand nombre de ménages sur l'ensemble du territoire français, il a été décidé de la rendre représentative au niveau du département, afin de pouvoir estimer des taux de prévalence de la perte d'autonomie à cette échelle.

Ainsi l'échantillonnage de l'enquête VQS avait deux principaux objectifs :

- Servir de base de sondage au volet « ménage » de l'enquête CARE qui est une enquête réalisée en face-à-face par des enquêteurs de l'INSEE. Cette base de sondage doit donc se limiter aux zones d'action enquêteurs (ZAE) de l'échantillon maître (EM) de l'INSEE.
- Assurer une représentativité départementale, ce qui suppose de couvrir également les communes situées en dehors des ZAE de l'EM.

La première partie de l'article présente les travaux mis en œuvre pour définir le plan de sondage de l'enquête VQS. Ces estimations se basent sur les résultats de la précédente édition des enquêtes VQS 2007 et Handicap-Santé Ménages 2008.

La suite de l'article concerne l'enquête CARE et présente les hypothèses prises pour définir le plan de sondage de l'enquête sans en connaître la base de sondage et ce afin de répondre aux besoins du nouveau Cadre d'Emploi des enquêteurs (NCEE).

Enfin, la dernière partie présente la méthodologie mise en œuvre pour le tirage de l'échantillon de CARE une fois la base de sondage disponible, sous contrainte du respect des allocations par ZAE livrées en 2014.

NB : Ce document a fait l'objet d'une présentation lors des XII^{es} Journées de méthodologie statistique à Paris (31 mars-2 avril 2015).

Mots-clés : Échantillonnage, collecte, équilibrage.

■ INTRODUCTION

Le débat national de 2011 sur la dépendance des personnes âgées a fait émerger de nombreuses questions quant à l'évolution de la dépendance et aux moyens financiers, humains et techniques qu'elle requiert. Les systèmes d'information existants, notamment les enquêtes Handicap Santé¹, ne permettaient pas d'apporter des réponses au niveau national sur ces thématiques. En particulier, il s'est révélé nécessaire de :

- lever une partie de l'incertitude qui entoure les hypothèses de développement de la dépendance. Ceci suppose de suivre l'évolution du nombre de personnes âgées en perte d'autonomie et, en particulier, de la comparer aux projections réalisées par la DREES en 2011.

Selon ces projections, 1,5 million de personnes seraient dépendantes en 2025 (au sens où elles bénéficieraient de l'allocation personnalisée d'autonomie – APA), contre près de 1,2 million en 2012, soit une hausse de 30 % (source : Groupe de travail présidé par J.M. Charpin « Perspectives démographiques et financières de la dépendance », 2011) ;

- mieux cerner les enjeux financiers auxquels font face les ménages confrontés à la dépendance, que la personne concernée réside à domicile ou en établissement. Ces questions de financement de la perte d'autonomie et, plus spécifiquement, de reste à charge sont cruciales, d'autant plus que le coût de la prise en charge pourrait s'accroître, du fait par exemple de la diversification des prestations, du renforcement des normes de sécurité et de qualité, des revalorisations des salaires et de l'augmentation du niveau de qualification des personnels intervenants. À l'heure actuelle, les besoins liés à la perte d'autonomie sont assurés à la fois par l'entourage (proches aidants), par le biais de prestations (telles que l'APA ou l'action sociale des caisses de retraite) et par les ressources propres des personnes âgées.

Les enquêtes Capacités, Aides et REssources des seniors (CARE) ont pour but de répondre à ces besoins d'information (voir encadré). Or, la population âgée dépendante est surtout présente aux très grands âges, dans des proportions qui restent en outre assez faibles : en 2011, 8 % des personnes âgées de 60 ans ou plus sont bénéficiaires de l'allocation personnalisée d'autonomie (APA) (Bérardier, 2014) et environ 10 % des personnes âgées de 80 ans ou plus.

La question qui se pose est donc : comment interroger suffisamment de personnes concernées dans l'enquête ?

Pour interroger un nombre suffisant de personnes dépendantes vivant en logement ordinaire, il a été nécessaire de procéder à une surreprésentation des personnes estimées dépendantes via une enquête filtre : Vie Quotidienne et Santé – VQS. Les répondants à VQS constituent la base de sondage de l'enquête CARE². La première partie met en évidence les motivations qui ont conduit à réaliser cette enquête filtre ainsi que les volumes nécessaires pour répondre aux objectifs de l'enquête CARE.

La deuxième partie de l'article présente les hypothèses mises en œuvre pour les prévisions d'allocations par zones d'action enquêteurs (ZAE), rendus nécessaires par le nouveau cadre d'emploi des enquêteurs (NCEE), et qui ont ici été réalisés alors même que la collecte de VQS n'avait pas encore commencé et que la base de sondage de l'enquête CARE n'était donc pas connue.

Une troisième partie s'intéressera à la méthodologie mise en œuvre pour le tirage de l'échantillon de CARE une fois la base de sondage disponible, sous contrainte du respect des allocations par ZAE prévues.

¹ Les enquêtes Handicap-Santé, qui ont un lien de parenté fort avec les enquêtes CARE, ont été conduites entre 2007 et 2009 et portaient sur l'ensemble de la population (tous âges), vivant en métropole et dans les départements d'outre-mer, en logement ordinaire ou en institution.

² Dans tout ce document, on fera référence à « l'enquête CARE » par simplification, même si cette référence renvoie plus précisément au seul volet « ménage » de l'enquête.

Encadré • Dispositif d'enquêtes CARE

Le dispositif d'enquêtes « CARE » pour Capacités, Aides et REssources des seniors, en logements ordinaires (2014-2015) et en institutions (2016) est proche de celui des enquêtes Handicap-Santé de 2007-2009. Il comprend :

- L'enquête Vie Quotidienne et Santé (VQS) en 2014-2015, enquête filtre du volet « seniors » de l'enquête CARE en ménages,
- L'enquête CARE en ménages en 2015, volets « seniors » et « aidants »,
- L'enquête CARE en institutions en 2016, volets « seniors » et « aidants ».

Pour le volet « seniors » de l'enquête en logements ordinaires, les unités statistiques enquêtées sont les individus âgés de 60 ans ou plus à la date de début de collecte de l'enquête (2 mai 2015) vivant en logements ordinaires (y compris en logement-foyer) en France métropolitaine.

Le champ a été défini par une limite d'âge à 60 ans car cela correspond à l'âge minimum pour prétendre à des droits spécifiques à la dépendance (allocation personnalisée d'autonomie – APA, action sociale des caisses de retraite, etc.). Parmi la population interrogée, l'intérêt est principalement porté à la population âgée dépendante, c'est-à-dire ayant des difficultés à réaliser seule des activités de la vie quotidienne et pour qui l'aide d'un tiers est nécessaire. Mais les personnes âgées en bonne santé sont aussi interrogées afin de disposer de résultats sur l'ensemble des seniors, de calculer des prévalences et de permettre des comparaisons entre les situations des personnes dépendantes et celles des personnes non dépendantes.

■ MOTIVATIONS D'UNE ENQUÊTE FILTRE

L'enquête CARE a, entre autres, pour objectif de mesurer l'évolution du nombre de personnes en perte d'autonomie. Il est donc nécessaire, pour mener des analyses robustes sur cette population, de disposer d'un nombre suffisant de personnes en perte d'autonomie dans l'échantillon de l'enquête CARE. C'est l'objectif principal de l'enquête VQS.

En outre, les départements étant les chefs de file de l'action sociale et médico-sociale auprès des personnes âgées, le besoin d'information sur la dépendance au niveau départemental est important. Le mode de collecte de l'enquête VQS a permis d'envisager un élargissement de l'échantillon VQS afin de disposer de résultats par département. Les modifications de l'échantillon de VQS pour obtenir un nombre de ménages suffisant pour réaliser ces analyses seront présentées dans un second temps.

Échantillonnage de l'enquête CARE

Bien que les prévisions estiment à la hausse le nombre de personnes dépendantes dans la population française (Lecroart *et al.*, 2013), celles-ci constituent toujours une population « rare » au sens statistique. L'enquête Handicap-Santé (INSEE-DREES, 2008) a estimé que 7 % de la population française des 60 ans ou plus vivant à domicile était dépendante au sens de la grille AGGIR³ (Eghbal-Téhérani et Makdessi, 2011). En supposant que cette proportion est identique en 2015, on peut alors estimer qu'environ 1,1 million d'individus seraient dépendants à domicile en 2015 (voir tableau 1). La même méthode a été appliquée pour les différentes sous-populations d'intérêt : personnes très dépendantes, personnes aidées, personnes ayant des troubles de la cohérence... Pour chaque sous-population, un nombre minimal de répondants a été défini (voir tableau 1). Cet effectif doit permettre de réaliser les analyses prévues pour répondre aux objectifs de l'enquête. L'objectif le plus contraignant est d'avoir au moins 600 répondants en GIR 1-2 (personnes les plus dépendantes).

Trois solutions étaient envisageables pour disposer de ce seuil minimal : soit tirer directement un large nombre de personnes pour constituer l'échantillon de l'enquête CARE, soit surreprésenter la population d'intérêt dans CARE en tirant un échantillon complémentaire de bénéficiaires de l'APA, soit surreprésenter la population d'intérêt dans CARE grâce à une enquête filtre.

Tableau 1 • Estimation de la population en 2015 et objectifs fixés pour CARE par variable d'intérêt

	Estimation nationale à la date de CARE-M sur le champ domicile	Objectif pour l'enquête CARE-M
<i>Personnes de 60 ans ou plus résidant à domicile</i>	16 000 000	10 000
<i>Personnes estimées en GIR 1-4</i>	1 120 000 (7 %)	2 000
<i>Personnes estimées en GIR 1-2</i>	320 000 (2 %)	600
<i>Personnes aidées par quelqu'un de l'entourage</i>	3 680 000 (23 %)	2 000
<i>Personnes aidées par un aidant professionnel</i>	2 400 000 (15 %)	2 000

³ Les groupes iso-ressources (GIR), au nombre de 6, permettent de classer les personnes en fonction de leur degré de perte d'autonomie. Le classement dans un GIR s'effectue à l'aide de la grille Aggir (Autonomie gérontologie-groupe iso-ressources) qui prend en compte différentes dimensions de la perte d'autonomie (cohérence, orientation, toilette, communication...). Les personnes en GIR 1 sont les plus dépendantes. Les personnes en GIR 5-6 sont considérées comme autonomes, alors que le classement en GIR 1 à 4 ouvre droit à l'allocation personnalisée d'autonomie (APA).

	Estimation nationale à la date de CARE-M sur le champ domicile	Objectif pour l'enquête CARE-M
Personnes présentant des troubles de la cohérence ⁴	480 000 (3 %)	600

Lecture > en 2015, 1 120 000 personnes de 60 ans ou plus seront estimées en GIR 1-4 et résidentes à domicile et en France métropolitaine, soit environ 7 % de la population de 60 ans ou plus.

Champ > Personnes de 60 ans ou plus résidant à domicile en France métropolitaine.

Sources > INSEE, DREES, VQS 2007, HSM 2008, Omphale - projection 01/01/2013, Recensement de la population ; Calculs DREES.

L'INSEE réalise le tirage de l'échantillon dans les sources fiscales. Compte tenu de la taille importante de l'échantillon souhaité, l'échantillon ne pouvait en effet pas être tiré dans la dernière EAR (enquête annuelle du recensement) sous Octopusse⁵. Par ailleurs, le mode de collecte de l'enquête VQS (postal, internet et téléphonique) aurait nécessité une phase de saisie des adresses. Il n'était pas non plus envisageable de tirer dans les cinq derniers EAR, car cela aurait engendré trop d'adresses non utilisables (« déchets »), étant donné que la population enquêtée est une population fragile qui peut être amenée à déménager, à partir en institution ou à décéder. Le choix s'est donc arrêté sur un tirage de l'échantillon dans les sources fiscales.

Le nouveau cadre d'emploi des enquêteurs de l'INSEE impose de connaître l'allocation par ZAE en septembre d'une année pour les enquêtes dont la collecte se déroule entre janvier et décembre de l'année suivante. Par ailleurs, les bases fiscales d'une année donnée ne sont disponibles qu'en mars de l'année suivante.

Les estimations de volumétrie réalisées par la suite utilisent ces informations, ainsi que les informations de suivi de collecte des enquêtes VQS 2007 et HSM 2008 au prorata de l'écart entre les bases de sondage et les collectes. En effet, la forte mobilité des personnes âgées (départ en institution, recohabitation avec les enfants, décès) conduit à être le plus précis possible concernant les hypothèses d'attrition de l'échantillon entre le tirage et la collecte.

Une option écartée : Enquête CARE sans enquête filtre

Pour une collecte de l'enquête CARE en 2015 sans enquête filtre, il aurait été nécessaire de tirer l'échantillon en mars 2014 dans une base datant de 2013. En considérant une collecte tout début 2015, cela implique une ancienneté de la base de tirage d'environ 19 mois⁶. Grâce aux enquêtes VQS et HSM, nous avons estimé qu'en 19 mois, 10 % des personnes âgées de 60 à 69 ans, 14 % des personnes âgées de 70 à 79 ans et 38 % des personnes âgées de 80 ans ou plus auront soit déménagé soit seront parties en institution ou décédées. Parmi la population restante, 20 % des 60-69 ans, 23 % des 70-79 ans et 26 % des 80 ans ou plus ne répondront pas à l'enquête. Le taux de réponse du tableau 2 a été estimé en prenant en compte cet écart et en lui appliquant les hypothèses d'attrition de l'échantillon entre le tirage et la collecte.

