

Recensement des données sur le vieillissement

Note méthodologique

Table des matières

| | | |
|------|---|---|
| I. | Contexte | 1 |
| II. | Périmètre du recensement et identification des sources | 2 |
| III. | Repérage de la population des personnes âgées | 2 |
| IV. | Champ des métadonnées | 3 |
| V. | Difficultés rencontrées | 5 |
| VI. | Présentation du document et des différents onglets | 6 |
| VII. | Annexes : | 7 |
| | A. Modèle de métadonnées utilisé par le recensement de la Drees : | 7 |
| | B. Descriptions des indicateurs | 8 |

I. Contexte

Le Conseil Scientifique du PPR Autonomie a déterminé l'opportunité de réaliser un recensement et une cartographie des données disponibles et portant sur l'autonomie de la personne (les situations de handicap et le vieillissement). Le travail ici présenté a pour objectif de faciliter l'identification et la réutilisation de données disponibles à des fins d'observation de la population des personnes âgées.

Il vient compléter [la cartographie de la Drees](#). Publiée initialement en juin 2022, celle-ci recense près de 140 bases de données (statistique publique, remontées administratives, enquêtes d'organismes de recherches) documentées à partir de 27 variables. Il s'agit de données collectées à des fins d'observations de la population en situation de handicap via des enquêtes spécifiques ou non. Les enquêtes non-spécifiques doivent contenir une possibilité d'identifier une sous-population concernée par handicap¹.

La version publiée par le PPR Autonomie est une version temporaire. D'abord, elle vise à être mise à jour régulièrement. De plus, le PPR étudie l'opportunité à développer un modèle de métadonnées plus complet et en accord avec des formats internationaux. Ceci permettrait d'améliorer l'interopérabilité de l'outil et de s'inscrire dans le mouvement FAIR (facile à trouver, accessible, interopérable et réutilisable).

¹ Dans le fichier produit par le PPR, un onglet est consacré aux données recensées par la Drees qui peuvent aussi renvoyer au vieillissement et à la perte d'autonomie.

II. Périmètre du recensement et identification des sources

Seules les données administratives et les données d'enquêtes de la statistique publique mises à disposition et dont la documentation est publique sont ici recensées, exception faite des très grandes cohortes (CONSTANCE, GAZEL).

31 enquêtes ont été identifiées à partir d'une fouille des catalogues du CASD (Centre d'accès sécurisés aux données), de celui de l'ADISP (data.progedo). Pour chaque jeu de données présent dans ces catalogues, nous avons appliqué une série de critères :

1. Les données doivent avoir été collectées après l'année 2000 et sur tout ou parties du territoire Français.
2. Les données doivent renvoyer au moins pour partie à des individus, et non des structures.
3. Les données ne doivent pas avoir été recensées par la Drees dans son travail de cartographie.
4. Les données doivent, soit porter directement sur des enjeux de vieillissement, soit s'intéresser à d'autres thématiques, mais comporter des indicateurs permettant d'identifier les personnes en situation de handicap ou de perte d'autonomie.

Pour certains jeux de données, les métadonnées ne permettaient pas de déterminer la présence ou non de variables permettant d'approcher le handicap ou la perte d'autonomie. L'absence d'un dictionnaire des variables ou d'un questionnaire rendait cette recherche impossible. Dans ce cas, les jeux de données n'ont pas été ajoutés au recensement.

Le recensement a consisté à réunir pour chaque jeu de données l'ensemble des informations disponibles dans les entrepôts de données puis à rechercher sur internet les documentations issues des producteurs. Nous avons ensuite fait coïncider ces informations avec un modèle de métadonnées issu du modèle suivi par la Drees et augmenté de 12 nouveaux champs.

Nous avons fait le choix de reprendre et de citer au maximum la documentation déjà existante pour documenter les jeux de données. Dans ce cadre, nous avons reporté les métadonnées avec des guillemets et en inscrivant entre parenthèses la structure autrice des métadonnées. Par conséquent, s'il nous a semblé que l'information ne se répétait pas entre deux sources, nous avons cité les différentes sources et informations dans un même champ de métadonnées.

