

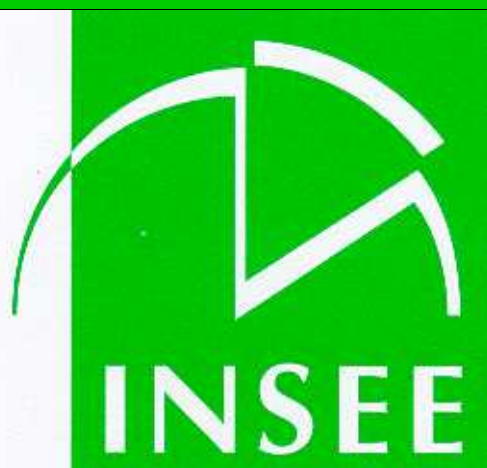
**Direction des Statistiques Démographiques et  
Sociales**

**F 1102**

**Le volet Institution de l'enquête  
Handicap-Santé :  
Présentation, calcul des poids**

Gérard Bouvier  
Insee, division Enquêtes et Etudes Démographiques

**Document de travail**



Institut National de la Statistique et des Etudes Economiques

**INSTITUT NATIONAL DE LA STATISTIQUE ET DES ÉTUDES ÉCONOMIQUES**

Série des Documents de Travail  
de la  
DIRECTION DES STATISTIQUES DÉMOGRAPHIQUES ET SOCIALES  
UNITES DES ÉTUDES DÉMOGRAPHIQUES ET SOCIALES  
DIVISION ENQUÊTES ET ÉTUDES DÉMOGRAPHIQUES

**NF1102**

**Le volet Institution de l'enquête Handicap-Santé :  
Présentation, calcul des poids.**

Gérard Bouvier  
Insee, division Enquêtes et Etudes Démographiques

Février 2011

Ces documents de travail ne reflètent pas la position de l'INSEE et n'engagent que leurs auteurs  
Working-papers do not reflect the position of INSEE but only their authors'views.

## Résumé

### **(Redressements et calages pour l'enquête Handicap-Santé, volet institution)**

L'enquête Handicap-Santé, volet institution, a été collectée auprès de personnes vivant en institutions en 2009. 9 104 personnes, vivant dans 1 519 institutions ont répondu à cette grande enquête structurelle comprenant de nombreuses questions sur l'état fonctionnel (maladies, déficiences, limitations), les restrictions de participation sociale (accès au travail, à la formation, aux loisirs, éléments de niveaux de vie) et les facteurs environnementaux (entourage familial ou social, aides techniques, aménagements du logement, accessibilité aux lieux de vie, discriminations ressenties). Ce volet complète un volet ménage collecté en 2008. Les institutions du champ comprennent des établissements pour personnes âgées, pour adultes handicapés, pour personnes en difficultés psychiques, ainsi que des centres d'hébergement pour réinsertion sociale. L'enquête est nationale (métropole et départements d'outre-mer).

Le document décrit la base de sondage, l'échantillonnage (à deux degrés), les traitements de non-réponse, enfin le calage permettant d'aboutir aux poids des personnes enquêtées. La base de sondage est essentiellement le fichier d'établissements de santé (FINESS), avec environ 12 000 institutions dans le champ. La stratification est décrite, puis les étapes d'une pré-enquête ciblant environ 2 200 institutions acceptant de participer à l'enquête, soit 94 % des institutions contactées. L'étape suivante consiste en un nouveau tirage d'institutions (parmi les 2200), puis, au sein de chaque institution, en un tirage d'individus. Le principe est de tirer les institutions proportionnellement à leur capacité puis un nombre fixe d'individus par institution, ce qui assure en théorie une équipondération de ceux-ci, au moins par strate. Les écarts à cette situation idéale sont décrits en lien avec les contraintes qui en sont à l'origine. A chaque tirage correspond un redressement pour non-réponse, intégrant les données utiles (strate, localisation, capacité, caractéristiques individuelles telles que genre, âge, présence d'un tuteur). Le taux global de réponse est de 97 % pour les institutions et de 91 % pour les individus. La contribution de chacun des traitements (échantillonnages et redressements) à la dispersion des poids est précisée. Le volet présenté ne concerne, en terme d'individus que la population durablement hébergée en institution. On calcule un poids pour chaque institution, puis, sachant son effectif en personnes durablement hébergées, on en déduit les effectifs de calage, étape finale des traitements.

## Summary

### **(Non-participation correction and adjustments of weights for the Institution section of the Health and Disability survey)**

The Institution section of the Health and Disability survey was collected among people living in institutions in 2009. 9 104 persons, living in 1 519 institutions, responded to this major structural survey, which includes many questions about respondents functional state (diseases, deficiencies, limitations), social participation restrictions (access to the labour market, educational attainments or leisure, aspects of standards of living), environmental factors (like familial or social network, assistive technologies, home fittings, accessibility, discriminations (self-declared)). This section is the complement of a household section, collected in 2008. The institutions in the field of the survey are institutions for the elderly, disabled adults, psychologically impaired adults and also shelter centres (for social reinsertion). The survey was collected in both the mainland France and overseas "départements". This document deals with the sampling frame (which is in two steps), non-participation corrections and adjustments leading to the computation of weights for the survey sample. The sampling frame is basically the file of health institutions (FINESS) with about 12 000 institutions in the field. Stratification is outlined first, secondly the steps of a preliminary survey which aims to target about 2 200 institutions agreeing to respond to the survey, and afterwards, we got in touch with 94 % of the targeted institutions. The next step is consisting of a new sampling of institutions and for each institution of a sampling of people. The principle is to sample an institution proportionally to its capacity and to sample for each institution a fixed number of people. Theoretically, this is securing equal weights for individuals, inside each stratum. The gap between the ideal situation and the results are described in relation with constraints causing this gap. Each sampling is followed by a non-participation correction based on data (stratum, geography, capacity, individuals data like sex, age and existence of legal tutors). The contribution from each sampling or correction to the weights dispersion is detailed. This section only deals with the population of individuals hosted for a long stay. The weights are computed for each institution, and afterwards, knowing the number of "long stay" hosted people, the elements for the calibration can be deduced, this calibration being the final step of the process.

# Le volet Institution de l'enquête Handicap-Santé :

## Présentation, calcul des poids.

Partie I : Vue d'ensemble .....	5
Partie II : Présentation de l'échantillon du volet institution .....	7
II-A : Contexte, éléments du champ .....	7
II-B : Adaptations au contexte des institutions.....	8
Partie III : Principes généraux de l'échantillonnage .....	8
III-A : L'échantillon, au vu des objectifs du volet institution .....	9
III-B : L'échantillon, tel que présenté au comité du label.....	9
III-C : Ajustements : la définition de l'échantillon (fin 2008-début 2009). .....	10
Partie IV : La pré-enquête : tirage et réponse .....	11
IV-A : Principes de tirage .....	11
IV-B : Tirage en France métropolitaine.....	12
IV-C : Tirage en départements d'outre-mer.....	14
IV-D : La pré-enquête : éléments sur la non-réponse .....	15
Partie V : L'échantillon pour l'enquête et le tirage (institution) .....	15
V-A : Les changements de strates.....	15
V-B : L'échantillon .....	16
VI : La non-réponse institution (au moment de l'enquête) .....	17
VI-A : Définition d'une institution répondante .....	17
VI-B : Analyse de la non-réponse .....	17
VI-C : Conclusion sur les poids des institutions.....	18
VII : Traitements au niveau individu .....	19
VII-A : Le tirage des individus .....	19
VII-B : La non-réponse individu .....	19
VII-C : Le calage .....	20
Annexe 1 : Définition détaillée des strates .....	22
Annexe 2 : Relation entre capacité et effectif de personnes hébergées durablement. ....	23
Annexe 3 : Dispersion du tirage de la pré-enquête.....	25
Annexe 4 : Dispersion due à l'écart entre « H » et « C » .....	27
Annexe 5 : Tableau des quantiles, poids par strates .....	29
Annexe 6 : Comparaison entre cible et résultats de collecte .....	29

## Partie I : Vue d'ensemble

Ce document résume les travaux effectués pour établir les poids des individus interrogés au titre du volet institution de l'enquête Handicap Santé<sup>1</sup>. Ces poids permettent l'exploitation du volet de façon autonome. En effet, l'échantillon d'institutions est réparti afin d'obtenir une taille d'échantillon suffisante pour certaines catégories d'intérêt, pour lesquelles on souhaite disposer à des fins d'études d'un échantillon plus important que la part qu'elles représentent dans l'ensemble des places en institutions (exemple : personnes en établissements psychiatriques). Il s'agit là de l'objectif secondaire<sup>2</sup> de l'enquête. Il convient de préciser le champ de ce volet. L'objectif principal de l'enquête Handicap Santé est de pouvoir faire des études sur l'ensemble de la population française en exploitant les deux volets, ménages et institutions. Les champs de ceux-ci doivent donc être disjoints et le volet institution ne concerne que les personnes hébergées durablement.

Une partie introductive ([partie II](#)) fournit une courte présentation de l'enquête. La suite du document est consacrée à la description des étapes de l'échantillonnage et du calcul des poids. Dans une première [partie III](#), on présente les principes généraux de l'échantillonnage. Puis les parties suivantes détaillent le tirage et le calcul des poids des trois étapes du volet institution : [pré-enquête](#) (sélection d'institutions dans la base de sondage), [enquête](#) phase institution puis [enquête](#) phase individus.

Le principe général de calcul des poids est de déterminer un poids de tirage, de corriger ce poids en fonction de la non-réponse, et enfin de caler sur une population de référence. En première approche, le tirage se déroule en trois temps : tirage d'institutions dans une base de sondage ([pré-enquête](#)), deuxième tirage d'institutions parmi celles ayant accepté l'enquête ([enquête](#) phase [institution](#)), puis, tirage d'individus ([enquête](#) phase [individu](#)).

Ce principe et cette approche conduisent à un travail de pondération que l'on détaillera de la façon suivante :

A partir d'une base de sondage comportant en principe de façon exhaustive les institutions du champ retenu, on tire d'abord (avec un taux d'environ 1/5) des institutions ([voir partie IV](#)). Une [pré-enquête](#) est effectuée<sup>3</sup>, permettant de recueillir des informations, confirmant ou précisant celles qui figurent dans la base de sondage et des références pour les personnes à contacter dans l'institution. En principe, une institution qui accepte l'enquête lors de cette phase devrait l'accepter dans la phase « individu » : la pré-enquête a aussi pour but de limiter la non-réponse, et donc d'échantillonner au plus juste lors de la phase individu. Les institutions dans le champ ont accepté l'enquête lors de cette phase avec un taux de 99 %<sup>4</sup>. La pré-enquête a également permis de repérer des hors-champ, environ 2% du total tiré. L'ensemble des institutions « acceptantes » est celui dans lequel le tirage ([enquête](#) phase [institution](#)), est effectué, ce qui est présenté en [partie V](#). Après celui-ci, 1 567 institutions sont retenues pour l'enquête<sup>5</sup>, et ce qui devait permettre d'interroger 10 020 personnes. Toutefois, pour permettre un ajustement en cours de collecte, en fonction de la non-réponse individuelle, il avait été prévu un échantillon complémentaire. En pratique, on a donc tiré un « gros-échantillon » puis dans celui-ci l'échantillon proprement dit, ce qui reste étant la réserve. Celle-ci n'a pas été utilisée. Une non-réponse institution est observée durant cette phase : 3 % de non-réponse (y compris hors-champ<sup>4</sup>), mais ceux-ci sont surtout dans les strates d'établissements psychiatriques. Les données et le traitement de cette non-réponse sont détaillés [partie VI](#). L'ensemble des poids calculés permet de déterminer un poids pour chaque institution répondante.