⁴ Ces personnes sont cotées B ou C sur l'axe 'cohérence' de la grille Autonomie Gérontologie Groupe iso-ressources (AGGIR) qui sert à évaluer le degré de perte d'autonomie des personnes dans le cadre d'une demande d'APA. Ces personnes cotées B ou C sur l'axe cohérence sont considérées comme potentiellement atteintes de la maladie d'Alzheimer ou de troubles assimilés.

⁵ Octopusse est un outil de tirage d'échantillon développé par l'INSEE. Il est basé sur un système d'Échantillon-Maître pour le tirage des échantillons dans la dernière enquête annuelle de recensement (Christine et Faivre, 2009).

⁶ Les déclarations fiscales sont à renvoyer courant mai de chaque année. Il y a donc un écart de 19 mois entre juin 2013 et janvier 2015.

Tableau 2 • Taux de réponse pour CARE 2015 estimés à partir d'HSM 2008, par tranche d'âge, sous l'hypothèse d'une enquête réalisée sans enquête filtre

	60-69 ans	70-79 ans	80 ans ou plus
<i>Taux de réponse total</i>	72 %	66 %	46 %

Lecture > Selon les taux de réponse et de mobilité observés par âge pour l'enquête HSM 2008, 72 % des individus âgés de 60 à 69 ans et échantillonnées pour l'enquête CARE-Ménages devraient y répondre.

Champ > Personnes de 60 ans ou plus résidant à domicile en France métropolitaine.

Source > INSEE, DREES, HSM 2008 ; Calculs DREES.

Bien que le taux de réponse des personnes plus âgées soit plus faible, la proportion de personnes peu autonomes y est plus importante : dans Handicap-Santé, 1 % des personnes âgées de 60-79 ans sont classées en GIR 1-2 estimé, alors que c'est le cas de 5 % des 80 ans ou plus. Une surreprésentation des plus âgés pourrait donc permettre d'atteindre les objectifs fixés précédemment. Ainsi, en considérant un taux de sondage huit fois plus grand pour les 70 ans et plus⁷ que pour les 60 – 69 ans, l'échantillon serait constitué de 42 700 personnes tirées en tout, dont 3 900 âgées de 60 à 69 ans, 19 200 âgées de 70 à 79 ans et 19 600 âgées de 80 ans ou plus. Parmi les 25 000 répondants estimés, 2 300 seraient des personnes dépendantes (estimation GIR 1 à 4) dont 600 personnes avec une estimation GIR 1 et 2.

Outre le coût d'une telle enquête en face-à-face, le tirage d'un échantillon de cette taille aurait épuisé l'échantillon-maître sur ces tranches de population et aurait pu difficilement être assuré uniquement par le réseau des enquêteurs de l'INSEE.

Cette solution a donc été écartée.

Autre option écartée : un échantillon complémentaire de bénéficiaires de l'APA dans l'enquête CARE

Un échantillon complémentaire de bénéficiaires de l'APA aurait eu pour objectif d'augmenter le nombre de personnes dépendantes dans l'échantillon et plus spécifiquement le nombre de bénéficiaires de l'APA. En effet, l'un des objectifs de l'enquête CARE est de mesurer le reste à charge lié à la dépendance des personnes et donc notamment de recenser les aides publiques (principalement l'APA) dont elles disposent.

Ce tirage comporte le risque de tirer deux fois un même individu. Pour pouvoir repérer les individus qui sont dans les deux bases de sondage et ainsi effectuer le partage des poids, la date de la remontée des données APA doit être proche de celle de la base de sondage principale. Dans le cas contraire, l'échec d'appariement pour double compte est accru du fait des déménagements entre la date de collecte de la base principale et celle de la date de référence de l'extraction des données APA (à partir des enquêtes VQS 2007 et HSM 2008 on estime à 6 % de déménagements en 9 mois environ). Ainsi, en complément des bases fiscales 2013 (validité au 1^{er} janvier), des remontées individuelles des données administratives des conseils départementaux en date du 31 mars⁸ 2013 pourraient être utilisées. Une base consolidée en date de mars 2013 serait disponible environ 6 mois plus tard – au mieux – soit en octobre 2013.

Pour une collecte de l'enquête CARE en 2015, il aurait été nécessaire de tirer l'échantillon en mars 2014. Les deux bases de sondage (base fiscale et bases administratives des conseils départementaux) dateraient du premier trimestre 2013. En considérant une collecte tout début 2015, cela implique une ancienneté de la base de tirage d'environ 19 mois, conduisant au même taux de réponse que précédemment pour l'échantillon tiré en population générale (voir tableau 2).

Environ un tiers des bénéficiaires de l'APA en 2011 sont entrés dans le dispositif au cours de l'année écoulée (d'après les remontées individuelles RI-APA de 2011⁹). Or, le nombre de bénéficiaires de l'APA est à peu près

⁷ Étant donné la chute du taux de réponse sur les 80 ans et plus, une surreprésentation de cette population seule ne saurait être efficace.

⁸ Les extractions de données issues des conseils départementaux se font par trimestre. Ainsi pour être au plus près des dates des déclarations fiscales, il est préférable de prendre le premier trimestre. Par ailleurs, un délai d'apurement des bases est aussi à prendre en considération.

⁹ Cette estimation est assez éloignée de celle de Debout (2010), soit 45 %. Ceci s'expliquerait principalement par le fait que les durées dans l'APA sont plus longues en 2011 qu'en 2007, du fait de la montée en charge de la prestation, créée en 2002.

stable, par conséquent environ un tiers sort¹⁰ également du dispositif chaque année. Ainsi, pour un échantillon constitué à partir de données 2013, on peut estimer qu'un tiers des bénéficiaires de l'APA seraient encore dans le dispositif en 2015. En ajoutant un taux de mobilité hors décès à l'intérieur du même département et un taux de non-réponse, environ un quart de l'échantillon de bénéficiaires de l'APA seraient répondants (voir tableau 3).

Tableau 3 • Taux de réponse de l'échantillon complémentaire APA estimés à partir d'HSM 2008 et des remontées individuelles RI-APA 2011, par tranche d'âge

	60-69 ans	70-79 ans	80 ans ou plus
Taux de réponse total	53 %	40 %	18 %

Lecture > Selon les taux de réponse observés par âge pour l'enquête HSM 2008 et la mobilité des individus bénéficiaires de l'APA, 53 % des bénéficiaires de l'APA âgés de 60 à 69 ans et tirés dans l'échantillon complémentaire de bénéficiaires de l'APA devraient répondre à l'enquête CARE-Ménages.

Champ > Personnes de 60 ans ou plus bénéficiaires de l'APA résidant à domicile en France métropolitaine.

Sources > INSEE, DREES, HSM 2008, Remontées individuelles RI-APA 2011 ; Calculs DREES.

La taille de l'échantillon tiré en population générale serait d'environ 15 000 pour 9 400 répondants, ce qui permettrait d'obtenir environ 140 personnes estimées en GIR 1-2. En considérant un échantillon complémentaire d'environ 4 000 bénéficiaires de l'APA, environ 500 répondants seraient en GIR 1-2. Au total, on aurait donc environ 640 personnes estimées en GIR 1-2 pour un échantillon de 19 000 enquêtés et 12 000 répondants.

Par ailleurs, les remontées individuelles RI-APA de 2011 réalisées par la DREES ne sont pas exhaustives sur la France entière. Elles n'étaient pas non plus un échantillon aléatoire puisque la remontée des données était basée sur le volontariat des départements. La remontée de ces données nécessite enfin un travail conséquent pour les conseils départementaux comme pour la DREES.

Plus fondamentalement, au-delà de la complexité et la lourdeur de la collecte de ces bases administratives, l'enquête CARE vise une acception large de la perte d'autonomie, que le bénéfice d'une allocation comme l'APA ne permet pas de recouvrer, ne serait-ce que parce que le bénéfice de cette prestation suppose une démarche de la part des personnes. Enfin, avec un échantillon complémentaire de bénéficiaires de l'APA l'analyse du recours ou non à cette prestation, via la mise en perspective du degré de dépendance (mesuré à partir du questionnaire CARE) et du bénéfice de l'APA (mesuré via un enrichissement avec les données administratives) serait caduque.

Aucune autre base administrative ne permet de repérer les personnes dépendantes, et la piste d'un suréchantillon à partir de données administratives a donc été écartée. Ainsi pour la surreprésentation de ces personnes, la reconduite d'une enquête filtre a donc été envisagée, comme cela avait été le cas en 2007.

L'expérience de l'enquête filtre VQS 2007

L'objectif de l'enquête Vie Quotidienne et Santé (VQS) de 2007 était d'assurer un échantillon de personnes handicapées de taille suffisante pour le volet ménages de l'enquête Handicap-Santé (2008). Les strates d'intérêt de VQS 2007 étaient déterminées selon l'âge et le groupe d'autonomie. Le degré d'autonomie était estimé sous la forme d'un score, qui utilisait les réponses à l'enquête VQS. La construction du score s'appuyait sur les relations observées entre VQS 1999 et HID 2000. Un groupe d'autonomie était ensuite alloué à chaque individu selon son âge et son score : le groupe I correspond aux personnes les plus autonomes (pour lesquelles le taux de sondage est faible) et le groupe IV aux personnes les plus handicapées (strate exhaustive, dans laquelle le taux de sondage était de 1 : toutes les personnes de ce groupe étaient tirées pour constituer l'échantillon HSM).

Pour estimer les taux de réponse aux enquêtes VQS 2014 et CARE 2015, nous utilisons ceux observés respectivement pour VQS 2007 et HSM 2008 modifiés de la façon suivante :

- Pour VQS 2014, nous enlevons les répondants en face-à-face de l'enquête VQS 2007 car ce mode de collecte n'était pas prévu pour VQS 2014, et les ménages sans individus de 60 ans ou plus. Par ailleurs, les taux de réponses sont estimés en considérant une base de sondage de 2013 et une collecte de l'enquête filtre en septembre 2014 (soit un délai de 15 mois environ). Cette modification de l'écart entre la base de sondage et la collecte modifie uniquement le taux de mobilité.

¹⁰ Les sorties de l'APA, telles qu'estimées à partir des données RI-APA 2011, correspondent principalement à des décès, mais aussi à des déménagements hors du département ou à des passages en établissement sous dotation globale.

- Pour CARE-M 2015, nous enlevons les individus de moins de 60 ans de l'enquête HSM et nous considérons une collecte en mai 2015 (soit 8 mois après VQS 2014), voir tableau 4.

Ces estimations sont très proches des taux observés sur les vraies collectes de VQS 2014 et CARE-ménages 2015.

Tableau 4 • Taux de réponse estimés vs. constatés après coup, par tranche d'âge

	Taux estimés à partir de VQS 2007 et HSM 2008				Taux observés
	60-69 ans	70-79 ans	80 ans ou plus	Ensemble	
Taux de réponse VQS 2014	64 %	62 %	47 %	58 %	57 %
Taux de réponse CARE-M seniors	69 %	66 %	55 %	73 %	71 %

Lecture > Selon les taux de réponse et de mobilité observés par âge pour les enquêtes VQS 2007 et HSM 2008, 73 % des individus âgés de 60 ans ou plus et tirés dans l'enquête CARE-ménages devraient y répondre. Le véritable taux de réponse sur cette enquête a été de 71 %.

Champ > Personnes de 60 ans ou plus résidant à domicile en France métropolitaine.

Sources > INSEE, DREES, VQS 2007, HSM 2008 ; Calculs DREES.

Un tirage de 120 000 ménages pour l'enquête filtre permettrait d'obtenir près de 70 000 ménages répondants grâce à une collecte postale et téléphonique. Parmi ces répondants, 15 000 seraient sélectionnés pour l'enquête CARE en face-à-face (1 personne par ménage), en surreprésentant les groupes les plus dépendants. On peut alors espérer 11 000 répondants dont 1 800 personnes estimées en GIR 1 à 4 et 515 personnes estimées en GIR 1-2, ce qui est en deçà de l'objectif visé.

Cependant, l'efficacité de l'enquête filtre peut être améliorée. En effet, le groupe d'autonomie de VQS 2007, bien que déterminé différemment selon l'âge, n'était pas spécifique à la population âgée et encore moins au repérage de la perte d'autonomie due au vieillissement. Or, contrairement à HSM, l'enquête CARE a un champ restreint aux personnes âgées de 60 ans ou plus, et il est donc nécessaire que la classification en groupes VQS capte plus spécifiquement les problèmes d'autonomie liés au vieillissement.

Il a donc été décidé d'amplifier le rôle de l'enquête filtre en adaptant le questionnaire et la stratification à la perte d'autonomie des seniors.

Enquête filtre VQS 2014

Pour le plan de sondage du volet ménages de Handicap-Santé (HSM), un score décrivant le degré de handicap avait été construit à partir des réponses à VQS 2007 (Midy 2010, Bouvier 2011). Ce score avait été défini en analysant les liens existant entre les réponses à VQS 1999 et le degré de handicap repéré dans HID concernant : les limitations et déficiences sensorielles, les limitations motrices, les limitations cognitives, les restrictions dans les activités essentielles de la vie quotidienne, les restrictions dans les activités instrumentales de la vie quotidienne et les restrictions de participation sociale (formation et emploi).

