

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

DREES

document
de *travail*

SÉRIE **SOURCES ET MÉTHODES**

N° 56 • SEPTEMBRE 2016

Les enrichissements prévus pour l'enquête
CARE-Ménages

MISE EN ŒUVRE, APPORTS ET CONTRAINTES

Amélie CARRÈRE



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

DREES

SÉRIE SOURCES ET MÉTHODES

N° 56 • SEPTEMBRE 2016

Les enrichissements prévus pour
l'enquête
CARE-Ménages

MISE EN ŒUVRE, APPORTS ET CONTRAINTES

Amélie CARRÈRE

Sommaire

■ RÉSUMÉ	7
■ INTRODUCTION	9
■ PRÉSENTATION GÉNÉRALE DES ENQUÊTES CARE.....	10
■ LES APPORTS DES ENRICHISSEMENTS PRÉVUS POUR L'ENQUÊTE CARE- MÉNAGES, VOLET SENIORS	13
Appariement avec les ressources fiscales et sociales (RFS)	14
Appariement avec les données des caisses de retraite (CNAV et CCMSA)	15
Appariement avec les données de la CNAM-TS.....	16
Appariement avec les données d'état civil	17
Enrichissement par les données des départements	18
Enrichissement par les données de la base permanente des équipements (BPE)	19
■ LES CONTRAINTES ET LES DIFFICULTÉS DES ENRICHISSEMENTS PRÉVUS POUR L'ENQUÊTE CARE-MÉNAGES VOLET SENIORS	21
Collecte du NIR et implications	21
Un décret en Conseil d'État	21
Deux demandes d'identification	22
Un haut niveau de sécurisation des échanges	23
Un travail dans la durée	26
Risque de ré-identification et questions de diffusion	27
Disponibilité partielle des données	27
Recouvrement partiel des périodes de référence	28
Erreurs d'appariement et travail d'imputation	30
Cas particulier : l'enrichissement par les données des départements	31
■ CONCLUSION	35
■ BIBLIOGRAPHIE	36
Annexe 1. Décret en Conseil d'État autorisant les enquêtes CARE et leurs enrichissements	38
Annexe 2. Liste des enquêtes mobilisables pour étudier la dépendance ou les aidants	42
Annexe 3. Liste des sigles utilisés.....	43

■ RÉSUMÉ

Sept enrichissements différents par des données administratives sont prévus pour l'enquête CARE (Capacités, Aides et REssources des seniors) – Ménages, volet seniors. Compléter les données d'enquêtes avec des données administratives permet de disposer de données fiables tout en limitant la charge d'enquête supportée par les personnes interrogées, mais il en résulte un travail important en amont et en aval de l'appariement : contact avec les acteurs, définition des variables à remonter et des modalités de transfert, tests, imputations des valeurs manquantes, annualisation des montants, etc. Ce document a pour objectif de présenter les enrichissements prévus de l'enquête CARE ménages, volet seniors, par des données administratives. Il met en évidence les objectifs de ce type d'enrichissements et les contraintes qu'ils impliquent.

■ 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 (Charpin, 2011). 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 d'évolution au fil du temps de la dépendance au sein de la population de personnes âgées. Ceci suppose de pouvoir estimer l'évolution du nombre de personnes âgées en perte d'autonomie sur le passé récent et de la comparer aux projections réalisées par la DREES, Direction de la recherche, des études, de l'évaluation et des statistiques (Lecroart et al., 2013). 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 % (Lecroart et al., 2013). Par ailleurs, il s'est avéré primordial de mieux cerner les enjeux financiers auxquels font face les ménages confrontés à la dépendance. 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 est appelé à s'accroître, notamment du fait d'une diversification des services aux personnes âgées dépendantes à domicile (accompagnement, courses, ménage, repas, soins du corps etc.), du renforcement des normes de sécurité et de qualité, ainsi que des revalorisations des salaires et de l'augmentation du niveau de qualification des personnels intervenants (aide à domicile et auxiliaire de vie).

Les enquêtes CARE (Capacités, Aides et REssources des seniors) ont pour but de répondre à ces besoins d'information. Or, y répondre uniquement par la collecte auprès des personnes interrogées peut s'avérer inefficace et lourd. Les personnes, surtout âgées, peuvent ne pas communiquer à l'enquêteur l'ensemble des aides qu'elles reçoivent et des coûts qu'elles supportent, même avec l'aide d'un tiers pour répondre à l'enquête, d'autant plus qu'elles peuvent bénéficier de remboursements ex-post.

La question qui se pose est : comment disposer de suffisamment de données pour répondre de façon fiable aux besoins d'information des pouvoirs publics sans alourdir la charge des enquêtés ?

Le code de bonnes pratiques de la statistique européenne rappelle, dans son principe 9, de ne pas soumettre une charge excessive pour les déclarants et, dans le cas où des sources administratives existent, l'article 7bis de la loi n° 51-711 du 7 juin 1951 prévoit la cession, à « des fins exclusives d'établissement de statistiques » des informations « recueillies dans le cadre de sa mission par une administration, une personne morale de droit public ou une personne morale de droit privé gérant un service public. » Par ailleurs, le Conseil national de l'information statistique (CNIS) fait état du souhait de disposer d'enquêtes en population générale sur la dépendance, mettant en relation les différents systèmes de gestion de la dépendance : organismes de sécurité sociale, conseils généraux, sources fiscales². Afin de respecter ce principe 9, et parce que les données administratives sont de qualité, il a été décidé de réaliser différents enrichissements par appariement des réponses à l'enquête à de telles données administratives.

Le dispositif d'enquêtes CARE sera présenté dans une première partie. La deuxième partie s'intéressera aux objectifs des enrichissements envisagés. Enfin, seront présentées les contraintes et les difficultés des enrichissements dans une dernière partie.

¹ Les enquêtes Handicap-Santé, qui ont un lien de parenté fort avec les enquêtes CARE (Capacités, Aides et REssources des seniors), ont été conduites entre 2007 et 2009 et portaient sur l'ensemble de la population, vivant en métropole et dans les DOM, en logement ordinaire ou en institution (Bouvier, 2011).

² Compte rendu n°28/D030 datant du 9 février 2012 de la réunion de la Commission services publics et services aux publics du 23 novembre 2011.

■ PRÉSENTATION GÉNÉRALE DES ENQUÊTES CARE

Le dispositif d'enquêtes CARE comprend deux enquêtes : une première auprès des seniors vivant à domicile (en 2015) et une seconde auprès des seniors vivant en institutions (en 2016), voir figure 1. Chacune de ces enquêtes comprend un volet auprès des seniors et un volet auprès des proches aidants déclarés comme tels par le senior. L'enquête à domicile a également été précédée d'une courte enquête intitulée « Vie Quotidienne et Santé » (VQS) en 2014 (voir encadré 1).

Les enquêtes CARE ont pour but d'éclairer le débat public sur la dépendance. Elles s'intéressent aux conditions de vie des personnes âgées, à leurs difficultés à réaliser les activités de la vie quotidienne et aux aides qu'elles reçoivent et visent à répondre à trois objectifs :

- suivre l'évolution de la perte d'autonomie depuis la seconde moitié des années 2000,
- estimer le reste à charge lié à la dépendance supporté par les personnes,
- et mesurer l'implication de l'entourage auprès de la personne âgée.

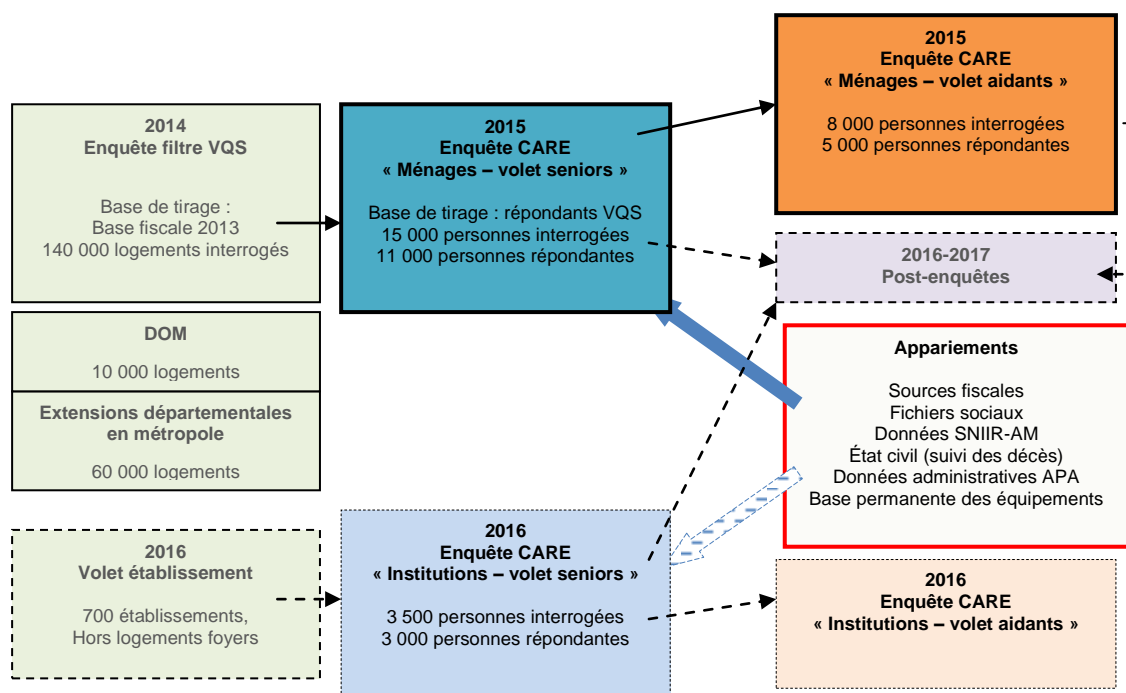
Depuis les enquêtes Handicap-Santé (2008-2009), les questions autour de la dépendance ont pris une part importante dans les débats publics. Les systèmes d'information actuels ne permettent pas d'apporter des réponses au niveau national sur l'évolution du nombre de personnes âgées dépendantes, au sens large donné à ce terme de dépendance – c'est-à-dire au-delà du simple bénéficiaire de la prestation APA (voir encadré 2). Le volet « seniors » des enquêtes CARE permettra de caractériser les personnes âgées selon plusieurs définitions de la perte d'autonomie. La prévalence de la dépendance pourra ainsi être estimée et comparée aux enquêtes Handicap-Santé (Dos Santos et Makdessi, 2010) afin d'obtenir des évolutions.

Par ailleurs, le débat national de 2011 sur la dépendance a révélé un manque d'information quant aux moyens financiers, humains et techniques qu'elle requiert. Les systèmes d'information actuels ne permettent pas d'évaluer dans leur globalité les moyens mobilisés par les ménages confrontés à la perte d'autonomie et leurs difficultés à y faire face. Les enquêtes CARE ont donc pour objectif d'estimer plus spécifiquement le reste à charge lié à la dépendance des ménages, en abordant les coûts de la dépendance et sa prise en charge publique et privée. En outre, l'enquête s'attache à comprendre comment les familles s'organisent pour faire face au reste à charge : aide dans les actes de la vie quotidienne apportée par l'entourage, paiement grâce aux revenus courants de la personne âgée, mais également vente du patrimoine, désépargne, transferts financiers issus de la famille ou renoncement à la prise en charge.

Les enquêtes CARE cherchent également à recenser la nature, la quantité et l'origine des aides humaines, financières et techniques existantes, ainsi que l'existence d'éventuels besoins non satisfaits. Cet objectif était déjà traité dans les enquêtes Handicap-Santé 2008-2009 (Soullier et Weber, 2011 ; Soullier, 2012a ; Soullier, 2012b ; Bérardier, 2012b), ce qui permettra un suivi des évolutions depuis la deuxième moitié des années 2000.

Ce document de travail s'intéresse uniquement à l'enquête CARE à domicile, volet seniors. L'enquête interroge 15 000 personnes nées avant le 02/05/1955 vivant en logements dits « ordinaires » en France métropolitaine et ayant répondu à l'enquête Vie Quotidienne et Santé 2014. L'enquête s'est déroulée du 2 mai au 17 octobre 2015. L'enquête en face-à-face durait environ une heure en moyenne. Ce questionnement, déjà long, ne permet que partiellement de répondre aux objectifs. Il a donc été nécessaire d'envisager différents enrichissements pour le compléter.

Figure 1. Dispositif d'enquêtes CARE (Capacités, Aides et REssources des seniors)



NOTE : LA DÉFINITION DE L'ENSEMBLE DES ACRONYMES EST DISPONIBLE EN ANNEXE 3. LES ÉTAPES ENCADRÉES EN POINTILLÉES CONCERNENT LA PARTIE EN INSTITUTIONS QUI N'EST PAS DÉTAILLÉE DANS CE DOCUMENT DE TRAVAIL. LES CHIFFRAGES SUR LE NOMBRE DE RÉPONDANTS ET LE NOMBRE DE PERSONNES INTERROGÉES AIDANTS SONT DES ESTIMATIONS AVANT COLLECTE.

ENCADRÉ 1 - L'ENQUÊTE VIE QUOTIDIENNE ET SANTÉ (VQS) 2014

L'enquête VQS a trois objectifs :

- connaître l'état de santé des seniors et des personnes vivant avec eux,
- disposer d'informations dans chaque département,
- permettre de sélectionner l'échantillon de l'enquête CARE (Capacités, Aides et REssources des seniors) volet seniors en logements ordinaires.

C'est une enquête au niveau ménage, remplie par tous les habitants d'un logement. Elle a été adressée à 210 000 ménages ayant au moins un individu qui aura 60 ans ou plus au 2/05/2015 (date de début de la collecte de l'enquête CARE-ménages) par voie postale, internet et par téléphone. Elle permet d'avoir de l'information sur l'état de santé de toutes les personnes du logement. Les informations contenues dans VQS sont essentielles pour l'analyse de la configuration d'aide et du réseau d'aidants potentiels. Elles permettent également de mettre en lumière la mutualisation d'aides au niveau du foyer.

La prise en charge sociale et médico-sociale auprès des personnes âgées étant décentralisée, le besoin d'information sur la dépendance au niveau départemental est important. Le format court de l'enquête VQS laisse la possibilité d'un large échantillon, qui peut facilement répondre à un objectif de représentativité départementale pour tous les départements, rendant possible les comparaisons entre territoires. Des résultats départementaux ont également été produits pour les DOM (hors Mayotte), territoires sur lesquels les données de santé sont habituellement relativement peu disponibles et exploitables (Enquête Vie Quotidienne et Santé 2014, Résultats départementaux). Les premiers résultats sont accessibles sur Data.Drees³ sous forme de cubes Beyond, permettant des dénombrements personnalisables.

L'enquête VQS est une "enquête filtre" destinée à préparer l'échantillon de l'enquête Capacités, Aides et REssources des seniors (CARE à domicile) plus approfondie sur la perte d'autonomie. Les personnes en situation de dépendance sont peu nombreuses relativement à la population générale : environ 10 % des personnes âgées de 80 ans ou plus sont bénéficiaires de l'Allocation personnalisée d'autonomie (APA) à domicile, et 1 % parmi les 60-79 ans (Bérardier, 2014). Il est donc important, pour disposer de résultats suffisamment robustes, de pouvoir les surreprésenter dans l'échantillon d'une enquête qui s'intéresse à leurs caractéristiques et à leur situation (Carrère et al., 2015).

³ www.data.drees.sante.gouv.fr

ENCADRÉ 2 – LA PLACE DE L'ENQUÊTE CARE PAR RAPPORT AUX SOURCES EXISTANTES SUR LA DÉPENDANCE

De nombreuses sources peuvent être mobilisées pour étudier la dépendance (voir liste des enquêtes mobilisables en annexe 2). Malgré cet éventail d'enquêtes, aucune ne répond aux objectifs et besoins de connaissance donnés aux enquêtes CARE (Capacités, Aides et REssources des seniors) et ce sur plusieurs points :

- Le champ : couverture de l'ensemble des personnes âgées de 60 ans ou plus, vivant à domicile ou en institution, en France métropolitaine
L'enquête SHARE (Survey on Health, Ageing and Retirement in Europe) s'intéresse spécifiquement à la population âgée de 50 ans ou plus, mais elle ne couvre que très imparfaitement personnes résidant en institution (voir ci-après). Les autres enquêtes ont une couverture incomplète du champ, car elles sont :
 - soit sur des données agrégées : enquêtes annuelles de la DREES sur les bénéficiaires et les dépenses de l'aide sociale départementale, enquêtes trimestrielles de la DREES sur l'APA (Allocation personnalisée d'autonomie) et la PCH (prestation de compensation du Handicap) ;
 - soit en population générale (pas de surreprésentation des personnes âgées, encore moins les personnes dépendantes) : enquête Santé et Protection Sociale (ESPS) de l'IRDES ; Baromètre santé de l'INPES ; baromètre de la DREES (Direction de la recherche, des études, de l'évaluation et des statistiques) ; enquête SRCV (Statistique sur les ressources et conditions de vie – INSEE) ; enquête santé et itinéraire professionnel (DREES – DARES) ;
 - soit sur certaines catégories d'individus seulement : bénéficiaires d'une prestation (remontées individuelles APA (DREES), enquête sur les bénéficiaires de minima sociaux (BMS – DREES)), utilisateurs d'un service (enquête permanente sur la prescription médicale - IRDES) ;
 - soit auprès d'un échantillon limité (cohortes).