---

<sup>1</sup> Une présentation générale de l'enquête Handicap-Santé est disponible aux e-adresses : <http://www.insee.fr/fr/methodes/default.asp?page=sources/ope-enq-handicap-sante-menages-hsm.htm> ;

[http://www.cnis.fr/ind\\_enquetes.htm](http://www.cnis.fr/ind_enquetes.htm) ; <http://www.sante-sports.gouv.fr/handicap-sante.html> ;

Ce dernier site présentant la documentation la plus complète sur l'ensemble du projet. L'enquête Handicap-Santé comporte un volet ménage et un volet institution. Un document analogue au présent document décrit les éléments relatifs au calcul des poids pour le volet ménage. ([http://www.insee.fr/fr/publications-et-services/docs\\_doc\\_travail/Docf1101.pdf](http://www.insee.fr/fr/publications-et-services/docs_doc_travail/Docf1101.pdf)).

Une note précisera les éléments relatifs aux poids permettant l'exploitation conjointe des deux volets.

<sup>2</sup> L'objectif principal est d'obtenir de données représentatives des personnes vivant en France, que ce soit en ménage ou en institution.

<sup>3</sup> Elle a été faite au printemps 2009.

<sup>4</sup> Le taux de hors champ-institution est de 2 % pour la pré-enquête, un peu inférieur pour l'enquête.

<sup>5</sup> Cette phase a été menée à l'automne 2009.

La **phase individu** est exposée en **partie VII**. Elle consiste en un tirage au sort de N (6,8 ou 9) individus par institution. Ce tirage est effectué, par institution, dans une liste de personnes hébergées durablement. Sur les 10 020 individus échantillonnés, on obtient 9104 « répondants ». La non-réponse varie suivant les strates et ses possibles déterminants sont analysés et intégrés si nécessaire. Il ne reste plus qu'à effectuer le calage annoncé. Connaissant le poids d'une institution et l'effectif des personnes hébergées durablement dans celle-ci, on en déduit les effectifs de ces populations par strate, ce qui constituera la population de calage. Des annexes précisant certains points complètent le document.

Pour visualiser :

**Tableau 1 : vue d'ensemble**

Base de sondage (Établissements)	<b>FINESS</b> (avec mises à jour : compléments en dom, étab psy)		
Stratification	Strate 1	.....	Strate 14
Tirage			
Pré-enquête	Échantillon PE-1	.....	Échantillon PE-14
Pré-enquête			
	Répondants PE-1	.....	Répondants PE-14
	<p>« Nettoyage » : exclusion des hors-champ, contrôle et correction des strates traitement de la non-réponse (refus) : constitution de : <b>Échantillon (d'institutions acceptant l') Enquête</b></p>		
Stratification	Strate 1	.....	Strate 14
Tirage Enquête	Échantillon E-1	.....	Échantillon E-14
Enquête	Répondants E-1	.....	Répondants E-14
	Traitement de la non-réponse institution par strate Une institution répondante est une institution qui donne la liste de personnes hébergées durablement (PHD)		
Pondération	Poids institutions Wist-1	.....	Poids institutions Wist-14
Échantillons individus	Liste des PHD-1	.....	Liste des PHD-14
	Avec les poids institutions : effectifs par strate de PHD (population de calage)		
Tirage Individus	Échantillon I-1	.....	Échantillon I-14
Enquête Individus	Répondants I-1	.....	Répondants I-14
	Traitement de la non-réponse individu par strate		
Pondération	Poids individus W-1 (nc <sup>6</sup> )	.....	Poids individus W-14 (nc)
Calage	Poids individus W-1		Poids individus W-14

<sup>6</sup> Non calés

## Partie II : Présentation de l'échantillon du volet institution

### II-A : CONTEXTE, ÉLÉMENTS DU CHAMP

En 1998-1999, l'enquête Handicap Incapacité Dépendance (HID) a permis, pour la première fois, d'évaluer les problèmes de handicap et de dépendance, ainsi que les difficultés rencontrées dans leur vie quotidienne par les personnes concernées. Cette enquête a été un grand succès (plus de 180 publications). L'enquête Handicap Santé a pour but de renouveler l'opération, en l'adaptant au nouveau contexte institutionnel et en tenant compte des enseignements de l'enquête HID.

Comme celle-ci, l'enquête Handicap-Santé comprend un volet « Ménages » et un volet « Institutions ». Le volet ménage est une enquête ayant eu lieu au 2<sup>ème</sup> trimestre 2008. Elle a permis de recueillir les informations pour un peu moins de 30 000 individus. Le volet institution a permis de recueillir les informations pour un peu plus de 9 000 individus, l'enquête ayant eu lieu au 4<sup>ème</sup> trimestre 2009.

L'objectif premier de l'enquête Handicap-Santé est de fournir des informations (d'abord sur les situations de handicap ou de dépendance, voir infra pour les thèmes de l'enquête), pour la population vivant en France, dans son ensemble. Toutefois, la chronologie des opérations statistiques et le volume des informations obtenues justifient des traitements séparés par volet, permettant donc d'une part des analyses pour la population vivant en ménages ordinaires, d'autre part des analyses pour la population vivant dans les types d'institutions enquêtées dans le volet éponyme.

Il y aura donc<sup>7</sup> :

- 1) un traitement des données du volet ménage seul ;
- 2) un traitement des données du volet institution seul.

**Le document présent est consacré aux travaux relatifs au volet institution seul**

Les répondants se répartissent par catégories d'institutions<sup>8</sup> :

	Répondants
<b>Établissements Hospitaliers pour Personnes Âgées Dépendantes</b>	2 597
<b>Maisons de retraite</b>	425
<b>Unités de Séjours de Longue Durée</b>	654
<b>Maisons d'accueil spécialisées, Foyers pour Adultes Handicapés</b>	1 475
<b>Autres établissements pour adultes handicapés</b>	1 487
<b>Établissements et unités psychiatriques</b>	1 435
<b>Centres d'Hébergement pour réinsertion sociale</b>	1 031
<b>Ensemble du volet institution</b>	<b>9 104</b>

*Champ : France métropolitaine et départements d'outre-mer.*

Les trois premières catégories regroupent des établissements pour personnes âgées, les deux suivantes, pour des personnes handicapées (de 20 à 60 ans). La dernière catégorie est celle des centres d'hébergements pour réinsertion sociale. Les catégories et strates sont détaillées plus loin. Tous ces établissements sont répertoriés dans un fichier (fichier FINISS), voir également plus loin.

L'enquête a été réalisée en partenariat par l'INSEE et la DREES, avec l'appui de plusieurs partenaires : l'Association nationale pour la gestion du fonds pour l'Insertion professionnelle des personnes handicapées (AGEFIPH), la Caisse nationale d'allocations familiales (CNAF), la Caisse d'assurance maladie des travailleurs salariés (CNAMTS), la Caisse nationale d'assurance vieillesse

<sup>7</sup> Apurement et pondérations du volet ménage menés d'octobre 2008 à septembre 2009, apurement et pondérations du volet institution menés de février 2010 à octobre 2010, travaux d'ajustements pour exploitation conjointe des deux volets effectués de septembre 2010 à mars 2011.

<sup>8</sup> Voir [annexe 1](#) pour le détail des strates (les 7 catégories d'institutions du tableau se divisent en 14 strates). Le nombre de répondants exact est 9 104, mais quelques tableaux ont été faits avec le chiffre provisoire de 9107.

(CNAV), la Caisse nationale de solidarité pour l'autonomie (CNSA), la Fédération nationale des mutuelles françaises (FNMF) et l'Institut de veille sanitaire (InVS).

## **II-B : ADAPTATIONS AU CONTEXTE DES INSTITUTIONS**

Le volet « Institution » de l'enquête Handicap Santé est proche de celui du volet « Ménages ». Quelques modifications ont cependant du être apportées, qui tenaient compte de contraintes de protocole et des spécificités de certaines personnes à enquêter.

L'approche d'une institution est en effet un peu différente de celle d'un ménage.

Il y a donc eu une pré-enquête qui devait essentiellement :

- 1) Tenter de diminuer les taux de refus mais aussi
- 2) S'assurer de l'exactitude (type d'institutions) des informations fournies par le fichier FINESS.
- 3) Obtenir un contact dans l'institution, notamment pour
- 4) Faciliter la tâche des enquêteurs (la partie « individu » est à leur charge, mais l'approche des institutions a été faite par les agents de l'Insee travaillant dans les divisions « enquêtes ménages » au sein des directions régionales (DR)).

Le questionnaire individu est divisé en trois sous-questionnaires :

- 1) Un questionnaire établissement : informations sur l'établissement, notamment le nombre de personnes hébergées durablement, mais aussi la distance de l'établissement à certains services (gare, poste, ...)
- 2) Le questionnaire individuel proprement dit ;
- 3) Toutefois, certaines questions (renseignements « administratifs » sur les allocations perçues) que les personnels administratifs des institutions connaissent probablement mieux que l'enquêté lui-même, a fortiori un proxy, ont été regroupées dans un petit questionnaire dit de « données de cadrage ».

L'enquête comporte alors :

La phase pré-enquête (deuxième trimestre 2009), menée par les directions régionales de l'Insee.

Puis la phase enquête (quatrième trimestre 2009), menée par les enquêteurs.

Cette phase commence par un premier passage dans l'institution : l'enquêteur peut poser le questionnaire établissement, et déposer le questionnaire « données de cadrage ». La liste des personnes hébergées durablement, demandée lors du contact, permet le tirage au sort des personnes à enquêter. Cette phase est également utile pour l'organisation concrète des interviews (lieux, horaires, proxys, ...).

Lors des passages suivants, l'enquêteur procède à l'interview des individus, et récupère le questionnaire « données de cadrage ».

Après une partie à peine adaptée du « tronc commun des ménages » (TCM) que l'INSEE utilise dans ses enquêtes auprès des ménages pour disposer d'informations harmonisées sur le logement et sur les personnes composant le ménage, le questionnaire de l'enquête proprement dit se déroule ensuite en abordant successivement les thèmes de « cœur d'enquête » : santé, repérage des handicaps, participation sociale et environnement. Les adaptations du questionnaire résultent de constats fait lors des tests : de nombreuses personnes à enquêter se trouvent fortement voire totalement limitées, d'où l'ajout d'un filtre (état végétatif ou coma) et une présentation alternative du module sur les restrictions d'activité. Dans les deux cas, il s'agit de ne pas poser des questions totalement hors de propos ou de les poser sans répétitions nombreuses donc lassantes.

En moyenne, le questionnaire individu est passé en environ 1 heure, il peut cependant y avoir de fortes variations de durées, notamment si la personne a de nombreux problèmes de santé (au sens large).

