

## **DONNEES ESPS 2006 : VERSION 1**

### **Présentation des fichiers**

#### ***Avertissement :***

*Ce document constitue une présentation synthétique des bases de données ESPS 2006 VERSION 1, et des règles d'apurement retenues.*

**Sommaire :**

Lexique - Descriptif des bases de données : .....	3
Règles d'apurement sur les variables sociodémographiques .....	5
Règles d'apurement questionnaire santé et soins médicaux .....	6
Règles d'apurement du questionnaire couverture complémentaire .....	13
Règles d'apurement sur les variables sociodémographiques du ménage.....	15

## Lexique

### **Assuré principal**

C'est la personne qui a fait l'objet du tirage et dont nous enquêtons le ménage. Sauf impossibilité ou refus explicite, c'est elle qui répond aux deux questionnaires (contacts 1 et 2) proposés au téléphone ou en face-à-face.

### **QP : questionnaire principal**

C'est le questionnaire qui est posé par téléphone ou en face-à-face lors des deux contacts avec les ménages enquêtés. Normalement, c'est l'assuré principal (personne ayant fait l'objet du tirage) qui y répond.

### **QS : questionnaire santé et soins médicaux**

C'est un questionnaire auto-administré qui doit être rempli par toutes les personnes du ménage.

## **Descriptif des bases de données :**

Les bases 2006 mises à disposition en version 1 :

**Table APPEL1** : elle contient les renseignements concernant uniquement la personne répondant à l'enquête obtenus lors de la 1<sup>ère</sup> visite ou du 1<sup>er</sup> contact téléphonique. Il y a une ligne par ménage enquêté.

**Table APPEL2** : elle contient les renseignements concernant uniquement la personne répondant à l'enquête obtenus lors de la 2<sup>ème</sup> visite ou du 2<sup>er</sup> contact téléphonique. Il y a une ligne par ménage enquêté.

**Table PERSONNE** : elle contient les renseignements individuels recueillis au cours de l'enquête. Il s'agit essentiellement des renseignements socio-éco-démographiques et des données de santé issues du questionnaire santé. Il y a une ligne par personne enquêtée.

**Table MENAGE** : elle contient les renseignements concernant le ménage, recueillis au cours des deux contacts. Il s'agit en particulier du revenu et de la situation familiale. Il y a une ligne par ménage.

**Table MALADIE** : elle contient les informations concernant l'ensemble des maladies déclarées par les individus ayant rempli le questionnaire santé ainsi que les maladies repérées par l'équipe de médecins au vu des informations collectées par l'enquête (ALD, motif de la dernière séance de médecin, motif des prises médicamenteuses de la veille) . Il y a une ligne par maladie déclarée. Ce sont les maladies prévalentes le jour de l'enquête. Font également parties de ce fichier, l'obésité (repérer par le calcul du BMI), le port de lunettes, de prothèses dentaires et autres prothèses. La provenance des maladies est repérée par la variable REMONTE.

**Table MUTUELLE** : elle contient la description des couvertures complémentaires déclarées dans le questionnaire principal. Il y a une ligne par contrat de couverture complémentaire et par personne.

**Table FJAUNE** : elle contient les informations recueillies dans le questionnaire auto-administré consacré aux couvertures complémentaires. Il y a une ligne par contrat de couverture complémentaire déclaré. La table FJAUNE est différente de la table MUTUELLE car elle contient plus d'informations. En revanche, elle décrit moins de contrats car certains questionnaires auto-administrés ne sont pas retournés. Les informations recueillies dans la table FJAUNE peuvent corriger celles recueillies par téléphone ou en face-à-face lors du 1<sup>er</sup> contact.

**Table PASMUT** : elle contient la description des non-bénéficiaires d'une couverture complémentaire. Il y a une ligne par personne qui ne bénéficie d'aucune couverture complémentaire santé.

**Table QGENE** : elle contient les informations recueillies dans le QS concernant la dernière séance de généraliste. Il y a une ligne par personne ayant renseigné cette partie du questionnaire santé et ayant déclaré une telle séance au cours des 12 derniers mois.

**Table QSPE** : elle contient les informations recueillies dans le QS concernant la dernière séance de spécialiste. Il y a une ligne par personne ayant renseigné cette partie du questionnaire santé et ayant déclaré une telle séance au cours des 12 derniers mois.