III. Repérage de la population des personnes âgées

Pour l'étude du vieillissement, la perte d'autonomie peut être déterminée par les mêmes indicateurs que la situation de handicap : le Gali, les questions du Washington Group ou le mini-module européen. Ceux-ci rapportent les limitations auto-déclarées que vivent le ou la répondante (un descriptif des indicateurs est disponible en annexe). Le **GIR (pour Groupe Iso-Ressource)** est un indicateur spécifique pour la perte d'autonomie du fait du vieillissement, il permet de rapporter différents niveaux de perte d'autonomie d'une personne âgée, et sert tant d'indicateur administratif –comme condition d'accès à certaines prestations sociales – que de variable de description au sein des enquêtes, par exemple au sein de l'enquête EHPA. Le GIR est défini par un médecin lors d'une évaluation médicale et se base sur la grille AGGIR. Celle-ci définit 10 activités discriminantes :

- Cohérence : la personne âgée est-elle capable de converser et/ou de se comporter de façon logique et sensée ?
- Orientation : la personne se repère-t-elle dans le temps, les moments de la journée et dans les lieux ?
- Toilette : quelle est sa capacité à assurer son hygiène corporelle du haut et du bas du corps ?
- Habillage : la personne peut-elle s'habiller, se déshabiller, choisir elle-même ses vêtements ?

- Alimentation : la personne peut-elle se servir les aliments préparés et les manger seule ?
- Hygiène de l'élimination : la personne souffre-t-elle d'incontinence urinaire ou fécale ?
- Transferts : quelle est sa capacité à se lever, se coucher et s'asseoir ?
- Déplacements à l'intérieur du logement ou de l'institution : la personne peut-elle se déplacer à l'intérieur, éventuellement aidée par une canne, un déambulateur ou un fauteuil roulant ?
- Déplacements à l'extérieur : la personne est-elle capable de se déplacer à l'extérieur, à partir de la porte d'entrée ?
- Communication à distance : quelle est sa capacité à utiliser les moyens de communication (téléphone, alarme, sonnette) ?

Néanmoins, de nombreux autres indicateurs peuvent aussi permettre d'identifier les limitations ou la perte d'autonomie. Les cohortes (Gazel ou Constance) peuvent accueillir d'autres indicateurs de limitations (comme le Medical Outcomes Study Short Form 36, aussi nommé SF-36). De plus, certaines questions, sans être standardisées, interrogent directement les situations de handicap ou de pertes d'autonomie, à l'image de la question "Etes-vous handicapé(e) ou avez-vous simplement quelques gênes ou difficultés dans la vie quotidienne ?" présente dans l'enquête Budget de Famille.

IV. Champ des métadonnées

Se focalisant sur la perte d'autonomie dû au vieillissement, nous avons pris la décision d'élargir les métadonnées recensées de 12 nouveaux champs.

| Champs de métadonnées du recensement PPR | Modalités | Description |
|--|---|---|
| Nom de l'enquête | Texte libre | |
| Acronyme | Texte libre | L'acronyme a été construit à partir du nom de l'enquête ainsi que de l'année de production. L'enjeu était de trouver une dénomination plus courte que le nom mais qui permettait d'être unique pour l'enquête et évocatrice. |
| N° de vague | Chiffre | |
| Types de données (d'enquêtes / administratives) | Données d'enquêtes Remontés administratives Données de cohortes | |
| Millésime(s) | Dates en années | Années à laquelle/auxquelles l'enquête, ou l'édition de l'enquête a été publiée |
| Producteur(s) | Texte libre | Institution(s) ayant produit l'enquête ou mené la fusion des bases de données initiales |
| Périodicité | Fréquence en texte libre | |
| Objectifs principaux | Texte libre | |
| Thématiques abordées | Texte libre | Thématiques issues des métadonnées de Progedo ou de l'INSEE. Les thématiques issues de |

| | | |
|---|---|--|
| | | PROGEDO renvoie aux "mots-clefs" de la plateforme. |
| Panel | Oui Non NSP | |
| Date de recueil / collecte | Dates Texte libre | |
| Champ géographique | France métropolitaine DOM Texte libre | |
| Méthode d'échantillonnage | Texte libre | |
| Méthode de collecte | Texte libre CATI CAWI CAPI | |
| Nombre d'unité | Texte libre | |
| Unité d'observation | Individu Ménage Etablissement Foyer Prestation texte libre | |
| Population Observée | Texte libre | Précision sur la population visée par l'étude. |
| Présence d'une population de moins de 18 ans | Oui Non Peut-être NSP | Est-ce que le jeu de données contient des personnes qui avaient moins de 18 ans lors de leur interrogation ? |
| Présence d'une population de plus de 65 ans | Oui Non Peut-être NSP | Est-ce que le jeu de données contient des personnes qui avaient plus de 65 ans lors de leur interrogation ? |
| Modes de repérage des populations âgées | Texte libre | |
| GALI | Oui Non | |
| Mini module européen | Oui Non | |
| Questions du Washington Group | Oui Non | |
| Module « groupe iso-ressources » (GIR) | Oui Non | |
| Reconnaissance administrative de la situation de handicap ou de la perte d'autonomie | Texte libre | |
| Indicateurs de la perte d'autonomie | Texte libre | |
| Autres indicateurs | Texte libre | |
| Autres questions et variables liées au handicap | Texte Libre | Qui ne sont pas des indicateurs standardisés. |