## **Partie III : Principes généraux de l'échantillonnage**



### **III-A : L'ÉCHANTILLON, AU VU DES OBJECTIFS DU VOLET INSTITUTION**

Le volet institution est ainsi présenté dans l'expression des besoins<sup>9</sup> (en 2006) : « Pour l'enquête en institutions, le protocole d'enquête proposé est identique à celui l'enquête HID de 1998. On utilisera comme base de sondage le Fichier National des Établissements Sanitaires et Sociaux du ministère de la Santé (FINESS). On procèdera d'abord au tirage stratifié (selon le type d'établissement) d'un échantillon d'établissements. Le tirage est effectué de façon à préserver dans chacune des strates un nombre d'interviews potentielles suffisant. (...) D'autre part, le tirage des établissements dans chaque strate est effectué à probabilités inégales, proportionnelles à la capacité (nombre de lits) de chaque établissement. La seconde phase du tirage est effectuée sur le terrain par chaque enquêteur. Elle consiste à tirer au sort 8 pensionnaires dans chaque établissement (pour ceux comptant au moins 8 personnes). On se restreint à 8 personnes pour limiter « l'effet de grappe ». Les pensionnaires d'un gros établissement ont moins de chances d'être tirés que ceux d'un petit, alors que l'établissement lui-même a d'autant plus de chances d'avoir été tiré qu'il est important. On parvient ainsi à un tirage équiprobable des individus dans chaque strate. En 1998 pour le premier passage de HID en institutions on avait échantillonné 16490 personnes (soit un taux moyen de sondage de 1/40). On propose de réduire la taille de l'échantillon à 10 000 personnes échantillonnées ».

Durant la phase de conception, l'objectif de 10 000 personnes n'est pas remis en cause. Les membres des maîtrises d'ouvrage<sup>10</sup> (à la DREES et à l'INSEE) et du groupe de conception<sup>11</sup> ont alors mené une réflexion sur les strates à enquêter.

Il fallait conserver un effectif suffisant dans chaque strate, dans la limite des 10 000, et, si possible rester proche de l'enquête HID. Certaines institutions ont donc été retirées du champ, avec l'idée que la population y résidant serait « captée » par le volet ménage. En effet, les personnes en ménages faisaient partie du champ de ce premier volet, même si elles étaient momentanément absentes ( y compris pour être hébergées dans une institution). *A contrario*, les personnes du champ institution se doivent d'être hébergées durablement dans celles-ci. D'autres institutions ont enfin été exclues, faute de pouvoir les intégrer tout en « tenant » les objectifs.

Ces types de structures conduisent aux sept catégories d'institutions présentées *supra*.

### **III-B : L'ÉCHANTILLON, TEL QUE PRÉSENTÉ AU COMITÉ DU LABEL<sup>12</sup>**

La population concernée par l'enquête est la population hébergée durablement en institution en France métropolitaine et dans les départements d'outre-mer, quel que soit son âge. Le tirage de l'échantillon est à deux degrés : tirage d'institutions puis d'individus parmi les personnes hébergées dans l'institution et éligibles à l'enquête.

La délimitation du champ des institutions concernées a été faite en tenant compte des coûts et de la complexité de la collecte, ainsi que de la spécificité et de l'importance numérique de la population concernée au regard des objectifs de l'enquête.

De ce fait, les prisons et les communautés religieuses, les cités universitaires, les foyers de travailleurs et les casernes sont hors champ<sup>13</sup>.

---

<sup>9</sup> Note N° 1098/DG75-F170 du 4 mai 2006

<sup>10</sup> L'enquête est pilotée, au sein de l'INSEE, par l'unité des Études démographiques et sociales, prenant la suite du département de la Démographie à la suite d'une réorganisation interne. Au sein de la DREES, elle est pilotée par la sous-direction de l'Observation de la santé et de l'assurance maladie et la sous-direction de l'Observation de la solidarité.

<sup>11</sup> Le groupe de conception associe des chercheurs (INED, INSERM, CNAV) et des médecins spécialisés en épidémiologie (IRDES, DREES, CNSA), ainsi que des psychiatres. Il y a eu des contributions des représentants des associations de personnes en situation de handicap (APF, UNAFAM, UNAPEI,...) et des représentants des financeurs de l'enquête (AGEFIPH, CNAF, CNAM, InVS, FNMF).

<sup>12</sup> Cette présentation a eu lieu le 20 octobre 2008.

<sup>13</sup> A la date du 20 octobre 2008, se posait la question de prendre en compte les populations hébergées dans les institutions citées par une repondération du volet "Ménages", lors de la construction du fichier commun aux deux volets et sous l'hypothèse que les personnes qui y sont hébergées ont des caractéristiques (au regard des thèmes de l'enquête) proches des personnes vivant en ménages. Les travaux sur l'opportunité et les modalités de tels traitements sont en cours au

Enfin, il s'agira d'éviter les doubles comptes statistiques. On rappelle que l'objectif principal du dispositif Handicap-Santé n'est pas de faire des enquêtes représentatives de la clientèle de chacune des institutions. L'objectif de l'enquête Handicap-Santé est en effet d'obtenir une vision d'ensemble des situations de handicap ou de l'état de santé pour toute la population résidant en France. L'objectif premier du volet "Institutions" est donc de connaître la situation des personnes qui ne peuvent pas être enquêtées via le volet "Ménages".

Sur un plan géographique, toutes les exploitations seront nationales ; il n'est pas prévu d'exploitation infranationale.

En revanche, on a réparti l'échantillon d'institutions afin d'obtenir une taille d'échantillon suffisante pour certaines catégories d'intérêt, pour lesquelles on souhaite disposer à des fins d'études d'un échantillon plus important que la part qu'elles représentent dans l'ensemble des places en institutions (exemple : personnes en établissements psychiatriques). Il s'agit là de l'objectif secondaire de l'enquête.

Afin de limiter la charge de collecte sur les personnels des institutions, le nombre de personnes interrogées sera de 6 individus dans les catégories d'institutions où l'on s'attend à un recours élevé au proxy. Il sera de 8 individus dans les autres types d'institutions. Les strates sont notamment constituées en tenant compte de ces catégories : on interroge le même nombre de résidents au sein de chacune des institutions d'une même strate. Le nombre de personnes interrogées dans chaque strate a été déterminé notamment en fonction des taux de proxys selon le type d'institution observés dans HID98. Il est de 9 individus pour toutes les institutions en département d'outre mer.

Le tirage des institutions sera réalisé proportionnellement à leur taille (capacité d'accueil en nombre de places).

On cherche à obtenir un échantillon de 9 600 individus résidents, répartis de la façon suivante selon la catégorie d'institution :

**Cible :**

1. Établissements Hospitaliers pour Personnes Âgées Dépendantes : 2 000 individus
2. Maisons de retraite : 1 500 individus
3. Unités de Séjours de Longue Durée : 600 individus
4. Maisons d'accueil spécialisées, Foyers pour Adultes Handicapés: 1 400 individus
5. Autres établissements pour adultes handicapés: 1 400 individus
6. Établissements et unités psychiatriques : 1 700 individus
7. Centres d'Hébergement pour réinsertion sociale : 1 000 individus

Ces catégories sont divisées en 14 strates<sup>14</sup>. Différents facteurs (statut administratif, particularité de l'organisation ou du public accueilli) justifient une différenciation à l'intérieur de ces catégories. Les effectifs au sein de chaque catégorie d'institutions (niveau 1) ont été ensuite répartis entre chaque strate (niveau 2) au prorata de la capacité totale des institutions de la strate.

### **III-C : AJUSTEMENTS : LA DÉFINITION DE L'ÉCHANTILLON (FIN 2008-DÉBUT 2009).**

Après la présentation du volet au comité du label, les caractéristiques de l'échantillon sont affinées. D'une part, la politique générale de l'INSEE prévoit d'enquêter en départements d'outre-mer et d'autre part une estimation plus précise des coûts conduit à adapter l'effectif de l'échantillon, ainsi que le protocole<sup>15</sup>.

La solution retenue est de préparer deux échantillons à l'issue de la pré-enquête :

- un échantillon principal d'établissements, correspondant à environ 8 500 personnes répondantes qui sera enquêté sur l'exercice budgétaire 2009, pendant la période de collecte fixée au programme de travail, d'octobre 2009 à décembre 2009.

---

premier trimestre 2011 et une synthèse en sera présentée dans la [note](#) annoncée en début de document.

<sup>14</sup> Voir [annexe 1](#).

<sup>15</sup> Notes n°2928/DG75-F420 du 5 décembre 2008 et DREES -BFHD N°6/2009 du 16 février 2009 pour l'ajustement du protocole (suppression des échantillons de réserves mais une pré-enquête) puis n° 938/DG75-F170 du 28 avril 2009 (Comité de pilotage : décision sur la taille de l'échantillon).

- un échantillon complémentaire, correspondant à 1 400 personnes répondantes, qui serait enquêté en janvier si la réalisation du programme d'enquêtes de l'Insee laissait apparaître une marge pour cela, tant en terme budgétaire qu'en terme de disponibilité du réseau des enquêteurs<sup>16</sup>.

## Partie IV : La pré-enquête : tirage et réponse

Les unités enquêtées dans le cadre de la pré-enquête sont les structures qui hébergent les résidents. Le tirage des institutions est stratifié selon la localisation des structures (départements d'outre-mer ou métropole) et selon le type d'institution. La base de sondage est constituée à partir du répertoire FINESS, enrichi des données de l'Enquête auprès des établissements psychiatriques 2009 et de données issues de l'enquête EHPA 2007 concernant les passages de maisons de retraite ou USLD en EHPAD.

### IV-A : PRINCIPES DE TIRAGE

Dans une strate, l'objectif est d'obtenir des poids à peu près égaux entre individus. Sachant qu'il y a un tirage d'institution suivi d'un tirage d'individu, on suivra le principe suivant :

- a) tirage d'institution avec probabilité proportionnelle à la capacité de l'institution
- b) dans l'institution, tirage équiprobable : plus la capacité de l'institution est grande, plus faible est la probabilité individuelle, mais le principe a) compense exactement le principe b).

On détermine donc les taux de sondages en suivant les étapes :

Ventilation des effectifs-cibles par strates (les 14) en partant des effectifs des 7 catégories d'institutions (voir [cible](#)). Cette ventilation est faite proportionnellement aux capacités de ces strates. Les effectifs cibles sont augmentés de 10 % pour anticiper une non-réponse (individu). Enfin, une nouvelle majoration de 40 % est appliquée, initialement pour permettre la constitution d'échantillons de réserve, finalement pour permettre un tirage dans l'effectif pré-enquêté en s'assurant qu'il y aura suffisamment d'institutions quelles que soient les strates et le taux de réponse anticipé. Ces principes sont corrigés à la marge pour tenir compte des contraintes plus fines : réduction de l'échantillon envisagé, tirage dans les départements d'outre-mer, glissements administratifs : durant la conception d'enquête avait lieu un mouvement de conventionnement des maisons de retraite qui les transforme en EPHAD : la majoration pour les strates 11,12, et 13 est de 60 %, il n'y pas de majoration pour les strates 21, 22 et 23.

### Tableau 2 : Estimation des taux cibles pour l'échantillon pour la pré-enquête en France métropolitaine.

---

<sup>16</sup> Jusqu'à l'automne 2008, il était prévu de constituer deux échantillons de réserve, à déclencher suivant l'avancement de la collecte. La faisabilité d'un suivi suffisamment rapproché et la nécessité d'estimer le coût au plus près ont conduit à la solution présentée ici. La pré-enquête a bien fonctionné : d'une part elle a bien permis de minimiser les taux de non-réponse pour l'enquête (phase institutions) et d'autre part elle a permis d'optimiser l'emploi des enquêteurs, en diminuant les temps et phases de contact au profit des interviews. C'est ce qui a permis de remonter d'un effectif visé de 8 500 répondants aux 9 104 répondants réels. Il était par ailleurs difficile de calibrer l'échantillon pour obtenir ces 8 500 répondants car les taux attendus de non-réponse ont été estimés à partir de l'observation des données d'HID, et varient suivant les strates (effet institution plus qu'individu).