**Table LESPOND** : variables de pondération (POND) ; quatre pondérations sont disponibles selon le fichier utilisé :

- Une pondération ménage
- Une pondération individuelle (table personne)
- Une pondération pour chaque partie du Questionnaire principal
- Une pondération spécifique pour l'échantillon standard complété du sur-« échantillon » CMUC

Ces pondérations sont réalisées *a posteriori* par calage sur marge ; elles reproduisent les répartitions des ménages ordinaires de France métropolitaine par âge, sexe, taille de ménage et régime d'Assurance maladie. Pour l'échantillon standard complété du sur-«échantillon » CMU-C, le taux de CMUistes a constitué une marge de calage supplémentaire. Il s'agit de pondérations relatives, centrées sur 1.

**Table ADDPHAR** : elle contient les consommations médicamenteuses de la veille. Il y a une ligne par médicament consommé pour chaque personne ayant rempli le tableau correspondant du QS.

**Table ORIGINE\_RELIG** : elle contient les informations sur les origines géographiques de la personne interrogée et de ses parents, ainsi que des informations sur ses croyances et sa pratique religieuse. Il y a une ligne par ménage enquêté.

Les tables **ADDPHAR** et **ORIGINE\_RELIG** ne sont disponibles que sur demande.

## **Règles d'apurement sur les variables sociodémographiques**

### Catégorie Socioprofessionnelle : la nouveauté SICORE.

Pour la première fois en 2006 nous avons utilisé le logiciel SICORE de l'INSEE pour expérimenter une codification automatique de la profession. Pour cela SICORE utilise les libellés de profession ainsi que des informations dites « annexes » telles que le statut dans l'emploi (salarié fonctionnaire ou non, indépendant...), la catégorie perçue (ouvrier, employé, cadre...), le nombre de salariés pour les indépendants...

Le bilan pour cette année concernant les 14 963 PCS à coder :

- \* 35 % des libellés ont été codés directement par Sicore.
- \* 50 % ont donné lieu à un codage en réussite partielle qui a été soumis à une validation de la part de la société de sondage. Les codes produits par SICORE ont été validés 8 fois sur 10.
- \* 15 % des libellés ont produit un échec de codage. Ces PCS ont été codées manuellement par la société de sondage avec les mêmes règles que dans les enquêtes précédentes.

### Traitement des modalités « autres, précisez » :

Le contenu des modalités « autres » est systématiquement relu.

S'il correspond à une modalité préexistante, la variable est recodée.

Si une modalité non prévue émerge, elle est créée. Ces nouvelles modalités sont identifiées par les codes 51 et suivants.

### Signification des modalités de non réponse.

98 : ne sait pas

99 : refus

### Autres apurements :

- Si des individus du ménage sont oubliés lors du 1<sup>er</sup> passage ou appel de l'enquêteur, ils sont insérés par la suite par la société d'enquête. Dans le cas où le dispositif de l'enquête ne permet pas de recueillir les informations individuelles les concernant (dans le questionnaire principal : place dans le ménage, situation au travail, niveau d'études, couverture complémentaire... et dans le questionnaire santé), ces dernières resteront manquantes.
- Lorsque des réponses manifestement erronées des enquêtés sont repérées par la société d'enquête, si elles interviennent au sein d'un bloc de questions dépendantes, l'ensemble de ces questions n'est pas reposé à l'enquêté, sauf pour une erreur constatée dans le bloc Sicore. Des propositions de corrections de ces erreurs sont transmises à l'IRDES. Si ces corrections sont validées, les questions dépendantes sont mises à vide.
- A partir de l'enquête 2004, nous avons décidé de conserver le plus souvent possible les informations déclarées par les enquêtés, sans les interpréter, même si elles peuvent parfois être sujettes à caution. En cas de doute, en particulier dans le bloc Sicore, nous avons choisi de reposer tout le questionnaire lors du contact suivant avec l'enquêté.

## Règles d'apurement questionnaire santé et soins médicaux (QS)

Le questionnaire Santé et soins médicaux est un questionnaire auto-administré qui est proposé à tous les membres des ménages enquêtés dans le cadre de l'enquête ESPS. Il est composé de trois parties : « Etat de santé » ; « Dernière séance de médecin généraliste » ; « Dernière séance de médecin spécialiste »

Nous détaillons ici les règles qui ont été adoptées pour l'apurement des informations recueillies par ces questionnaires.