La construction du score se fait par incrémentation : chaque réponse ajoute 0, 1, 3 ou 6 points au score selon que la modalité représente l'absence de limitation, une faible limitation, une limitation moyenne ou une forte limitation. Ce score est ensuite divisé par le score maximal possible pour chaque individu, afin de tenir compte de la non-réponse partielle : le score maximal n'est pas incrémenté si la personne n'a pas répondu à la question. Le score est enfin normalisé : il est compris entre 0 et 100.

La classification dans un groupe de handicap était définie par des seuils de score variant selon l'âge de la personne (voir tableau 5).

Tableau 5 • Détermination du groupe d'autonomie selon le score et l'âge dans VQS 2007

	0 – 19 ans	20 – 59 ans	60 – 79 ans	80 ans ou plus	Groupe VQS
Score	0	[0 – 3]	[0 – 4]	[0 – 5]	I
	[1 – 9]	[4 – 11]	[5 – 24]	[6 – 39]	II
	[10 – 24]	[12 – 29]	[25 – 44]	[40 – 64]	III

	0 – 19 ans	20 – 59 ans	60 – 79 ans	80 ans ou plus	Groupe VQS
	[25 – 100]	[30 – 100]	[45 – 100]	[65 – 100]	IV

Lecture > Dans l'enquête VQS 2007, un individu âgé de 60 à 79 ans ayant un score VQS de 28 (compris entre 25 et 44) était classé dans le groupe III.

Champ > Personnes résidant à domicile en France métropolitaine.

Source > INSEE, DREES, VQS 2007.

Afin d'améliorer le rôle filtre de l'enquête VQS, une étude des liens entre VQS 2007 et HSM est effectuée. Deux objectifs sous-tendent cette analyse :

- recentrer la construction du score sur la perte d'autonomie liée au vieillissement et ainsi améliorer le pouvoir discriminatoire des groupes VQS par rapport aux objectifs de l'enquête CARE ;
- prendre en compte la précision accrue du questionnaire VQS, qui a, en 2007 et en 2014, une modalité supplémentaire par rapport à 1999 pour qualifier les difficultés.

La population d'intérêt est définie sous trois angles : l'âge, le niveau de dépendance et le bénéfice de l'APA. Ces trois variables définissent des populations qui se recoupent, mais qui ne se superposent pas :

- les plus âgés sont plus dépendants mais ils ne le sont pas tous ;
- les plus dépendants bénéficient plus souvent de l'APA, mais il existe du non-recours.

Par conséquent, la stratification doit prendre en compte ces trois éléments, qui seront soit collectés dans l'enquête (degré de dépendance et bénéfice de l'APA) soit présents dans la base de sondage (âge). Les strates tiendront compte du degré d'autonomie résumé par le groupe VQS et estimé sous la forme d'un score (du même type que pour VQS 2007) à partir des réponses à l'enquête filtre.

Les variables d'intérêt choisies pour représenter la perte d'autonomie liée au vieillissement sont le GIR estimé à partir des réponses à l'enquête HSM (regroupé en trois catégories : 1-2, 3-4, 5-6) ainsi qu'un indicateur de troubles de la cohérence (également en 3 modalités : A, B ou C). Les analyses sont effectuées via une modélisation logistique ordinaire (avec les trois modalités mentionnées) et séparément sur les moins de 80 ans et les 80 ans ou plus, afin de prendre en compte les spécificités de ces sous-populations quant à la perte d'autonomie. Dans les analyses, il a été constaté que la perte d'autonomie était différente selon ces tranches d'âge : plus de troubles physiques avant 80 ans contre plus de troubles cognitifs après 80 ans par exemple. Les variables explicatives sont les réponses aux questions de VQS 2007.

La modélisation est utilisée pour choisir les variables qui seront retenues ou non dans la construction du score VQS, selon qu'elles jouent significativement ou non sur l'une au moins des deux variables d'intérêt retenues et pour estimer leur importance afin de connaître le nombre de points à incrémenter dans le score VQS (voir annexe 1). L'analyse montre que les variables relatives aux déficiences sensorielles et au handicap déclaré ne sont pas discriminantes pour les variables d'intérêt de CARE et sont donc exclues du score. Par ailleurs, les restrictions sur l'activité essentielle de la vie quotidienne « se laver » sont importantes dans la définition de la perte d'autonomie due au vieillissement : la variable correspondante de HSM a été ajoutée pour simuler l'effet de l'ajout de cette question dans VQS 2014. De la même manière, une question sur le bénéfice de l'APA est ajoutée dans VQS 2014 afin de mieux cibler cette population d'intérêt.

Comme pour VQS 2007, le score est incrémenté d'un certain nombre de points selon la modalité répondue (ici de 1 à 12 points, voir annexe 2) ; il est ensuite ramené à un score sur 100 en prenant en compte la non-réponse partielle.

La distribution du nouveau score différant de celle du score défini pour VQS 2007, les seuils définissant les groupes ont été modifiés (voir tableau 6).

Tableau 6 • Détermination du groupe d'autonomie selon le score et l'âge dans VQS 2007 et 2014

	60 – 79 ans		80 ans ou plus		Groupe VQS
	VQS 2007	VQS 2014	VQS 2007	VQS 2014	
Score	[0 – 4]	[0 – 7]	[0 – 5]	[0 – 9]	I
	[5 – 24]	[8 – 23]	[6 – 39]	[10 – 34]	II
	[25 – 44]	[24 – 39]	[40 – 64]	[35 – 59]	III
	[45 – 100]	[40 – 100]	[65 – 100]	[60 – 100]	IV

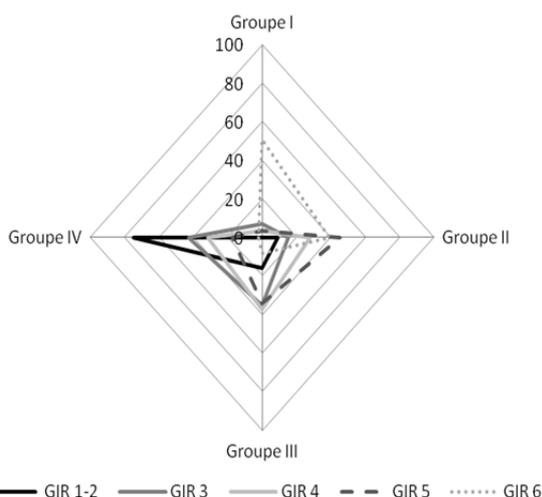
Lecture > Dans l'enquête VQS 2014, un individu âgé de 60 à 79 ans ayant un score VQS de 28 (compris entre 24 et 39) était classé dans le groupe III.

Champ > Personnes âgées de 60 ans ou plus résidant à domicile en France métropolitaine.

Sources > INSEE, DREES, VQS 2007 et 2014.

Ces nouveaux seuils ont été choisis de manière à avoir une bonne discrimination sur le critère « GIR estimé ». Le nouveau score permet ainsi d'avoir une meilleure discrimination dans les groupes : le groupe IV capte plus majoritairement des individus dépendants (GIR 1-4) et le groupe I plus exclusivement des individus autonomes (GIR 5-6) (voir graphiques 1 et 2). Cela permet d'avoir une meilleure optimisation du plan de sondage sur ce critère.

Graphique 1 • Répartition des répondants à HSM selon le GIR estimé et le groupe d'autonomie VQS déterminé par l'ancien score

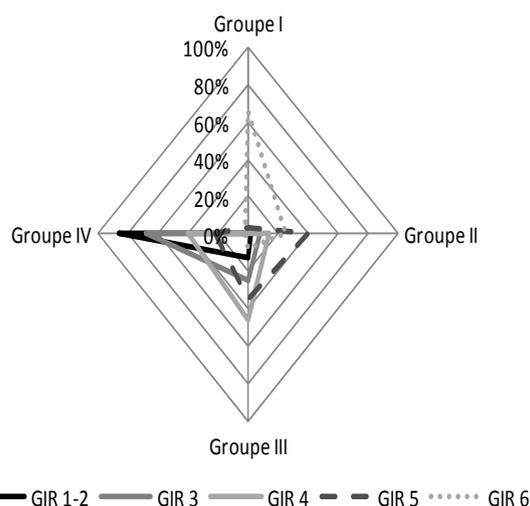


Lecture > 75 % des répondants à HSM estimés en GIR 1-2 sont classés dans le groupe IV avec l'ancien score.

Champ > Personnes âgées de 60 ans ou plus résidant à domicile en France métropolitaine.

Sources > INSEE, DREES, VQS 2007 et HSM 2008 ; Calculs DREES.

Graphique 2 • Répartition des répondants à HSM selon le GIR estimé et le groupe d'autonomie VQS déterminé par le nouveau score



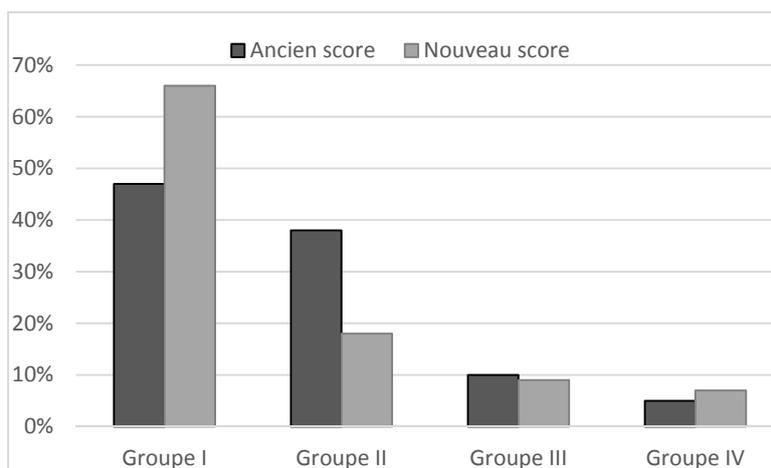
Lecture > 86 % des répondants à HSM estimés en GIR 1-2 sont classés dans le groupe IV avec le nouveau score.

Champ > Personnes âgées de 60 ans ou plus résidant à domicile en France métropolitaine.

Sources > INSEE, DREES, VQS 2007 et HSM 2008 ; Calculs DREES.

Cette optimisation modifie la répartition par groupe d'autonomie des répondants à VQS – qui est déterminée d'une part par les cibles visées pour la répartition par GIR et d'autre part, par la correspondance entre GIR et groupe VQS. Par exemple, la meilleure correspondance entre « GIR 6 » et « Groupe I » permet de réduire le nombre d'individus estimés en GIR 6 dans les groupes II, III ou IV et donc augmente l'effectif attendu du groupe I (voir graphique 3). Mais les effectifs des populations d'intérêt n'en sont pas altérés, car la meilleure discrimination assurée par le nouveau score confère une meilleure efficacité au plan de sondage.

Graphique 3 • Répartition des répondants à VQS 2007 selon le groupe, avec l'ancien et le nouveau score



Lecture > Dans l'enquête VQS 2007 et avec l'ancien score VQS, 47 % des répondants sont classés dans le groupe I, ils sont 66 % avec le nouveau score VQS.

Champ > Personnes âgées de 60 ans ou plus résidant à domicile en France métropolitaine.

Sources > INSEE, DREES, VQS 2007 ; Calculs DREES.

À taille d'échantillon égale (15 000 seniors), à plan de sondage équivalent et avec les taux de réponses attendus et supposés identiques dans chaque groupe qui sont mentionnés au tableau 4, les effectifs attendus par sous-population croissent sensiblement avec le nouveau score : 35 % de personnes estimées en GIR 1-2 supplémentaires et 28 % de personnes estimées en GIR 1-4 supplémentaires. En effet, les GIR 1-4 sont plus concentrés dans le groupe IV qu'avant, ainsi en sélectionnant de manière exhaustive ce groupe, on obtient plus d'individus dépendants. Le nouveau score optimisé par rapport au GIR permet également d'avoir des effectifs de répondants suffisants pour les autres populations d'intérêt, car celles-ci sont bien corrélées à l'indicateur GIR (voir tableau 7).

La solution de l'enquête filtre comporte le risque d'apparition de « faux négatifs », en cas d'écart temporel trop important entre les enquêtes VQS et CARE. Un faux négatif correspond à un individu classé en bonne santé à l'issue de VQS mais qui s'avèrerait en mauvaise santé dans CARE. Lors de l'enquête HID 1999, l'échantillon des personnes en bonne santé était trop petit. Le poids des personnes en bonne santé était alors élevé et un « faux négatif » a fortement perturbé les analyses (Bouvier, 2012). Pour limiter ce risque, un plus grand nombre de personnes en bonne santé a été tiré dans Handicap-Santé volet ménages (2008) et le rapport des poids entre les strates a été limité. Ces précautions sont également mises en œuvre pour le tirage de CARE seniors 2015, afin de se prémunir d'une influence excessive des « faux négatifs ».

La conduite d'une enquête VQS avec un filtre amélioré par rapport à 2007 a donc été retenue. L'estimation du nombre de répondants visé par population d'intérêt pour chacune des solutions est présentée dans le tableau 7.

Afin de se prémunir d'un taux de réponse en dessous des estimations, une réserve de 20 000 ménages a été ajoutée à cet échantillon principal.