Estimations du nombre d'institutions à enquêter (pré-enquête, métropole)	Institutions			Personnes		Institutions		
	N° de strate	Nombre N	Capacité C	Cibles				
				effectifs prévus				
				répondants	enquêtés	A enquêter	A pré-enquêter	
EHPAD	11	3 181	268 240	1 080	1 188	198	317	318
	12	1 859	131 683	530	583	97	155	156
	13	1 464	96 917	390	429	72	114	115
Maisons de retraite	21	170	13 881	607	668	83	83	78
	22	322	11 452	501	551	69	69	64
	23	260	8 978	392	432	54	54	50
USLD	30	191	14 131	600	660	110	154	154
Adultes Handicapés 1	40	989	33 073	1 400	1 540	257	359	359
Adultes Handicapés 2	50	2 045	62 258	1 400	1 540	193	270	269
Établissements psychiatriques	61	423	29 393	858	944	157	220	220
	62	354	15 871	463	510	85	119	118
	63	153	11 207	327	360	60	84	84
	64	37	1 770	52	57	9	13	14
CHRS	70	667	24 947	1 000	1 100	138	193	192
<b>Total (toutes strates)</b>		<b>12 115</b>	<b>723 801</b>	<b>9 600</b>	<b>10 560</b>	<b>1 581</b>	<b>2 205</b>	<b>2 191</b>

Note de lecture : on souhaite avoir 1080 répondants dans la strate 11, on prévoit donc d'en enquêter 1188 (+10 %). Sachant qu'on interroge 6 personnes par institution dans cette strate, il faut échantillonner 198 ehpad de strate 11. Chiffre porté à 317 pour les besoins de la pré-enquête et (dans ce cas) les possibilités de glissement d'institutions de la strate 21 dans la strate 11.

#### IV-B : TIRAGE EN FRANCE MÉTROPOLITAINE

On tire 2 191 structures dans le fichier des 12 115 institutions.

Tableau 3 : Répartition par strate de l'échantillon pour la pré-enquête en France métropolitaine

Niveau 1	N° de strate	Nombre d'institutions	Taille de l'échantillon pré enquête par strate	% d'établissements échantillonnés
EHPAD	11	3 181	318	10%
	12	1 859	156	8%
	13	1 464	115	8%
Maisons de retraite	21	170	78	46%
	22	322	64	20%
	23	260	50	19%
USLD	30	191	154	81%
Adultes Handicapés 1	40	989	359	36%
Adultes Handicapés 2	50	2 045	269	13%
Établissements psychiatriques	61	423	220	52%
	62	354	118	33%
	63	153	84	55%
	64	37	14	38%
CHRS	70	667	192	29%
<b>Total</b>		<b>12 115</b>	<b>2 191</b>	

On définit donc une première probabilité  $M_p$  (et son inverse, le premier poids  $W_p$  : le poids de tirage dans la base d'une institution pour la pré-enquête). Cette première probabilité est donc celle qu'une institution a d'être tirée pour la pré-enquête, sachant qu'elle figure dans la base de sondage.

Les données (et notations) sont alors les suivantes :

capacité d'une institution donnée (indexée par  $i$ ) :  $C_i$   
 capacité totale de la strate  $C$ ,  
 nombre d'institutions dans la strate  $N$ ,  $C$  est donc la somme des  $N$  valeurs  $C_i$ ,  
 nombre d'institutions à tirer :  $n$ , celui-ci étant déterminé comme cela vient d'être exposé.

Le tirage est proportionnel à la capacité donc (pour une institution) :  $M_p = a \cdot C_i / C$  ;  
 $a$  se détermine en sommant les  $M_p$  sur toutes les institutions de la strate : il y aura un total de  $n$   
 institutions tirées donc il faut que  $a = n$ .

$$M_p = n \cdot C_i / C \quad (\text{donc } W_p = (1/n) \cdot (C / C_i) ).$$

Cependant, lorsque la strate est trop petite, les capacités trop dispersées et le taux de sondage trop fort, la quantité  $n \cdot C_i / C$  peut être supérieure à 1. Plus précisément, ceci se produit si des institutions ont une capacité  $C_i$  supérieure à  $C / n$ . Cette valeur est donc une première valeur de seuil : si des institutions de la strate ont des capacités supérieures à ce seuil, il y aura tirage exhaustif de ces grosses institutions. Mais leurs probabilités de tirages ne sont égales qu'à 1, alors que  $n \cdot C_i / C$  excède 1. On doit donc réévaluer les probabilités de tirage des institutions restantes (comme le nombre d'institution à tirer est fixé, cela signifie que l'on va tirer de façon moins probable les plus grosses institutions, donc de façon plus probable les petites institutions).

On tire donc exhaustivement les institutions suffisamment grandes pour que  $n \cdot C_i / C > 1$ .

A l'issue de ce processus, on a donc tiré  $m$  institutions dont la capacité totale est  $B$ . On est alors ramené au problème précédent, mais avec seulement  $n - m$  institutions à tirer pour une capacité totale de  $C - B$ .

Le processus décrit est itéré autant de fois que nécessaire. On modifie ainsi progressivement les valeurs de seuil jusqu'à ce que tout se passe comme s'il y avait décomposition de l'échantillon :

Sous-strate 1 : grosses institutions, c'est-à-dire de capacité supérieure à un seuil final (qui dépend de la strate) : tirage exhaustif (probabilité = 1).

Sous-strate 2 : autres institutions, tirage proportionnel, comme conçu initialement. Les valeurs de  $n$  et  $C$  sont à modifier en fonction du nombre d'institutions et de la capacité totale de la sous-strate 1.

Dans les strates 11, 12 et 13, les capacités sont inférieures au seuil défini par le ratio  $C / n$ .

**Tableau 4 : détails du tirage de la pré-enquête, France Métropolitaine**

Métropole		Nombre	Capacité	à tirer	Seuils		Type tirage	
		N	C	n	initial	final	1	2
EHPAD	11	3 181	268 240	318	844	844	0	318
	12	1 859	131 683	156	844	844	0	156
	13	1 464	96 917	115	843	843	0	115
Maisons de retraite	21	170	13 881	78	178	164	21	57
	22	322	11 452	64	179	184	1	63
	23	260	8 978	50	180	225	2	48
USLD	30	191	14 131	154	92	62	93	61
Adultes Handicapés 1	40	989	33 073	359	92	96	9	350
Adultes Handicapés 2	50	2 045	62 258	269	231	288	2	267
Établissements psychiatriques	61	423	29 393	220	134	56	98	122
	62	354	15 871	118	135	115	20	98
	63	153	11 207	84	133	125	13	71
	64	37	1 770	14	126	145	1	13
CHRS	70	667	24 947	192	130	124	19	173
		12 115	723 801	2 191			279	1 912

*Note de lecture : dans la strate 21, le tirage proportionnel à la capacité ne peut avoir lieu, des institutions ayant une capacité supérieure au seuil initial C / n (ici 178 places). Dans cette strate, 78 institutions sont à tirer. Il y a d'une part tirage exhaustif (premier type de tirage) des institutions ayant plus de 164 places, soit 21 institutions. Pour les 57 institutions à tirer restantes, le tirage est proportionnel à la capacité (deuxième type de tirage).*

#### IV-C : TIRAGE EN DÉPARTEMENTS D'OUTRE-MER

L'enquête a été réalisée en Guadeloupe, Martinique, et Réunion. Pour ces trois départements, l'enquête était exhaustive et concernait 102 établissements fournis initialement par le fichier FINNESS de la DREES, auxquels sont venus se greffer 13 établissements guadeloupéens repérés sur place.

**Tableau 5 : Base de sondage DOM (Guadeloupe, Martinique, Réunion) (issue de la pré-enquête)**

Niveau 1	Niveau 2	Nombre d'institutions	Capacité (nombre de places)
EHPAD	public	22	1 784
	privé non lucratif	12	1 057
	privé lucratif	1	66
Maisons de retraite	public	5	213
	privé non lucratif	6	212
	privé lucratif	11	179
USLD	USLD	4	139
Adultes Handicapés 1	MAS-FAM	18	571
Adultes Handicapés 2	Autres établissements pour PH	10	190
Établissements psychiatriques	CHS-HPP	11	604
	Autres établissements publics ou privés non lucratifs	4	162
	Établissements privés lucratifs	3	131
	Post cure	0	0
CHRS	CHRS	8	169

On obtient donc une base de sondage de 115 établissements. Le tirage sera exhaustif, on affecte une probabilité **Mp** égale à 1 à ces établissements.

#### **IV-D : LA PRÉ-ENQUÊTE : ÉLÉMENTS SUR LA NON-RÉPONSE**

Il y a donc un total de 2 306 institutions pré-enquêtées. Une acceptation de la participation à la future enquête est obtenue dans 2 173 institutions (97 %). Les motifs de non-acceptation sont d'abord la discordance entre fichier et terrain (code de réponse « inconnu » : 57 cas) puis le hors champ (51 cas) Reste les échecs pour mauvais<sup>17</sup> adressage (17 cas) et refus direct dans 8 cas.

Une analyse de la non-réponse prise globalement a été faite, sans qu'il ressorte de grandes variations suivant les critères considérés : strate, taille, domiciliation (soit suivant la « zeat »).

On considère les 25 (soit 8 refus et 17 problèmes d'adresses) institutions comme non-répondantes, le choix de la probabilité de réponse est donc fixé à  $Ma = 2\,173 / 2\,198$  (ou le poids  $Wa$  à  $2\,198 / 2\,173$ ).

#### **Partie V : L'échantillon pour l'enquête et le tirage (institution)**

L'ensemble des institutions ayant accepté l'enquête constitue une nouvelle base de sondage. C'est dans cet ensemble que l'on va tirer l'échantillon des institutions à interroger. Plus précisément, on tire d'abord un échantillon de 1744 institutions en métropole. Puis, parmi ces 1744 institutions, on en tire à nouveau 1487. Les 1 487 vont constituer l'échantillon à interroger. Les  $1\,744 - 1\,487 = 257$  constituent un échantillon de réserve (non utilisé). Le principe suivi dans les départements d'outre-mer est précisé plus loin. On opère une stratification comme lors de la 1<sup>ère</sup> phase, si ce n'est que cette nouvelle stratification prend en compte les informations issues de la pré-enquête. Il y a donc des institutions qui changent de strate. L'essentiel des changements résulte du passage des maisons retraite en « EHPAD », du fait du processus administratif en cours au moment de l'enquête. Cela induit également la décision de regrouper les EHPAD et les maisons de retraite (par sous-strates analogues : la 11 et la 21, la 12 et la 22, enfin la 13 et la 23) après la pré-enquête, pour effectuer le tirage « institution ».

#### **V-A : LES CHANGEMENTS DE STRATES**

129 établissements changent de strates.

En détail :

Les deux établissements de strate 50 qui changent vont en strate 40.

Les 57 établissements de strate 21 qui changent vont en strate 11 sauf un qui va en strate 30.

Les 29 établissements de strate 22 qui changent vont en strate 12 sauf un qui va en strate 30.

Les 30 établissements de strate 23 qui changent vont en strate 13.

Les 11 établissements de strate 30 qui changent vont en strate 11.

Les effectifs par strates sont donc modifiés :

---

<sup>17</sup> L'institution existe mais pas à l'adresse indiquée.