La majorité n'est réalisée qu'auprès des personnes vivant dans un logement ordinaire. Il est à noter que l'enquête SHARE permet un suivi des personnes en institution, mais ce suivi est très partiel, l'échantillon en institution est donc quasi inexistant. Seules les enquêtes Handicap-Incapacités-Dépendance (HID) et Handicap-Santé (HS) de l'INSEE et la DREES couvrent l'ensemble des personnes âgées de 60 ans ou plus, vivant à domicile ou en institution, en France métropolitaine.

- L'objectif de disposer de suffisamment d'observations pour étudier les personnes âgées dépendantes
De nombreuses enquêtes surreprésentent la population malade (personnes atteintes d'une ALD (Affection de longue durée) ou les personnes en institution, mais aucune ne surreprésente spécifiquement la population dépendante, tout en interrogeant la population non dépendante. La représentation de cette population est donc souvent limitée. C'est en particulier le cas de l'enquête SHARE où le nombre de personnes dépendantes dans l'échantillon français est faible. Inversement, les remontées individuelles APA sont spécifiquement sur les bénéficiaires de l'APA mais ne permettent pas une comparaison avec des personnes dépendantes et non bénéficiaires de l'APA et des personnes non dépendantes. Seules les enquêtes HID et HS répondent à cet objectif.
- Les thèmes traités
En règle générale, les questionnaires renseignent peu sur la dépendance des individus. Les enquêtes HID puis HS sont les seules qui abordent de façon détaillée la perte d'autonomie, permettent d'estimer différents indicateurs de dépendance (indicateur Colvez, Katz, pseudo-GIR) et s'intéressent aux aidants de l'entourage. Quant aux restes à charge, ceux-ci sont majoritairement déclaratifs, et ne permettent pas de d'identifier les ressources et les dépenses qui sont liées spécifiquement à la perte d'autonomie. C'est notamment l'un des manques importants des enquêtes HID et HS. Les remontées individuelles APA (Bérardier, 2015 ; Bérardier, 2014 ; Bérardier, 2012a ; Bérardier et Clément, 2011 ; Bérardier et Debout, 2011) permettent d'avoir des éléments détaillés et fiables sur les plans d'aide dont bénéficient les personnes puisque les données sont directement issues des conseils départementaux qui financent l'APA. Mais, cette remontée de données administratives se limite aux informations issues des conseils départementaux, c'est-à-dire uniquement la partie qu'ils financent dans le cadre du plan d'aide personnalisé APA.
- La fiabilité des données
Les seuls appariements réalisés sur les enquêtes passées les plus similaires à CARE (enquête HID en 1998-1999 et HSI-HSM en 2008-2009) jusqu'à maintenant le sont avec les données de consommation de soins (SNIIRAM – système national d'information inter régimes de l'assurance maladie). Or, l'appariement permet d'alléger les questionnaires et de disposer de données plus fiables car non déclaratives.

Parmi les nombreuses sources à disposition, les enquêtes se rapprochant le plus des thèmes abordés par les enquêtes CARE sont les enquêtes HID 1998-1999 et HS 2008-2009 et les enquêtes SHARE, puisqu'elles s'intéressent à la population âgée et à l'évolution de la dépendance. Mais, les enquêtes HID et HS ne permettent pas de répondre à l'objectif de calculer un reste à charge lié à la dépendance et leurs données sont anciennes. Par ailleurs, l'enquête SHARE étant une enquête européenne et en panel, la taille de la population française très âgée interrogée reste faible (moins de 700 personnes âgées de 75 ans ou plus), tout comme le nombre de répondants français déclarant au moins une restriction d'activité (moins de 600 personnes). Enfin, aucun appariement n'est prévu pour cette enquête et les questions sur le reste à charge y sont limitées.

■ LES APPORTS DES ENRICHISSEMENTS PRÉVUS POUR L'ENQUÊTE CARE-MÉNAGES, VOLET SENIORS

Conformément au décret (en Conseil d'État) n° 2015-343 du 26 mars 2015 (cf. annexe 1), sept enrichissements sont prévus pour l'enquête (voir figure 2) :

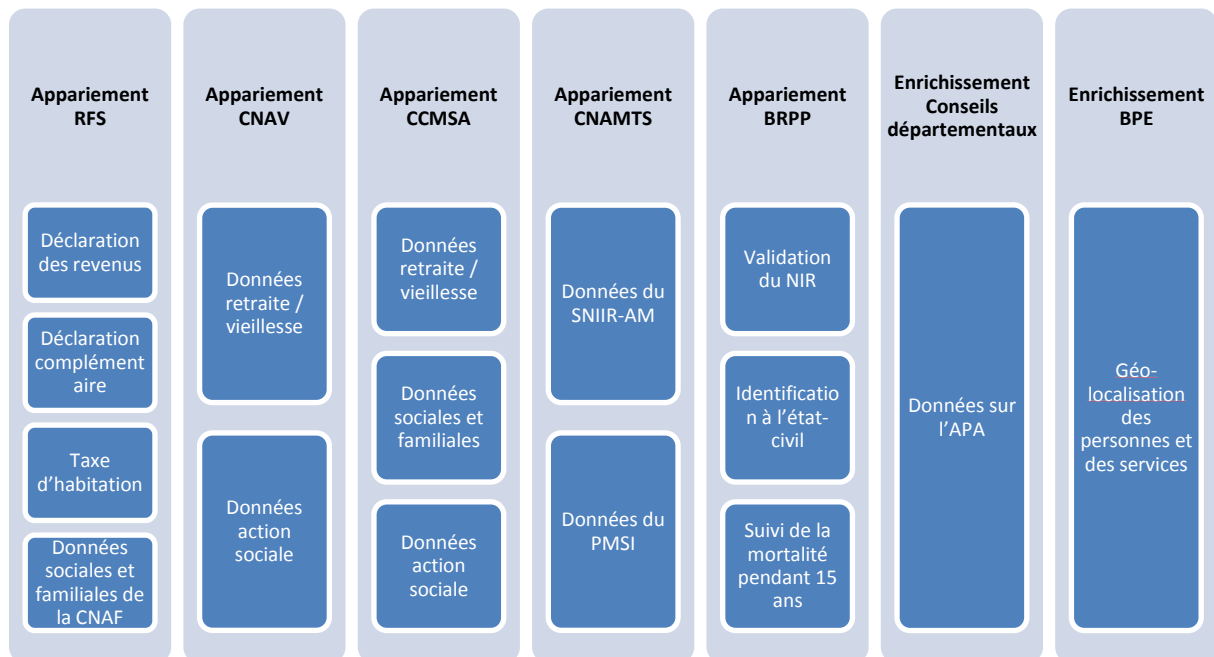
- un appariement avec les ressources fiscales et sociales (RFS) qui permet de disposer de données sur les ressources issues des déclarations fiscales de revenus, sur les impôts tels que la taxe d'habitation (TH), l'impôt sur le revenu (IR) et l'impôt de solidarité sur les fortunes (ISF) et sur les prestations sociales et familiales versées par la CNAF (Caisse Nationale des allocations familiales), voir Bilan de l'enquête revenus fiscaux et sociaux (ERFS) 2012.
Les prestations sociales et familiales versées par la CCMSA et la CNAV seront obtenues par l'appariement spécifique avec les fichiers des deux caisses et non par l'appariement RFS (voir point suivant). En effet, cette remontée d'information *via* l'appariement RFS demandait un délai plus important au pôle RFD (Répertoire et fichiers démographiques) de l'INSEE, en charge de sa mise en application, pour un taux d'appariement peu satisfaisant.
- deux appariements avec des données sur les prestations sociales, familiales et vieillesse : celles fournies par la CNAV, celles fournies par la CCMSA – qui permettent de disposer, outre des revenus sociaux⁴, d'une information sur le bénéfice éventuel de l'action sociale proposée par ces régimes de retraite ;
- un appariement avec les données de la CNAM-TS : données du SNIIR-AM (Système national d'information inter-régimes de l'Assurance maladie)⁵, permettant de disposer de données de consommation médicale détaillées ;
- un appariement avec la BRPP (Base de répertoire des personnes physiques), afin de reconstituer les numéros de sécurité sociale (NIR – numéro d'inscription au répertoire) des personnes ou de les valider puis, afin de suivre la mortalité des répondants après l'enquête pendant quinze ans ;
- un enrichissement avec les données sur l'APA (Allocation personnalisée d'Autonomie) issues des conseils départementaux (CD) ;
- et, enfin, un enrichissement avec les données de la base permanente des équipements (BPE) permettant, à partir d'informations sur la localisation géographique des logements des enquêtés, d'estimer les distances d'accès à divers services.

Ils répondent à différents objectifs qui sont décrits ci-après.

⁴ Minimum-vieillesse, ASPA (Allocation de solidarité aux personnes âgées), MTP (Majoration Tierce Personne), ASI (Allocation supplémentaire d'invalidité), Majoration de pension pour enfants, Pension ex-invalidé.

⁵ Le SNIIR-AM regroupe l'ensemble des informations issues des remboursements effectués par l'ensemble des régimes d'assurance maladie pour les soins du secteur libéral. Le SNIIR-AM inclut aussi le PMSI (Programme de médicalisation des systèmes d'information) relatif aux hôpitaux et cliniques.

Figure 2 : Les sept enrichissements de l'enquête CARE-ménages, volet seniors



NOTE : LA DÉFINITION DE L'ENSEMBLE DES ACRONYMES EST DISPONIBLE EN ANNEXE 3.

Appariement avec les ressources fiscales et sociales (RFS)

L'appariement des données de l'enquête « CARE ménages – volet seniors » avec les sources fiscales de la DGFIP (Direction générale des finances publiques) et les données sociales et familiales de la CNAF (Caisse nationale des allocations familiales) permettra de mieux connaître les ressources dont disposent les ménages pour faire face à la dépendance.

En effet, la déclaration des revenus dans les enquêtes souffre d'un biais de sous-déclaration. Ainsi, dans Handicap-Santé Ménages (HSM), le revenu moyen ainsi que tous les déciles de revenus sont inférieurs de 20 % aux revenus estimés à partir des sources fiscales. Les revenus observés par l'enquête HSM sont ainsi de 1 512 € par mois en moyenne, contre 1 946 € pour les ERFS, et de 480 € par mois pour le 1^{er} décile dans HSM, contre 633 € pour le 1^{er} décile des ERFS.

Par ailleurs, cet appariement permettra de disposer d'informations sur le patrimoine imposable des personnes. Les taux de non-réponse dans les enquêtes en face-à-face pour ces questions sont très élevés. Par exemple, pour l'enquête CARE-Ménages volet seniors, près de 13 % des personnes interrogées n'ont pas répondu à la question sur le type de placements détenus. Au final, seule la moitié des personnes interrogées a déclaré un montant en tranche des placements détenus.

L'appariement avec les sources fiscales et sociales permettra de disposer des revenus détaillés, de l'ensemble des membres du ménage de la personne enquêtée, à la fois imposables et (pour une grande partie d'entre eux) non imposables, c'est-à-dire, pour les fichiers fiscaux :

- des revenus d'activité et de remplacement (traitements, salaires, revenus du chômage, préretraites, retraites, pensions alimentaires reçues, rentes viagères, revenus des professions non salariées) ;
- des revenus du patrimoine déclarés (dont revenus fonciers) ;
- des revenus exceptionnels (dont gains de levée d'options, plus-values) ;
- des revenus dits accessoires (revenus tirés d'activités dont les résultats sont imposables dans la catégorie des bénéficiaires industriels et commerciaux non professionnels) et les revenus perçus à l'étranger ;
- des pensions alimentaires versées ;

- de la Prime Pour l'Emploi ;
- des impôts directs (impôt sur le revenu, taxe d'habitation, prélèvement libératoire sur les valeurs mobilières, CSG et CRDS sur les revenus d'activité, de remplacement et du patrimoine, impôt de solidarité sur la fortune) ;

Et pour les fichiers sociaux :

- des prestations familiales versées par la CNAF (allocations familiales, complément familial, allocation de rentrée scolaire, allocation d'éducation enfant handicapé, allocation de soutien familial, primes de naissance et d'adoption, allocation de base PAJE, CLCA, COLCA, compléments libre choix de mode de garde, allocation journalière de présence parentale) ;
- des minima sociaux (AAH (Allocation aux adultes handicapés) et compléments, RSA, primes de Noël) ;
- des allocations logement versées par la CNAF (ALF – allocation de logement familiale, APL – aide personnalisée au logement, ALS – allocation de logement sociale).

L'appariement avec ces sources permet de couvrir pratiquement l'intégralité du champ de l'enquête. En effet, pour les sources fiscales, il couvre 100 % des revenus imposables. Les CAF (Caisses des allocations familiales) gèrent pour la métropole le versement de 93 % des prestations familiales, minima sociaux et allocations logement, et la CCMSA gère quasiment tout le reste (appariement spécifique avec la CCMSA, voir plus bas). L'appariement de l'enquête CARE-ménages volet seniors avec les données de la CNAF a permis de retrouver 20 % des seniors, proportion qui est cohérente avec la couverture des 60 ans ou plus par une prestation familiale.

Cet appariement est réalisé par l'INSEE (Institut national de la statistique et des études économiques) grâce à différentes données d'état civil et aux identifiants fiscaux⁶. L'INSEE procède en effet régulièrement à des appariements de ces sources sociales et fiscales avec des données d'enquêtes, selon un circuit de circulation des données désormais bien établi, par exemple pour l'Enquête Emploi, appariée aux sources fiscales afin de constituer les ERFS (Enquêtes sur les Revenus Fiscaux et Sociaux), ou pour le panel européen SILC-SRCV (Statistics on Income and Living Conditions – Enquêtes Sur les Ressources et les Conditions de Vie des ménages). Pour l'enquête CARE-ménages volet seniors, ces appariements se justifient surtout par l'intérêt d'avoir des données économiques fiables qui permettront de calculer les ressources du ménage pouvant servir au financement du reste à charge lié à la dépendance.

Appariement avec les données des caisses de retraite (CNAV et CCMSA)

L'appariement RFS (revenus fiscaux et sociaux) permet en théorie de remonter des informations sur les revenus sociaux versés par les caisses de retraite (CNAV et CCMSA). Cependant, sur certaines ressources, le taux d'appariement s'avère très faible. Par exemple, les montants collectés couvrent seulement 29 % du minimum vieillesse. Il a semblé important d'avoir des informations plus fiables sur cette allocation. La CNAV gère le versement de 70 % des sommes relatives au minimum vieillesse (ASPA), la CCMSA 5 % (le reste relevant d'autres caisses). Il a été décidé de contacter ces deux caisses qui représentent 75 % du champ pour l'ASPA, pour disposer des informations sur les prestations qu'elles versent et qui ne seraient pas disponibles dans les fichiers fiscaux et sociaux de l'INSEE. Le taux d'appariement sera meilleur que celui sur les RFS puisqu'il pourra être réalisé directement auprès des caisses de retraite, grâce au numéro de sécurité sociale (NIR), renseigné par les répondants à l'enquête ou reconstitué grâce à l'état civil.

Par ailleurs, l'action sociale n'est pas une donnée remontée par l'appariement RFS. Chaque année, une dotation est attribuée par les différents régimes de retraite aux caisses locales pour financer l'action sociale en faveur des

⁶ L'enquête est tirée dans la source fiscale, les identifiants fiscaux (variables *dirnosek* et *dirindik*) ont donc été conservés par l'Insee afin de réaliser cet appariement.

retraités. L'action sociale s'adresse aux retraités en GIR⁷ 5 et 6, donc relativement peu dépendants mais socialement fragilisés (isolement géographique ou social, situation de rupture sociale). Pour ce faire, l'évaluation globale des besoins du retraité est réalisée à son domicile afin de prendre en compte tous les aspects de sa situation pour déterminer ses besoins prioritaires et lui apporter les réponses les plus appropriées. Les caisses de retraite ne peuvent évaluer grâce à leurs bases administratives les effets de l'action sociale qu'elles opèrent puisqu'elles ne disposent de données que sur leurs bénéficiaires. L'enquête « CARE ménages – volet seniors » permettra une première évaluation de son impact, par la comparaison des personnes enquêtées bénéficiaires de l'action sociale des caisses de retraite à un groupe témoin composé de personnes semblables qui n'en bénéficient pas.

Outre l'information sur le bénéfice du minimum vieillesse (ASPA) et sur l'action sociale, l'appariement avec les données de deux caisses de retraites partenaires de l'enquête – la CNAV et la CCMSA – *via* le numéro de sécurité sociale (NIR) permet de recueillir directement des informations sur d'autres prestations servies par celles-ci : Majoration Tierce Personne (MTP), allocation supplémentaire d'invalidité (ASI), Majoration de pension pour enfants, pension d'ex-invalide...

Enfin, il a été décidé de remonter entièrement les prestations sociales et familiales versées par la CCMSA uniquement par l'appariement spécifique avec la CCMSA et non par l'appariement RFS. En effet, cette remontée d'information *via* l'appariement RFS demandait un délai plus important au pôle RFD (Répertoire et fichiers démographiques) de l'INSEE, chargé de l'appariement, pour un taux d'appariement moins satisfaisant.