Tableau 7 • Synthèse des résultats : taille d'échantillon et nombre de répondants estimés selon l'échantillonnage retenu

	Sans enquête filtre	Avec échantillon APA	Avec score VQS 2007	Avec score VQS 2014
Échantillon VQS (ménages)	-	-	120 000	120 000
Répondants VQS (ménages)	-	-	69 700	69 700
Échantillon CARE (individus)	42 700	19 000	15 000	15 000
Répondants CARE (individus)	24 500	10 300	10 700	10 700
GIR 1-4 (individus)	2 270	3 000	1 900	2 400
GIR 1-2 (individus)	620	640	515	700
Personne aidée par l'entourage	7 200	3 000	4 200	5 600

	Sans enquête filtre	Avec échantillon APA	Avec score VQS 2007	Avec score VQS 2014
<i>Personnes aidées par un professionnel</i>	5 200	2 300	3 000	4 000
<i>Personnes présentant des troubles de la cohérence</i>	930	710	720	910

Lecture > En tirant 120 000 ménages dans l'échantillon VQS 2014, et 15 000 individus dans l'échantillon CARE dont les plus dépendants repérés grâce au score VQS 2014, il est estimé que 10 700 individus répondront à l'enquête CARE-ménages dont 2 400 classés en GIR 1-4.

Champ > Personnes âgées de 60 ans ou plus résidant à domicile en France métropolitaine.

Sources > INSEE, DREES, VQS 2007, HSM 2008 ; Calculs DREES.

Représentativité départementale

La gestion de l'APA est décentralisée au niveau départemental et les bases de bénéficiaires de l'APA ne sont disponibles qu'au sein des conseils départementaux. La DREES envisageait, dès la conception de l'enquête CARE, des extractions de ces bases départementales pour les apparier avec l'enquête, et ainsi bénéficier d'une information plus précise sur la population des bénéficiaires de l'allocation. Ce type de travaux (nécessitant la participation des conseils départementaux) mobilise des moyens importants. En contrepartie de cet investissement, les conseils départementaux participants peuvent souhaiter un retour d'information permettant d'estimer des résultats sur leur territoire.

De façon générale, le besoin d'information sur la dépendance est important au niveau départemental, les départements étant chefs de file sur l'action sociale et médico-sociale auprès des personnes âgées. Les indicateurs disponibles concernant les personnes âgées repris pour certains parmi les Indicateurs socio-départementaux, ou ISD renvoient généralement au bénéfice d'une allocation ou prestation (APA, ACTP, PCH, ASH). Une mise en regard de ces indicateurs et de l'estimation de l'état de santé, et en particulier de la perte d'autonomie, de la population au niveau local permettrait d'éclairer les acteurs locaux.

Une première solution serait de réaliser des extensions départementales dans l'enquête CARE. Cette solution ne pourrait pas être mise en place pour tous les départements car cela serait trop coûteux. Pour une estimation au niveau local, la conduite d'extensions sur CARE demanderait la mise en place d'une approche « petits domaines ». Pour que celle-ci soit possible, il faut un échantillon suffisamment important dans chaque région, ce qui implique une stratification géographique (Le Guenec, 2012). L'échantillon VQS devrait aussi être plus conséquent. Pour réaliser une extension départementale avec un objectif d'environ 1 500 individus répondants pour un département, il faudrait disposer dans l'échantillon CARE de 2 200 individus, soit 9 500 logements à tirer sur le département dans VQS. En considérant que l'échantillon CARE est réparti de façon homogène sur le territoire métropolitain, on peut considérer qu'environ 1 300 logements sont tirés par département pour VQS. De ce fait, une extension sur un seul département conduirait à tirer environ 8 700 logements supplémentaires sur VQS et 2 000 individus supplémentaires sur CARE (voir Tableau 8).

Tableau 8 • Échantillon par département dans le cas d'extensions départementales dans l'enquête CARE

	Sans extension sur CARE*	Avec extensions sur CARE	Différence
<i>Échantillon VQS</i>	1 300	9 500	8 200
<i>Répondants VQS</i>	740	5 500	4 800
<i>Échantillon CARE</i>	160	2 200	2 040
<i>Répondants CARE</i>	120	1 500	1 420

* > Estimation moyenne en considérant un département dont la population est de 1/100 celle de la population française.

Lecture > Pour obtenir au moins 1 500 répondants par département dans l'enquête CARE-ménages, il faudrait en sélectionner 9 500 dans chaque département dans l'échantillon VQS.

Champ > Personnes âgées de 60 ans ou plus résidant à domicile en France métropolitaine.

Sources > INSEE, DREES, VQS 2007, HSM 2008 ; Calculs DREES.

Par ailleurs, des extensions sur CARE sur quelques départements se seraient avérées peu « rentables » compte tenu de l'écart important entre la complexité du questionnaire et les souhaits des conseils départementaux. En effet, lors du bilan des extensions départementales menées dans l'enquête Handicap-Santé en ménages, les

conseils départementaux ont relaté que le coût engendré par la collecte de cette enquête longue en face-à-face était plus important que le bénéfice qu'ils en avaient tiré. Notamment, la complexité du questionnaire a rendu son exploitation difficile et, conjuguée à un échantillon qui ne permettait pas d'isoler chaque sous-population d'intérêt, n'a que partiellement répondu à leurs préoccupations.

La deuxième solution était donc de réaliser ces extensions sur l'enquête VQS. Le format court de l'enquête VQS a l'avantage de pouvoir être réalisé sans face-à-face¹¹, ce qui en diminue le coût et laisse la possibilité d'un large échantillon. Cela permet de répondre facilement à un objectif de représentativité départementale pour tous les départements, permettant ainsi des comparaisons entre territoires. Le questionnaire simple garantit, de plus, une utilisation optimale par les acteurs locaux. Il explore divers aspects de la perte d'autonomie et permet d'estimer des prévalences sur chaque territoire.

L'introduction d'extensions au niveau de l'enquête filtre semble donc être un bon compromis et la solution la plus rentable pour tous les acteurs. En effet, l'enquête filtre peut aisément être stratifiée géographiquement, et le coût d'un enquêté supplémentaire est moins élevé pour cette enquête courte. De plus, le nombre d'interrogés pour l'enquête CARE, en face à face, ne sera pas modifié, puisque les extensions locales peuvent n'être réalisées qu'au niveau de l'enquête filtre et ne pas être prises en compte dans la base de tirage de CARE.

L'effectif enquêté dans VQS a donc été élargi et stratifié de sorte à assurer la représentativité de chaque département.

L'objectif pour ces extensions départementales était de disposer d'une moyenne de 2 000 individus de 60 ans ou plus répondants par département sur le territoire métropolitain¹², avec un minimum de 500 répondants par département.

En utilisant les taux de réponses du tableau 4 (supposé uniforme dans tous les départements) et le nombre de personnes de 60 ans ou plus par logement, le nombre de logements supplémentaires devant être tirés est de 60 000 (voir tableau 9).

Par ailleurs, 10 000 logements ont été tirés dans les DOM (hors Mayotte), afin d'obtenir un nombre de répondants de 500 individus par DOM, et produire des résultats également sur ces territoires.

¹¹ Une analyse des réponses de VQS 2007 sans face-à-face et avec face-à-face a permis de mettre en évidence qu'il ne risquait pas d'y avoir de biais de sélection sans ce mode de collecte (voir Annexe 3).

¹² Le tirage a ensuite pris en compte les disparités géographiques, en tirant un échantillon plus important pour les départements accueillant plus de personnes âgées de 60 ans ou plus.

Tableau 9 • Estimation de l'échantillon VQS permettant une représentativité départementale dans l'enquête VQS

	Nombre total France métropolitaine	Nombre moyen par département
<i>Échantillon total VQS</i>	200 000	2 080
<i>Dont principal hors réserve</i>	120 000	1 250
<i>Dont réserve</i>	20 000	210
<i>Dont extensions</i>	60 000	620
<i>Nombre de ménages répondants à VQS</i>	116 200	1 210
<i>Dont principal hors réserve</i>	69 700	730
<i>Dont réserve</i>	11 600	120
<i>Dont extensions</i>	34 900	360
<i>Nombre d'individus répondants à VQS</i>	186 600	1 940
<i>Dont principal hors réserve</i>	111 900	1 170
<i>Dont réserve</i>	18 700	190
<i>Dont extensions</i>	56 000	580

Lecture > Avec un échantillon de 200 000 ménages dans l'enquête VQS 2014, il a été estimé que 186 600 individus y répondraient à cette enquête soit en moyenne 1 940 individus par département.

Champ > Ménages comprenant au moins une personne âgée de 60 ans ou plus résidant à domicile en France métropolitaine.

Sources > INSEE, DREES, VQS 2007, HSM 2008 ; Calculs DREES.

■ PLAN DE SONDAGE DE VQS

Constitution de la base de sondage de VQS

Le champ des enquêtes VQS et CARE-Ménages correspond aux logements où est présente au moins une personne âgée de 60 ans et plus à la date de l'enquête CARE¹³. Les communautés, dont les maisons de retraite et les EHPAD¹⁴, sont exclues du champ. Une deuxième enquête, dite « CARE institutions » est menée par la DREES auprès des personnes âgées de 60 ans ou plus séjournant de manière permanente dans des établissements d'hébergement pour personnes âgées. Il était donc nécessaire de disjoindre au mieux les bases de sondage de ces deux enquêtes différentes dont les résultats pourront être agrégés. En revanche, les logements foyers non EHPAD (rebaptisé depuis « résidences autonomie »), assimilés à des communautés au sens du répertoire statistique des logements (RSL), sont inclus car ces logements sont équipés d'une cuisine qui en faisaient, du point de vue de l'enquête CARE, des logements dans lesquels la personne peut vivre de façon autonome.

La base de sondage a été constituée à partir des sources fiscales (taxe d'habitation, impôt sur le revenu et fichier d'imposition des personnes) et du RSL (Haag O., Loonis V., Mamari C., Vinet L., 2013). Les différentes étapes de constitution ont été les suivantes :

Restriction aux personnes de 60 ans ou plus

En recherchant les personnes nées avant le 2 mai 1955, le nombre d'individus présents dans le champ s'élève à 17 156 678 personnes. Ces individus peuvent faire partie d'une même unité de tirage : un logement fiscal. Un logement fiscal est un local répondant aux trois caractéristiques suivantes :

- le local est occupé ;
- le local est un appartement, une maison, une pièce indépendante ou une chambre domestique ;
- la section cadastrale du local est renseignée dans la base fiscale.

Ainsi, les 17 156 678 personnes nées avant le 2 mai 1955 appartiennent à 12 065 306 unités de tirage. Il est à noter que lorsqu'on considère les communautés (établissements d'hébergement pour personnes âgées, foyers de travail, casernes etc.), une unité de tirage équivaut souvent à un individu (plusieurs unités de tirage sont donc identifiées à une même adresse). Lorsqu'on considère les logements ordinaires, une unité de tirage équivaut plutôt à un logement et donc souvent une adresse (plusieurs individus appartiennent à une même unité de tirage).

Conservation des logements-foyers

La DREES a fourni deux bases d'adresses : la première contient les adresses de logements foyers (« résidences autonomie » depuis 2016) qui ne sont pas des communautés ou des EHPAD, et qu'elle souhaite donc conserver ; la deuxième contient une liste d'adresses d'EHPAD qui doit être au contraire retirée.

Un appariement à l'adresse postale (numéro de voie, indice de répétition, libellé de voie, commune) a été fait pour conserver les individus présents dans la base d'adresses de logements foyers. 29 276 individus ont ainsi été repérés à 807 adresses. La base de logements foyers contenant 2 132 adresses, le taux de recouvrement est de 37,9 %.

Cette base de logements-foyers (29 276 unités) est mise de côté. Les traitements suivants se font sur le reste des unités de tirage soit 12 036 030.

¹³ Soit l'ensemble des individus ayant au moins 60 ans le 2 mai 2015.

¹⁴ Établissement d'hébergement pour personnes âgées dépendantes.

Retrait des personnes vivant dans des communautés hors champ de l'enquête : constitution d'une base de « logements-ordinaires »

- le premier retrait concerne les maisons de retraite et EHPAD identifiables directement dans les sources fiscales. Il s'agit des individus pour lesquelles le complément de l'adresse contient les mots clés « MDR », « RETRAITE », ou « EHPA ». Les individus sont retirés de la base, soit ici 100 746 unités de tirage ;
- le deuxième retrait concerne les maisons de retraite et EHPAD identifiés par la DREES d'après le répertoire FINESS¹⁵. Comme pour les logements foyers, un appariement à l'adresse postale est fait, pour retirer les adresses communes. 2 801 adresses ont aussi été repérées (le fichier en contient 8 053, soit un taux d'appariement de 34,8 %). À ces adresses sont rattachées 48 319 unités de tirage, qui sont donc retirées ;
- le troisième retrait concerne les foyers hors champ. Cette fois-ci on recherche les mots « ADOMA » et « SONACOTRA ». Cela fait enlever 3 738 unités supplémentaires ;
- le quatrième retrait concerne des unités de tirage repérées comme appartenant à une communauté selon le RSL. Elles étaient au nombre de 222 005.

Nous appelons cette base la base des logements ordinaires, elle est composée de 11 661 222 unités de tirage.

Mise en commun de la base des logements-foyers et des « logements-ordinaires »

La base des logements-foyers (29 276 unités) est associée à celle des logements ordinaires (11 661 222 unités) soit une base de 11 690 498 unités. Malgré toutes les précautions prises dans les filtres précédents, il peut rester des communautés non identifiées. Ainsi si un logement¹⁶ hébergeait au moins 5 personnes âgées d'au moins 75 ans au 01/05/2015, ces individus ont été supprimés. 239 unités de tirage ont été retirées de la sorte.