Tableau 6 : changement de strate à l'issue de la pré-enquête

		Effectifs (institutions)			
		avant	inchangé	changé	après
EHPAD	11	328	328	0	395
	12	165	165	0	193
	13	115	115	0	145
Maisons de retraite	21	78	21	57	21
	22	58	29	29	29
	23	53	23	30	23
USLD	30	152	141	11	143
Adultes Handicapés 1	40	365	365	0	367
Adultes Handicapés 2	50	262	260	2	260
Établissements psychiatriques	61	218	218	0	218
	62	110	110	0	110
	63	85	85	0	85
	64	12	12	0	12
CHRS	70	172	172	0	172
		<b>2 173</b>	<b>2 044</b>	<b>129</b>	<b>2 173</b>

Champ : Métropole et départements d'outre-mer.

### V-B : L'ÉCHANTILLON

Le tirage de l'échantillon s'est donc effectué en deux temps. D'abord un « gros-échantillon » destiné à inclure l'échantillon proprement dit et l'échantillon-réserve a été tiré (tirage simple). Ensuite, l'échantillon a été tiré dans le « gros-échantillon ».

On aura donc deux probabilités (**Mg** et **Ms**) (respectivement les poids **Wg** et **Ws**) pour le tirage : probabilité d'être dans le « gros échantillon » sachant que l'on est dans la base déterminée après la pré-enquête, puis pour le tirage : probabilité d'être dans l'échantillon de l'enquête sachant que l'on est déjà dans le « gros échantillon ». Le produit est alors la probabilité **Mt** (respectivement le poids **Wt**) d'être dans l'échantillon de l'enquête sachant que l'on est dans la base déterminée après la pré-enquête.

Tableau 7 : Probabilités de tirage, enquête (institution), métropole.

De la pré-enquête à l'enquête :		Tirage "gros"		Tirage "final"		Probas.	
Effectifs (institutions, métropole) et probabilités de tirage ( <b>Mg</b> puis <b>Ms</b> )		Base effectifs	Probas.	Tirées effectifs	Probas.		A enquêter
EHPAD	11	391	0,716	280	0,850	223	0,609
	12	205	0,756	155	0,852	111	0,644
	13	157	0,745	117	0,846	91	0,631
Maisons de retraite	21		0,716		0,850	15	0,609
	22		0,756		0,852	21	0,644
	23		0,745		0,846	8	0,631
USLD	30	139	0,978	136	0,853	116	0,835
Adultes Handicapés 1	40	351	0,846	297	0,848	252	0,718
Adultes Handicapés 2	50	253	0,889	225	0,849	191	0,755
Établissements psychiatriques	61	208	0,889	185	0,849	157	0,755
	62	108	0,917	99	0,848	84	0,778
	63	82	0,878	72	0,833	60	0,732
	64	12	1,000	12	0,917	11	0,917
CHRS	70	166	1,000	166	0,886	147	0,886
		<b>2072</b>		<b>1744</b>		<b>1487</b>	



Dans les départements d'outre-mer, le tirage étant fait par département et par strate, on ne donne pas le détail des probabilités ici. Il y a 80 institutions enquêtées. Cela conduit donc à 1 567 institutions (en France) qui feront l'objet de l'enquête.

## VI : La non-réponse institution (au moment de l'enquête)

### VI-A : DÉFINITION D'UNE INSTITUTION RÉPONDANTE

Tableau 8 : les codes résultats d'enquête, toutes institutions.

			ERV	ABA	ERP	IHC	REF
EHPAD	11	242	240		1		1
	12	120	116		2		2
	13	92	91				1
Maisons de retraite	21	16	16				
	22	25	23		1	1	
	23	15	15				
USLD	30	120	118		1		1
Adultes Handicapés 1	40	264	262				2
Adultes Handicapés 2	50	196	193		1	2	
Établissements psychiatriques	61	165	157		2	4	2
	62	86	70	1	4	7	4
	63	63	51		1	3	8
	64	11	10			1	
CHRS	70	152	144		5	3	
		<b>1 567</b>	<b>1 506</b>	<b>1</b>	<b>18</b>	<b>21</b>	<b>21</b>

Champ : Métropole et départements d'outre-mer.

Il y a donc un abandon (ABA), 21 hors champ (IHC) et 21 refus (REF). La distinction entre ERV (entretiens validés) et ERP (entretiens partiels) va être ajustée.

En effet, on définit la non-réponse par rapport aux questionnaires. Le questionnaire d'enquête se décompose en trois sous-questionnaires. Le premier est un questionnaire institution, avec des renseignements sur les effectifs hébergés. Ces renseignements sont indispensables pour les calculs qui suivent. Le sous-questionnaire suivant est le questionnaire individu proprement dit, le troisième n'a pas d'incidence sur les travaux de pondérations.

On définit donc une variable (*quest*), suivant que l'on a :

- le questionnaire institution et au moins un questionnaire individu : 1511 institutions (la variable *quest* est codée 1) ;
- seulement le questionnaire établissement : 8 institutions (code 2) ;
- ni l'un, ni l'autre (code 9) : 48 institutions.

Pour ces 1511+8 = 1519 établissements, on dispose de la variable **Hi**, soit l'effectif de personnes hébergées durablement dans l'institution (et pour ceux-là seulement)<sup>18</sup>.

### VI-B : ANALYSE DE LA NON-RÉPONSE

Le taux global de réponse est de 96,9 % . Ce taux s'établit à 98,3 % dès lors que l'on retire les hors-champ<sup>19</sup>. Une analyse par strate est menée. Les critères étudiés sont la taille (capacité) et la domiciliation (département-région-zeat). Dans la plupart des strates le nombre d'institution en échec est donc trop petit pour relever d'un traitement particulier. Seules les strates 62 et 63 pourraient

<sup>18</sup> Voir également [annexe 2](#).

<sup>19</sup> L'un des buts de la pré-enquête était de limiter au maximum, si possible d'éviter les hors champ pour cette phase : on constate que l'approche « terrain » (contact physique de l'enquêteur avec une personne de l'institution) complète l'approche « distante » (contact téléphonique). Cependant, il faut relever que le travail accompli lors de la pré-enquête a été salué par les enquêteurs comme facilitant grandement l'approche « terrain », notamment en déterminant une personne contact au sein de l'institution.

relever d'un traitement plus spécifique. Bien que la non-réponse soit un peu plus forte pour les petites institutions, il n'apparaît pas nécessaire de définir des poids différents (intra-strates) à ce stade. La probabilité de répondre **Me** (respectivement le poids **We**) à cette phase, est donc calculée par simple ratio, mais par strate.

**Tableau 9 : la non-réponse (institution) par strates, probabilités de réponse Me.**

Ensemble		répond	échec	hors-champ	enquêtées	Me (%)
EHPAD	11	240	2	0	242	99,2
	12	118	2	0	120	98,3
	13	91	1	0	92	98,9
Maisons de retraite	21	16	0	0	16	100,0
	22	24	0	1	25	100,0
	23	15	0	0	15	100,0
USLD	30	119	1	0	120	99,2
Adultes Handicapés 1	40	262	2	0	264	99,2
Adultes Handicapés 2	50	194	0	2	196	100,0
Établissements psychiatriques	61	158	3	4	165	98,1
	62	72	7	7	86	91,1
	63	51	9	3	63	85,0
	64	10	0	1	11	100,0
CHRS	70	149	0	3	152	100,0
		1519	27	21	1567	98,3

Champ : Métropole et départements d'outre-mer .

#### VI-C : CONCLUSION SUR LES POIDS DES INSTITUTIONS

Au niveau institution, on a donc un poids **Wp.Wa.Wt.We**. On pose **Wist = Wp.Wa.Wt.We**.

Le produit **Wp.Wt** est le produit des poids de tirage, le produit **Wa.We**, le poids de redressement.

En terme de probabilités, **Mp.Ma.Mt.Me** est bien la probabilité qu'une institution soit répondante (sachant qu'elle est dans la base de sondage initiale). On pose aussi **Mist = Mp.Ma.Mt.Me**.

Les valeurs moyennes par strate des poids **Wist** reflètent le taux de sondage dans cette strate. Cette valeur moyenne est ainsi nettement plus élevée (environ 200) dans la première strate (ehpad) que dans les autres. La dispersion intra-strate est aussi assez importante :

**Tableau 10 : dispersion par strate du poids institution (Wist = Wp.Wa.Wt.We) :**

Dispersion sur 1519 institutions, Métropole et DOM			
	N°strate	effectif	D9/D1
EHPAD	11	240	14,0
	12	118	12,2
	13	91	10,6
Maisons de retraite	21	16	8,7
	22	24	11,4
	23	15	13,1
USLD	30	119	2,4
Adultes Handicapés 1	40	262	3,5
Adultes Handicapés 2	50	194	4,4
Établissements psychiatriques	61	158	3,4
	62	72	7,3
	63	51	3,0
	64	10	5,2
CHRS	70	149	6,1

ratios inter-déciles pour les 1519 institutions :	
	D9/D1
Tirage pré-enquête	11,5
(non-réponse)	1,0
Tirage enquête	1,5
Non-réponse (enquête)	1,1
Poids institutions	13,9

Champ : Métropole et départements d'outre-mer.

Note de lecture : Le rapport inter-décile des poids **Wist** est de 14,0 dans la strate 11. Celle-ci contient 240 institutions répondantes.

Considérées globalement (toutes strates confondues), les distributions de poids ont pour rapports interdéciles respectifs : Poids de tirage de la pré-enquête : 11,5 ; poids de tirage de l'enquête (phase institution) : 1,5 ; poids **Wist** : 13,9.

L'essentiel de la dispersion provient du tirage de la pré-enquête. Celle-ci est toutefois assez différente selon les strates, notamment parce que les strates 1 et 2 sont bien plus dispersées que les autres.

## VII : Traitements au niveau individu

### VII-A : LE TIRAGE DES INDIVIDUS

On tire au sort parmi les personnes éligibles, soit hébergées durablement dans l'institution. L'effectif de ces personnes est noté **Hi**. On tire N personnes par institution, N = 6 ou 8 suivant la strate en métropole, et N = 9 en départements d'outre-mer. Si l'effectif des personnes hébergées durablement est inférieur à N, toutes les personnes sont donc tirées.

La probabilité de tirage individu est alors **Mh** : **Mh** égal à  $N / Hi$  ; (et le poids **Wh** =  $Hi / N$ )

A noter qu'il y a encore, à l'issue des vérifications menées en début d'enquête proprement dite (partie « établissement » du questionnaire) 24 institutions répondantes qui changent de strate (dont 5 de strate 11 en strate 30 et autant dans l'autre sens).

### VII-B : LA NON-RÉPONSE INDIVIDU

L'ensemble des individus tirés à ce niveau de l'enquête forme un échantillon de 10 020 personnes. Après examen des questionnaires reçus et des non-réponses, un ensemble de 9 104 répondants est défini (91 %). On mène alors une analyse classique de la non-réponse.

Lorsqu'on considère les taux de réponses pour l'ensemble des personnes échantillonnées, suivant leurs caractéristiques, on constate des différences, mais peu marquées :

Effet de genre : les femmes répondent un peu plus (92 % ) que les hommes (90 %).

Effet d'âge : les 80-100 ans répondent un peu plus (93 % ) que les autres. Pas d'effet tuteur<sup>20</sup> si l'on raisonne sur l'ensemble des échantillonnés.