Lorsque la création d'une variable résulte du résultat d'un calcul, qu'elle a fait l'objet d'une modification par rapport à la déclaration de départ ou qu'elle est créée en raison d'un nombre d'occurrences importantes non proposées au départ, elle sera systématiquement suivie de la mention \_GV.

Règle générale : Certaines questions sont filtrantes et ouvrent vers une sous-question. S'il y a des incohérences entre les réponses à une question filtrante et à la sous-question correspondante, l'ensemble est mis à vide et est alors considéré comme de la non-réponse. Si la question filtre est vide et que la sous question est renseignée, nous cochons la filtrante pour obtenir la cohérence.

### I. La partie Santé

Dans la table PERSONNE : 22 150 observations.

\* Suppression des QS des individus qui ne résident pas habituellement dans le foyer au moment de l'enquête.

\* Il y a parfois une inversion des QS entre les membres du ménage. Certains remplissent pour eux des questionnaires santé qui ne leurs sont pas adressés. Les prénoms et dates de naissance renseignés en début de QS permettent de réattribuer les QS aux bonnes personnes.

\* Création de la variable QS : QS=1 → la personne a renvoyé son questionnaire santé et soins médicaux et celui-ci n'est pas totalement vide (au moins une réponse). Sinon QS=2

\* Création de la variable QS\_QUI : la personne remplit elle-même son questionnaire Santé et Soins Médicaux ?

1 : oui / 2 : non / 0 : nous ne savons pas qui l'a rempli.

1) CALCUL DE L'IMC = indice de masse corporelle.

- Apurement de la taille

Si l'individu déclare une taille égale à 1, au vu de la courbe de taille, nous estimons que s'il a entre 3 et 5 ans, il s'agit d'une erreur d'écriture et nous lui attribuons une taille de 100 cm (1 mètre). Dans les autres cas, nous lui attribuons une valeur manquante.

- Création des variables IMC - ETATIMC

L'IMC se calcule par formule suivante :  $IMC = \text{poids (en kg)} / \text{taille}^2(\text{en mètre})$

ETATIMC est construit à partir de seuils de l'IMC qui dépendent de l'âge et du sexe de la personne. ETATIMC renseigne sur l'état physiologique de la personne.

- = 0 : Dans la norme
- = 1 : En Surpoids
- = 2 : Obèse
- = 3 : Maigre

Ces variables ne sont pas renseignées pour les individus n'ayant pas déclaré leur poids ou leur taille.

Pour les adultes (18 ans et plus), qu'ils soient hommes ou femmes, les seuils définis par l'OMS sont les suivants :

L'obésité correspond à un IMC supérieur ou égal à 30 kg/m<sup>2</sup> (ETATBMI=2)

Le surpoids à un IMC supérieur à 24,9 kg/m<sup>2</sup> et inférieur à 30 kg/m<sup>2</sup> (ETATBMI=1)

La maigreur correspond à un IMC inférieur à 18,5 kg/m<sup>2</sup> (ETATBMI=3)

Pour les enfants et adolescents (entre 2 et 18 ans compris) les normes utilisées tiennent compte de l'âge et du sexe et sont publiées dans les références suivantes :

- Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ* 2000;320:1-6.

- Rolland-Cachera MF, Cole TJ, Sempé M, Tichet J, Rossignol C, Charraud A. Body Mass Index variations : Centiles from birth to 87 years. *European Journal of Clinical Nutrition* 1991;45:13-21.

La variable ETATIMC n'est pas renseignée pour les enfants de moins de 2 ans.

L'âge utilisé pour évaluer ETATIMC des 18 ans est calculé de la façon suivante :

1. calcul de l'âge exact de la personne par différence entre la date de l'interview et sa date de naissance (un enfant né le 15 octobre 2000 et le 15 juin 2006 aura 6 ans et 8 mois)
2. troncature à l'année (on considère qu'il a 6 ans)
3. ajout de 6 mois (ce qui lui fera 6 ans et demi).

Dans ESPS, la variable « AGE » donne l'âge arrondi de la personne (7 ans pour notre exemple).

Pour un enfant d'1 an et dix mois, l'âge arrondi d'ESPS sera de 2 ans alors que l'âge pour calculer l'IMC sera de 1,5 an, ETATIMC ne sera donc pas évalué.