Caractéristiques de la base finale de sondage

Au final, nous avons donc 11 690 259 unités de tirage (voir récapitulatif dans le tableau 10) réparties comme suit :

- 11 424 001 logements en résidence principale (logement fiscal¹⁷ soumis à la taxe d'habitation en résidence principale) ;
- 21 079 logements en résidence non principale (logement fiscal soumis à la taxe d'habitation mais pas en qualité de résidence principale) ;
- 245 179 foyers fiscaux non rattachés à un logement fiscal.

Tableau 10 • Constitution de la base de sondage VQS.

Unités...	Dont logements-foyers	Dont autres logements
Avec au moins une personne née avant le 2 mai 1955	29 276	12 036 030
Retrait des adresses contenant les mots clés « MDR », « RETRAITE », ou « EHPA »		– 100 746
Retrait des maisons de retraite et EHPAD identifiés par la DREES d'après le répertoire FINESS		– 48 319
Retrait des adresses contenant les mots clés « ADOMA » et « SONACOTRA »		– 3 738
Retrait des communautés au sens du RSL		– 222 005
Base des « logements-ordinaires »		11 661 222

¹⁵ Le fichier national des établissements sanitaires et sociaux (FINESS) recense les structures et équipements des domaines sanitaire, médico-social, social et de formation aux professions de ces secteurs; soumis à autorisation préalable en application des dispositions du code de la santé publique ou du code de l'action sociale et des familles.

¹⁶ Dans les faits, il s'agissait soit d'un logement au sens de la taxe d'habitation (TH) soit d'un foyer fiscal non rattaché à un logement au sens de la TH.

¹⁷ Un logement fiscal est un appartement ou une maison occupé disposant d'une référence cadastrale.

Unités...	Dont logements-foyers	Dont autres logements
Retrait des logements hébergeant au moins 5 personnes âgées d'au moins 75 ans		- 239
Base de sondage finale de l'enquête VQS		11 690 259*

* > 11 661 222 (« logements-ordinaires ») + 29 276 (« logements-foyers ») – 239 (logements avec plus de 5 personnes âgées).

Lecture > La base finale de tirage de l'échantillon VQS 2014 comporte 11 690 259 unités de tirage (logement fiscal).

Champ > Logements fiscaux comprenant au moins une personne âgée de 60 ans ou plus résidant à domicile en France métropolitaine.

Sources > INSEE, sources fiscales (taxe d'habitation, impôt sur le revenu et fichier d'imposition des personnes) et RSL 2013.

Tirage de l'échantillon

L'échantillon de VQS est construit comme la réunion de deux sous-échantillons :

- un échantillon principal, au sein duquel sera tiré l'échantillon de l'enquête CARE ;
- un échantillon complémentaire, servant uniquement à augmenter l'échantillon de VQS de façon à assurer la représentativité départementale (selon les critères discutés ci-avant).

Calcul des allocations de l'échantillon principal de 140 000 logements

Comme l'échantillon principal doit servir de base de sondage pour CARE, il a été décidé de le tirer comme tout échantillon d'enquête dans l'échantillon maître (EM)¹⁸, à savoir en cherchant à obtenir une équipondération des logements interrogés (Favre, 2009). Bien que l'échantillon soit tiré dans les sources fiscales, il est nécessaire de sélectionner l'échantillon principal (c'est-à-dire celui servant de base d'échantillonnage à l'enquête CARE-M) au sein des zones géographiques faisant parties de l'échantillon-maître pour que les individus répondants de cet échantillon soient couverts par le réseau des enquêteurs de l'INSEE et puissent être interrogés en face-à-face lors de la collecte de l'enquête CARE-M. Aucun minimum par zones d'action enquêteurs (ZAE) n'a été défini.

L'allocation par ZAE est donc la suivante :

$$n_{ZAE} = \frac{n_{princ} * N_{ZAE}}{\pi_{ZAE} * \sum_{ZAE \in EM} \frac{N_{ZAE}}{\pi_{ZAE}}}$$

Où :

n_{princ} est la taille de l'échantillon principal ;

N_{ZAE} est le nombre de logements dans le champ de l'enquête, pour chaque ZAE ;

π_{ZAE} est la probabilité d'inclusion de la ZAE dans l'EM ;

$\sum_{ZAE \in EM} \frac{N_{ZAE}}{\pi_{ZAE}}$ est l'estimateur du nombre total de logements du champ de l'enquête¹⁹.

L'allocation par département s'en déduit, par sommation des allocations des ZAE du département²⁰.

Il est à noter que 3 départements ne sont pas présents dans l'EM²¹ :

- Alpes-de-Haute-Provence (04) ;
- Lozère (48) ;
- Haute-Marne (52).

¹⁸ Un échantillon maître est un échantillon à partir duquel des sous-échantillons peuvent être sélectionnés afin de réaliser des enquêtes. L'échantillon-maître de l'INSEE couvre le territoire français répartis en zones sur lesquelles peuvent intervenir des enquêteurs pour réaliser des enquêtes en face-à-face, ce que nous nommons des Zones d'Action Enquêteurs (ZAE).

¹⁹ La base de sondage compte a priori 12 162 622 logements dans le champ et cet estimateur vaut 11 958 087, soit un écart de 1,7 %.

²⁰ Il est à noter qu'une ZAE a été affectée au département de sa commune pivot. Or il existe des ZAE multi départementales. Les comptes obtenus ici diffèrent donc à la marge de ceux réellement observés sur l'échantillon final.

²¹ La représentativité départementale ne faisait pas partie des contraintes lors de la définition de l'EM. Le tirage des ZAE de l'EM n'a été équilibré qu'au niveau régional.

Échantillon complémentaire de 60 000 logements

Cet échantillon permet de compléter l'échantillon principal pour assurer une diffusion départementale de l'ensemble de l'échantillon VQS et ne nécessite pas un tirage dans l'échantillon-maître puisque sans contrainte d'une collecte en face-à-face. En outre, il n'était pas possible de tirer l'intégralité de l'échantillon VQS dans l'échantillon-maître, d'une part, puisque l'échantillon-maître ne couvre pas tous les départements comme indiqué plus haut et d'autre part, pour ne pas épuiser l'échantillon-maître (contrainte de ne pas interroger un même individu sur des enquêtes différentes la même année).

Il doit permettre d'avoir au moins 1 000 logements de l'échantillon VQS par département. Cette contrainte se traduira bien entendu in fine par une dispersion des poids plus importante que pour celle observée dans les enquêtes ménages « traditionnelles ».

Comme il ne donnera pas lieu à une interrogation en face à face, il a été décidé de le disjoindre le plus possible de l'EM.

Il n'était toutefois pas possible de le tirer totalement en dehors de l'EM car le complémentaire de l'EM ne contenant aucune très grande commune, ne peut être considéré comme représentatif du territoire français.

Ainsi, la base de sondage contient :

- l'ensemble des ZAE ayant une probabilité d'inclusion à l'EM supérieure à 0,5²² ;
- les ZAE hors EM ayant une probabilité d'inclusion à l'EM strictement inférieure à 0,5.

À partir du minimum départemental (fixé à 1 000 logements échantillonnés), le reliquat R qu'il restait à répartir par département en fonction du minimum à tirer par département et de l'allocation départementale de l'échantillon principal a été calculé de la façon suivante :

$$R = 200\,000 - 96 * n_{dep}^{min} - \sum_{dep} \max(0 ; n_{dep}^{princ} - n_{dep}^{min})$$

Où :

n_{dep}^{min} est le minimum de logements à tirer par département (1 000 dans notre cas) ;

n_{dep}^{princ} est le nombre de logements tirés dans l'échantillon principal pour le département.

Ce reliquat est ensuite réparti au prorata du nombre de logements dans le champ par département.

L'allocation départementale à tirer dans l'échantillon complémentaire devient alors :

$$n_{dep}^{compl} = \text{arrondi} \left(\max(n_{dep}^{princ} ; n_{dep}^{min}) + R * \frac{N_{dep}}{\sum_{dep} N_{dep}} - n_{dep}^{princ} \right)$$

Où N_{dep} est le nombre de logements dans le champ, par département.

On déduit ensuite l'allocation par ZAE de cette allocation départementale de la façon suivante :

$$n_{ZAE}^{compl} = \frac{n_{dep}^{compl} * N_{ZAE}}{(1 - \pi_{ZAE}) * k_{\pi_{ZAE} < 0,5} * \sum_{ZAE \in bds_{compl}} \left(\frac{N_{ZAE}}{(1 - \pi_{ZAE}) * \mathbb{1}_{\pi_{ZAE} < 0,5}} * \mathbb{1}_{ZAE \in dep} \right)}$$

Où :

$$\mathbb{1}_{ZAE \in dep} = \begin{cases} 1 & \text{si la ZAE est dans le département "dep"} \\ 0 & \text{sinon} \end{cases}$$

$$k_{\pi_{ZAE} < 0,5} = \begin{cases} 1 & \text{si } \pi_{ZAE} < 0,5 \\ \frac{1}{1 - \pi_{ZAE}} & \text{sinon} \end{cases}$$

²² Il s'agit des grandes communes. Ce seuil de 0,5 a été retenu afin d'avoir une bonne représentativité de la base de sondage tout en évitant de tirer trop de FA dans les grandes communes. En effet, ce nombre est proportionnel à $\frac{1}{1 - \pi_{ZAE}}$.

Calcul des poids

Deux jeux de poids sont calculés :

- un pour l'échantillon principal seul. Ce poids sera utile pour le calcul du poids de sondage des individus tirés dans CARE ;
- un sur l'ensemble de l'échantillon VQS (principal + complémentaire) qui sera utile pour l'exploitation de l'enquête VQS.

Poids à utiliser pour CARE

Il s'agit du poids traditionnel d'un échantillon de logements en 2 phases tiré dans l'EM.

On note $\omega_{log\ ech\ p}^{ZAE}$ le poids d'un logement échantillonné dans l'échantillon principal dans une ZAE. Ce poids est égal à :

$$\omega_{log\ ech\ p}^{ZAE} = \frac{N_{ZAE}}{\pi_{ZAE} * n_{ZAE}^{princ}}$$

Où :

N_{ZAE} est le nombre de logements du fichier TH dans le champ, pour la ZAE ;

π_{ZAE} est la probabilité d'inclusion de la ZAE dans l'EM ;

n_{ZAE}^{princ} est le nombre de logements échantillonnés dans l'échantillon principal pour la ZAE.

Poids à utiliser pour VQS

L'échantillon complémentaire n'est pas tiré dans une base totalement disjointe de celle de l'échantillon principal. Il est donc nécessaire de procéder à un partage des poids pour le calcul du poids VQS définitif.

Le poids d'un logement de l'échantillon complémentaire, noté, $\omega_{log\ ech\ c}^{ZAE}$, vaut :

$$\omega_{log\ ech\ c}^{ZAE} = \frac{N_{ZAE}}{n_{ZAE}^{compl}} \text{ pour la strate des ZAE telles que } \pi_{ZAE} > 0,5 ;$$

$$\omega_{log\ ech\ c}^{ZAE} = \frac{N_{ZAE}}{(1-\pi_{ZAE}) * n_{ZAE}^{compl}} \text{ pour les autres.}$$

La technique du partage des poids permet ensuite de réunir ces deux échantillons (principal et complémentaire) et pondérer l'ensemble obtenu pour pouvoir inférer sans biais à la population France métropolitaine (ou sur toutes sous-populations, comme par exemple pour des études départementales).

Le principe est le suivant : $\omega_{log\ ech\ tot}^{ZAE} = \alpha * \omega_{log\ ech\ p}^{ZAE} + (1 - \alpha) * \omega_{log\ ech\ c}^{ZAE}$

Comme la variance des estimations est inversement proportionnelle à la taille de l'échantillon, le coefficient α optimum introduit dans le partage des poids peut être approximé par le ratio suivant : nombre de logements via l'EM (soit 140 000) divisé par le nombre de logements total de l'échantillon (soit 200 000).

■ PRINCIPE DU CALCUL ANTICIPÉ DES ALLOCATIONS POUR CHESTER²³

L'enquête VQS était réalisée par voie postale, par internet ou par téléphone, n'importe quel enquêteur pouvait donc réaliser l'enquête téléphonique de n'importe quel enquêté. La répartition de l'échantillon entre enquêteurs était donc facilitée. Par contre, pour CARE, il était nécessaire d'estimer la charge d'enquête pour chaque enquêteur. L'échantillon d'individus CARE a été enquêté à partir de mai 2015. Dans le cadre du nouveau cadre d'emploi des enquêteurs (NCEE), il a été nécessaire de fournir pour le mois de septembre 2014 une estimation du temps total d'enquête par ZAE, que nous nommerons par la suite « allocations par ZAE ». Cette estimation permet de calculer le temps total de toutes les enquêtes de l'INSEE réalisées par des enquêteurs qui auront lieu en 2015, afin de planifier l'activité des enquêteurs pour l'année à venir.

Comme présenté ci-dessus, cet échantillon de 15 000 individus est particulier car il a été tiré parmi les ménages répondant à l'enquête VQS. La « base de tirage » de l'enquête n'a donc été connue qu'en début d'année 2015. Et, contrairement aux enquêtes plus traditionnelles, on ne disposait pas au moment du calcul des allocations d'une base de sondage proche de celle attendue.