Effet département ou région : il sont peu clairs au niveau de l'ensemble. Après différents essais, les tests ont consisté à découper en trois zones : départements d'outre mer, Île-de-France, reste de la France, avec les regroupements possibles.

Par contre, les taux de réponses par strates diffèrent notablement.

On mène donc l'analyse de la non-réponse, strate par strate. Les effets de genre et d'âge ne « ressortent » plus, par contre, dans certaines strates, un effet tuteur existe bien. Enfin, il arrive que la « taille » de l'établissement joue. Cette taille est définie par la capacité **Ci** ou l'effectif des personnes hébergées durablement **Hi**<sup>21</sup>. Le taux de réponse est en général un peu meilleur pour les institutions de taille moyenne (soit le deuxième quartile (en partant des plus petites structures)). Ce constat est vrai, que l'on définisse la taille par l'une ou l'autre façon. Suivant la taille des strates, on a donc introduit dans le modèle de réponse une indicatrice de quartiles (parfois simplement la position par rapport à la médiane), avec les deux variantes (capacité ou effectifs de personnes durablement hébergées).

On a donc : des éléments sur l'institution : strate, localisation et taille  
des éléments individus : sexe - âge  
un élément clairement important mais « entre les deux » : la présence de tuteur.

<sup>20</sup> Un tel effet se traduirait par un taux de réponse différent entre les personnes sous tutelle d'une part, sans tutelle d'autre part. Le protocole d'enquête suppose naturellement l'accord du tuteur pour les personnes de première catégorie. Le tuteur n'est pas (en général) présent dans l'institution au moment de l'enquête. Son consentement a donc été recueilli par courrier. Dans certains cas, ce sont des associations qui sont à contacter. Il pouvait donc y avoir un effet spécifique, l'attitude d'une association conduisant à un refus systématique. Cela a en effet constaté au moment de l'enquête.

<sup>21</sup> Voir annexe 2 : les deux grandeurs ne sont pas tout à fait les mêmes.

Dans chaque strate, on teste donc quelques modèles suivant la codification des variables (sexe, âge, taille d'établissement, présence de tuteur, région où est l'établissement). On retient le meilleur modèle (probabilité de réponse **Mr**, poids de réponse **Wr**).

**Tableau 11 : non-réponse (individu), suivant la strate.**

Répondants (individus)		Oui	Non	total		Pour mémo :	
<i>Métropole et DOM</i>					taux	effectifs "LABEL"	
EHPAD	11	1 389	96	1 485	93,5		1 034
	12	704	37	741	95,0		544
	13	504	51	555	90,8		423
Maisons de retraite	21	132	13	145	91,0		749
	22	172	6	178	96,6		456
	23	121	6	127	95,3		294
USLD	30	654	72	726	90,1		600
Adultes Handicapés 1	40	1 475	123	1 598	92,3		1 400
Adultes Handicapés 2	50	1 487	71	1 558	95,4		1 400
Établissements psychiatriques	61	790	148	938	84,2		915
	62	333	89	422	78,9		338
	63	256	62	318	80,5		428
	64	56	4	60	93,3		18
CHRS	70	1 031	138	1 169	88,2		1 000
		<b>9 104</b>	<b>916</b>	<b>10 020</b>			<b>9 599</b>

Champ : Métropole et départements d'outre-mer.

**Tableau 12 : non-réponse, modèles par strate, puis rapport interdécile.**

Modèles de redressement pour la non-réponse individu		variables retenues					D9/D1
		Tuteur	Région	Taille	Sexe	Rien	
EHPAD	11	oui	oui				1,05
	12	oui	oui	Hbg			1,20
	13	oui	oui				1,10
Maisons de retraite	21			Hbg			1,12
	22					oui	1,00
	23					oui	1,00
USLD	30	oui					1,11
Adultes Handicapés 1	40	oui		Cap			1,09
Adultes Handicapés 2	50	oui					1,07
Établissements psychiatriques	61	oui		Cap			1,33
	62					oui	1,00
	63					oui	1,00
	64					oui	1,00
CHRS	70			Hbg	oui		1,13

Champ : Métropole et départements d'outre-mer.

Note de lecture : dans la strate 12, le modèle de non-réponse retenu comprend les explicatives « effet tuteur », « effet région » et « effet taille d'établissement ». Dans ce cas l'effet taille est mieux pris en compte en utilisant la variable « effectifs des personnes hébergées durablement ». Le coefficient de redressement (il s'agit du poids **Wr**) est dispersé, l'indicateur « D9/D1 », soit le rapport interdécile, prenant la valeur de 1,20.

De façon générale, la dispersion induite par ce redressement est assez faible, et pour la plupart des strates, sauf les plus grosses, les ratios D9 / D1, Q3 / Q1, voire C99 / C01 sont quasiment les mêmes.

### VII-C : LE CALAGE

On dispose à cette étape d'un poids individu pour tout répondant. Par ailleurs, les calculs faits conduisent aussi à une estimation, par strate, de l'effectif des personnes hébergées durablement dans l'ensemble des institutions : il suffit de sommer le produit **Wist\*Hi** par institution, soit le poids de l'institution par l'effectif que cette institution héberge durablement.

**Tableau 13 : Effectifs par strate de la population hébergée durablement.**

<b>Population de calage</b>		
<b>EHPAD</b>	<b>11</b>	<b>255 527</b>
	<b>12</b>	<b>122 794</b>
	<b>13</b>	<b>94 381</b>
<b>Maisons de retraite</b>	<b>21</b>	<b>4 565</b>
	<b>22</b>	<b>4 179</b>
	<b>23</b>	<b>2 468</b>
<b>USLD</b>	<b>30</b>	<b>13 424</b>
<b>Adultes Handicapés 1</b>	<b>40</b>	<b>33 550</b>
<b>Adultes Handicapés 2</b>	<b>50</b>	<b>56 071</b>
<b>Établissements psychiatriques</b>	<b>61</b>	<b>17 543</b>
	<b>62</b>	<b>8 656</b>
	<b>63</b>	<b>6 320</b>
	<b>64</b>	<b>1 195</b>
<b>CHRS</b>	<b>70</b>	<b>15 716</b>
<b>Total</b>		<b>636 389</b>

*Champ : Métropole et départements d'outre-mer .*

Il suffit finalement d'effectuer un calage (simple, par strate) en ramenant la somme des poids au total de la population hébergée durablement. Ce calage ne change pas la dispersion des poids dans une strate donnée.

## Annexe 1 : Définition détaillée des strates

Les 7 catégories puis les 14 strates sont définies :

Au niveau des catégories d'institutions :

- 1) Établissements d'Hébergement pour Personnes Âgées Dépendantes (EHPAD).
- 2) Maisons de retraite.
- 3) Unités de Soins de Longue Durée (USLD).
- 4) Établissements pour les personnes handicapées (Maisons d'Accueil Spécialisées (MAS), Foyers d'accueil médicalisés (FAM).
- 5) Autres établissements pour personnes handicapées.
- 6) Établissements psychiatriques et services de psychiatrie.
- 7) Établissements pour les personnes en difficulté sociale (Centres d'Hébergement et de Réinsertion Sociale (CHRS).

Plus en détail :

Catégories d'institutions	Strates	Numéro
EHPAD	Public	11
	Privé non lucratif	12
	Privé lucratif	13
Maisons de retraite	Public	21
	Privé non lucratif	22
	Privé lucratif	23
USLD		30
Adultes Handicapés 1	MAS-FAM	40
Adultes Handicapés 2	Autres établissements pour PH	50
Établissements psychiatriques	CHS-HPP ( <i>ie</i> en hôpitaux)	61
	Autres établissements publics ou privés non lucratifs	62
	Établissements privés lucratifs	63
	Post cure	64
CHRS		70

## Annexe 2 : Relation entre capacité et effectif de personnes hébergées durablement.

On dispose (dans la base de sondage) pour toutes les institutions de sa capacité **Ci**.

On dispose pour les 1519 institutions définies comme répondantes à l'issue de l'enquête en institution de l'effectif de personnes hébergées durablement **Hi**.

Il est intuitif de considérer que les deux valeurs sont proches, **Hi** devant être un peu inférieur à **Ci**, voire nettement inférieur à **Ci** dans les institutions de catégories 6 et 7 (établissements psychiatriques et centre de réinsertion), où les personnes ont moins vocation à rester durablement. On s'attend évidemment à quelques écarts.

Les capacités interviennent comme paramètre dans le tirage, au moment du tirage de la pré-enquête.

Les effectifs de personnes hébergées durablement interviennent comme paramètre dans le tirage des individus.

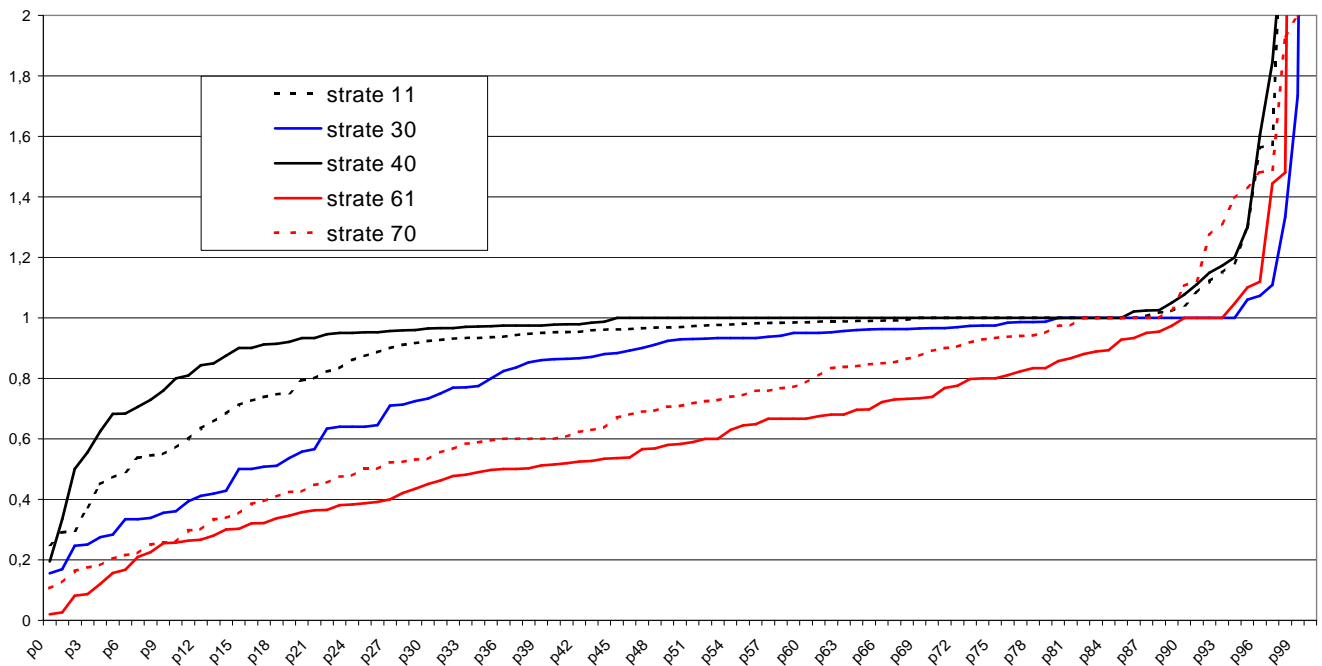
Distributions des ratios **Hi / Ci** par strate :

1519 institutions			Quantiles de Hi / Ci				
			D1	Q1	M	Q3	D9
EHPAD	11	238	0,57	0,87	0,97	1,00	1,04
	12	119	0,69	0,92	0,97	1,00	1,04
	13	92	0,79	0,87	0,95	1,00	1,01
Maisons de retraite	21	18	0,50	0,62	0,92	0,99	1,03
	22	22	0,40	0,58	0,82	0,93	1,00
	23	15	0,52	0,78	0,96	1,08	1,50
USLD	30	119	0,36	0,64	0,93	0,98	1,00
Adultes Handicapés 1	40	261	0,80	0,95	1,00	1,00	1,08
Adultes Handicapés 2	50	195	0,57	0,82	0,97	1,00	1,08
Établissements psychiatriques	61	158	0,26	0,39	0,58	0,80	1,00
	62	71	0,21	0,37	0,53	0,67	0,95
	63	52	0,33	0,41	0,51	0,63	0,90
	64	10	0,45	0,53	0,84	0,92	1,00
CHRS	70	149	0,26	0,50	0,71	0,93	1,11

Champ : Métropole et départements d'outre-mer.