Appariement avec les données de la CNAM-TS

L'appariement avec les données du SNIIR-AM (Système national d'information inter régimes de l'assurance maladie) de la CNAM-TS (Caisse nationale d'assurance maladie – des Travailleurs salariés) permettra de disposer, pour les personnes enquêtées, de données de consommation médicale détaillées et très complètes. Trois types de données seront recherchées dans le SNIIR-AM appariées aux données de l'enquête CARE-ménages volet seniors : les consommations médicales (consultations de soins de ville, prescriptions, soins hospitaliers, séances à l'hôpital et autres prestations), les pathologies (affections de longue durée et pathologies reconstituées par la CNAM-TS à partir des consommations), et les coûts des équipements techniques acquis par les enquêtés. Le champ couvert est celui de la consommation en ville et à l'hôpital, tous régimes confondus⁸. Il est à noter qu'il existe une base de données nommée Resid-Ehpad dont dispose la CNAMTS qui permet d'avoir des informations sur les séjours en EHPAD (Établissement d'hébergement pour personnes âgées dépendantes) des personnes. Le décret en Conseil d'État pour CARE prévoit uniquement l'accès au SNIIR-AM, ce qui n'inclut pas les bases Resid-Ehpad. La DREES aurait pu accéder à ces tables détaillées en faisant une déclaration à la CNIL (Commission nationale de l'informatique et des libertés) mais elle y a renoncé du fait que l'enquête porte sur les personnes résidant à domicile.

Les informations issues de l'appariement avec les fichiers de la CNAM-TS permettront de disposer de données détaillées sur les dépenses de santé des personnes (Calvet et Montaut, 2013), excepté les dépenses non prises en charge par l'Assurance maladie obligatoire (psychologues par exemple) et les consommations sans prescription (automédication). La dépense de soins engagée ainsi que le reste à charge mesurés de façon plus certaine grâce à cet appariement⁹ permettront de mieux évaluer les différentiels de dépenses de santé entre la

⁷ Groupe Iso-Ressource : évaluation administrative du niveau de dépendance calculé grâce à la grille AGGIR (Autonomie Gérontologie Groupes Iso-Ressources), qui détermine le degré d'autonomie au regard de plusieurs critères.

⁸ Certains indicateurs ne seront pas disponibles pour les régimes autres que le régime général, notamment les pathologies reconstituées à partir des consommations.

⁹ Le reste à charge calculé par la CNAMTS est un reste à charge après remboursement de l'assurance maladie obligatoire. Il faudra, pour reconstruire le reste à charge de la personne, modéliser les remboursements des organismes complémentaires, grâce au modèle de microsimulation OMAR (Lardellier et al. 2012). Des questions ont été incluses dans le questionnaire en face-à-face (recours à une assurance complémentaire et montant de prime) afin d'estimer au mieux les remboursements des organismes complémentaires.

population dépendante et la population non dépendante au regard de leurs caractéristiques sociodémographiques et économiques et d'avoir une vision plus globale du reste à charge : reste à charge médical et médico-social. Ceci est indispensable pour évaluer le surcoût médical des personnes dépendantes et donc anticiper les politiques publiques adéquates, notamment le besoin de financement de la protection sociale face à l'augmentation prévue du nombre de personnes âgées dépendantes. Le recueil par questionnaire de ce type de consommation aurait été non seulement fastidieux en termes de protocole d'enquête, pour l'enquêteur comme pour l'enquêté¹⁰, mais surtout peu fiable en raison de la mémorisation très incertaine des consommations médicales (Auvray et Le Fur, 2002). L'expérience de l'enquête HSM montre un biais de « désirabilité sociale » sur le recours au dentiste et un biais de mémoire sur le recours au généraliste (Lengagne et *al.*, 2015 ; Penneau et *al.*, 2015).

L'appariement entre l'enquête Handicap-Santé 2008-2009 et le SNIIR-AM a montré que même la déclaration d'une hospitalisation est entachée d'un important biais lors d'un recueil par enquête, par rapport à ce que l'on peut retrouver dans les données de remboursement. Selon Montaut (2012), 12 % des personnes ont été hospitalisées au cours des douze derniers mois (même proportion dans les données déclaratives de l'enquête Handicap-Santé Ménages et dans les données administratives du SNIIR-AM), mais seuls 7 % ont été hospitalisés au cours des douze derniers mois à la fois selon l'enquête (c'est-à-dire les données déclaratives) et selon les données administratives. Par ailleurs, il est probable que le biais de mémoire n'est pas homogène entre la population estimée dépendante et la population estimée non dépendante. Un appariement avec le SNIIR-AM permettra donc d'améliorer la qualité de l'information recueillie.

Enfin, les données du SNIIR-AM permettront de connaître les dépenses d'aides techniques (matériels tels que fauteuils roulants, cannes, chaises percées, lits médicalisés...). Le questionnaire CARE-ménages volet seniors renseigne sur les aides techniques utilisées par les personnes. Il avait été envisagé de demander le coût supporté par la personne pour acquérir chaque aide technique. Or lors du test papier de l'enquête, les enquêteurs ont indiqué que certains enquêtés se souviennent des montants lorsque ceux-ci sont importants mais oublient parfois de déduire les remboursements, ce qui génère un biais dans l'estimation du reste à charge. De plus, la plupart des personnes âgées ne sont plus au fait de ces questions. Nous avons donc décidé, en concertation avec le groupe de conception, de supprimer ces questions sur les coûts des aides techniques et aménagements du logement (excepté la téléalarme) car la déclaration de leur coût est approximative et parce que le coût d'acquisition de certains appareils médicaux (ceux qui font l'objet d'un remboursement par l'assurance maladie (AM)) sera disponible dans les données du SNIIR-AM.

Appariement avec les données d'état civil

Cet appariement répond à trois objectifs :

- valider le NIR collecté ;
- reconstituer le NIR à partir de l'état civil des personnes ;
- suivre la mortalité des personnes pendant quinze ans.

Dans le questionnaire CAPI¹¹ (questionnaire informatique) du volet « seniors » de l'enquête CARE-ménages était introduit un contrôle lors de la collecte du numéro de sécurité sociale (NIR) des personnes afin de vérifier que la clef du NIR était correcte. Cependant, il est possible que des personnes aient donné le NIR de quelqu'un d'autre (celui de leur mari dans le cas d'ouvrants-droits par exemple) ou que le NIR collecté ait été erroné (le contrôle sur la clef du NIR n'était pas bloquant)¹². L'appariement avec les données de la Base des répertoires des personnes

¹⁰ Montaut (2012) a évalué à 25 le nombre de minutes d'interrogation en face-à-face économisées grâce à cet appariement.

¹¹ Computer Assisted Personal Interviewing.

¹² La clef des NIR Corses est calculée différemment de celle des NIR continentaux. Ces cas n'avaient pas été prévus, le contrôle était donc déclenché par erreur sur les NIR des personnes nées en Corse.

physiques (BRPP) de l'INSEE a permis dans un premier temps de valider ces NIR. 4 % des seniors ayant donné leur numéro de sécurité sociale n'ont pas été identifiés dans la BRPP à partir de leur NIR.

De nombreux appariements sont réalisés grâce au numéro de sécurité sociale des individus. Il est donc nécessaire de pouvoir disposer du numéro de sécurité sociale d'un maximum de personnes même si elles n'ont pas donné leur NIR ou si elles ont indiqué un NIR erroné. Plus de 8 % des personnes CARE n'ont pas donné de NIR. Tous les individus répondants et leur ouvrants-droits (sauf deux pour lesquels les données manquantes étaient trop importantes) ont donc été ajoutés à l'identification état civil afin de reconstituer leur NIR grâce à leur nom, prénom, date de naissance, sexe et lieu de naissance.

Grâce aux deux identifications précédentes (validation du NIR et reconstitution du NIR grâce à l'état civil), le CASD (centre d'accès sécurisé aux données) procède à un travail de sélection du bon NIR (car il est possible qu'à l'issue de ces deux identifications deux NIR soient attribués à un même individu) qui sera utilisé pour tous les appariements utilisant le NIR, dont le suivi de la mortalité.

Comme cela avait été fait à la suite de l'enquête Handicap-Incapacités-Dépendance de 1998-2000 (Ankri et Mormiche, 2002), et à la suite de l'enquête Handicap-Santé de 2008-2009 (Bouvier, 2011), il est prévu de suivre la mortalité des personnes interrogées, *via* un appariement avec les données issues de l'état civil. Ceci permet de mesurer la mortalité différentielle des personnes dépendantes dont la connaissance est, entre autres, nécessaire à la projection du nombre de personnes âgées dépendantes.

Cette analyse suppose d'avoir assez de recul pour disposer d'un nombre d'observations suffisant concernant le décès des personnes enquêtées. Le NIR des répondants sera conservé pendant quinze ans après la fin de l'enquête, ou jusqu'à leur décès, afin de pouvoir effectuer une recherche au RNIPP (Répertoire National d'Identification des Personnes Physiques) sur leur décès éventuel.

Enrichissement par les données des départements

Les seniors enquêtés dans CARE-ménages sont des personnes âgées de 60 ans ou plus qui peuvent être en perte d'autonomie. Ainsi, certaines d'entre elles ont droit à l'allocation personnalisée d'autonomie (APA), allocation spécifique pour la perte d'autonomie. Les informations déclarées dans l'enquête sur cette allocation sont limitées pour ne pas alourdir le questionnaire. De plus les données collectées par questionnaire sur le bénéficiaire d'une prestation peuvent s'avérer peu fiables, les personnes concernées n'étant pas nécessairement au courant de la source de financement de l'aide qu'elles reçoivent. Il est donc nécessaire d'enrichir les données déclarées avec des données administratives des départements. Cet enrichissement répond à cinq objectifs :

- Améliorer le repérage des bénéficiaires de l'APA ;
- Améliorer l'évaluation du reste à charge lié à la dépendance des personnes âgées ;
- Connaître les dépenses en aides techniques et aménagements du logement ;
- Valider les pondérations calculées sans enrichissement ;
- Disposer de données longitudinales sur la perte d'autonomie au sens de l'APA.

L'enquête CARE-ménages volet seniors a pour objectif de mesurer la perte d'autonomie des personnes âgées. L'enrichissement permettra de confronter le niveau de dépendance administratif de la grille AGGIR (GIR) et le niveau de dépendance déclaratif disponible grâce aux données d'enquête. L'expérience d'Handicap-Santé montre un important biais de déclaration lors d'un recueil par enquête par rapport à ce que l'on peut retrouver dans les données administratives telles que l'enquête trimestrielle APA de la DREES : les informations déclarées dans l'enquête sous-estiment de 27 % le nombre de bénéficiaires de l'APA à domicile : 533 500 bénéficiaires déclarés dans Handicap-Santé Ménage contre 677 500 bénéficiaires selon l'enquête trimestrielle APA au 31 décembre 2008 (Eghbal-Téhérani et Makdessi, 2011).

Un autre objectif de l'enquête CARE-ménages volet seniors est d'évaluer le reste à charge lié à la dépendance supporté par les personnes âgées. Cet enrichissement permettra de disposer des plans d'aide notifiés et versés

pour les bénéficiaires de l'APA. Les données administratives sont de meilleure qualité que les données déclaratives d'enquête sur la question des prestations et permettent de disposer de plus d'informations.

Par ailleurs, ces données permettront de connaître les dépenses d'aides techniques (matériels tels que protections absorbantes, téléalarmes...) et d'aménagements du logement. Le questionnaire CARE-ménages volet seniors renseigne sur les aides techniques utilisées par les personnes. Il avait été envisagé de demander le coût supporté par la personne pour acquérir chaque aide technique. Or, les données déclaratives comportent des biais importants (ce qu'a confirmé le test papier de l'enquête), il a donc été décidé, en concertation avec le groupe de conception, de supprimer ces questions sur les coûts des aides techniques et aménagements du logement (excepté la téléalarme). Le coût d'acquisition de certaines aides techniques et aménagements du logement, lorsque le département participe à leur financement, sera disponible dans les données de cet enrichissement.

L'enrichissement permettra de disposer d'information sur les non-répondants à l'enquête CARE-ménages volet seniors. Cette information sera utilisée pour valider les pondérations de CARE-ménages volet seniors. En effet, les variables utilisées pour traiter la non-réponse sont des variables issues de la base de sondage (réponses à l'enquête VQS et données de la taxe d'habitation). Si un biais de réponse des personnes bénéficiaires de l'APA existe, il pourra être redressé grâce à cet enrichissement. On pourrait par exemple utiliser le nombre de bénéficiaires de l'APA, connu par ailleurs, comme marge de calage, et ajuster les poids des répondants, bénéficiaires de l'APA ou non, en fonction de ces marges.

L'APA s'adresse aux personnes âgées de 60 ans ou plus résidant à domicile ou en établissement et confrontées à des situations de perte d'autonomie. Les données remontées seront regroupées en deux parties :

- Une partie « instruction » qui comprendra les informations sur :
 - Le statut de la demande d'APA à trois dates (début de la collecte de l'enquête, fin de la collecte, date de l'extraction),
 - Le niveau et l'évolution de la dépendance (évolutions du GIR, historique des dates d'évaluation, type d'APA...),
 - Les montants et caractéristiques des plans d'aides APA notifiés au moment de l'enquête (taux du ticket modérateur, montant payé par le conseil départemental, montant en aide humaine, nombre d'heures d'aide humaine...).
- Une partie « versement » qui comportera les consommations et versements des plans d'aide APA (montants versés en aide humaine par type d'intervention et par type de tarif, montants versés en aides techniques et aménagements du logement).

L'enrichissement permettra en théorie de remonter des données historiques – sous réserve d'une expertise, qui pourra être réalisée après appariement, de la qualité des informations remontées. Il sera alors possible d'étudier l'évolution de la perte d'autonomie au sens de l'APA, grâce à l'historique du GIR.

Enrichissement par les données de la base permanente des équipements (BPE)

La BPE, dont l'INSEE est responsable, contient la liste des équipements et des services rendus à la population à l'échelle locale¹³. Elle couvre ou devrait couvrir d'ici fin 2016 les équipements suivants :

- Les services administratifs (banques, postes et services funéraires, gendarmeries, tribunaux...);
- Les commerces (hypermarchés, supermarchés, supérettes, épiceries, boulangeries, boucheries, charcuteries, produits surgelés et poissonneries...);

¹³ La liste des équipements est disponible sur le site de l'INSEE : [liste des équipements](#).

- Les services de soins (établissements de santé, hôpitaux, psychiatrie, soins à domicile, professionnels généralistes, professionnels spécialistes, professionnels sanitaires, professionnels auxiliaires, pharmacies, services spécialistes des personnes âgées...);
- Les transports (gares, aéroports...);
- Les équipements sportifs, culturels et de loisirs (piscines, stades, théâtres, cinémas, gymnases...).

Ces données sont disponibles par région, département, commune ou quartier IRIS (Ilots Regroupés pour l'Information Statistique)¹⁴. Depuis 2013, certains types d'équipements sont géo-localisés, c'est-à-dire qu'ils sont fournis avec leurs coordonnées dites « XY »¹⁵, mais de manière incomplète¹⁶. Les équipements sont ainsi géo-localisés dans les grandes communes ; en revanche, dans les petites communes, les coordonnées de la mairie sont reportées. Par ailleurs, tous les secteurs ne sont pas géo-localisés : ainsi les supermarchés et petits commerces ne sont actuellement pas géo-localisés. Ils pourraient l'être dans un avenir proche. Les coordonnées des établissements de santé et autres services de soins sont disponibles directement dans les fichiers de la DREES (FINESS par exemple). Le principe retenu pour la diffusion des coordonnées XY est de couvrir au moins 85 % des équipements d'un domaine. Par domaine on entend, par exemple, « les transports » : il faut donc que 85 % des équipements relevant des transports dans leur ensemble soient couverts, même si à un niveau fin les sous-secteurs ne sont pas bien couverts.

Les fichiers contiennent également des informations sur la qualité des données, et notamment de la géo-localisation. Des enquêtes qualité sont menées et permettent d'estimer les écarts entre les informations du fichier et la réalité du terrain¹⁷.

La base de sondage de l'enquête CARE-ménages volet seniors, c'est-à-dire la source fiscale, permet la géo-localisation des logements enquêtés. Grâce aux coordonnées géographiques des logements, la base de données d'enquêtes sera enrichie des informations contenues dans la BPE. Cela a permis d'alléger significativement le questionnaire, en évitant de poser à la personne enquêtée une série de questions portant sur l'accessibilité de ces équipements et services depuis son logement. Bien que l'enrichissement ne permettra pas de répondre aux mêmes questions que le faisait la partie du questionnaire de l'enquête Handicap-Santé dédiée à ces thèmes (notion de participation sociale notamment), il permettra de mesurer la présence ou l'absence d'un équipement, la densité d'un équipement ou la disponibilité d'un équipement du point de vue des habitants. À ce jour, cet enrichissement est encore en phase exploratoire. Pour l'instant, deux types d'indicateurs pourraient être construits par la DREES :

- Un nombre d'équipements par commune (pour tous les équipements disponibles dans la BPE).
 - Des distances par la route entre un équipement et le lieu d'habitation du répondant.
- Ces indicateurs ne seraient pas disponibles pour tous les équipements. Les équipements de santé et d'action sociale, d'enseignement, de transports, de tourisme, de sport, de loisir et de culture sont actuellement disponibles ; les services postaux et bancaires devraient être aussi disponibles d'ici à la date de l'enrichissement.