## 2) Question concernant le port de prothèse et de lunettes

### a- Le codage des AUTRES PROTHESES (Q6)

- Au vu du grand nombre de prothèses déclarées dans la catégorie « Autres prothèses », nous avons créé les variables suivantes sur le même mode que les autres :

Prothèse de hanche(s) (prothanch\_GV)

Prothèse de genou(x) (protgenou\_GV)

Prothèse de Coeur (protcoeur\_GV) → stimulateur cardiaque

Pour ces prothèses absentes de la liste initiale, nous considérons que les personnes qui n'en ont pas fait état dans la modalité « Autres prothèses » n'en sont pas porteuses. Nous leur

affectons donc la valeur 2 « Non » aux variables prothanch\_GV, protgenou\_GV et protoeur\_GV.

#### b- Le port de lunettes

- Pour cette question, en ce qui concerne le port de lunette, il convient de noter que les données manquantes sont plus importantes que pour les autres du fait de la formulation de la question. On considère les personnes comme n'ayant pas répondu, celles qui ont coché la modalité « aucune prothèse » ou qui n'ont pas fait état d'aucune autres prothèses ou lunettes pour les autres modalités. Aussi, du fait sans doute d'une mauvaise compréhension/formulation -les lunettes ne semblent pas avoir été considérées comme une prothèse- de la question, beaucoup de personnes n'ont rien coché, ils sont donc manquant alors qu'ils devraient être à 2 a priori.

#### 3) La variable de SANTE PERCUE (Q1)

Si la note déclarée est supérieure à 10, nous considérons la réponse comme manquante.

#### 4) LE MODULE INCAPACITE (Q9-Q14)

Il s'agit de questions qui permettent d'identifier les incapacités (vue, audition, mobilité) dont souffrent les personnes et qui font l'objet d'une sous-question pour savoir si les personnes en souffrent toujours avec un équipement adapté (lunettes, appareil auditif...).

Pour chaque bloc de questions :

- si l'individu n'a pas répondu à la question principale permettant d'identifier l'incapacité, nous vidons l'éventuelle réponse à la sous-question.
- si l'individu a répondu 1 (Oui, sans difficulté) à la question et qu'il a répondu 3 (je n'ai pas d'appareillage), nous conservons la réponse à la question et nous vidons la réponse à la sous question.
- si l'individu a répondu 1 (Oui, sans difficulté) à la question et qu'il a répondu 1 ou 2 (Oui sans difficulté ou non, j'ai toujours des difficultés), nous vidons les réponses à la question et à la sous question.

#### 5) LE RENONCEMENT AUX SOINS (Q16)

Il s'agit d'une question qui permet d'identifier les personnes qui ont renoncé à voir un médecin ou des soins médicaux au cours des 12 derniers mois et d'une sous question qui permet de relever les motifs des renoncements.

- Nous supprimons les incohérences entre la question et la sous question. Si la personne a répondu non à la question et que la sous question est renseignée, alors tout est mis à vide. Si la sous question est renseignée alors que la question ne l'est pas, cette dernière est mise à « 1 : Oui », la personne a renoncé. Si la question est à « 1 : oui » et la sous question n'est pas renseignée, nous n'intervenons pas.

- Concernant les motifs de renoncement, nous codons ou nous remontons les réponses en clair déclarées en modalité 6 « Pour une autre raison, précisez :..... » dans AUTRENONCMED



Au vu du nombre de cas important de certains motifs non proposés explicitement, nous créons les variables suivantes sur le même mode que les autres :

Renoncement par manque de temps (RENON\_PASTPS\_GV)

Renoncement par négligence (RENON\_NEGLI\_GV)

Renoncement à cause d'un problème de transport (RENON\_TRANSP\_GV)

## 6) LE BLOC POUR LES FEMMES DE 16 ANS et PLUS (Q17-Q21)

- Mise à vide des réponses des individus non concernés (femmes de moins de 16 ans et hommes)

- Avec les mêmes règles que précédemment, correction des incohérences des questions sur la mammographie (Q20. bloc MAMMO) et sur le frottis cervico-utérin (Q21. bloc FROTTIS).

- Bloc sur la mammographie

MAMMO : avez-vous déjà eu une mammographie ?

MAMMOGRAT : MAMMO dans le cadre du dépistage organisé.

MAMMOSYMPT : MAMMO en raison de symptômes.