| | | |
|---|--------------------------|---|
| Information sur l'environnement physique des personnes | Texte libre | |
| Taux de réponse | Nombre et texte libre | |
| Proportion de personnes en situation de handicap | Nombre et texte libre | Soit donné en clair dans l'explorateur des variables de Progedo, soit calculé à partir des informations disponibles dans l'explorateur. |
| Proportion des personnes en grand âge | Nombre et texte libre | Soit donné en clair dans l'explorateur des variables de Progedo, soit calculé à partir des informations disponibles dans l'explorateur. |
| Appariement(s) effectués | Texte libre | |
| Entrepôt de données | CASD Progedo Autre | |
| Liens pour accéder à l'entrepôt | Lien et texte libre | |
| Autres pages internet disponibles pour les données | Liens | |
| Problèmes de documentations | Texte libre | |
| Divers | Texte libre | |
| DOI | DOI | |

V. Difficultés rencontrées

Au cours du recensement des données, des difficultés ont pu apparaître :

- La documentation des jeux de données est éclatée entre divers supports. Les données de la statistique publique peuvent être mises à disposition sous trois formats différents : en open access, sous des formats anonymisés dans les portails de la statistique publique (data.gouv et data.drees) ; en format fortement pseudonymisés sur data.progedo ; en format faiblement pseudonymisé et parfois sans pseudonymisation via le CASD. Enfin, les producteurs de données peuvent consacrer une page de leur site internet à décrire l'enquête et leurs données, par exemple pour [l'enquête auprès des établissements d'hébergement pour personnes âgées EHPA](#). Par conséquent, il a été choisi de répertorier et de puiser autant que possible dans les différentes ressources disponibles.
- Certaines documentations fournies par l'entrepôt différaient de celles exposées par le producteur. La décision a été prise de privilégier les métadonnées produites par le producteur.
- Dans le cas des jeux de données les plus anciens, certaines informations manquaient, certains liens vers des pages de documentations étaient des liens-morts, et certains fichiers de documentations (par exemple le dictionnaire des variables) sont aujourd'hui inaccessibles. Dans ce cadre, si l'information manquante n'empêchait pas de sélectionner le jeu de données (procédure évoquée dans la partie II), le champ correspondant porte la marque « NSP » pour « Ne Sait Pas » ou est laissé inoccupé.
- Les modèles de métadonnées utilisés par les différents portails ou sites de producteurs de données ne correspondent pas toujours entre eux. Il a ainsi fallu procéder à des fusions de champs différents. La catégorie du recensement « Thématiques abordées » peut correspondre aux thèmes évoqués dans le champ « Description des données » dans la documentation de l'INSEE. Elle évoque les différents sujets de recherches pouvant être saisi via le jeu de données ou bien la manière dont a été

construite l'enquête ou découpée le questionnaire². Cependant, la catégorie « Thèmes » du portail data.progedo est très sommaire et ne comporte que 8 modalités différentes. De fait, il nous a semblé nécessaire de conserver les thèmes évoqués par le producteur des données, et à défaut de sélectionner non le thème, mais les « mots-clefs » de la plateforme data.progedo. Ceux-ci semblent plus pertinents pour décrire les sujets abordés par une enquête.

- Certains jeux de données disponibles peuvent être interprétés comme des compilations de données issues de plusieurs sources, par exemple sur des bases géographiques. Il a été décidé de ne pas les inclure et de privilégier les jeux de données originaux.
- L'explorateur des variables présent pour certains jeux de données sur la plateforme data.progedo permet d'afficher le tri à plat des variables présentes dans le jeu de données. Cependant, afin de ne pas attenter à l'anonymat des informations affichées, il ne permet pas de croiser des variables. Par conséquent seul le GALI et l'âge, qui ne nécessitent qu'une seule question, ont été mobilisés pour approcher la part de personnes en situation de perte d'autonomie.

VI. Présentation du document et des différents onglets

Le recensement se présente par le biais d'un classeur Excel et possède 3 onglets différents. Le premier expose le dictionnaire des variables utilisé pour décrire les jeux de données. Le second est le tableau du recensement. Le troisième onglet présente une sélection des jeux de données recensés par la Drees qui peuvent servir à l'étude de la perte d'autonomie due au vieillissement. Ces jeux de données doivent pouvoir exposer une population de plus de 65 ans.