Cette partie a pour objet de présenter la méthodologie et les hypothèses faites pour la mise en œuvre de ces calculs d'allocations.

Les allocations à livrer à Chester

La base de sondage de CARE constituée des personnes répondantes à VQS âgées de 60 ans ou plus au 2 mai 2015 était disponible à la fin de l'enquête VQS en février 2015. Or, pour planifier le travail des enquêteurs, deux allocations différentes au niveau ZAE ont dû être calculées en septembre 2014. En effet, 2 temps de passation du questionnaire en face-à-face différents selon le degré de dépendance des individus ont été estimés à partir des durées d'interrogation du test CAPI pour le calcul de la charge de travail des enquêteurs. Ainsi, les allocations suivantes ont été fournies :

- une allocation des individus des groupes VQS I et II (les personnes les plus autonomes et pour lesquelles le temps d'interrogation provisionné dans chester est moins important) ;
- une allocation des individus des groupes VQS III et IV (les personnes les plus dépendantes).

Choix de la base de sondage pour les calculs

Deux solutions étaient envisageables :

- partir de l'échantillon VQS, simuler la non-réponse pour obtenir une base de ménages répondants, sélectionner un individu par ménage répondant et calculer les allocations sur cette base en cherchant à minimiser la dispersion des poids des individus par strate ;
- partir du champ complet de l'enquête CARE issu des fichiers fiscaux²⁴, sélectionner un individu par ménage comportant plusieurs individus dans le champ et calculer les allocations sur cette base en cherchant à minimiser la dispersion des poids des individus par strate.

C'est finalement cette deuxième solution qui a été mise en œuvre. Elle permet de limiter le nombre de paramètres à estimer dans nos modèles de calcul des allocations car elle ne nécessite pas de connaître les taux de réponse à VQS. En effet, on connaît mal les déterminants de la non-réponse à VQS et dans ces conditions il devient difficile d'estimer une probabilité de réponse fiable par ménage.

²³ Outil de planification pour la répartition des charges entre les enquêteurs.

²⁴ Il s'agit des individus de 60 ans ou plus au 2 mai 2015.

Les hypothèses mises en œuvre

On ne dispose pas du groupe d'autonomie des individus dans la base de sondage de CARE issue des sources fiscales (cette information n'est en effet disponible qu'à l'issue de la collecte de VQS). Les seules données dont on dispose sont les résultats des enquêtes VQS/HSM de 2007-2008, qui fournit au niveau national une évaluation de la répartition de la population par groupe VQS et par tranches d'âge (60-69 ans, 70-79 ans et 80 ans ou plus). On suppose ici que ces proportions sont toujours valables en 2015. Ces proportions sont les suivantes :

Tableau 11 • Répartition des individus selon leur âge et leur groupe VQS

	60-69 ans	70-79 ans	80 ans ou plus	Total
Groupe VQS I (autonomes)	76 %	58 %	30 %	59 %
Groupe VQS II	18 %	25 %	38 %	25 %
Groupe VQS III	4 %	10 %	21 %	10 %
Groupe VQS IV (dépendants)	2 %	8 %	11 %	6 %

Lecture > 30 % des individus de 80 ans ou plus sont classés dans le groupe VQS (nouveau score) I en 2007.

Champ > Individus âgés de 60 ans ou plus résidant à domicile en France métropolitaine.

Source > INSEE, DREES, VQS 2007 ; Calculs DREES.

En outre, pour la sélection d'un individu par ménage, comme on ne connaît pas a priori le résultat du processus de sélection de l'individu CARE au sein des ménages répondants à VQS, on prend comme référence l'âge de la personne la plus âgée dans le ménage, de façon à avoir la prévision la plus prudente possible en termes de temps d'enquête²⁵. En effet, comme le montre le tableau ci-avant, la proportion de personnes très dépendantes augmente avec l'âge.

Ainsi, le nombre de personnes dans le champ pour chaque ZAE*strate est obtenu en appliquant la répartition des groupes VQS par tranche d'âge (du tableau 11 ci-dessus) au nombre d'individus²⁶ par tranche d'âge de la ZAE.

Le calcul anticipé des quotités par ZAE pour la planification de l'activité 2015

Pour un groupe VQS et une tranche d'âge donnés, l'allocation optimale par ZAE est calculée à partir de l'allocation totale à tirer pour cette strate et du nombre de personnes « potentiellement interrogeables » dans ce groupe VQS estimé dans chaque ZAE. L'objectif de l'échantillonnage est de minimiser la dispersion des poids par ZAE pour une strate donnée.

Ainsi, la formule de calcul de l'allocation par ZAE est la suivante :

$$n_h^{ZAE} = \frac{N_h^{ZAE} * n_h}{N_h * \pi_{ZAE}}$$

Avec :

h le groupe VQS ;

π_{ZAE} la probabilité d'inclusion de la ZAE dans l'échantillon maître ;

n_h le nombre d'individus à tirer dans la strate « h ». Ce nombre a été fourni par la DREES ;

N_h^{ZAE} le nombre d'individus (i.e. uniquement le plus âgé du logement) de la base de sondage de la ZAE appartenant à la strate « h » ;

$$N_h = \sum_{ZAE} N_h^{ZAE}$$

²⁵ Cette décision a été prise afin de se placer dans la situation la plus prudente vis à vis du calcul des quotités dans chester vu que le temps d'entretien des individus les plus âgés sera plus long. L'objectif était d'éviter autant que possible de se trouver en sous-quotité dans une ZAE où le hasard du tirage Kish ferait qu'on sélectionne souvent la personne la plus dépendante.

²⁶ Seul l'individu de 60 ans ou plus le plus âgé d'un ménage est retenu pour ces calculs.

Afin d'éviter tout biais, une borne minimum d'un individu par strate est nécessaire à partir du moment où la strate contient au moins un individu au départ. Ceci permet en effet d'éviter que des individus aient une probabilité nulle d'être dans l'échantillon.

Tableau 12 • Répartition de l'échantillon alloué à partir des spécifications de la DREES

	60-69 ans	70-79 ans	80 ans ou plus	Total
Groupe VQS I (autonomes)	1 639	809	353	2 802
Groupe VQS II	1 120	982	1 275	3 378
Groupe VQS III	836	1 387	2 533	4 755
Groupe VQS IV (dépendants)	728	1 498	1 839	4 065
Total	4 323	4 677	6 000	15 000

Lecture > Selon les hypothèses de réponse pour VQS 2014 et le taux de sondage utilisé pour sélectionner l'échantillon CARE-ménages, 1 639 individus âgés de 60 à 69 ans appartenant au groupe VQS I seront tirés dans l'échantillon CARE-Ménages.

Champ > Individus âgés de 60 ans ou plus résidant à domicile en France métropolitaine.

Source > INSEE, DREES, VQS 2007, HSM 2008 ; Calculs DREES.

Enfin, pour le calcul de la charge de travail par enquêteur, deux temps de remplissage du questionnaire sont prévus selon le niveau de dépendance des personnes (personnes des groupes VQS I et II versus III et IV). Ainsi, les allocations fournies par ZAE seront les suivantes :

$$n_1^{ZAE} = \sum_{\text{groupe VQS}=I}^{II} n_{\text{groupe VQS}}^{ZAE}$$

$$n_2^{ZAE} = \sum_{\text{groupe VQS}=III}^{IV} n_{\text{groupe VQS}}^{ZAE}$$

En fournissant ces allocations, on fixe donc le temps de traitement de CARE à allouer par ZAE. Ce temps vaudra :

$$T_{ZAE}^{CARE} = T^{rep} * (n_1^{ZAE} + n_2^{ZAE}) + n_1^{ZAE} * T_1 + n_2^{ZAE} * T_2$$

Où :

T^{rep} est le temps de repérage et de prise de contact ;

T_1 est le temps de passation des questionnaires des personnes fortement dépendantes (y compris le temps de passation des aidants en face-à-face) ;

T_2 est le temps de passation des questionnaires des personnes moins dépendantes (y compris le temps de passation des aidants en face-à-face).

Pour l'allocation finale, cette contrainte de temps total par ZAE sera respectée. Sous cette contrainte, et au vu des répondants à VQS, les allocations n_1^{ZAE} et n_2^{ZAE} pourront être modifiées par rapport au calcul prévisionnel initial pour limiter au maximum la dispersion des poids, tout en respectant bien entendu en parallèle les allocations par strate fournies par la DREES.

■ TIRAGE DE L'ÉCHANTILLON CARE

Le principe général

L'objectif recherché est toujours d'essayer autant que faire se peut d'aboutir in fine à une équipondération des individus. La difficulté ici est en plus de respecter les deux contraintes suivantes :

- le nombre d'individus par groupe VQS fourni par la DREES
- le temps total alloué dans chester par ZAE. En revanche, le respect stricto sensu des allocations n_1^{ZAE} et n_2^{ZAE} calculées de façon anticipée pour la prévision des quotités (cf. partie III) n'était pas nécessaire.

Dans les faits, ces allocations plus fines ont été respectées quand ceci était possible.

Dans le cas de CARE, le poids d'un individu dépend :

- du poids de sondage de son logement dans l'échantillon VQS ;
- du poids de sondage issu de la sélection de l'individu parmi les répondants de VQS.

Ainsi, afin de se rapprocher de l'équipondération des individus tirés, l'idée est de sélectionner l'individu dans CARE par le biais d'un tirage à probabilité inégale.

L'objectif d'équipondération s'écrit directement au niveau individu, en introduisant les notations suivantes : π_{BS}^i qui est la probabilité que l'individu i soit présent dans la base de sondage CARE, et qui est le produit de la probabilité que le logement soit tiré et réponde à VQS par la probabilité que l'individu i soit sélectionné dans son logement ; et $\pi_{CARE/BS}^i$ qui est la probabilité que l'individu soit tiré pour CARE sachant qu'il est dans la base de sondage. Dans ce cas :

$$\omega^i = \frac{1}{\pi_{BS}^i} * \frac{1}{\pi_{CARE/BS}^i} = \frac{N_{groupe j}}{n_{groupe j}}$$

D'où on en déduit que :

$$\pi_{CARE/BS}^i = \frac{\hat{\omega}^i * n_{groupe j*ZAE}}{\sum_{i \in ZAE} \hat{\omega}^i * \mathbb{1}_{i \in groupe j}}$$

Où finalement :

$\hat{\omega}^i$ est le poids de l'individu i répondant à VQS et qui a été sélectionné dans CARE. C'est-à-dire le poids du logement de VQS après traitement de la non-réponse que multiplie le nombre d'individus du champ de CARE de ce logement ;

$n_{groupe j*ZAE}$ est l'allocation par strate calculée dans la partie précédente.

On a donc bien in fine un tirage proportionnel au poids des individus répondants à VQS.

Les différentes étapes de ce tirage et notamment l'estimation de $\hat{\omega}^i$ sont décrites en détail dans les sous-sections suivantes.

L'estimation du poids des logements répondants à VQS

Afin de disposer d'un poids des logements répondants à l'enquête VQS, poids qui sera utilisé pour le tirage à probabilité inégale décrit ci-après, la non-réponse dans VQS a été traitée par repondération selon la méthode des Groupes de Réponse Homogène (GRH) (Caron N., 2005).

Une première régression logistique a permis de déterminer les variables les plus explicatives de la non-réponse parmi les variables auxiliaires disponibles²⁷. Six variables ont finalement été retenues :

- la nature du logement ;
- le revenu total du foyer (en tranche) ;

²⁷ Il s'agit de diverses variables issues de la base de sondage.

- l'âge de la personne la plus âgée ;
- la tranche d'unité urbaine ;
- la région ;
- le type d'occupation.

Ce modèle a permis d'estimer une probabilité de réponse de chaque logement parmi les 15 000 que contient la base de sondage. Cette dernière a ensuite été triée selon cette probabilité de réponse. Les 10 groupes de réponse homogènes (GRH) sont ensuite constitués en regroupant les 1 400 premiers logements de la base triée dans le 1^{er} GRH, les 1 400 suivants dans le 2^e etc.

Tableau 13 • Médiane, moyenne et écart-type de la probabilité de réponse estimée pour chacun des 10 groupes homogènes de réponse

GRH	Moyenne	Médiane	Écart-type
1	0,3759	0,3913	0,0548
2	0,4681	0,4689	0,0165
3	0,5151	0,5150	0,0122
4	0,5520	0,5517	0,0090
5	0,5850	0,5855	0,0089
6	0,6129	0,6116	0,0083
7	0,6407	0,6413	0,0073
8	0,6665	0,6657	0,0086
9	0,6946	0,6945	0,0080
10	0,7357	0,7313	0,0186

Lecture > La probabilité de réponse du premier groupe homogène de réponse (GHR) est estimée à 37,59 %.

Champ > Individus âgés de 60 ans ou plus résidant à domicile en France métropolitaine.

Source > INSEE, DREES, Base de sondage de VQS 2014, VQS 2014 ; Calculs INSEE.