MAMMOSURV : MAMMO pour surveillance.

Quand au moins l'une des trois variables ci-dessus est renseignée, nous affectons la valeur NON (2) à celles qui ne sont pas renseignées. Souvent les enquêtés oublient de cocher le NON.

## 7) LE MODULE SANTE MENTALE (Q25)

Calcul des scores pour les individus de 16 ans et plus qui ont répondu à au moins trois des cinq questions du module.

### Règles de calcul du score

- Nous affectons un poids à chacune des modalités de 1 à 6 (du pire vers le meilleur). En permanence = 1 jusqu'à jamais = 6.

Nous inversons l'ordre des modalités pour CALME et HEUREUX pour les prendre en compte dans le calcul.

- Pour les individus n'ayant répondu qu'à 3 ou 4 questions sur les 5, nous affectons aux valeurs manquantes la moyenne des valeurs moyenne des personnes ayant répondu à ces questions.

Création des scores :

- score d'origine : somme des réponses aux 5 questions du module.
- score transformé :  $((\text{score d'origine} - 5)/25)*100$

## 8) LA PARTIE RESERVEE AUX 16 ANS ET PLUS (Q27-Q33)

Suppression des incohérences dans les questions à filtre avec les règles déjà énoncées.

Mise à vide des réponses des personnes non concernées par ces questions (les moins de 16 ans)

### 9) LE MODULE TABAC (Q28)

- Suppression des incohérences dans les questions à filtres avec les règles déjà énoncées.
- Pour le nombre d'unités fumées, si au moins un des types est renseigné (cigarettes, pipes, cigares), nous mettons ceux qui ne sont pas renseignés à 0.  
Ex : déclaration de 3 cigarettes fumées par jour et aucune information pour la consommation de cigare/cigarillo et pipe, nous admettons alors que l'individu fume 0 cigarillo ou cigare et 0 pipe.

### 10) LE MODULE SANTE RESPIRATOIRE (Q34-Q38)

Les informations sont laissées telles qu'elles sont déclarées par les enquêtés. L'apurement est donc laissé au libre arbitre des utilisateurs de ce module.

## **II. La partie « dernier recours » à un médecin généraliste (Q39-Q54)**

Il s'agit d'un ensemble de questions posées dans le QS pour décrire la dernière séance chez le généraliste de la personne concernée par le questionnaire.

Dans la table QGENE : 14 326 observations

Nous ne conservons que les fiches des individus ayant déclaré avoir consulté un généraliste au cours des 12 derniers mois et ayant répondu au moins à une question.

### 1) SUPPRESSION DES SEANCES HORS CHAMPS

- Les consultations aux urgences.
- Les séances antérieures à 2005.
- Les consultations pour visite médicale du travail.

### 2) LA QUESTION « ETAIT-CE VOTRE MEDECIN TRAITANT ? » (Q42)

- Suppression des réponses des individus de moins de 16 ans car ils n'ont pas à faire de déclaration de médecin traitant.
- Création de la variable NONREEMPLAC\_GV pour les individus ayant vu le remplaçant de leur médecin traitant.

### 3) LE MOTIF DE LA CONSULTATION (Q45)

Il existe deux niveaux de lecture pour le motif de consultation.

- Soit par « maladie » : un code CIM a été affecté pour chacun des motifs médicaux indiqués par la personne dans la modalité 1 (RAISPBSANTE) et éventuellement dans les modalités 2

(RAISANXIEU) et 6 (RAISAUT). Pour les modalités fermées, un code CIM a été attribué (exemple).

- Soit par modalités proposées comme réponse à la question.  
Si au moins un motif est coché, nous affectons la valeur 2 « NON » à tous les autres.

#### 4) LA PRESCRIPTION (Q51)

Si au moins un motif est coché, nous affectons la valeur 2 (NON) à tous les autres.

### III. La partie « dernier recours » à un médecin spécialiste (Q55-Q75)

Il s'agit d'un ensemble de questions posées dans le QS pour décrire la dernière séance chez un spécialiste de la personne concernée par le questionnaire.

Dans la table QSPE : 8 081 observations

Nous ne conservons que les fiches des individus ayant déclaré avoir consulté un spécialiste au cours des 12 derniers mois et qui ont coché au moins une spécialité dans le tableau.

#### 1) LE CODAGE DES AUTRES SPECIALITES (Q55 – modalité 14)

Remontée des spécialités citées dans « Autre » correspondant à des spécialités présentes dans la liste proposée.