² Ainsi l'enquête Emploi Continue de 2021 est-elle structurée en différents thèmes : « la description du ménage et des personnes qui le composent », « le statut d'activité au sens du BIT » par exemple.

VII. Annexes :

A. Modèle de métadonnées utilisé par le recensement de la Drees :

| Modalités DREES | Modalités |
|---|--|
| Nom de l'enquête | Texte libre |
| Types de données (d'enquêtes / administratives) | Données d'enquêtes / Données Administratives / Données d'enquêtes et administratives |
| Date d'observation - millésime(s) | date et texte libre |
| Date de recueil / collecte | date et texte libre |
| Producteur(s) | Texte libre |
| Périodicité | dates ou fréquences |
| Panel | Oui/Non + texte libre |
| Taille (nbre d'unités) | Nombre avec dates / Texte libre |
| Échantillonnage (si enquête) | Texte libre |
| Unité d'observation | Individu / ménage / Allocataires / Etablissement / Contrats / Affaires ou encore Texte Libre |
| Population statistique | Texte libre |
| Moins de 18 ans (O/N) | Oui / Non + Texte libre |
| Plus de 65 ans (O/N) | Oui / Non + Texte libre |
| Méthode de collecte | Texte libre + CATI /CAWI / CAPI |
| Champ géographique | France métropolitaine / DOM / Texte libre |
| Objectifs principaux | Texte Libre |
| Thématiques abordées | Texte Libre |
| Questions / variables liées au handicap | Texte Libre |
| GALI | Oui / Non + Texte libre |
| Mini module européen | Oui / Non + Texte libre |
| Questions du WG | Oui / Non + Texte libre |
| Reconnaissance administrative du handicap | Texte Libre / Aucune |
| Autres indicateurs | Texte libre |
| Appariement(s) | Texte libre |

| | |
|--------------------------------|--------------------|
| Diffusion / Conditions d'accès | Lien + texte libre |
| Sites internet | Lien + texte libre |
| Divers | Texte libre |

B. Descriptions des indicateurs

Le *Gali*, pour *Global Activity Limitation Indicator* consiste en une unique question :

- "Êtes-vous limité(e), depuis au moins six mois, à cause d'un problème de santé, dans les activités que les gens font habituellement ?" Oui, fortement limité(e) / Oui, limité(e), mais pas fortement / Non, pas limité(e) du tout. »

Le *Mini-module Européen* est composé de trois questions dont le *Gali* :

- "Comment est votre état de santé en général ?" Très bon/ Bon / Assez bon / Mauvais / Très mauvais
- "Avez-vous [X] a-t-il(elle)] une maladie ou un problème de santé qui soit chronique ou de caractère durable ?" Oui / Non
- " Êtes-vous limité(e), depuis au moins six mois, à cause d'un problème de santé, dans les activités que les gens font habituellement ?" Oui, fortement limité(e) / Oui, limité(e), mais pas fortement / Non, pas limité(e) du tout

Le *Washington Group on Disability Statistics*, comprend six questions sur les limitations :

- "Avez-vous des problèmes de vue, même avec vos lunettes ? Diriez-vous..." Pas du tout/ Un peu/ Beaucoup/ Je ne vois rien du tout/ Refuse de répondre/ Ne sait pas
- "Éprouvez-vous des difficultés à entendre, même avec une ou des prothèse(s) auditive(s) ? Diriez-vous..." Pas du tout/ Un peu/ Beaucoup/ Je ne vois rien du tout/ Refuse de répondre/ Ne sait pas
- "Éprouvez-vous des difficultés à marcher ou à monter les escaliers ? " Pas du tout/ Un peu/ Beaucoup/ Je ne vois rien du tout/ Refuse de répondre/ Ne sait pas
- "Éprouvez-vous des difficultés à vous rappeler certaines choses ou à vous concentrer ? Diriez-vous..." Pas du tout/ Un peu/ Beaucoup/ Je ne vois rien du tout/ Refuse de répondre/ Ne sait pas
- "Éprouvez-vous des difficultés à communiquer dans votre langue habituelle, à comprendre les autres ou à vous faire comprendre, par exemple ? " Pas du tout/ Un peu/ Beaucoup/ Je ne vois rien du tout/ Refuse de répondre/ Ne sait pas
- Éprouvez-vous des difficultés à prendre soin de vous, à vous laver ou à vous habiller, par exemple ? Pas du tout/ Un peu/ Beaucoup/ Je ne vois rien du tout/ Refuse de répondre/ Ne sait pas