¹⁴ Les communes d'au moins 10 000 habitants et la plupart des communes de 5 000 à 10 000 habitants sont découpées en IRIS. Ce découpage constitue la maille de base de la diffusion de statistiques infra-communales, il partitionne ces communes en "quartiers" dont la population est de l'ordre de 2 000 habitants. La France compte environ 16 000 IRIS dont 650 pour les DOM. Chaque commune non découpée en IRIS est assimilée à un IRIS. Ce découpage est construit à partir de critères géographiques et statistiques et, autant que possible, chaque IRIS doit être homogène du point de vue de l'habitat.

¹⁵ Les coordonnées XY sont des coordonnées géographiques exprimées en latitude (X) et longitude (Y). Cette information permet de positionner un objet (une personne, service etc.) sur une carte.

¹⁶ Les difficultés de la géo-localisation sont détaillées sur les fichiers de présentation de la BPE : [Géo-localisation](#).

¹⁷ La qualité des données est disponible sur le site de l'INSEE : [Qualité des données](#).

■ LES CONTRAINTES ET LES DIFFICULTÉS DES ENRICHISSEMENTS PRÉVUS POUR L'ENQUÊTE CARE-MÉNAGES VOLET SENIORS

Collecte du NIR et implications

Pour procéder à des enrichissements par des sources utilisant comme identifiant le NIR (« numéro de sécurité sociale »), trois solutions sont envisageables :

- Soit collecter l'état civil de la personne (nom, prénom, date de naissance, etc.) et reconstituer son NIR grâce à un appariement avec le répertoire national d'identification des personnes physiques (RNIPP), administré par l'INSEE ;
- Soit collecter directement le NIR (ainsi que l'état civil) auprès des personnes ;
- Soit collecter les deux.

La troisième solution est privilégiée car elle donne lieu à des taux d'appariement meilleurs du fait de la stabilité du NIR contrairement à l'état civil et permet de reconstituer le NIR dans les cas où celui-ci n'a pas pu être collecté ou l'a été de façon erronée. Par ailleurs, l'expérience de l'enquête Handicap-Santé (voir encadré 3) montre que la collecte de l'état civil uniquement rend l'appariement plus complexe. Toutefois, la troisième solution (collecte du NIR et de l'état civil auprès des personnes enquêtées) donne lieu à de nombreuses contraintes.

Un décret en Conseil d'État

Tout d'abord, la collecte du NIR lors d'une enquête n'est rendue possible que par la publication d'un décret en Conseil d'État. Le texte du décret doit être accompagné d'un dossier technique décrivant l'ensemble des données de collecte, des données servant aux enrichissements, des données issues des différents enrichissements, les acteurs ayant accès à chaque type de données, les procédures de sécurité mises en place, et les modalités de transmission des données. Produire ces deux documents et faire les démarches juridiques associées (CNIL, puis Conseil d'État) aura nécessité à la DREES 22 mois en tout : depuis les premières réflexions jusqu'à la publication du décret au journal officiel en mars 2015. Cette durée provient principalement des délais importants pour chaque soumission aux différentes instances. Le suivi de l'avancée du projet a été un travail de longue haleine, qui a demandé un lourd travail de coordination entre les différentes entités actrices du projet.

ENCADRÉ 3 - RAPPEL DE L'EXPÉRIENCE D'HANDICAP-SANTÉ

Pour l'enquête HSM (Bouvier, 2011, Montaut, 2012, Calvet et Montaut, 2013), le décret en Conseil d'État n'est pas paru avant le début de la collecte. Cela a nécessité une adaptation du questionnaire et des instructions aux enquêteurs à la dernière minute afin de récupérer les données nécessaires à une reconstitution du NIR. La récupération des codes des communes de naissance a été mauvaise (voire inexistante), ce qui a dans un premier temps augmenté le taux d'échec car il n'y a pas eu de recherche à partir du libellé (qui était bien renseigné). La correction mise en œuvre dans un deuxième temps a permis d'augmenter sensiblement le taux d'appariement. Le nom patronymique des femmes n'était pas toujours disponible et cela a également affecté l'appariement.

L'échange des données s'est fait en « double aveugle », conformément aux règles strictes élaborées pour le dossier CNIL, et cela tout au long du processus (y compris pour les fichiers en cours de production). Aucun acteur n'a alors eu de visibilité sur la chaîne d'appariement dans son ensemble, ce qui a empêché de repérer et solutionner rapidement les problèmes. Il a de plus été impossible de distinguer les NIR non reconstitués de ceux non retrouvés dans les fichiers de la CNAM-TS pour cause de non-consommation sur la période. La construction de la pondération s'en est trouvée compliquée.

Les échecs d'appariement ont été plus fréquents pour les ayants droit puisqu'il fallait retrouver deux NIR au lieu d'un. Ces ayants droit sont plutôt des femmes, des jeunes et des inactifs, ce qui a nécessité de redresser l'échantillon pour qu'il reste représentatif. La comparaison du chiffrage des dépenses obtenu à partir de l'appariement avec celui obtenu des données exhaustives de la CNAM-TS a permis de valider la méthodologie.

En début de réflexion, en avril 2013, l'équipe conceptrice a rencontré la CNIL afin de présenter le projet d'enquête, les différents enrichissements envisagés et les circuits de données possibles, et de connaître les formalités juridiques à suivre. Les premières versions de texte de décret et de dossier technique ont été produites au second semestre 2013. En mars 2014 et avant la soumission du dossier à la CNIL, les Conseils d'Administration ou direction des différents partenaires de ces appariements (INSEE, CNAV, CCMSA, CNAM-TS) ont dû être consultés, *via* la Direction de la sécurité sociale (DSS) pour ce qui est des Caisses, afin qu'ils donnent formellement leur avis sur le projet de décret (documents nécessaires à la déclaration CNIL).

Une fois tous les documents préparés, la CNIL a été saisie et le projet de décret est passé en commission CNIL en juillet 2014.

Lors de la réception de l'avis CNIL, le projet de décret et son dossier technique sont envoyés à chaque ministère cosignataire afin de collecter leur accord de principe sur le texte. La réception des accords de principe permet de lancer la demande d'audiencement en Conseil d'État.

Deux mois plus tard, le rapporteur du Conseil d'État a convoqué pour une réunion de travail les différents commissaires au gouvernement proposés par la DREES. Quinze fiches explicatives de différents éléments sur l'enquête ont été demandées au cours de la séance pour la semaine suivante. L'audiencement en Conseil d'État a eu lieu le 17 février 2015. Un mois et 10 jours plus tard, le décret était signé et publié au journal officiel. Le décret est disponible en annexe 1 de ce document.

Deux demandes d'identification

Même si le NIR est collecté, il a fallu concevoir un module autonome de recueil des états civils (nom de naissance, nom d'usage, prénom, sexe, date, commune et pays de naissance). Ces informations sont utiles pour reconstituer le NIR des personnes ne l'ayant pas donné, ou ayant donné un NIR incorrect. Par ailleurs, l'identifiant dans les données du SNIIR-AM est un identifiant non signifiant calculé sur le NIR de l'ouvrant droit, la date de naissance du bénéficiaire et le sexe du bénéficiaire. Il est donc nécessaire de demander à la fois l'état civil du bénéficiaire et les NIR de l'ouvrant droit et du bénéficiaire.

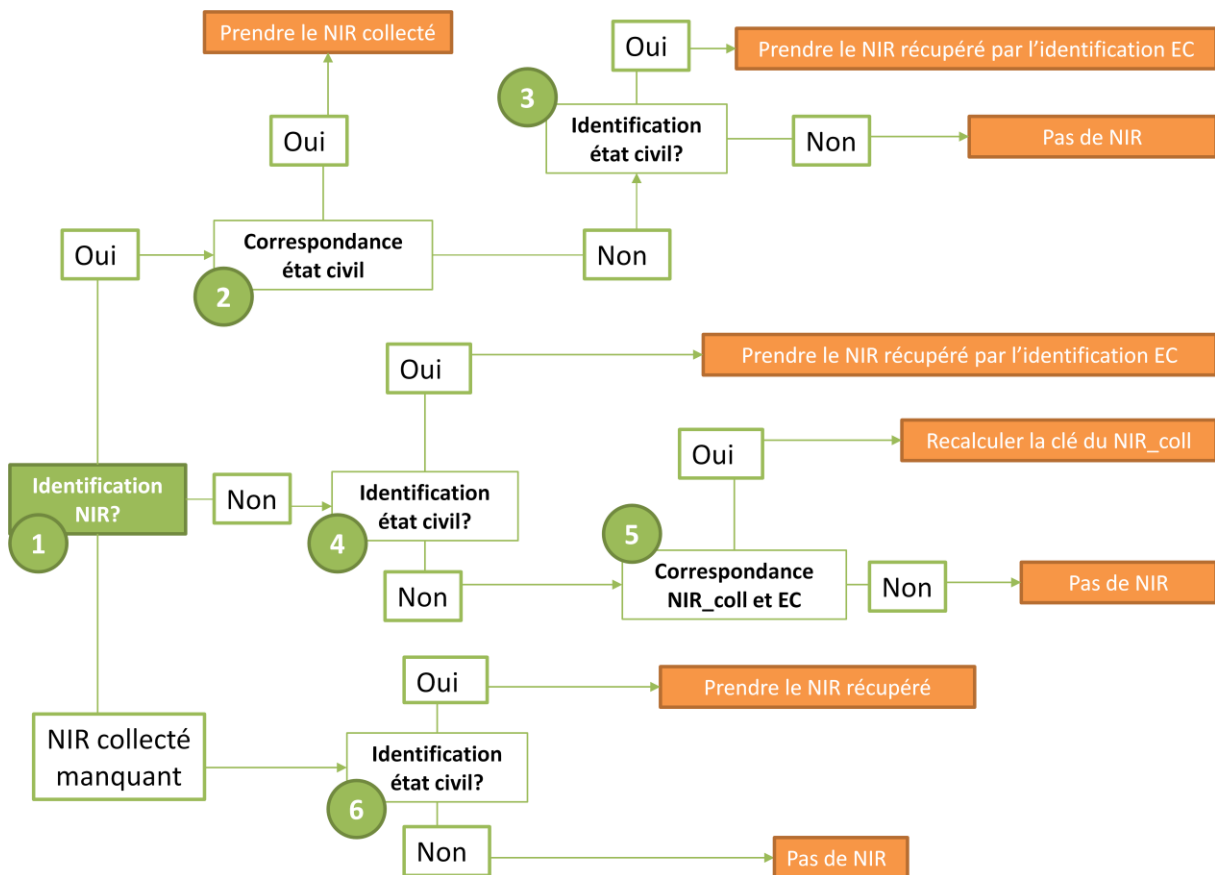
Mis à part le NIR et l'état civil des personnes interrogées (les bénéficiaires), il est aussi nécessaire, pour l'appariement avec les données du SNIIR-AM, de collecter ces informations pour les ouvrants-droits (i.e. tutelle, conjoint, etc.), afin de pouvoir mettre en place l'appariement dans le cas où la personne interrogée serait uniquement affiliée à l'assurance-maladie en tant qu'ayant droit d'un autre assuré (femmes n'ayant jamais travaillé et ayants droit de leur conjoint, par exemple).

Dans le cadre de l'enquête CARE-ménages volet seniors, le NIR de l'ouvrant droit (qui peut être la personne elle-même) a été demandé à chaque personne enquêtée. Il lui a été également demandé d'indiquer les coordonnées de son ouvrant droit si celui-ci n'assistait pas à l'interview afin de l'informer de la collecte d'informations à caractère personnel qui le concernent et des droits dont il disposait. Les enquêtés ont été informés de cette collecte dans la lettre-avis et le dépliant adressés avant l'interrogation. Il a également été demandé à l'enquêté les date et lieu de naissance en sus des nom de naissance, prénom de leur ouvrant droit afin que l'INSEE puisse en rechercher le NIR au sein du RNIPP. Les personnes peuvent faire valoir leur droit d'accès et de rectification. L'enquête étant obligatoire le droit de suppression ne peut être exercé.

Contrairement à ce que l'on pourrait croire, les enquêtés acceptent généralement de communiquer leur NIR ou des éléments d'état civil, dès lors que le thème de l'enquête le justifie. Quoique très identifiant, le NIR semble toutefois peu explicite, certains enquêtés ressentent même parfois moins de gêne à le donner que des éléments d'état civil plus explicites. Lors de l'enquête CARE-ménages volet seniors, plus de 90 % des répondants ont accepté de donner leur NIR.

Deux procédures d'identification sont réalisées par le Pôle RFD de l'INSEE. La première est une demande d'identification sur l'état civil des seniors interrogés et de leurs ouvrants-droits. La seconde est une demande d'identification sur le NIR des seniors interrogés et de leurs ouvrants-droits. Chacune utilisant des informations différentes, ces deux procédures d'identification ont été réalisées en même temps, au premier semestre 2016. Tous les individus répondants à l'enquête CARE ainsi que leur ouvrant droit (excepté deux) ont pu être envoyés au pôle RFD pour au moins une des deux identifications (NIR ou état civil). À l'issue d'une procédure de comparaison des informations réalisée par le CASD, un seul NIR sera conservé pour les appariements utilisant le NIR. Il a donc fallu spécifier au CASD des règles pour choisir le NIR à conserver (voir figure 3).

Figure 3 : Règles de décision sur le numéro de sécurité sociale (NIR) à garder



NOTE : LA DÉFINITION DE L'ENSEMBLE DES ACRONYMES EST DISPONIBLE EN ANNEXE 3. (NIR_COLL= NIR COLLECTÉ)
 LES NUMÉRATIONS CORRESPONDENT À L'ORDRE DANS LE CHOIX DU NIR. D'ABORD, LE CASD REGARDE SI L'IDENTIFICATION NIR A ABOUTI. IL Y A TROIS POSSIBILITÉS : SOIT LE NIR N'ÉTAIT PAS COLLECTÉ (PASSAGE À L'ÉTAPE 6), SOIT LE NIR A ÉTÉ IDENTIFIÉ (PASSAGE À L'ÉTAPE 2), SOIT LE NIR N'A PAS ÉTÉ IDENTIFIÉ (PASSAGE À L'ÉTAPE 4). LES DÉCISIONS QUANT AU NIR GARDÉ SONT MENTIONNÉES DANS LES ENCADRÉS ORANGES.

Un haut niveau de sécurisation des échanges

L'utilisation du NIR renforce le besoin d'un très haut degré de sécurité, notamment concernant les échanges (détaillés dans la figure 4 ci-dessous). Ce point a fait l'objet de nombreuses questions et demandes de précisions de la part de la CNIL, au moment de son examen du texte proposé pour le décret en Conseil d'État, puis du Conseil d'État lui-même.

Tout d'abord, l'utilisation de données directement identifiantes suppose le recours à un tiers de confiance, afin qu'aucun acteur n'ait l'intégralité des données. Le CASD joue ce rôle de tiers de confiance pour les appariements

de l'enquête CARE-ménages, volet seniors. La procédure d'échange des données est dite en double aveugle, de sorte qu'aucun des acteurs ne peut avoir en sa possession l'ensemble des fichiers, ce qui permettrait de rompre l'anonymat (Gensbittel et Riandey, 2011, Golberg *et al.*, 2008). Ainsi, les données d'enquête ne sont disponibles qu'à la DREES et au SIN (Service informatique national) de Lille de l'INSEE ; les numéros de sécurité sociale et états civils ne sont disponibles que pour le SIN de Lille, le pôle RFD, le CASD et les partenaires réalisant les appariements (CNAM-TS, CCMSA, CNAV, pôle RFD, prestataire pour l'enrichissement par les données des départements) ; les données issues des appariements ne sont disponibles qu'au CASD, à la DREES et chez chacun des partenaires réalisant les appariements (CNAM-TS, CCMSA, CNAV, pôle RFD, prestataire pour l'enrichissement par les données des départements).

Pour la transmission des données entre le CASD et la CNAM-TS, les fichiers dans leur ensemble seront chiffrés *via* le logiciel GNU-PG, qui correspond à l'état de l'art actuel en termes de cryptage. Les fichiers seront cryptés grâce à un système de double clé : une clé publique connue des deux partenaires de la transmission (CASD et CNAM-TS), et une clé privée connue de chacun des partenaires et de lui seul. La transmission se fera *via* la plateforme PETRA¹⁸ de la CNAM-TS, qui dispose d'une interface Web. La convention DREES-CNAM-TS identifie la DREES comme responsable de l'opération. Une fois l'appariement réalisé, la CNAM-TS enverra donc directement les données à la DREES *via* la plateforme PETRA (ces données ne comporteront aucune information directement identifiante). Puis, la DREES les transmettra au CASD *via* la plateforme sécurisée d'échange MIOGA¹⁹ pour que le CASD puisse réaliser les contrôles demandés.