Suppression des consultations de praticiens non médecins spécialistes. Si après cette suppression il ne reste plus de consultation pour une personne donnée, alors nous vidons Q55 et nous considérons que cette personne n'a pas répondu à la question.

#### 2) LE CODAGE DE LA SPECIALITE DU DERNIER SPECIALISTE CONSULTE (Q56)

- Suppression des consultations de praticiens non médecins spécialistes.
- Suppression des consultations aux urgences.
- Suppression des séances antérieures à 2005.

Dans ces différents cas, nous considérons que la personne n'a pas répondu à la question. Il s'agit alors d'une séance hors champ que nous ne conserverons pas. L'ensemble des informations concernant cette séance est mis à vide.

- Si la spécialité signalée en Q56 n'avait pas été cochée en Q55, nous la cochons.

#### 3) LE CODAGE DE « QUI A CONSEILLE LA CONSULTATION ? » (Q59)

Création des variables :

- QUI\_TRAVECOL\_GV : Médecine du travail / école / crèche.
- QUI\_PROCHE\_GV : Un proche.
- QUI\_SUIVI\_GV : Suivi régulier pour soins, prévention.
- QUI\_PARA\_GV : Personnel paramédical, social.

#### 4) TEMPS D'ATTENTE ENTRE LA PRISE DE RENDEZ-VOUS ET LA CONSULTATION (Q60)

Création d'une seule variable

- Nous considérons le 0 comme une valeur manquante
- Nous tolérons les doubles déclarations si elles sont cohérentes : 1 semaine + 7 ou 8 j → nous codons 7 jours ou 8 jours / 2 semaines + 14 ou 15 j → nous codons 14 ou 5 jours / 1 mois + 30 j → nous codons 30 jours
- Nous éliminons les autres doubles ou triples déclarations  
⇒ nous créons une variable en jours

#### 5) LE MOTIF DE LA CONSULTATION (Q62)

Il existe deux niveaux de lecture pour le motif de consultation.

- Soit par « maladie » : un code CIM a été affecté pour chacun des motifs médicaux indiqués par la personne dans la modalité 1 (MOTPBSANT) et éventuellement dans les modalités 3 (MOTORDO) et 4 (MOTAUT).
- Soit par modalités proposées comme réponse à la question.  
Si au moins un motif est coché, nous affectons la valeur 2 « NON » à tous les autres.

#### 6) LE NOMBRE TOTAL DE SPECIALISTES CONSULTES (Q55)

Cette information est essentielle pour attribuer une pondération à la consultation retenue pour chaque personne qui est inversement proportionnelle au nombre de consultations.

- Dans le fichier PERSONNE : renseigné uniquement pour les individus ayant déclaré un nombre de séances pour chaque spécialité auxquelles il a déclaré avoir recourues au cours des 12 derniers mois. Dans le cas contraire cette information est mise à vide.
- Dans le fichier QSPE : lorsqu'une personne n'a pas déclaré le nombre de consultations pour une spécialité donnée, nous lui affectons alors le nombre médian observé pour cette spécialité sur l'ensemble de l'échantillon.

## **Règles d'apurement du questionnaire couverture complémentaire**

### 1 – Le Recueil de l'information

Nous recueillons l'information sur le fait que la personne est couverte ou non par une couverture complémentaire (variable CC) lors de l'interview du 1<sup>er</sup> contact (cf. QP Q47-Q53). De cette information découle l'envoi d'autant de questionnaires auto-administrés (FJAUNE) qu'il y a de contrats de couverture complémentaire déclarés par l'interviewé dans le ménage.

### 2 – Enrichissement de l'information par la FJAUNE

Après relecture des questionnaires, la société d'enquête peut noter, au vu des renseignements trouvés sur les FJAUNE, des modifications dans les listes de bénéficiaires des contrats telles qu'elles avaient été déclarées lors du 1<sup>er</sup> contact ou bien dans l'identification de la personne assurée. Ces corrections modifient le statut de certaines personnes vis-à-vis de la couverture complémentaire : assuré / bénéficiaire/ ayant-droit.