L'hypothèse faite sur la valeur moyenne du ratio est donc assez raisonnable. Cependant, elle n'est pas parfaitement vérifiée (**Hi** ne devrait pas être plus grand que **Ci** !). La dispersion constatée est donc signe d'une concordance modérément bonne.

Pour préciser, voici les fonctions de répartition pour 5 strates, choisies pour représenter la correspondance  $H_i / C_i$  la meilleure, la moins bonne et des situations intermédiaires :



*Note de lecture : dans la strate 61 (établissements psychiatriques), un peu plus de 50 % des établissements ont un ratio  $H_i / C_i$  inférieur à 0,6 : moins de 60 % des pensionnaires sont hébergés durablement.*

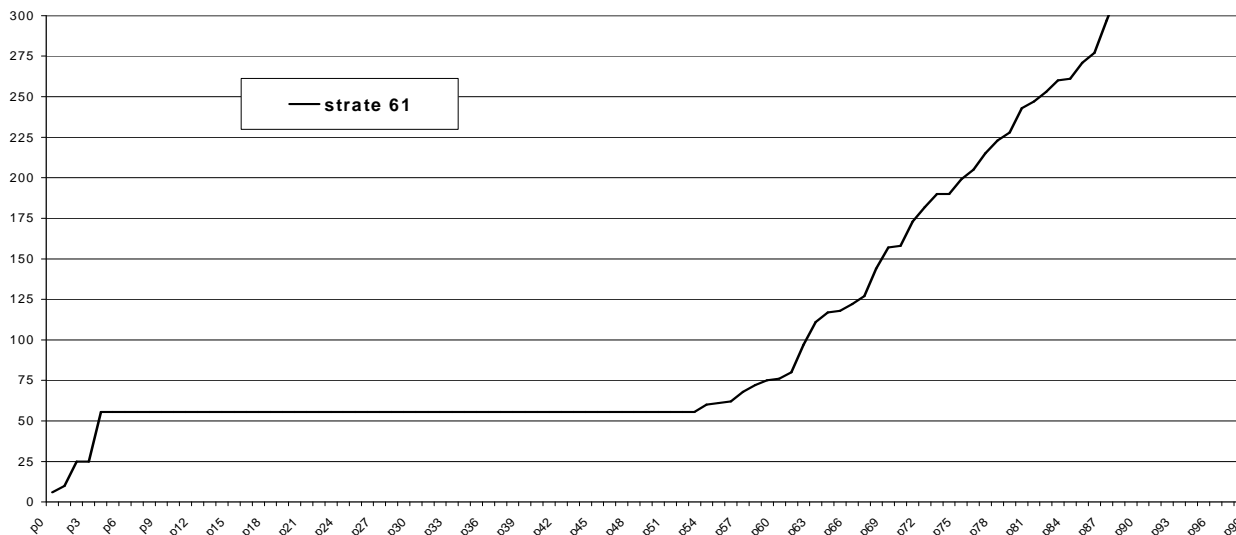


### Annexe 3 : Dispersion du tirage de la pré-enquête

Pour apprécier la dispersion des poids de tirage de la pré-enquête, on étudie en fait la distribution de la quantité **Wp.Ci**.

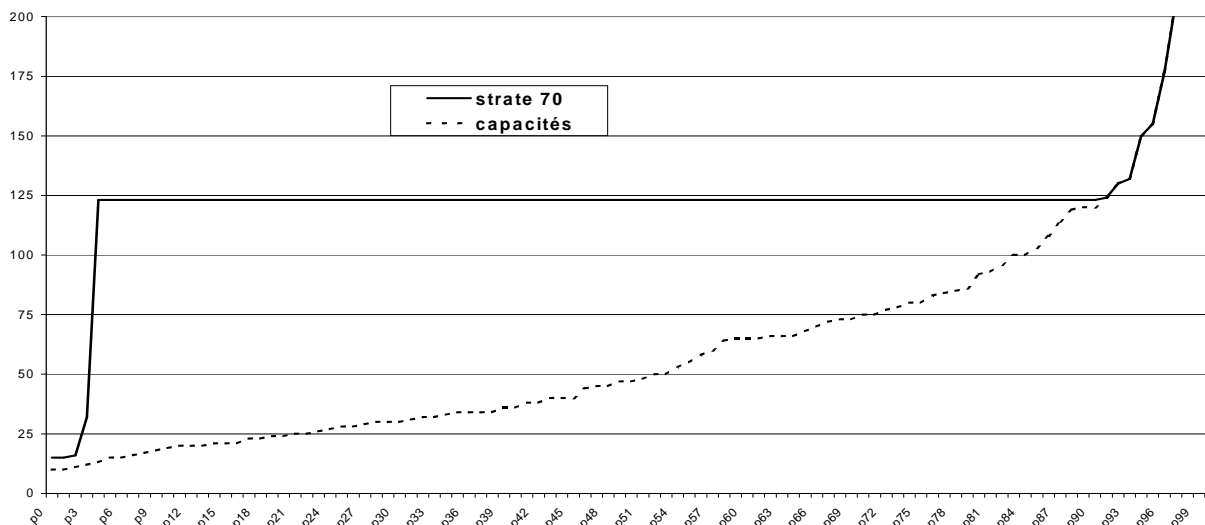
On s'attend à ce que cette quantité soit constante par strate. Il n'est est évidemment rien.

Pour une strate assez exemplaire, la dispersion de la quantité **Wp.Ci**, strate 61 se visualise :



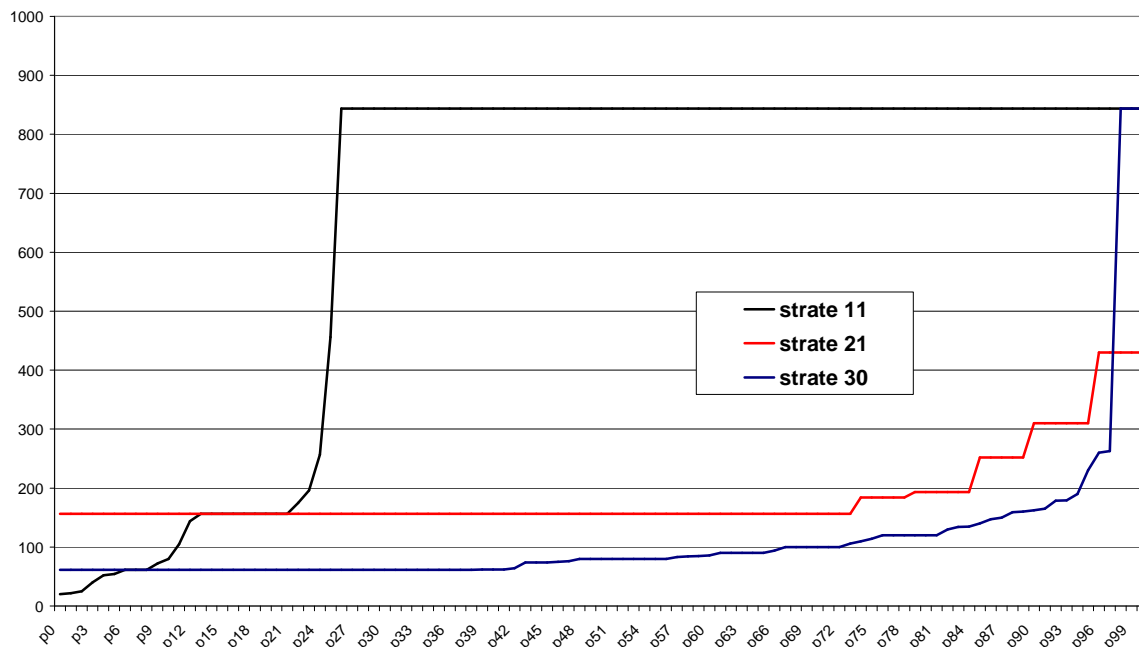
La déformation des poids « à droite » est attendue. Pour la strate 61, dès le 55<sup>ème</sup> centile de la distribution de la grandeur **Wp.Ci**, s'observe non plus une constante mais une croissance assez régulière. Cela résulte de la contrainte de tirage exposée en [partie IV-B](#) : au-delà d'un certain seuil S (ici un peu au-dessus de 50), la probabilité de tirage proportionnelle à la capacité eut été supérieure à 1. Elle est donc bornée par 1 (tirage exhaustif), le poids étant donc minoré par cette valeur. On multiplie le poids par la capacité : lorsque celle-ci dépasse le seuil, la dispersion reproduit donc celle des capacités.

Ceci est également bien illustré dans la strate 70 par exemple, où l'on représente la distribution des capacités (**Ci**, pointillé) en accompagnement de celle des **Wp.Ci** (trait plein) :



La déformation des poids « à gauche », soit pour les plus petites valeurs de  $Wp.Ci$ , résulte des changements de strates : on insère une dispersion avec les effectifs et la valeur moyenne déterminée par une strate telle que définie en pré-enquête dans une autre dispersion.

Ceci est bien illustré dans la strate 11 (qui « absorbe » une grande partie de la strate 21 et une petite partie de la strate 30) :



## Annexe 4 : Dispersion due à l'écart entre « H » et « C »

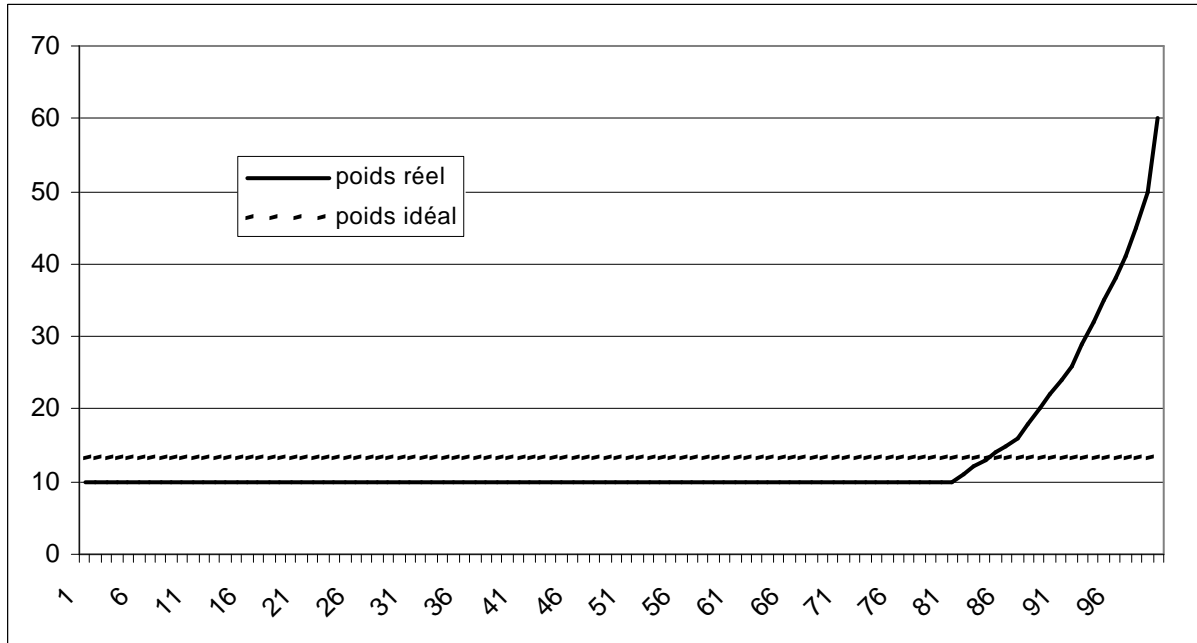
La présentation faite ici est « un cas d'école ».