Pour la transmission des données entre le CASD et les autres caisses (CNAV et CCMSA), les fichiers dans leur ensemble seront cryptés par un chiffrement symétrique : AES-256 (en utilisant 7-zip). Chacune des caisses recevra une archive auto-extractible protégée par un mot de passe convenu entre le CASD et la caisse par téléphone. La transmission se fera *via* la plateforme d'échange du CASD (<https://xchangefile.ensae.fr>). Une fois l'appariement réalisé, chaque caisse renverra les données enrichies et cryptées en AES-256 au CASD *via* la plateforme d'échange du CASD. Une fois les contrôles réalisés et les données directement nominatives supprimées, le CASD renverra ces données enrichies et cryptées en AES-256 à la DREES *via* la plateforme sécurisée d'échange MIOGA.

Il est à noter que l'appariement CCMSA se fera en deux temps :

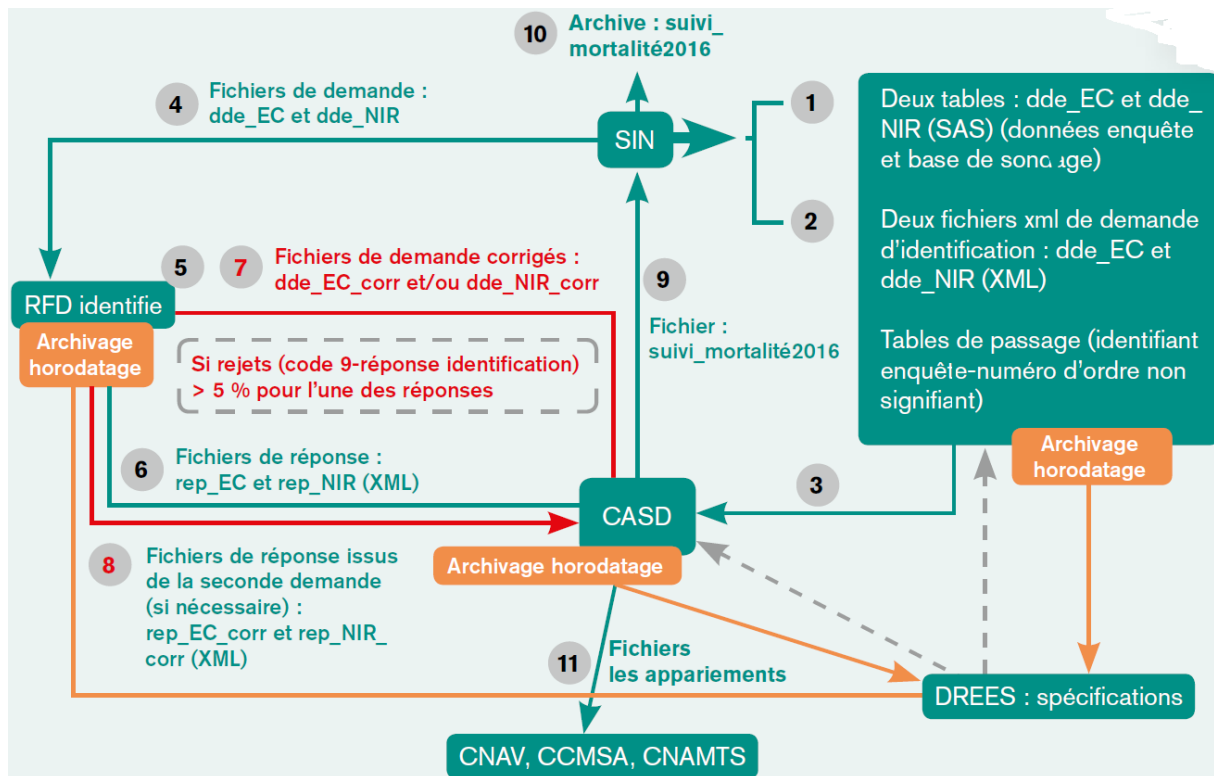
- Appariement Famille et retraite :
La direction des statistiques de la CCMSA réalisera l'appariement famille et retraite (sur l'ensemble de l'échantillon CARE : salariés agricoles et non-salariés agricoles) qu'elle renverra au CASD selon le mode de transmission défini plus haut.
- Appariement Action sociale :
La direction des statistiques de la CCMSA créera un fichier comportant les données d'identité (NIR, nom, prénom, date de naissance, identifiant MSA) des personnes retrouvées dans ses fichiers lors des appariements famille et retraite et le scindera par caisse locale (35 caisses MSA). La direction des statistiques de la CCMSA enverra à chaque caisse régionale le fichier ainsi constitué, contenant les identifiants des personnes relevant de la caisse. Chaque caisse recherchera les individus présents dans le fichier et renseignera les variables action sociale demandées (soit grâce à une extraction informatique, soit grâce à une saisie manuelle car il y aura peu d'individus concernés dans chaque caisse) et les renverra à la direction des statistiques de la CCMSA. La direction des statistiques de la CCMSA enverra l'appariement action sociale au CASD selon le mode de transmission défini plus haut.

Comme indiqué dans l'article 5 du décret en Conseil d'État, l'ensemble des échanges (envois et réceptions) devra être tracé. Il s'agira d'horodater l'ensemble des transferts de tables contenant des données identifiantes.

¹⁸ La plateforme sécurisée PETRA permet de transférer des fichiers, par envoi d'un lien sécurisé (https) et temporaire, à l'adresse mèl du seul référent de l'étude.

¹⁹ La plateforme MIOGA (My Internet Open Groupware Application) est un site collaboratif et de partage de fichiers, avec accès restreint (<https://mioqa.finances.gouv.fr>).

Figure 4. Circulation des fichiers directement identifiants



NOTE : LA DÉFINITION DE L'ENSEMBLE DES ACRONYMES EST DISPONIBLE EN ANNEXE 3.

LES NUMÉRATIONS CORRESPONDENT À L'ORDRE DU CIRCUIT DES FICHIERS.

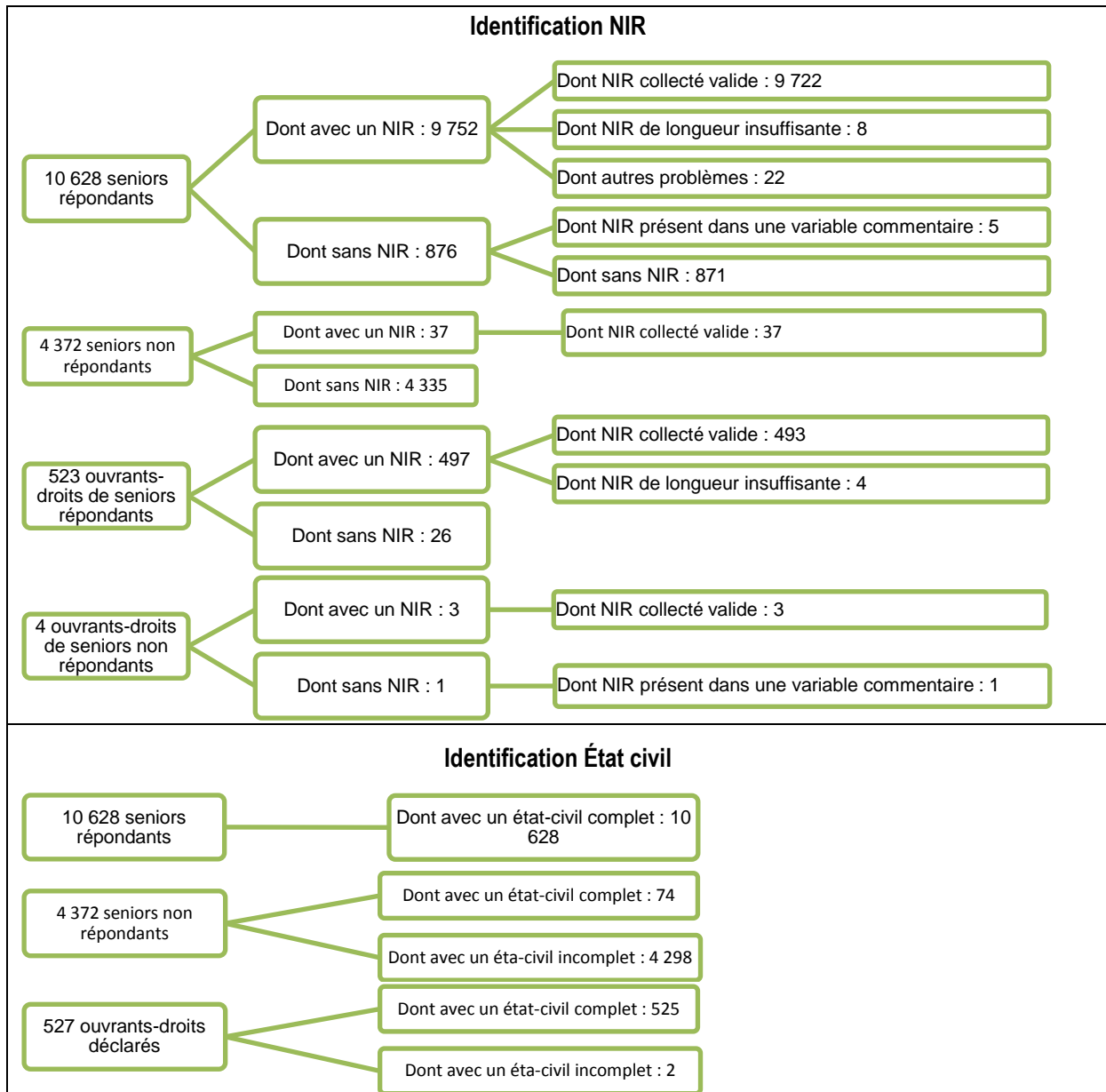
LE SCHEMA COMMENCE PAR LE SIN DE LILLE QUI CONSTRUIT LES TABLES DE DEMANDE D'IDENTIFICATION (NUMÉROS 1 ET 2) ET ENVOIE CES FICHIERS AU CASD (3) ET AU PÔLE RFD (4). LE PÔLE RFD PROCÈDE À L'IDENTIFICATION DES DEMANDES (5) ET RENVOIE LES FICHIERS DE RÉPONSES AU CASD (6). EN CAS DE REJETS TROP NOMBREUX (TAUX SUPÉRIEUR À 5%), UNE IDENTIFICATION CORRECTIVE SERA DEMANDÉE PAR LE CASD AU PÔLE RFD (ÉTAPES 7 ET 8). LE CASD PROCÈDE AU CONTRÔLE DES RÉPONSES, CONSTRUIT LES FICHIERS POUR LES APPARIEMENTS UTILISANT LE NIR ET LES ENVOIE AUX PARTENAIRES (11).

Il est à noter que la DREES ne sera jamais destinataire de données directement identifiantes (NIR, nom de naissance et nom de famille, adresse...) Le fait que la DREES ne puisse disposer des données directement identifiantes rend le suivi de ce type d'opération délicat puisque le destinataire des données ne peut en contrôler directement la construction et la qualité. C'est pour cette raison qu'un test à partir de données personnelles de l'équipe conceptrice (et donc, auxquelles elle peut avoir accès) a été mené entre les différents acteurs. Il a permis de mettre en place, et de tester :

- Les échanges des fichiers entre le SIN de Lille, le pôle RFD, le CASD, la CNAM-TS, la CNAV, la CCMSA et la DREES ;
- Les formats des variables utiles aux appariements ;
- Les travaux du CASD concernant les vérifications et les constructions de tables.

À l'issue de la collecte de l'enquête Care-Ménages, volet seniors, les effectifs envoyés pour identification sont décrits dans la figure 5.

Figure 5. Individus envoyés à l'identification



NOTE : LA DÉFINITION DE L'ENSEMBLE DES ACRONYMES EST DISPONIBLE EN ANNEXE 3. SUR LES 10 628 SENIORS QUI ONT RÉPONDU INTÉGRALEMENT À L'ENQUÊTE, 9 752 ONT DONNÉ UN NIR LORS DE LA COLLECTE. IL ÉTAIT VALIDE POUR 9 722 D'ENTRE EUX.

Un travail dans la durée

L'un des objectifs de la collecte du NIR est de pouvoir réaliser pendant quinze ans un suivi de la mortalité. Le suivi de la mortalité sera effectué sur la base du NIR validé, en collectant l'information sur le statut vital des personnes figurant au RNIPP tous les ans, pendant quinze ans. Le suivi de la mortalité des enquêtes HSM (de 2008 à 2023) et HID (de 1999 à 2014) ayant connu des interruptions, un suivi spécifique et une définition précise du rôle de chacun des acteurs doivent être mis en place.

Chaque premier trimestre de chaque année jusqu'en 2031, le chef du bureau « Handicap-Dépendance » de la DREES enverra une notification au service informatique national de l'INSEE à Lille – SIN de Lille – (service qui archive les NIR validés) pour initier les travaux à réaliser. Suite à cette notification, le SIN de Lille procédera à la mise aux normes du fichier contenant les NIR pour une demande d'identification sur le NIR et l'enverra au pôle

RFD. Ce dernier réalisera une identification sur le NIR et renverra les réponses au SIN. En vertu de l'article 7 du décret en Conseil d'État de l'enquête, la conservation du NIR par l'INSEE est autorisée jusqu'à l'obtention de la mention de décès. Le SIN traitera donc le fichier identifié en le comparant à celui de l'année précédente afin de supprimer le NIR et les autres informations directement identifiantes une fois l'obtention de la mention de décès « certaine ». Un décès est considéré comme « certain » dès lors que la mention de décès apparaît deux années de suite dans le fichier BRPP. Au deuxième trimestre de chaque année, le SIN de Lille transmettra à la DREES le fichier comportant uniquement l'identifiant anonyme de l'individu dans l'enquête CARE-ménages volet seniors et la mention du décès, à savoir la date et le lieu du décès. L'ensemble de ces traitements a été consigné dans une note produite par la DREES qui sera envoyée chaque année en même temps que la notification au SIN de Lille.

Risque de ré-identification et questions de diffusion

Les différents enrichissements permettront de disposer de données si détaillées et si nombreuses qu'elles deviendront indirectement identifiantes. La DREES n'est donc pas autorisée à fournir l'ensemble des bases détaillées au réseau Quételet, dans l'optique d'une mise à disposition aux chercheurs qui souhaiteraient y avoir accès.

La DREES s'assurera que les bases diffusées à Quételet ne sont pas indirectement identifiantes ; pour cela elle procédera à une anonymisation des bases. L'ensemble des données diffusées (données d'enquêtes, du SNIIR-AM, RFS, CNAV, CCMSA, départementales, de la BPE...) seront donc agrégées et partielles.

Concernant les données du SNIIR-AM, des montants agrégés seront fournis sur le même modèle que ceux fournis pour l'enquête Handicap-Santé de 2008. Pour les données issues de la base permanente des équipements, la DREES devra construire elle-même les indicateurs au plus près des besoins des chercheurs car elle ne pourra pas fournir les données géo-localisées. Les revenus, qu'ils proviennent de l'appariement RFS, CNAV ou CCMSA, seront agrégés au niveau ménage et regroupés par catégorie de revenus. Les allocations spécifiquement liées à dépendance (enrichissement par les données des conseils départementaux et données de l'action sociale de la CNAV et de la CCMSA) donneront des agrégats d'aide humaine, d'aides techniques etc. Enfin, concernant le suivi de la mortalité, seuls le mois et l'année du décès sera diffusée aux chercheurs.

L'article 6 du décret en Conseil d'État permet toutefois que « Les [données enrichies soient] communiquées à des fins de statistique publique ou de recherche scientifique ou historique sur décision de l'administration des archives prise après avis du comité du secret statistique dans les conditions prévues à l'article 17 du décret du 20 mars 2009 susvisé. ». Les chercheurs qui souhaitent étudier un thème précis pourront faire une demande d'accès à certains fichiers détaillés. Cette demande nécessitera une demande d'autorisation de la CNIL et un passage devant le comité du secret statistique. Ce comité est appelé à se prononcer sur toute question relative au secret en matière de statistiques. Il donne son avis sur les demandes de communication de données individuelles collectées par voie d'enquête statistique ou transmises au service statistique public, à des fins d'établissement de statistiques.

Disponibilité partielle des données

D'un appariement à l'autre, les données remontées ne seront pas disponibles pour toute la population concernée.

Concernant l'appariement avec les données du SNIIR-AM, un travail important a été réalisé par la DREES, la CNAM-TS et l'IRDES (Institut de Recherche et Documentation en Économie de la Santé) pour que la majorité des indicateurs soient construits pour toute la population. Il en restera toutefois qui ne seront disponibles que pour les personnes affiliées au régime général (les "tops pathologies" par exemple et les variables agrégées issues des "datamarts" construits par la CNAMTS).

Concernant les données des départements, la DREES a consulté les conseils départementaux en amont de la collecte afin de connaître les variables disponibles dans leurs bases de gestion. L'expression des besoins a donc

été définie de sorte à limiter les manques pour certaines variables d'un département à l'autre. Par ailleurs, le recours aux éditeurs de logiciels de gestion utilisés par les départements pour la construction des variables permettra une homogénéisation des variables remontées. Il restera néanmoins des vides pour certaines variables, parce que certains départements ont des pratiques différentes (saisie partielle des données, paiement des prestataires en CPOM²⁰ ou des établissements en dotation globale²¹, versements disponibles dans une autre base que celle des données de gestion²²).