Une fois ces groupes homogènes de réponse constitués, les probabilités de réponse sont supposées indépendantes d'un groupe à l'autre et constantes au sein de chaque groupe²⁸. La probabilité de réponse $\hat{\pi}_i^{rep}$ d'un logement i appartenant à un groupe h donné est alors estimée à partir des poids de sondage initiaux $\omega_{log\ ech\ p}^{ZAE}$ (voir précédemment), par :

$$\hat{\pi}_i^{rep} = \hat{\pi}_h^{rep} = \frac{\sum_{i=1}^{140\ 000} (\omega_{log\ ech\ p}^{ZAE} * \mathbb{1}_{i \in ZAE} * \mathbb{1}_{i \in h} * \mathbb{1}_{i \in \text{répondant}})}{\sum_{i=1}^{140\ 000} (\omega_{log\ ech\ p}^{ZAE} * \mathbb{1}_{i \in ZAE} * \mathbb{1}_{i \in h})}$$

Avec :

$$\mathbb{1}_{i \in ZAE} = \begin{cases} 1 & \text{si } i \in ZAE \\ 0 & \text{sinon} \end{cases}$$

$$\mathbb{1}_{i \in h} = \begin{cases} 1 & \text{si } i \in \text{strate } h \\ 0 & \text{sinon} \end{cases}$$

$$\mathbb{1}_{i \in \text{répondant}} = \begin{cases} 1 & \text{si } i \text{ a répondu à VQS} \\ 0 & \text{sinon} \end{cases}$$

À l'issue de cette étape, les poids des logements répondants ont été ajustés de façon à prendre en compte le mécanisme de non-réponse. La pondération d'une unité i donnée suite à cette correction de la non-réponse totale (CNRT) vaut donc :

$$\hat{\omega}_i^{VQS} = \frac{\omega_{log\ ech\ p}^{ZAE} * \mathbb{1}_{i \in ZAE}}{\hat{\pi}_i^{rep}} * \mathbb{1}_{i \in \text{répondant}}$$

²⁸ De façon à limiter la sensibilité des estimations au modèle de non-réponse.

Le poids des individus de la base de sondage CARE

Une des contraintes était de ne pas interroger pour CARE plus d'une personne dans un même logement. En effet, le protocole CARE prévoit d'interroger en plus d'une personne de 60 ans ou plus du logement tous ses aidants de l'entourage, dont certains peuvent résider dans le même logement. Interroger plusieurs personnes d'un même logement aurait pu conduire à des doubles interrogations de personnes d'une part et à un temps de collecte extrêmement long d'autre part.

Pour respecter cette contrainte, la première étape du tirage de l'échantillon a été de sélectionner, à probabilité égale, dans chaque logement un individu parmi les individus répondants à VQS et faisant partie du champ de l'enquête CARE. On obtient alors une base d'individus et non plus de logements.

À ce niveau on peut considérer que la probabilité de tirage d'un individu à l'issue de cette première phase du tirage est :

$$\pi_i^{log} = \frac{1}{\sum_{j \in log} (\mathbb{1}_{j \in \text{répondant VQS}} * \mathbb{1}_{j \in \text{champ CARE}})}$$

Où :

$$\mathbb{1}_{j \in \text{répondant VQS}} = \begin{cases} 1 & \text{si l'individu } j \text{ du ménage a répondu à VQS} \\ 0 & \text{sinon} \end{cases}$$
$$\mathbb{1}_{j \in \text{champ CARE}} = \begin{cases} 1 & \text{si l'individu } j \text{ du ménage appartient au champ de CARE} \\ 0 & \text{sinon} \end{cases}$$

On en déduit donc le poids des individus de la base de sondage de CARE, qui sont obtenus en faisant le produit du poids estimé du logement répondant à VQS par le nombre d'individus du champ de CARE dans ce logement. En reprenant les notations précédentes, on obtient :

$$\hat{\omega}^i = \frac{\hat{\omega}_{log}^{VQS}}{\pi_i^{log}}$$

Afin de se rapprocher de l'équipondération des logements, le tirage de CARE se fera donc proportionnellement à ce poids dans chaque strate de tirage.

Les modifications à la marge des allocations théoriques

Comme, on l'a vu dans la troisième partie, l'allocation de l'échantillon a été calculée à partir d'une base de sondage qui n'est pas celle dans laquelle est réalisé le tirage et ce afin de répondre aux contraintes du NCEE.

Cette contrainte a deux conséquences importantes et plutôt négatives quant à l'objectif d'équipondération :

- l'allocation à respecter n'est plus optimale par rapport à celle qui aurait été calculée avec la base de sondage définitive ;
- il existe des strates pour lesquelles la taille de l'échantillon à tirer est supérieure au nombre d'individus dans la strate.

Si le premier problème n'a pas été corrigé²⁹, le second, lui, l'a été de la façon suivante.

Sur les 2 268 strates de tirage (567 ZAE * 4 groupes d'autonomie), 226 étaient sujettes à ce problème. Ce manque d'individus dans la base de sondage ne concernait que des strates d'individus dépendants (groupes VQS III et IV).

Ces strates peuvent être classées en deux catégories :

- les ZAE pour lesquelles un seul groupe VQS était déficitaire et pour lesquelles le groupe équivalent au sens chester³⁰ permettait de combler ce déficit (136 cas). Ces ZAE seront appelées « cas simple » par la suite ;

²⁹ Il était en effet très difficile de converger vers une allocation optimale qui permette à la fois de respecter les contraintes chester par ZAE et la taille de l'échantillon par groupe d'autonomie voulue par la DREES.

³⁰ Il s'agissait du groupe III (respectivement IV) si le groupe IV (respectivement III) était déficitaire.

- les autres ZAE (68 ZAE dont 22 qui avaient un déficit pour les 2 groupes VQS). Ces ZAE seront appelées « cas complexes ».

Les ZAE « cas simples »

Pour résoudre le problème de ces ZAE, l'idée était :

- de tirer dans l'échantillon l'ensemble des individus du groupe VQS à problème ;
- d'augmenter l'allocation du groupe VQS équivalent au sens chester de la ZAE afin de respecter la contrainte chester ;
- de sélectionner une ZAE dite partenaire, « sans problème » a priori, dans le même département si possible sinon dans la même région et de faire le traitement inverse (i.e. augmenter l'allocation du groupe VQS à problème de la ZAE à problème du déficit observé dans la première ZAE et de réduire d'autant l'allocation du groupe chester équivalent). Ce traitement permet de conserver l'allocation par groupe d'autonomie « France entière » imposé par la DREES. Était choisie comme ZAE partenaire celle ayant l'allocation la plus forte pour le groupe VQS à problème. Cette ZAE est choisie de façon à réduire les poids extrêmes. Cependant dans certains cas où le déficit était trop grand (au-delà de 3 pour une strate donnée), plusieurs ZAE étaient sélectionnées afin de ne pas déséquilibrer trop fortement l'une d'elles et entraîner une forte augmentation des poids. La répartition des allocations se faisait alors équitablement entre les ZAE de soutien.

Exemple :

ZAE à problème dans un cas simple

	Groupe III	Groupe IV	Total
<i>Nombre d'individus dans la base de sondage</i>	5	10	15
<i>Allocation chester initiale</i>	7	6	13
<i>Allocation tirée</i>	5 (-2)	8 (+2)	13

ZAE « partenaire » sans problème a priori

	Groupe III	Groupe IV	Total
<i>Nombre d'individus dans la base de sondage</i>	15	17	32
<i>Allocation chester initiale</i>	9	8	17
<i>Allocation tirée</i>	11 (+2)	6 (-2)	17

Pour les deux ZAE

	Groupe III	Groupe IV	Total
<i>Allocation chester initiale</i>	16	14	30
<i>Allocation tirée</i>	16	14	30

Les ZAE « cas complexes »

Pour résoudre le problème de ces ZAE, l'idée était :

- de tirer dans l'échantillon l'ensemble des individus des deux groupes d'autonomie III et IV ;
- d'augmenter l'allocation des groupes I et II dans la ZAE afin de respecter la charge de travail totale pour la ZAE. Sachant que pour Chester, en termes de temps d'enquêtes, 2 questionnaires des groupes III et IV sont équivalents à 3 questionnaires des groupes I et II, c'est cette clé de répartition qui a été mise en œuvre pour définir le nombre de questionnaires à rajouter dans les groupes I et II. Entre les groupes I et II, le choix du nombre d'unités à rajouter s'est fait au prorata du poids de ces groupes VQS dans la base de sondage ;

- de sélectionner une ZAE dite partenaire, « sans problème » a priori, dans le même département si possible, sinon dans la même région, et de faire le traitement inverse (i.e. augmenter l'allocation des groupes d'autonomie à problème du déficit observé dans la première ZAE et réduire l'allocation des groupes I et II afin de conserver la charge de travail prédéfinie). Ce traitement permet de conserver l'allocation par groupe VQS « France entière » imposée par la DREES. Était choisie comme ZAE partenaire celle ayant la somme des marges la plus grande entre les allocations pour les groupes à problème et le nombre de répondants dans les strates correspondantes. Cependant dans certains cas où le déficit était trop grand (au-delà de 3 pour une strate donnée), plusieurs ZAE étaient sélectionnées afin de ne pas déséquilibrer trop fortement l'une d'elles et entraîner une explosion des poids. La répartition des allocations se faisait alors équitablement entre les ZAE partenaires. Enfin, la réduction du nombre d'individu dans les groupes I et II pour les ZAE partenaires a également été contrôlée afin de ne pas trop augmenter leur poids.

Exemple :

ZAE à problème dans un cas complexe

	Groupe I	Groupe II	Groupe III	Groupe IV	Total
<i>Nombre d'individus dans la base de sondage</i>	53	74	3	3	133
<i>Allocation chester initiale</i>	3	3	5	4	15
<i>Allocation tirée</i>	5 (+2)	6 (+3)	3 (-2)	3 (-1)	17

ZAE « partenaire » sans problème *a priori*

	Groupe I	Groupe II	Groupe III	Groupe IV	Total
<i>Nombre d'individus dans la base de sondage</i>	41	35	11	8	95
<i>Allocation chester initiale</i>	4	5	6	5	20
<i>Allocation tirée</i>	2 (-2)	2 (-3)	8 (+2)	6 (+1)	18

Pour les deux ZAE

	Groupe I	Groupe II	Groupe III	Groupe IV	Total
<i>Allocation chester initiale</i>	7	8	11	9	35
<i>Allocation tirée</i>	7	8	11	9	35

Le tirage proprement dit

Le tirage s'est ensuite fait à probabilité inégale au sein de chaque strate.

La probabilité théorique de sélection d'un individu dans la strate ZAE*groupe VQS j était la suivante :

$$\pi_{CARE/BS}^i = \frac{\hat{\omega}^i * n_{\text{groupe } j * ZAE}}{\sum_{i \in ZAE} \hat{\omega}^i * \mathbb{1}_{i \in \text{groupe } j}}$$

Toutefois, dans certains cas, cette probabilité dépassait 1. Dans ces cas, l'individu était d'office dans l'échantillon avec une probabilité de tirage égale à 1. Il y avait 1 096 individus dans ce cas.

De même, comme on l'a vu au paragraphe précédent, certaines strates étaient sélectionnées de façon exhaustive. Dans ces cas aussi, l'individu était d'office dans l'échantillon avec une probabilité de tirage égale à 1. Il y avait 3 328 individus dans ce cas.

Au final, la probabilité de tirage était donc de :

$$\begin{cases} 1 \text{ pour les individus de l'exhaustif} \\ \frac{\hat{\omega}^i * n_{groupe\ j * ZAE \setminus exh}}{\sum_{i \in ZAE} \hat{\omega}^i * \mathbb{1}_{i \in groupe\ j}} \end{cases}$$

Les résultats en termes de dispersion des poids

Tableau 14 • Distribution des poids de sondage CARE par groupe VQS

Groupe VQS	Taille de l'échantillon	Fractiles de la distribution des poids								
		Minimum	5 ^e percentile	1 ^{er} décile	1 ^{er} quartile	Médiane	3 ^e quartile	9 ^e décile	95 ^e percentile	Maximum
I	2 804	1 063	2 350	2 646	3 060	3 550	4 062	4 653	5 202	9 907
II	3 377	295	557	647	839	1 005	1 193	1 397	1 639	7 079
III	4 323	99	148	168	219	297	381	476	548	2 135
IV	4 496	99	137	151	180	234	305	376	413	2 583

Lecture > Le poids de sondage médian des individus CARE appartenant au groupe VQS I est de 3 550.

Champ > Individus âgés de 60 ans ou plus résidant à domicile en France métropolitaine.

Source > INSEE, DREES, Base de sondage de VQS 2014, VQS 2014 ; Calculs INSEE.

Comme on peut le constater l'équipondération des individus CARE sélectionnés par groupe VQS n'est pas totalement atteinte. Toutefois la distribution des poids obtenue est proche de celle d'une enquête ménage traditionnelle tirée dans l'échantillon maître par le biais de l'outil Octopusse. Tous les traitements mis en œuvre pour se rapprocher de l'équipondération ont donc été importants car sans eux la dispersion finale aurait été bien plus importante.

La principale raison qui explique le non-respect de l'équipondération vient bien entendu de la différence entre la base de sondage théorique utilisée pour le calcul des allocations et la base de sondage définitive issue de la collecte de VQS. Ainsi, les allocations par strate de tirage (ZAE*groupe VQS) calculées sur la base de sondage définitive, sont souvent différentes de celles estimées pour Chester. Dans 63 % des cas l'écart par strate de tirage entre les 2 allocations est certes inférieur à 1 individu mais pour 5 % des strates cet écart dépasse les 5 individus. Ceci a un impact direct sur les poids finaux. En outre, les modifications d'allocations nécessaires pour respecter les contraintes Chester et décrites ci-dessus ne font qu'amplifier ces écarts de poids entre ZAE.