Dans un premier temps, on considère une distribution de poids stylisée, telle qu'elle aurait été obtenue par le processus décrit, mais :

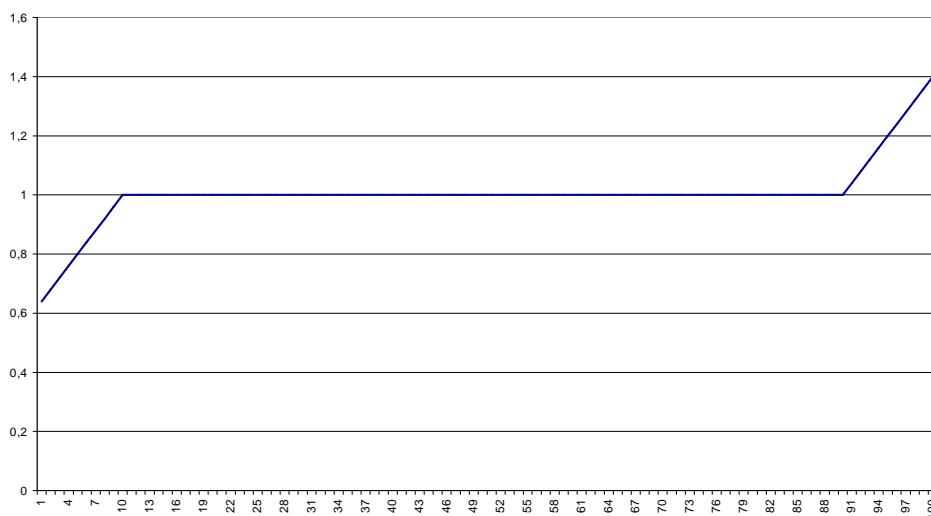
Les taux de non-réponse sont égaux à 1, pas de changement de strate,  $H_i = C_i$ .

Seul le procédé de tirage de la pré-enquête est susceptible de disperser les poids (dans le cas où les capacités des institutions conduisent à un tirage exhaustif d'une partie de celles-ci).

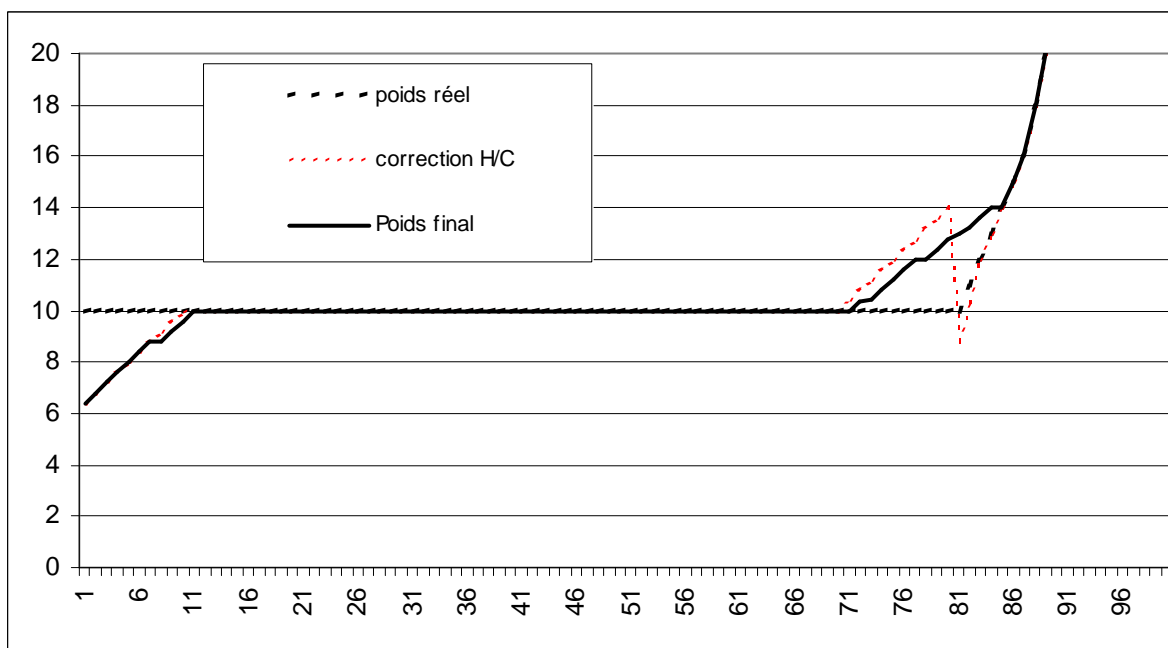
On a donc une distribution qui s'apparente à :



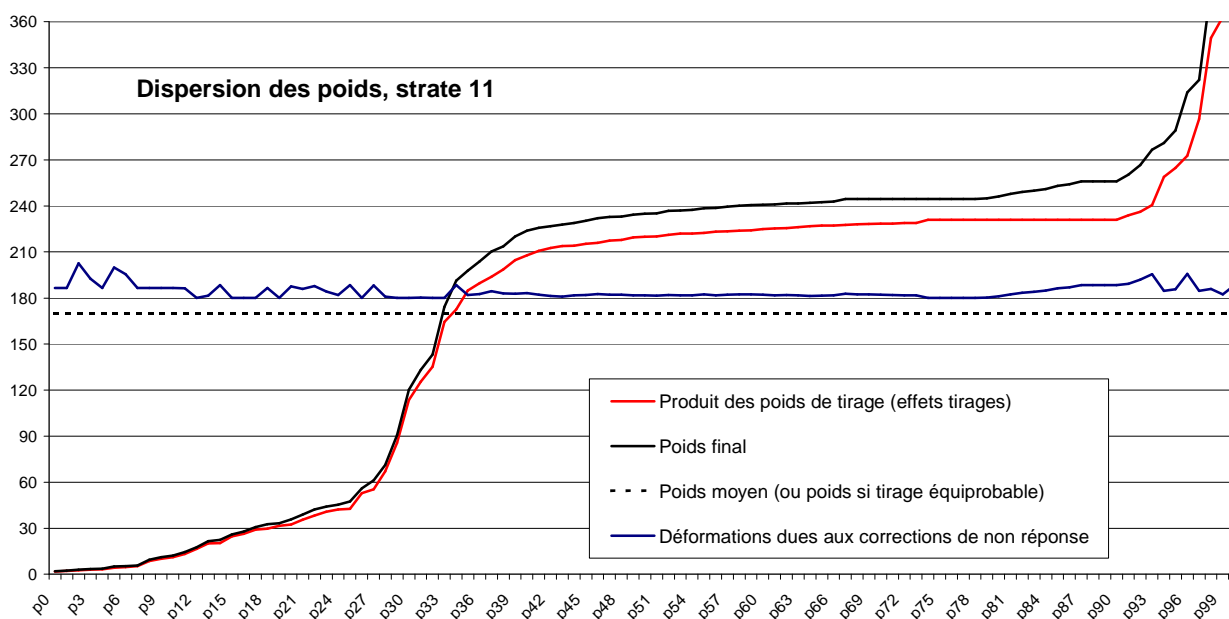
On introduit maintenant une distribution simplifiée du ratio  $H_i / C_i$  :



D'où la combinaison des deux effets :



Une distribution réelle est la suivante :



Note de lecture : Le tracé noir plein donne la répartition des poids (pondération finale) individus. En noir pointillé, le poids de tirage moyen qui est aussi celui qu'on aurait obtenu dans une situation « idéale » : dispersion des capacités modérée, donc tirage proportionnellement à celles-ci, sans effet de saturation (ou déformation de la répartition pour tenir compte du fait que les plus grosses institutions doivent être exhaustivement tirées), sans particularité géographique (tirages spécifiques en départements d'outre-mer), sans changement de strate entre base et enquêtée et correspondance entre capacité et population durablement hébergée. En rouge, la répartition des poids de tirage individus. En bleu la dispersion des poids induite par les traitements de non-réponse.

## Annexe 5 : Tableau des quantiles, poids par strates

9107 individus			Quantiles des poids				
			D1	Q1	M	Q3	D9
EHPAD	11	1 390	12,12	47,39	235,01	244,56	256,02
	12	704	13,91	54,34	222,36	232,40	249,28
	13	504	43,46	70,83	225,12	241,63	257,94
Maisons de retraite	21	132	4,19	21,68	32,70	37,71	71,05
	22	172	4,01	15,00	29,06	33,34	37,50
	23	121	1,97	2,30	14,90	34,79	36,30
USLD	30	654	6,38	10,93	14,42	20,41	29,30
Adultes Handicapés 1	40	1 475	13,52	21,38	22,58	23,41	27,01
Adultes Handicapés 2	50	1 487	21,29	32,91	39,03	41,04	45,20
Établissements psychiatriques	61	791	5,12	7,86	12,32	26,60	50,78
	62	333	12,06	15,35	23,03	29,43	41,88
	63	256	12,33	17,14	21,71	28,39	42,58
CHRS	64	56	12,27	13,97	24,05	24,88	26,58
	70	1 032	5,22	9,79	14,80	19,33	24,24

Champ : données pour l'ensemble des individus répondants, en métropole et dans les départements d'outre-mer.

## Annexe 6 : Comparaison entre cible et résultats de collecte

Comparaison en effectifs (individus) par catégories d'établissements : personnes échantillonnées, personnes répondantes, et effectifs ciblés (présentation au Comité du Label, octobre 2008).

	tirées	"bon"	taux	cible	taux
<b>Strates "personnes âgées"</b>	3 957	3 675	<b>92,9</b>	4 100	<b>89,6</b>
dont EHPAD	2 781	2 596	<b>93,3</b>		
dont Maisons de retraite	450	425	<b>94,4</b>		
dont USLD	726	654	<b>90,1</b>		
<b>Strates "adultes handicapés"</b>	3 156	2 961	<b>93,8</b>	2 800	<b>105,8</b>
<b>Etablissements psychiatriques</b>	1 738	1 436	<b>82,6</b>	1 700	<b>84,5</b>
<b>CHRS</b>	1 169	1 032	<b>88,3</b>	1 000	<b>103,2</b>
<b>Total</b>	10 020	9 104	<b>90,9</b>	9 600	<b>94,8</b>

Champ : données pour l'ensemble des individus répondants, en métropole et dans les départements d'outre-mer.