Par ailleurs, il a été décidé de ne pas réaliser d'appariement avec d'autres caisses de retraite que la CNAV et la CCMSA (RSI²³, fonctionnaires civils et militaires, collectivités locales, régimes spéciaux, régimes non-salariés...). Tout d'abord parce que les pensions seront disponibles grâce à l'appariement RFS. Ces appariements supplémentaires auraient permis de remonter des informations sur l'action sociale et les prestations non imposables (ASPA, MTP, ASI...). Or seulement 15 % des répondants de l'enquête CARE-ménages, volet seniors déclarent être affilié à l'un de ces régimes. Le nombre de bénéficiaires de ces prestations retrouvés grâce à l'appariement aurait donc été très faible, pour un coût élevé. Il a donc été décidé d'utiliser les données d'enquête ainsi que les barèmes en vigueur pour réaliser des imputations sur ces prestations.

Recouvrement partiel des périodes de référence

L'enquête CARE-ménages volet seniors a été collectée entre le 2 mai et le 17 octobre 2015. Afin d'éviter les biais de mémoire, c'est la situation au moment de leur interrogation qui est demandée aux seniors. Par contre, les données issues des sources fiscales (RFS) ne peuvent couvrir que des années calendaires. Pour disposer d'une année complète en cas de décès du bénéficiaire durant l'année 2015, et pour utiliser des données disponibles plus rapidement, l'appariement avec des données relatives à l'année 2014 a été préféré à l'appariement avec des données de l'année 2015, en ce qui concerne les données fiscales du moins.

La plupart des personnes interrogées seront retraitées. Les retraites étant peu fluctuantes, les montants de retraite en 2014 seront semblables à ceux de 2015²⁴. Par contre, pour les personnes actives ou celles ayant changé de situation en 2015 (passage à la retraite par exemple), seules les données collectées donneront une information sur les montants de retraite perçus.

Concernant les données « retraite » de la CNAV, comme la CNAV ne peut pas remonter l'ensemble des montants mensuels de 2014 et/ou 2015, les données remontées seront celles au 31/12/2014 et au 31/12/2015. Ce choix a été fait pour plusieurs raisons :

- Détection des changements de situation entre 2014 et 2015, ce qui permettra d'imputer des montants de revenu 2015 ;
- Estimation plus fiable des revenus au cours des douze derniers mois précédant la collecte CARE par rapport à l'unique montant mensuel au 31/12/2014 ;
- Disponibilité d'au moins une année complète en cas de décès du bénéficiaire durant l'année 2015.

Concernant les données « retraite » et « famille » de la CCMSA, la période de référence couvrira les années 2014 et 2015. Ce choix a été fait pour plusieurs raisons :

- pour disposer de données temporellement cohérentes avec toutes les sources (RFS, données d'enquête, données du SNIIR-AM, données des départements) ;
- pour avoir des données sur les douze derniers mois avant la date d'enquête, quel que soit le mois de collecte ;

²⁰ Contrat Pluriannuels d'Objectif et de Moyens.

²¹ 68 % des départements ayant répondu à la sollicitation de la DREES indiquent que tous les montants du plan sont individualisables et 31 % indiquent qu'une partie seulement des montants est individualisable.

²² 82 % des départements ayant répondu à la sollicitation de la DREES indiquent que les montants versés se trouvent dans la même base que les montants notifiés.

²³ Régime Social des indépendants

²⁴ Du fait notamment du gel des pensions de retraite entre 2014 et 2015, dans un contexte d'inflation très faible.

- pour aider à déceler des changements de situation en 2015 et pouvoir imputer les montants de revenus 2015 ;
- pour disposer d'une année complète en cas de décès du bénéficiaire durant l'année 2015.

Concernant les données « action sociale » de la CNAV, la période de référence couvrira les années 2014 et 2015, ainsi que certaines variables au moment de l'appariement. Ce choix a été fait pour les mêmes raisons que précédemment.

Concernant les données « action sociale » de la CCMSA, l'appariement nécessitant un travail de la part de chaque caisse locale MSA, il a été décidé de restreindre le nombre de variables remontées et donc la période de l'appariement. Seuls la date d'entrée, la date de sortie, les derniers montants et le dernier nombre d'heures d'aide seront remontés. Les montants attribués étant peu fluctuants, cela permettra de reconstruire les montants des douze derniers mois précédant l'enquête.

La période de référence pour la prise en compte des consommations médicales couvre les douze mois précédant l'enquête, soit du 1^{er} mai 2014 au 2 mai 2015, comme le stipule le décret en Conseil d'État. Cette période ne correspondra pas exactement aux douze mois précédant la collecte pour tous les individus, comme la collecte s'étale sur plusieurs mois. Cela pose des difficultés lorsque l'on cherche à croiser le SNIIR-AM et les données de l'enquête rétrospectives sur douze mois (« Avez-vous dans les douze derniers mois... »). Pour tenir compte des délais de liquidation parfois longs, notamment pour certaines dépenses hospitalières, et être sûr de prendre en compte l'ensemble des consommations au cours de la période de référence, l'appariement sera effectué sur le fichier SNIIRAM 1^{er} mai 2014 jusqu'à la date d'extraction ; mais, conformément au décret, seuls les actes établis lors de la période de référence seront pris en compte.

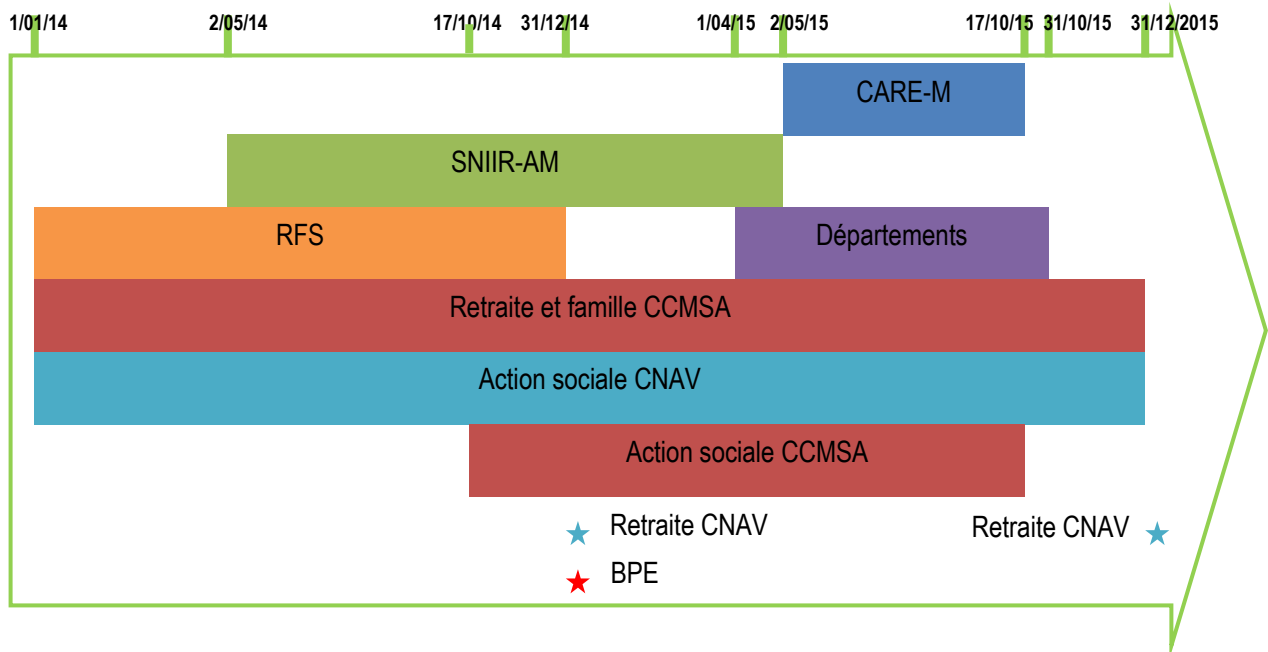
Pour l'enrichissement par les données des départements, un enrichissement sur des données d'avril 2015 à octobre 2015 a été préféré. La restriction à sept mois seulement a pour objectif principal de limiter le nombre de variables à remonter par les départements. Par ailleurs, ces dates permettront de disposer pour chaque individu des montants du plan d'aide du mois précédant la collecte, quel que soit le mois durant lequel celle-ci a eu lieu, et d'être ainsi en cohérence avec la situation et l'organisation de l'aide décrite dans le questionnaire. Par contre, pour les aides ponctuelles (et donc plus rares), il a été décidé de remonter l'ensemble des paiements entre le 1/04/2014 et le 31/10/2015. Cela permettra de disposer de l'ensemble des aides ponctuelles reçues (aides techniques et aménagements du logement) par la personne les douze mois précédant l'enquête.

La BPE est mise à jour annuellement, avec comme date de référence le 1^{er} janvier. La BPE 2014 est actuellement disponible, celle de 2015 le sera au second semestre 2016. Il est cependant à craindre que certains petits commerces n'aient disparu ou ne se soient créés entre le 1^{er} janvier 2015 et la date de collecte de l'enquête.

Les périodes de référence sont rappelées dans la figure 6.

Ces différentes périodes de références demanderont un important travail d'imputation pour pouvoir analyser ensemble les données issues des différents enrichissements.

Figure 6. Période de référence pour chaque enrichissement



NOTE : LA DÉFINITION DE L'ENSEMBLE DES ACRONYMES EST DISPONIBLE EN ANNEXE 3.
LECTURE : L'APPARIEMENT RFS COUVRE LA PÉRIODE 1/01/14-31/12/2014.

Erreurs d'appariement et travail d'imputation

Les données remontées grâce aux enrichissements nécessitent de gros travaux d'imputation notamment parce que certaines variables sont partiellement remontées, que certaines périodes ne sont pas couvertes par les enrichissements (voir *supra*) et que certaines unités n'ont pas été appariées. Pour l'appariement SNIIR-AM d'Handicap-Santé, le nombre d'individus non appariés était important. Le choix avait alors été fait de calculer de nouvelles pondérations plutôt que d'imputer des consommations médicales aux individus non appariés.

Pour CARE-ménages volet seniors, étant donné le nombre d'enrichissements différents, nous privilégierons les imputations plutôt que le calcul de nouvelles pondérations propres à chaque enrichissement. Cela permettra d'utiliser l'ensemble des enrichissements pour les analyses et de ne pas avoir à recalculer des jeux de poids relatifs à chaque enrichissement, complexes à réaliser et à utiliser.

La collecte du NIR et sa validation *via* deux demandes d'identification (voir plus haut) améliorera significativement les appariements utilisant le NIR par rapport à l'expérience de l'enquête Handicap-Santé. Les erreurs d'appariements devraient donc être en nombre limité.

Dans un premier temps, il sera nécessaire de distinguer les erreurs d'appariements des individus non consommateurs/non bénéficiaires. Les valeurs des variables issues de l'appariement, pour les individus non appariés du fait d'une erreur d'appariement, devront alors être imputées.

Ce travail de distinction des individus non appariés du fait d'une erreur d'appariement et des non-consommateurs sera réalisé grâce aux données d'enquête. Il sera facilité pour l'appariement avec les données du SNIIR-AM, car la CNAM-TS recherchera dans les archives de ses référentiels bénéficiaires les individus CARE-ménages volet seniors et remontera une variable prenant les modalités suivantes : 1 si l'individu a consommé au cours des douze derniers mois, 0 si l'individu n'a pas consommé au cours des douze derniers mois mais a été retrouvé dans les archives du SNIIR-AM, 9 si l'individu n'a pas été retrouvé. Les individus non retrouvés seront donc

assimilés à une erreur d'appariement²⁵ et seront imputés grâce aux données d'enquête soit en individu consommant soit en individu non consommant.

L'appariement avec les sources fiscales utilisera comme clé d'appariement l'identifiant fiscal puisque l'échantillon a été tiré dans ces sources. L'appariement sera donc facilité. À titre indicatif, pour la constitution de l'ERFS 2009, l'appariement de l'enquête Emploi et des données fiscales permet de retrouver 93 % des individus et 94 % des ménages attendus. En outre, pour les données sociales remontées par l'appariement RFS, les montants couvrent : 68 % de la masse financière des prestations familiales, 61 % des minima sociaux hors minimum vieillesse, 29 % du minimum vieillesse et 75 % des allocations logement pour l'année 2009 (données pondérées, source : ERFS 2009). Rappelons que, comme les appariements RFS ne fournissent pas une information exhaustive, la DREES réalisera en parallèle des appariements avec les données de la CNAV et de la CCMSA, pour le minimum vieillesse notamment.

Les montants toujours manquants seront complétés par des imputations :

- de revenus fiscaux individualisables manquants pour les individus qui n'ont pas été appariés aux données fiscales (individus manquants dans des ménages appariés, individus appariés au fichier CNAF mais pas aux fichiers fiscaux) ;
- des différentes prestations pour les familles éligibles à ces prestations (définies comme telles selon la configuration familiale, ou qui se déclarent comme telles dans l'enquête CARE-ménages volet seniors) et non appariées aux fichiers sociaux.

Concernant l'enrichissement par les données des départements, il est possible que certains départements ne soient pas en capacité de remonter ces informations. À titre indicatif, lors d'une opération de remontées de données individuelles sur l'APA réalisée par la DREES en 2011, 66 départements avaient accepté de participer. Dans le cas de CARE-ménages volet seniors, le nombre de départements participant devrait être plus élevé pour l'enrichissement puisque certains départements ont donné leur accord pour réaliser toute l'opération eux-mêmes ou une partie de l'opération. Il ne sera pas forcément nécessaire de calculer de nouvelles pondérations suite à l'enrichissement, et ce même si certains départements ne répondent pas à la demande d'enrichissement. En fonction des résultats d'une analyse comparant les départements répondants et non répondants, les variables pour les individus des départements non enrichis pourraient être imputées comme s'il s'agissait simplement de non-réponse partielle à une enquête – méthode *hot deck*, plus proche voisin, etc. – ce qui permettra ensuite d'exploiter les variables issues de l'enrichissement pour la production de statistiques sur l'ensemble du champ de l'enquête. Les estimations issues de l'imputation seront comparées aux marges connues, pour évaluer la qualité de l'imputation.

Cas particulier : l'enrichissement par les données des départements

Cet enrichissement n'ayant jamais été réalisé, il n'existe pas de procédure classique d'enrichissement. Par ailleurs, comme il n'existe pas de base harmonisée au niveau national – mais uniquement des bases département par département – et pour limiter la charge de travail supportée par les départements, il sera fait appel à un prestataire qui travaillera en collaboration avec les éditeurs de logiciel de gestion de l'APA pour récupérer les informations. Pour déterminer le protocole le plus adapté et rédiger le cahier des charges du prestataire, cinq conseils départementaux ont été rencontrés. Ces rencontres ont permis d'aborder le fonctionnement de la gestion de l'APA à domicile dans chaque département et de visualiser le contenu des bases de gestion. Par ailleurs, un questionnaire a été envoyé à la fin de l'année 2015, à tous les départements afin de collecter le même type d'information et d'avoir leur avis concernant le protocole envisagé. 85 départements ont

²⁵ Les archives SNIIR-AM couvrent 3 années. Or, peu d'individu de 60 ans ou plus sont non-consommateurs trois années consécutives.

répondu favorablement sur les 94²⁶ dans lesquels résident des personnes enquêtées. Une partie des départements a indiqué qu'elle préférerait réaliser l'ensemble de l'opération elle-même. Une autre partie des départements a indiqué qu'elle préférerait avoir recours à un prestataire pour une fraction de l'opération. Le reste des départements a indiqué qu'il préférerait avoir recours à un prestataire pour l'ensemble de l'opération (voir tableau 1).

Tableau 1. Répartition des départements selon leur implication dans l'opération d'enrichissement

Préférence des départements	Individus à rechercher	Dont : bénéficiaires estimés de l'APA	Nombre de départements
<i>Groupe 1 : Extraction et recherche faite par le département</i>	13 263	2 283	60
<i>Groupe 2 : Recherche faite par le prestataire et extraction faite par le département</i>	1 617	269	6
<i>Groupe 3 : Recherche faite par le département et extraction faite par un éditeur de logiciel de gestion</i>	1 312	235	5
<i>Groupe 4 : Extraction faite par un éditeur de logiciel de gestion et recherche faite par le prestataire</i>	4 072	687	22
<i>Dont départements n'ayant pas répondu à la consultation de la DREES</i>	1 346	224	8
<i>Non consulté</i>	2	0	1
Total général	20 270	3 474	94

NOTE : LE NOMBRE DE PERSONNES RETROUVÉES DANS LES BASES DES BÉNÉFICIAIRES DE L'APA, I.E. À EXTRAIRE, EST ESTIMÉ GRÂCE AU SCORE DE DÉPENDANCE CALCULÉ DANS LA PRÉ-ENQUÊTE À CARE-MÉNAGES VOLET SENIORS : L'ENQUÊTE VQS. UNE CORRESPONDANCE ENTRE LE GROUPE DE DÉPENDANCE, APPELÉ GROUPE VQS ET LE GIR ESTIMÉ EST DISPONIBLE GRÂCE AUX DONNÉES DE LA PRÉCÉDENTE ENQUÊTE HANDICAP-SANTÉ. PUIS, UN TAUX DE NON-RECOURS DE 20 % A ÉTÉ APPLIQUÉ POUR ABOUTIR AU NOMBRE ESTIMÉ DE BÉNÉFICIAIRES EFFECTIFS. POUR LES COHABITANTS, COMME LE GROUPE VQS N'EST PAS CONNU, UN TAUX DE 8 % A ÉTÉ APPLIQUÉ (PRÉVALENCE DES BÉNÉFICIAIRES DE L'APA DANS LA POPULATION 60 ANS OU PLUS).
LES DÉPARTEMENTS N'AYANT PAS RÉPONDU À LA CONSULTATION ONT ÉTÉ RÉPARTIS DANS CHAQUE GROUPE SELON LE LOGICIEL DE GESTION QU'ILS UTILISENT.