■ BIBLIOGRAPHIE

- Bérardier M., « Les bénéficiaires de l'allocation personnalisée d'autonomie à domicile et leurs ressources en 2011 », *Études et Résultats*, n°876, DREES, février 2014.
- Bouvier G., « L'enquête Handicap-Santé : présentation générale », *Document de travail*, N° F1109, INSEE, 61 p. (RAP), 2011.
- Bouvier G., « Les faux négatifs du volet ménage de l'enquête Handicap-Santé 2008 », *Actes des Journées de Méthodologies Statistiques*, INSEE, 2012.
- Groupe de travail présidé par Charpin J.M. « Perspectives démographiques et financières de la dépendance », 2011.
- Christine M. et Faivre S., « OCTOPUSSE : un système d'Échantillon-Maître pour le tirage des échantillons dans la dernière EAR », *Journées de Méthodologie Statistique*, Paris, 2009.
- Debout C. « Durée de perception de l'allocation personnalisée d'autonomie (APA) », *Dossier de travail, Série Sources et Méthodes*, n°15, DREES, juillet 2010.
- Debout C. « La durée de perception de l'APA : 4 ans en moyenne », *Études et résultats*, n° 724, DREES, avril 2010.
- Eghbal-Téhérani S. et Makdessi Y., « Les estimations GIR dans les enquêtes Handicap-Santé 2008-2009 », *Dossier de travail, Série Sources et Méthodes*, n°26, DREES, septembre 2011.
- Lecroart A., Froment O., Marbot C. et Roy D., « Projection des populations âgées dépendantes : deux méthodes d'estimation », *Dossier Solidarité Santé*, n°43, DREES, septembre 2013.
- Le Guennec J. « Estimations locales du handicap dans l'enquête Handicap-Santé 2008 », *Document de travail*, INSEE, H2012/02, 200p. (RAP), 2012.
- Midy L., « Présentation de l'enquête « Vie quotidienne et Santé » de 2007 », *Document de travail*, INSEE, F1001, 55 p. (RAP), 2010.
- Haag O., Loonis V., Mamari C., Vinet L., « La source fiscale : une base de sondage qui s'impose pour les enquêtes auprès des logements et des individus », Recueil du Symposium de Statistique Canada « Produire des estimations fiables à partir de bases imparfaites », octobre 2013.
- Faivre S., « Le projet OCTOPUSSE de nouvel Échantillon-Maître de l'Insee », actes des journées de méthodologie statistique 23, 24 et 25 mars 2009.
- Caron N., « La correction de la non-réponse par repondération et par imputation », chapitre IV du document de travail M0502, 2005.

Annexe 1. Importance des variables dans VQS 2007-HSM 2008 : rang selon la modélisation

Variable	Moins de 80 ans			80 ans ou plus			Poids dans score VQS 2007
	Logit GIR*	Logit Cohérence*	ACM	Logit GIR*	Logit Cohérence*	ACM	
<i>Aide</i>	3		7	1	6	5	6
<i>Aide technique</i>	11			9			1
<i>Aménagement du logement</i>				7			2
<i>Difficultés à attraper un objet</i>	4		12			12	3
<i>Difficultés pour entendre</i>							3
<i>Difficultés pour comprendre</i>				6		14	3
<i>Difficultés pour se concentrer</i>		2	10	3	4	7	3
<i>État général de santé</i>	9		2			6	6
<i>Sentiment d'un handicap</i>			5			13	6
<i>Impossibilité à faire au moins une activité</i>	7		11			10	6
<i>Difficultés à prendre des initiatives</i>	10	4	9	8		4	3
<i>Limitation depuis plus de 6 mois</i>	2		1	5		1	6
<i>Difficultés pour se servir de ses mains</i>				4		15	3
<i>Maladie chronique</i>							1
<i>Difficultés pour marcher</i>			4		5	9	3
<i>Difficultés pour se souvenir de certaines choses</i>					2	11	3
<i>Difficultés pour parler</i>	5						3
<i>Difficultés à résoudre les problèmes quotidiens</i>		1	8		1	2	3
<i>Difficultés pour se pencher</i>	6		6			8	3
<i>Reconnaissance officielle d'un handicap</i>	8					3	6
<i>Sexe</i>							0
<i>Difficultés pour sortir du logement</i>	1	3	3	2		3	3
<i>Difficultés pour voir de loin</i>							3
<i>Difficultés pour voir de près</i>							3

Note > Le rang n'est pas reporté pour les variables non significatives au seuil de 5 %.

* > pondéré avec poids normalisés, seuils entrée et sortie = 0,05.

Lecture > La variable difficultés pour sortir du logement est la plus importante dans le modèle estimant le GIR chez les moins de 80 ans (rang = 1). Cela signifie que cette variable est celle ayant le plus grand pouvoir prédictif du GIR chez les moins de 80 ans.

Champ > Individus âgés de 60 ans ou plus résidant à domicile en France métropolitaine.

Sources > INSEE, DREES, VQS 2007, HSM 2008 ; Calculs DREES.

Annexe 2. Incrémentation de l'ancien et du nouveau score par variable du questionnaire

Variable VQS	Ancien score	Nouveau score	
		Moins de 80 ans	80 ans ou plus
<i>Aide de l'entourage</i>	6 / 3	6	6
<i>Aide d'un professionnel</i>	6 / 3	6	6
<i>Aide technique</i>	1	1	1
<i>Aménagement du logement</i>	2	0	1
<i>Difficultés à attraper un objet</i>	3 / 1	3 / 1	1
<i>Difficultés pour entendre</i>	3 / 1	0	0
<i>Difficultés pour comprendre</i>	3 / 1	3 / 1	3 / 1
<i>Difficultés pour se concentrer</i>	3 / 1	3 / 1	6 / 3
<i>État général de santé</i>	6 / 3 / 1	3 / 1	3 / 1
<i>Sentiment d'un handicap</i>	6	0	0
<i>Impossibilité à faire au moins une activité</i>	6	3	1
<i>Difficultés à prendre des initiatives</i>	3 / 1	0	0
<i>Limitation depuis plus de 6 mois</i>	6 / 3	6 / 3	6 / 3
<i>Difficultés pour se servir de ses mains</i>	3 / 1	0	3 / 1
<i>Maladie chronique</i>	1	0	0
<i>Difficultés pour marcher</i>	3 / 1	1	3 / 1
<i>Difficultés pour se souvenir de certaines choses</i>	3 / 1	0	6 / 3
<i>Difficultés pour parler</i>	3 / 1	0	0
<i>Difficultés à résoudre les problèmes quotidiens</i>	3 / 1	12 / 6	9 / 4
<i>Difficultés pour se pencher</i>	3 / 1	3 / 1	1
<i>Reconnaissance officielle d'un handicap</i>	6	0	0
<i>Difficultés pour sortir du logement</i>	3 / 1	6 / 3	6 / 3
<i>Difficultés pour voir de loin</i>	3 / 1	0	0
<i>Difficultés pour voir de près</i>	3 / 1	0	0
<i>Difficultés pour se laver</i>	0	6 / 3 / 1	6 / 3 / 1
<i>Bénéfice de l'APA</i>	0	6	6

Lecture > Le score VQS 2007 est incrémenté de :

- 6 points supplémentaires pour les personnes déclarant avoir beaucoup d'aide de l'entourage,
- 3 points supplémentaires pour celles déclarant un peu d'aide de l'entourage.
- 6 points supplémentaires pour celles déclarant un très mauvais état de santé.
- 3 points supplémentaires pour celles déclarant un état de santé moyen.
- 1 point supplémentaire pour celles déclarant un bon état de santé.

Le score VQS 2014 est incrémenté de :

- 6 points supplémentaires pour les personnes déclarant avoir beaucoup ou un peu d'aide de l'entourage.
- 3 points supplémentaires pour celles déclarant un très mauvais état de santé.
- 1 point supplémentaire pour celles déclarant un mauvais état de santé.

Champ > Individus âgés de 60 ans ou plus résidant à domicile en France métropolitaine.

Sources > INSEE, DREES, VQS 2007, HSM 2008 ; Calculs DREES.

Annexe 3. Structure des réponses à VQS 2007 avec et sans face-à-face

	Avec face-à-face		Sans face-à-face	
	N	% (hors manquant)	N	% (hors manquant)
Aide humaine dans la vie quotidienne				
Non	42420	82%	36269	82%
Oui, un peu	6116	12%	5325	12%
Oui, beaucoup	3410	6%	2797	6%
Aide technique				
Non	43740	85%	37378	85%
Oui	7625	15%	6435	15%
Aménagements du logement				
Non	48669	94%	41496	94%
Oui	2861	6%	2481	6%
Difficultés pour lever le bras				
Non	39005	74%	33180	74%
Oui, un peu	9171	18%	8163	18%
Oui, beaucoup	4202	8%	3482	8%
Difficultés pour entendre (conversation)				
Non	32996	63%	27694	62%
Oui, un peu	14596	28%	12954	29%
Oui, beaucoup	4742	9%	4136	9%
Difficultés pour comprendre				
Non	45920	88%	39026	87%
Oui, un peu	4794	9%	4332	10%
Oui, beaucoup	1485	3%	1299	3%
Difficultés pour se concentrer				
Non	43163	83%	36744	82%
Oui, un peu	6775	13%	6007	14%
Oui, beaucoup	2152	4%	1801	4%
État de santé général				
Très bon	4379	8%	3729	8%
Bon	21631	41%	18727	42%
Moyen	20266	39%	17318	39%
Mauvais	4964	10%	4133	9%
Très mauvais	1076	2%	859	2%
Considère avoir un handicap				
Non	38507	76%	32904	76%
Oui	12382	24%	10453	24%
Impossibilité totale de faire les activités citées				

Non	46704	90%	40358	91%
Oui	5288	10%	4079	9%
Difficultés pour prendre des initiatives				
Non	43103	83%	36645	82%
Oui, un peu	6263	12%	5559	13%
Oui, beaucoup	2761	5%	2381	5%
Limité dans les activités de la vie quotidienne				
Non	28184	54%	24098	54%
Oui, un peu	14853	28%	12999	29%
Oui, beaucoup	9125	18%	7511	17%
Difficultés pour se servir de ses mains				
Non	40706	78%	34510	77%
Oui, un peu	8488	16%	7622	17%
Oui, beaucoup	3184	6%	2693	6%
Maladie(s) chronique(s)				
Non	23306	45%	20012	45%
Oui	28615	55%	24362	55%
Difficultés pour monter un étage ou marcher 500m				
Non	31651	60%	27060	60%
Oui, un peu	12154	23%	10594	24%
Oui, beaucoup	8525	16%	7122	16%
Difficultés pour se souvenir				
Non	36765	71%	31056	70%
Oui, un peu	12410	24%	11011	25%
Oui, beaucoup	2917	6%	2475	6%
Difficultés pour parler				
Non	49485	95%	42297	95%
Oui, un peu	2122	4%	1858	4%
Oui, beaucoup	690	1%	587	1%
Difficultés pour résoudre les problèmes de la vie quotidienne				
Non	43847	84%	37388	84%
Oui, un peu	5544	11%	4874	11%
Oui, beaucoup	2824	5%	2406	5%
Difficultés pour se pencher				
Non	31061	59%	26313	59%
Oui, un peu	14161	27%	12527	28%
Oui, beaucoup	7087	14%	5916	13%
Reconnaissance officielle d'un handicap				
Non	45865	89%	39088	89%
Oui	5404	11%	4640	11%
Difficultés pour sortir du domicile				
Non	43359	83%	37105	83%
Oui, un peu	4512	9%	3918	9%
Oui, beaucoup	4229	8%	3525	8%
Difficultés pour voir de loin				
Non	46814	89%	40068	90%

Oui, un peu	3968	8%	3417	8%
Oui, beaucoup	1532	3%	1277	3%
Difficultés pour voir de près				
Non	39069	75%	33009	74%
Oui, un peu	9600	18%	8606	19%
Oui, beaucoup	3623	7%	3127	7%

Lecture > Parmi les personnes ayant répondu en face-à-face à l'enquête VQS 2007, 82 % déclarent ne pas recevoir d'aide humaine.

Champ > Individus âgés de 60 ans ou plus résidant à domicile en France métropolitaine.

Source > INSEE, DREES, VQS 2007 ; Calculs DREES.

Annexe 4. Lexique

AGGIR : Autonomie gérontologie-groupe iso-ressources
APA : Allocation personnalisée d'autonomie
CARE : Capacités, Aides et REssources des seniors
CNRT : Correction de la non-réponse totale
EAR : Enquête annuelle du recensement
EHPAD : Établissement d'hébergement pour personnes âgées dépendantes
EM : Échantillon maître
FA : Fiche-adresse
GIR : Groupe Iso-Ressources
HSM : Handicap-Santé Ménages
MDR : Maison de retraite
NCEE : Nouveau cadre d'emploi des enquêteurs
VQS : Vie Quotidienne et Santé
ZAE : Zone d'action des enquêteurs

Les dossiers de la DREES

N° 43 • octobre 2019

Échantillonnage des enquêtes VQS et CARE-ménages

Directeur de la publication
Jean-Marc AUBERT

Responsable d'édition
Souphaphone Douangdara

ISSN
2495-120X



Ministère des Solidarités et de la Santé

Direction de la recherche, des études, de l'évaluation et des statistiques (DREES)

14 avenue Duquesne - 75 350 paris 07 SP

Retrouvez toutes nos publications sur drees.solidarites-sante.gouv.fr et nos données sur www.data.drees.sante.fr