Outre son coût, une telle opération nécessite un important travail de coordination entre les différents acteurs : prestataire, éditeurs de logiciel de gestion, départements, DREES, CASD, et un suivi continu.

L'appel d'offres lancé en fin d'année 2015 pour recruter un prestataire en charge de l'opération s'est avéré infructueux (aucune réponse reçue). La DREES a dû revoir à la marge la procédure d'enrichissement. En effet, la DREES avait envisagé que le prestataire recruté se rendrait dans chaque département ne faisant pas l'opération lui-même pour réaliser le travail de recherche des individus CARE et d'extraction des données les concernant. Il semble que c'est principalement cette partie « déplacements » qui a freiné les prestataires potentiels. La DREES a, de ce fait, décidé d'avoir recours aux éditeurs de logiciel de gestion qu'utilisent les départements afin que l'opération d'extraction puisse se faire à distance. La DREES a lancé une procédure négociée avec un prestataire et lui a demandé de contractualiser une co-traitance avec les éditeurs de logiciel. Avant de signer la négociation, il a fallu de nombreux échanges entre les différents acteurs pour définir le rôle et les échéances que chacun devra respecter.

Lors du lancement de l'opération, le Centre d'accès sécurisé aux données (CASD), tiers de confiance de l'opération, enverra au prestataire, grâce à un serveur FTP, un fichier Excel par département comprenant les données nominatives de chaque individu CARE, comme précisé dans le décret en Conseil d'État (décret n° 2015-343 du 26 mars 2015), dont l'article 5 précise que :

²⁶ Un département n'a pas été consulté car seules deux personnes de l'enquête y résidaient.

« les noms, prénoms, sexe, date et lieu de naissance des personnes enquêtées sont utilisés pour :

(...)

2° effectuer un rapprochement des informations collectées avec les données concernant les allocations et prestations présentes dans les fichiers de départements ».

L'avis du comité du secret statistique a dû être obtenu afin d'autoriser le prestataire (ainsi que les départements qui font eux-mêmes l'opération) à recevoir du CASD ces listes nominatives d'individus CARE.

Pour chaque groupe de départements (défini dans le tableau 1), le protocole est adapté.

Groupe 1 : Extraction et recherche faite par le département

Dès réception des listes des « individus CARE », le prestataire les enverra selon le mode de transmission sécurisé défini avec les départements réalisant eux-mêmes l'opération. Le prestataire n'enverra à chaque département que la liste le concernant. Ainsi, chaque département disposera uniquement d'informations relatives à ses propres administrés. Le département devra réaliser une recherche dans ses bases de gestions APA de chaque « individu CARE » (soit par requête soit de façon manuelle grâce aux champs noms, prénoms, sexe, date et lieu de naissance des individus). Lorsqu'un individu sera trouvé, le département devra remonter toutes les variables demandées par la DREES. Chaque département renverra au prestataire par envoi informatique sécurisé les données individuelles souhaitées (un dessin de fichier sera fourni par le prestataire, que les départements devront compléter). Le prestataire assurera un suivi et un appui aux départements pour la réalisation des extractions.

Groupe 2 : Recherche faite par le prestataire et extraction faite par le département

Le prestataire enverra le dessin de fichier à compléter, ainsi que la liste des dates de naissance des personnes à rechercher au département. Le département procédera à l'extraction de l'ensemble des données demandées pour l'ensemble des personnes présentes dans la base de gestion du département qui sont nées aux dates fournies par le prestataire. Chaque département concerné renverra au prestataire les données individuelles souhaitées par envoi informatique sécurisé. Le prestataire assurera un suivi et un appui aux départements pour la réalisation des extractions. Le prestataire sera chargé de rechercher les « individus CARE » dans les bases reçues grâce à un algorithme de recherche sur les noms, prénoms, date de naissance et sexe.

Groupe 3 : Recherche faite par le département et extraction faite par un éditeur de logiciel de gestion

Dès réception des listes des « individus CARE », le prestataire les enverra selon le mode de transmission défini avec les départements. Le prestataire n'enverra à chaque département que la liste des personnes le concernant. Ainsi, chaque département disposera uniquement d'informations relatives à ses propres administrés. Le département devra rechercher chaque « individu CARE » dans les bases de gestions APA, soit par requête soit de façon manuelle grâce aux champs noms, prénoms, sexe, date et lieu de naissance des individus. Le prestataire assurera un suivi et un appui aux départements pour la réalisation de la recherche. Lorsqu'un individu sera trouvé, le numéro de dossier du logement et le numéro individuel seront collectés et ils serviront à la seconde phase, réalisée par l'éditeur de logiciel. Chaque département renverra les identifiants individuels des « individus CARE retrouvés » à son éditeur de logiciel par un envoi informatique sécurisé. L'éditeur de logiciel sera chargé de faire l'extraction des variables demandées par la DREES uniquement pour les identifiants fournis par le département. L'éditeur de logiciel renverra au prestataire les données selon le mode de transmission sécurisé défini conjointement.

Groupe 4 : Extraction faite par un éditeur de logiciel de gestion et recherche faite par le prestataire

Le prestataire enverra le dessin de fichier à compléter, ainsi que la liste des dates de naissance des personnes à rechercher à chaque éditeur de logiciel. L'éditeur de logiciel procédera à l'extraction de l'ensemble des données demandées pour l'ensemble des personnes présentes dans la base de gestion du département qui sont nées aux dates fournies par le prestataire. Chaque éditeur renverra au prestataire les données individuelles souhaitées

par envoi informatique sécurisé. Le prestataire sera chargé de rechercher les « individus CARE » dans les bases reçues grâce à un algorithme de recherche sur les noms, prénoms, date de naissance et sexe.

À la suite de ces deux phases, le prestataire devra vérifier le contenu et la cohérence des fichiers collectés. Il s'assurera que les fichiers comportent l'ensemble des informations demandées et que le nombre de bénéficiaires trouvés est cohérent avec les estimations de la DREES. Par ailleurs, le prestataire devra procéder à un ensemble de contrôles de cohérence. Une fois les vérifications faites, le prestataire procédera à une estimation du nombre de bénéficiaires par département grâce aux pondérations de l'enquête CARE-ménages volet seniors afin de relever d'éventuelles incohérences entre cette estimation et les données de cadrage publiées par la DREES pour l'année correspondante.

Une fois la base finale construite, validée et anonymisée, le prestataire la transmet à la DREES grâce à un serveur sécurisé.

■ CONCLUSION

Les enrichissements des enquêtes par des appariements avec des données administratives représentent un facteur important d'amélioration de la qualité de la statistique publique. Le code de bonnes pratiques de la statistique européenne rappelle en effet, dans son principe 9, l'objectif de ne pas soumettre les déclarants à une charge excessive d'interrogation. La très longue durée d'interrogation en face-à-face lors de l'enquête Handicap-Santé a poussé l'équipe conceptrice de l'enquête CARE-ménages à envisager des moyens alternatifs pour récupérer suffisamment d'information. Les enrichissements par des sources administratives sont la solution la plus efficace pour limiter la charge supportée par les enquêtés et disposer de données plus riches, plus fiables et plus détaillées.

Les différents enrichissements améliorent notablement la qualité des données mais vont de pair avec de nombreuses difficultés, qui représentent autant de contraintes pour les équipes productrices : démarches juridiques, travaux d'apurement et d'imputation, coordination de différences services, mises en place de procédures très sécurisées, restriction dans la diffusion des données etc.

Par ailleurs, un arbitrage est à faire entre deux éléments importants de la qualité des statistiques publiques : d'une part « exactitude et fiabilité » (principe 12 du code de bonnes pratiques) et d'autre part, « actualité et ponctualité » (principe 13). En effet, les enrichissements nécessitent une préparation d'environ deux ans, ce qui empêche de répondre à de nouvelles questions dans un temps limité (actualité). De plus, ils demandent un certain nombre de post-traitements prenant beaucoup de temps (ponctualité). La collecte de l'enquête CARE-ménages volet seniors s'est terminée le 27 octobre 2015 et la DREES ne dispose des données de collecte que depuis mars 2016. En effet, la collecte est suivie de nombreux traitements aval réalisés par l'INSEE et la DREES : vérification de la cohérence des données, construction des tables, préparation des tables pour l'appariement prioritaire avec le RNIPP, reprise manuelle des catégories socioprofessionnelles (CSP)... Les apurements et pondérations des données, réalisés par la DREES, permettront de disposer, puis de diffuser des données propres utilisables (sans enrichissement) début 2017. À cette même période, la DREES recevra les données enrichies grâce aux données administratives et ces données seront corrigées puis diffusées d'ici fin 2017 – début 2018. Il faut donc prévoir un écart au minimum d'un an entre la collecte et l'exploitation des données non enrichies et de deux ans entre la collecte et l'exploitation des données enrichies.

Rappelons que les données d'enquêtes seules ou les données administratives seules ne se suffisent pas à elles-mêmes, du fait de non-réponses partielles (pour les enquêtes) ou de trous d'appariement (pour les données administratives), ou du fait de réponses ou enregistrements erronés. Cela nécessite donc un travail important d'imputation, qui implique de ne pas se reposer uniquement sur un seul des deux modes de collecte des données. Ainsi, dans l'enquête CARE-ménages volet seniors, un questionnement résiduel sur les ressources des ménages est par exemple conservé afin de s'assurer d'avoir des informations minimales sur les revenus et de faciliter le travail d'imputation au cas où les revenus n'auraient pas pu être retrouvés dans les sources fiscales. Surtout, un module important permettant de recenser les prestations liées à la dépendance a été maintenu. Il permet d'avoir des informations déclaratives sur les aides que reçoivent les personnes du fait de leur état de santé ou de leur âge : APA, action sociale, ACTP (Allocation compensatrice de tierce personne), PCH... Les informations recueillies dans ce module pourront également être comparées aux informations collectées par appariements, pour les individus ayant les deux sources d'information ; ceci présente, en soi, un intérêt méthodologique (mesurer l'écart entre la déclaration d'une aide et sa perception effective) et permettra d'imputer des aides aux personnes pour lesquelles l'appariement n'aura pas pu se faire.

■ BIBLIOGRAPHIE

Ankri J. et Mormiche P., 2002, « Incapacités et dépendance de la population âgée française : apport de l'enquête Handicaps, incapacités, dépendance (HID) », *Santé, Société et Solidarité*, 2, 25-38.

Auvray L. et Le Fur P.A., 2002, « Améliorer la mesure de la consommation pharmaceutique : une nouvelle méthode de recueil », *Questions d'économie de la santé IRDES* n° 54, juillet.

Bérardier M. et Clément É., 2011, « L'évolution de l'allocation personnalisée d'autonomie (APA) de 2002 à 2009 », *Études et Résultats*, DREES, n° 780, octobre.

Bérardier M. et Debout C., 2011, « Une analyse des montants des plans d'aide accordés aux bénéficiaires de l'APA à domicile au regard des plafonds nationaux applicables », *Étude et Résultats*, DREES, n° 748.

Bérardier M., 2012a, « Allocation personnalisée d'autonomie à domicile : quels restes à charge pour les bénéficiaires ? », *Revue française des affaires sociales*, 2012/2-3 (n° 2-3), pp. 197-217.

Bérardier M., 2012b, « Vieillir chez soi : usages et besoins des aides techniques et des aménagements du logement », *Études et résultats*, DREES, n° 823, 6 p.

Bérardier M., 2014, « Les bénéficiaires de l'allocation personnalisée d'autonomie à domicile et leurs ressources en 2011 », *Études et Résultats*, DREES, n°876, février.

Bérardier M., 2015, « Aide sociale à l'hébergement en 2011 : des bénéficiaires moins dépendants et plus seuls que ceux de l'allocation personnalisée d'autonomie, en établissement », *Études et Résultats*, DREES s, n°909, mars.

Bouvier G., 2011, « L'enquête Handicap-Santé, Présentation Générale », *Document de travail N°F1109*, INSEE, octobre.

Calvet L. et Montaut A., 2013, « Dépenses de soins de ville des personnes âgées dépendantes à domicile : des dépenses comparables en institution et à domicile », *Dossier Solidarité Santé n° 42*, DREES, août.

Carrère A., Haag O. et Soullier N., 2015, « Enquêtes VQS et CARE : ou comment prendre en compte le tirage d'échantillons dépendants dans le cadre du NCEE », *Acte de la XIe édition des Journées de méthodologie statistique*, INSEE, avril.

Charpin J.M., 2011. « Rapport du groupe de travail "Perspectives démographiques et financiers de la dépendance" », juin.

Dos Santos S. et Makdessi Y., 2010, « Une approche de l'autonomie chez les adultes et les personnes âgées », *Études et résultats*, DREES, n°718, 8 p.

DREES, 2016, « Enquête Vie Quotidienne et Santé 2014, Résultats départementaux » à paraître.

Eghbal-Téhérani S. et Makdessi Y., 2011, « Les estimations GIR dans les enquêtes Handicap-Santé 2008-2009 », *Série Sources et méthodes* N° 26, septembre.

Enquête Revenus fiscaux et sociaux (ERFS) 2012, *Sources et méthodes*, Version du 15 décembre 2014.

Gensbittel M.-H. et Riandey B., 2011, « Appariements sécurisés et statistiques (2000-2011) : Une décennie d'expériences » *Courrier des statistiques* n°131, septembre.

Goldberg, M., Quantin, C., Guéguen, A., et Zins, M. 2008, « Bases de données médicoadministratives et épidémiologie: intérêts et limites ». *Courrier des Statistiques*, n°124.

Lardellier R., Legal R., Raynaud D. et Vidal G., 2012, « Dépenses de santé et restes à charge des ménages : le modèle de microsimulation Omar », *Document de Travail, série Sources et Méthodes* n° 34, DREES, août.

Lecroart A., Froment O., Marbot C. et Roy D., 2013, « Projection des populations âgées dépendantes : deux méthodes d'estimation », *Dossier Solidarité Santé n° 43*, DREES, septembre.

Lengagne P., Penneau A., Pichetti S. et Sermet C., 2015, « L'accès aux soins courants et préventifs des personnes en situation de handicap en France, Tome 1 – Résultats de l'enquête Handicap-Santé volet Ménages », juin. 130 pages.

Montaut A., 2012, « L'appariement Handicap-Santé et données de l'Assurance Maladie », *Actes des Journées de Méthodologie Statistique de l'INSEE*, janvier.

Penneau A., Pichetti S. et Sermet C., 2015, « L'accès aux soins courants et préventifs des personnes en situation de handicap en France, Tome 2 – Résultats de l'enquête Handicap-Santé volet Institutions », juin. 148 pages.

Soullier N. et Weber A., 2011, « L'implication de l'entourage et des professionnels auprès des personnes âgées à domicile », *Études et résultats*, DREES, n°771, 8 p.

Soullier N., 2012a, « Aider un proche âgé à domicile : la charge ressentie », *Études et résultats*, DREES, 799, 8 p.

Soullier N., 2012b, « L'aide humaine auprès des adultes à domicile : l'implication des proches et des professionnels », *Études et résultats*, DREES, 827, 6 p.

Annexe 1. Décret en Conseil d'État autorisant les enquêtes CARE et leurs enrichissements

JORF n°0074 du 28 mars 2015 page 5626
texte n° 42

DECRET

Décret n° 2015-343 du 26 mars 2015 portant création d'un traitement automatisé de données à caractère personnel dénommé « capacités, aides et ressources des seniors (CARE) »

NOR: AFSE1425897D

ELI: <http://www.legifrance.gouv.fr/eli/decret/2015/3/26/AFSE1425897D/jo/texte>

Alias: <http://www.legifrance.gouv.fr/eli/decret/2015/3/26/2015-343/jo/texte>

Publics concernés : personnes et entourage susceptibles de participer à une enquête statistique des services publics.

Objet : modalités de mise en œuvre d'une enquête statistique sur l'autonomie des seniors.

Entrée en vigueur : le texte entre en vigueur le lendemain de sa publication.

Notice : le présent décret définit les modalités de mise en œuvre du dispositif d'enquêtes « CARE » destiné à l'étude statistique de la perte d'autonomie des personnes âgées. Cette enquête concernera 20 000 personnes ainsi que, le cas échéant, leur entourage. Elle sera réalisée au cours des années 2015 et 2016 par le service statistique du ministère des affaires sociales, de la santé et des droits des femmes. Ses résultats feront l'objet de publications.

Le Premier ministre,

Sur le rapport de la ministre des affaires sociales, de la santé et des droits des femmes,

Vu le [code du patrimoine](#), notamment son article L. 212-4 ;

Vu le [code de la sécurité sociale](#), notamment son article L. 161-28-1 ;

Vu la [loi n° 51-711 du 7 juin 1951](#) modifiée sur l'obligation, la coordination et le secret en matière de statistiques ;

Vu la [loi n° 78-17 du 6 janvier 1978](#) modifiée relative à l'informatique, aux fichiers et aux libertés, notamment le I de son article 27 ;

Vu le [décret n° 2009-318 du 20 mars 2009](#) relatif au Conseil national de l'information statistique, au comité du secret statistique et au comité du label de la statistique publique ;

Vu l'avis du conseil de la Caisse nationale de l'assurance maladie des travailleurs salariés en date du 25 mars 2014 ;

Vu l'avis du conseil d'administration de la Caisse nationale d'assurance vieillesse en date du 2 avril 2014 ;

Vu l'avis du conseil central d'administration de la Mutualité sociale agricole en date du 3 avril 2014 ;

Vu l'[avis de la Commission nationale de l'informatique et des libertés en date du 3 juillet 2014](#) ;

Le Conseil d'État (section sociale) entendu,

Décète :

Article 1

Le ministre chargé des affaires sociales (direction de la recherche, des études, de l'évaluation et des statistiques) est autorisé à mettre en œuvre un traitement automatisé de données à caractère personnel relatif à un dispositif d'enquêtes à finalité statistique portant sur la perte d'autonomie des personnes âgées de 60 ans ou plus, dénommé « capacités, aides et ressources des seniors » ou CARE.

Ce traitement de données à caractère personnel a pour finalités :

1° De mesurer la perte d'autonomie de ces personnes et d'en suivre l'évolution ;

- 2° D'estimer le reste à charge lié à la perte d'autonomie ;
- 3° De recenser les aides humaines, financières et techniques nécessaires à la prise en charge de la perte d'autonomie, y compris l'aide apportée à ces personnes par leur entourage.

Article 2

Une première collecte des informations par enquête aura lieu auprès d'un échantillon représentatif d'environ 15 000 personnes âgées de 60 ans ou plus vivant à domicile.

Une seconde collecte aura lieu auprès d'un échantillon représentatif d'environ 5 000 personnes âgées de 60 ans ou plus séjournant de manière permanente dans des établissements d'hébergement pour personnes âgées.

Des enquêtes qualitatives ultérieures sous forme d'entretiens autour des thèmes abordés dans l'enquête CARE seront mises en œuvre dans un délai de deux ans auprès des personnes enquêtées ayant donné leur accord lors de l'une ou l'autre des collectes mentionnées au premier et au deuxième alinéa du présent article. Ces enquêtes feront l'objet de formalités préalables auprès de la Commission nationale de l'informatique et des libertés.

Les personnes enquêtées sont informées de la tenue de l'enquête par l'envoi d'une lettre-avis accompagnée d'un dépliant d'informations. Une lettre de relance leur est adressée en cas de non-réponse.

L'enquête se déroule dans le respect des dispositions relatives aux majeurs protégés.

Article 3

I. - Les catégories de données à caractère personnel collectées auprès des personnes enquêtées et enregistrées dans le traitement autorisé par l'article 1er sont :

- 1° L'identité : nom de naissance, nom d'usage, prénoms, sexe, date et lieu de naissance ;
- 2° Le numéro d'inscription au répertoire national d'identification des personnes physiques (NIR), dans les conditions précisées à l'article 4 ;
- 3° Le cas échéant, l'identité (nom de naissance, prénoms, date et lieu de naissance) et le NIR des assurés sociaux dont les personnes enquêtées sont ayants droit, dans les conditions précisées à l'article 4 ;
- 4° L'état de santé et le recours aux soins ;
- 5° Les limitations fonctionnelles, les restrictions d'activité ;
- 6° Les différentes formes d'aides techniques nécessaires et/ou utilisées ;
- 7° La situation familiale ;
- 8° Les revenus et prestations, les dépenses contraintes ;
- 9° Les différentes formes d'aide humaine reçues, professionnelle et/ou de l'entourage et leur coût ;
- 10° Le cas échéant, l'identité (nom d'usage, prénom) et les coordonnées (adresse postale et numéros de téléphone) des personnes de leur entourage qui leur apportent une aide, afin de permettre la réalisation d'une enquête spécifique auprès de cette population, à la suite de l'enquête principale et rattachée à celle-ci.

II. - Ces données seront complétées par les informations suivantes issues des systèmes de gestion des organismes de protection sociale, des départements, de la direction générale des finances publiques et de l'Institut national de la statistique et des études économiques :

- 1° L'adresse postale du sujet de l'enquête ;
- 2° Ses revenus fiscaux et sociaux ;
- 3° Sa consommation médicale et son statut vital.

III. - Les données permettant l'identification directe sont conservées séparément des autres informations issues de l'enquête et des systèmes d'information susmentionnés.

Article 4

Le numéro d'inscription au répertoire national d'identification des personnes physiques (NIR) est recueilli par les enquêteurs auprès des personnes enquêtées, qui reçoivent, sous la responsabilité de la direction de la recherche, des études, de l'évaluation et des statistiques, une information détaillée sur les conditions d'utilisation et de conservation de cette donnée par l'envoi des documents mentionnés à l'avant-dernier alinéa de l'article 2. Les mêmes conditions sont applicables au recueil du NIR des assurés sociaux dont les personnes enquêtées sont ayants droit.

Les enquêteurs ayant recueilli le NIR auprès des personnes enquêtées, de l'Institut national de la statistique et des études économiques (INSEE) pour la partie en ménages, du prestataire pour la partie en institutions sont

seuls habilités à l'enregistrer dans le système d'information de recueil des données de l'enquête. A l'issue de la collecte, aucune information n'est conservée par l'enquêteur.

L'INSEE, chargé de la gestion du répertoire national d'identification des personnes physiques, vérifie le NIR des personnes enquêtées et le reconstitue, le cas échéant.

Article 5

I. - Le numéro d'inscription au répertoire national d'identification des personnes physiques (NIR) des personnes enquêtées et des assurés sociaux dont les personnes enquêtées sont ayants droit est utilisé pour :

1° Effectuer un rapprochement des informations collectées avec les données de consommation médicale présentes au cours des douze mois précédant le début de la collecte des données, dans le système national d'information inter régimes de l'assurance maladie (SNIIR-AM) visé à l'[article L. 161-28-1 du code de la sécurité sociale](#) concernant :

- a) Des données relatives au bénéficiaire (régime, couverture maladie universelle) ;
- b) Des données sur les pathologies, notamment les affections de longue durée ;
- c) Des données sur les dépenses de santé et le recours aux soins ;

2° Effectuer un rapprochement des informations collectées avec les données présentes dans les fichiers des organismes d'assurance vieillesse concernant l'action sociale et les éléments de pension et d'allocations versées aux titulaires de pension de retraite et de réversion ;

3° Suivre la mortalité des personnes enquêtées grâce à la consultation régulière du répertoire national d'identification des personnes physiques.

II. - Les noms, prénoms, sexe, date et lieu de naissance des personnes enquêtées sont utilisés pour :

1° Effectuer un rapprochement des informations collectées avec les données fiscales et sociales présentes dans les fichiers fournis par la direction générale des finances publiques et la Caisse nationale des allocations familiales à l'INSEE ;

2° Effectuer un rapprochement des informations collectées avec les données concernant les allocations et prestations présentes dans les fichiers des départements.

III. - Afin d'effectuer le rapprochement des informations collectées par enquête avec les données administratives, le fichier contenant les données identifiantes est, pour garantir la confidentialité lors des transferts, chiffré avant transmission aux partenaires concernés. Pour le rapprochement avec les données de consommation médicale, ce chiffrement sera suivi par un double hachage du NIR.

IV. - A l'issue des rapprochements d'informations précités, les données directement identifiantes sont séparées des données rapprochées. La traçabilité des données directement identifiantes est assurée par la direction de la recherche, des études, de l'évaluation et des statistiques.

V. - Sont transmis à la direction de la recherche, des études, de l'évaluation et des statistiques des fichiers de données individuelles issues des opérations décrites au présent article ne permettant aucune identification directe des personnes auprès desquelles l'enquête a été réalisée. Ces fichiers contiennent les données indiquées aux 4°, 5°, 6°, 7°, 8° et 9° du I de l'article 3 ainsi que les données indiquées aux 2° et 3° du II du même article.

Article 6

Les données individuelles non directement identifiantes issues du traitement visé à l'article 1^{er} et mentionnées au V de l'article 5 sont soumises aux [dispositions de l'article L. 212-4 du code du patrimoine](#).

A l'issue du versement mentionné au premier alinéa, le service mentionné au V de l'article 5 procède sur les données qu'il détient mentionnées au même article à un traitement d'anonymisation complète.

Les mêmes données peuvent être communiquées à des fins de statistique publique ou de recherche scientifique ou historique sur décision de l'administration des archives prise après avis du comité du secret statistique dans les conditions prévues à l'[article 17 du décret du 20 mars 2009 susvisé](#).

Article 7

Les noms, prénoms, date et lieu de naissance, et NIR des personnes enquêtées ainsi que les noms, prénoms, date et lieu de naissance, et NIR des éventuels assurés sociaux dont les personnes enquêtées sont ayants droit sont conservés pendant cinq ans après la fin de la collecte d'informations afin de permettre les rapprochements d'information mentionnés à l'article 5.

Afin de suivre la mortalité des personnes concernées, le NIR de toutes les personnes enquêtées sera conservé à l'INSEE dans un fichier distinct et jusqu'à l'obtention de la mention du décès de la personne ou pendant quinze ans au plus. Un numéro d'ordre non significatif permettra de faire le lien entre ce fichier et les données issues des opérations décrites à l'article 5.

Les adresses des personnes enquêtées sont conservées pendant cinq ans après la fin de chacune des deux collectes d'informations mentionnées à l'article 2 afin de permettre la réalisation des enquêtes qualitatives mentionnées au troisième alinéa de l'article 2 et les rapprochements décrits à l'article 5.

Les nom, prénom et adresse des personnes mentionnées au point 10° de l'article 3 sont conservés pendant cinq ans après leur collecte afin de permettre la réalisation des enquêtes qualitatives mentionnées au troisième alinéa de l'article 2.

A l'issue des durées de conservation précitées, tous les intervenants du système (INSEE, GENES, CNAM-TS, CNAV, CCMSA) supprimeront les données ainsi que les tables de passage.

Article 8

Le droit d'accès et de rectification prévu par les articles 39 et 40 de la loi du 6 janvier 1978 modifiée susvisée s'exerce :

1° Auprès de la direction générale de l'INSEE ou des directions régionales ou interrégionales de l'INSEE pour les personnes enquêtées vivant à domicile ;

2° Auprès du prestataire sous la responsabilité de la direction de la recherche, des études, de l'évaluation et des statistiques pour les personnes enquêtées en établissement. Le service auprès duquel les personnes enquêtées pourront exercer leurs droits sera identifié sur la lettre-avis qui leur sera adressée.

Le droit d'opposition prévu à l'[article 38 de la loi du 6 janvier 1978 susvisée](#) ne s'applique pas au traitement mentionné à l'article 1er.

En application de l'[article 7 de la loi du 7 juin 1951 susvisée](#), les personnes concernées par la collecte mentionnée au premier alinéa de l'article 2 seront informées dans la lettre-avis qui leur sera adressée du caractère obligatoire de leur participation ; celles concernées par la collecte mentionnée au deuxième alinéa de l'article 2 seront informées dans la même lettre du caractère obligatoire ou non de leur participation.

Article 9

La direction de la recherche, des études, de l'évaluation et des statistiques garantit la sécurisation des transferts de données par des protocoles assurant la confidentialité des données ainsi que la restriction de l'accès aux données aux seules personnes habilitées.

Article 10

La ministre des affaires sociales, de la santé et des droits des femmes est chargée de l'exécution du présent décret, qui sera publié au Journal officiel de la République française.

Fait le 26 mars 2015.

Manuel Valls

Par le Premier ministre :

La ministre des affaires sociales, de la santé et des droits des femmes,

Marisol Touraine

Annexe 2. Liste des enquêtes mobilisables pour étudier la dépendance ou les aidants

- Baromètre de la DREES ;
- *Cohortes* : Paquid (INSERM), 3C (3Cités) (équipe de recherche de l'INSERM, du CNRS et de différentes Universités), Haute Normandie (INSERM) et AMI (Agrica, MSA et IFR99) ;
- Enquêtes annuelles sur les bénéficiaires et les dépenses de l'aide sociale départementale (DREES) ;
- Enquête « comportements de santé » (ou Baromètre santé) de l'INPES ;
- Enquêtes EHPA (DREES) ;
- Enquête emploi du temps (EdT) (INSEE) ;
- Enquêtes ES Handicap (DREES) ;
- Enquête permanente trimestrielle sur la prescription médicale (EPPM) de l'IRDES ;
- Enquête santé et itinéraire professionnel (SIP) (DREES-DARES) ;
- Enquête « Services de soins infirmiers à domicile » (SSIAD) de la DREES ;
- Enquête sur les conditions de vie des bénéficiaires de minima sociaux (BMS) (DREES) ;
- Enquêtes trimestrielles sur la PCH (DREES) ;
- Enquêtes trimestrielles sur l'APA (DREES) ;
- ESPS (Enquête Santé et Protection Sociale) de l'IRDES ;
- EU-SILC (European Union Statistics on Income and Living Conditions) ou Statistiques sur les ressources et les conditions de vie (SRCV) (INSEE) ;
- Enquête Handicap-Incapacités-Dépendance (HID) (INSEE) ;
- Enquêtes Handicap-Santé (HSM/HSI/HSA) (INSEE-DREES) ;
- PATHOS sur l'échantillon « ERNEST » (DREES) ;
- Remontées individuelles APA (DREES) ;
- SHARE (Survey of Health, Ageing and Retirement in Europe) (IRDES puis Université Paris-Dauphine).

Annexe 3. Liste des sigles utilisés

AAH : Allocation aux adultes handicapés
ACTP : Allocation compensatrice de tierce personne
AGGIR : Autonomie Gérontologie Groupes Iso-Ressources
ALD : Affection de longue durée
ALF : Allocation de logement familiale
ALS : Allocation de logement sociale
AM : Assurance maladie
APA : Allocation personnalisée d'autonomie
APL : Aide personnalisée au logement
ASI : Allocation supplémentaire d'invalidité
ASPA : Allocation de solidarité aux personnes âgées
BPE : Base permanente des équipements
BRPP : Base des répertoires des personnes physiques
CAF : Caisse d'allocations familiales
CAPI : Computer Assisted Personal Interviewing
CARE : Capacités, Aides et REssources des seniors
CASD : Centre d'accès sécurisé aux données
CCMSA : Caisse centrale de la mutualité sociale agricole
CD : Conseils Départementaux
CLCA : Complément de libre choix d'activité
CNAF : Caisse nationale d'allocations familiales
CNAM-TS : Caisse nationale de l'assurance maladie des travailleurs salariés
CNAV : Caisse nationale d'assurance vieillesse
CNIL : Commission nationale de l'informatique et des libertés
CNIS : Conseil national de l'information statistique
COLCA : Complément optionnel de libre choix d'activité
CPOM : Contrat pluriannuels d'objectif et de moyens
CRDS : Contribution pour le remboursement de la dette sociale
CSG : Contribution sociale généralisée
CSP : Catégorie socioprofessionnelle

DGFIP : Direction générale des Finances publiques

DOM : Département d'outre-mer

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

DSS : Direction de la sécurité sociale

EC : État civil

EHPAD : Établissement d'hébergement pour personnes âgées dépendantes

ERFS : Enquête revenus fiscaux et sociaux

EU-SILC : European Union Statistics on Income and Living Conditions

FINESS : Fichier national des établissements sanitaires et sociaux

FTP : File Transfer Protocol

GIR : Groupes iso-ressources

GnuPG : GNU Privacy Guard

HID : Handicaps-incapacités-dépendance

HSA : Handicap-Santé Aidants

HSI : Handicap-Santé Institutions

HSM : Handicap-Santé Ménages

INPES : Institut national de prévention et d'éducation pour la santé

INSEE : Institut national de la statistique et des études économiques

IR : Impôt sur le revenu

IRDES : Institut de recherche et documentation en économie de la santé

IRIS : Îlots regroupés pour l'information statistique

ISF : Impôt de solidarité sur les fortunes

MIOGA : My Internet Open Groupware Application

MSA : Mutualité sociale agricole

MTP : Majoration pour tierce personne

NIR : Numéro d'inscription au répertoire

PAJE : Prestation d'accueil du jeune enfant

PCH : Prestation de compensation du handicap

PMSI : Programme de médicalisation des systèmes d'information

RFD : Répertoire et fichiers démographiques

RFS : Revenus fiscaux et sociaux

RNIPP : Répertoire national d'identification des personnes physiques

RSA : Revenu de solidarité active

RSI : Régime social des indépendants

SHARE : Survey on Health, Ageing and Retirement in Europe

SILC-SRCV : Statistics on Income and Living Conditions - statistique sur les ressources et conditions de vie

SIN : Service informatique national

SRCV : Statistique sur les ressources et conditions de vie

SNIR-AM : Système national d'information inter-régimes de l'Assurance maladie

TH : Taxe d'habitation

VQS : Vie Quotidienne et Santé

DOCUMENT DE TRAVAIL

ISSN : 1621-4358

Reproduction autorisée sous réserve de la mention des sources