3 – Grâce à la relance «Etes-vous certain que [PRENOM] n'est pas couvert par une mutuelle ?», nous pouvons être amenés à modifier les informations de la table MUTUELLE, soit en ajoutant une personne à la liste des bénéficiaires d'un contrat déjà existant, soit en ajoutant un nouveau contrat si [PRENOM] est couvert par une assurance complémentaire qui n'avait pas été déclarée précédemment.

Certaines réponses au questionnement : « Précisez pour quelle autre occasion [PRENOM] a-t-il perdu cette couverture complémentaire ? » nous ont également conduit à ajouter des personnes protégées dans les listes de bénéficiaires, celles bénéficiant en fait toujours de la complémentaire.

## **La notion de prise en charge par une couverture complémentaire santé**

La notion de prise en charge par une couverture complémentaire santé est présente dans le fichier PERSONNE au travers de deux variables : CC (pour la prise en charge par une complémentaire privée) et CMU (pour la prise en charge par la CMU complémentaire).

### La couverture complémentaire privée (CC)

La variable CC est une variable générée à partir des tables MUTUELLE et PASMUT. Dans la table MUTUELLE figurent les individus couverts par une complémentaire privée. MUTUELLE contient une ligne par contrat. Un individu peut donc se retrouver comme assuré ou bénéficiaire dans plusieurs lignes de la table mutuelle. S'il est présent dans au moins une ligne, il est considéré comme pris en charge par une couverture complémentaire privée → CC='1'.

Dans la table PASMUT figurent les individus qui ne sont pas dans le fichier MUTUELLE donc non pris en charge par une complémentaire privée → CC='2'.

Certaines personnes déclarent à l'enquêteur, ne pas savoir si elles sont couvertes ou non. Elles ne figurent ni dans la table MUTUELLE, ni dans le fichier PASMUT → CC='3'.

La couverture maladie universelle (CMU)

La **variable CMU** correspond à ce que l'enquêté a déclaré dans l'interview du premier contact à la question « **Actuellement, avez-vous la CMU complémentaire appelée aussi couverture maladie universelle complémentaire ?** » (1 = oui / 2 = non / 3 = ne sait pas).

Un questionnaire auto-administré FJAUNE est à remplir pour chaque contrat de couverture complémentaire privé. Certaines personnes confondent assurance privée et CMU et remplissent donc FJAUNE a tort. De ce fait, après relecture, certains individus basculent du statut de pris en charge par une couverture complémentaire privée au statut de non pris en charge.

## Règles d'apurement sur les variables sociodémographiques du ménage

### ▪ Les caisses de tirage

Un ménage entre dans l'enquête ESPS parce que l'un de ses membres a été tiré dans une de l'un des trois principaux régimes de Sécurité sociale : 1. la Cnamts pour les travailleurs salariés (hors secteur agricole), 2. la MSA pour les salariés ou les exploitants agricoles et 3. le RSI pour les travailleurs indépendants.

La variable CAISSE est créée par le CENTI. Elle permet de repérer les doublons en historique. Elle identifie les caisses parmi lesquelles un individu a été tiré depuis le début de l'enquête ESPS en 1989. Par exemple, un individu qui aura été tiré dans les fichiers de la CNAMTS depuis le début de l'enquête sera codé « 1 : CNAMTS ». S'il a été tiré dans les fichiers de la CNAMTS jusqu'en 2002, puis dans les fichiers de la MSA à partir de 2006, il sera codé « 4 : Doubleton CNAMTS-MSA ».

La variable REG\_TIRES repère les doublons une année donnée. La valeur « 4 : Doubleton CNAMTS-MSA » dans le fichier de l'enquête 2006 signifie que le ménage a été tiré à la fois dans les fichiers de la CNAMTS et de la MSA cette année là.

### ▪ La variable CHEF :

Cette variable donne le numéro d'individu (NIND) qui a été sélectionné comme « personne de référence » dans le ménage.

Elle sert à déterminer la personne dans le ménage dont nous prenons l'OP (occupation principale) → OPC, et la PCS (catégorie socioprofessionnelle) → PCSC pour décrire le milieu social du ménage.

Au début du 1<sup>er</sup> contact, nous recueillons des informations (tableau Q5-Q12) sur chaque individu du ménage qui permettent de catégoriser le ménage et de choisir une personne de référence (PR) selon des règles définies par l'INSEE.

CAS 0 : Ménage d'une seule personne dans le ménage.

PR = cette personne

CAS 1 : Ménage d'un seul couple.

PR = Conjoint masculin du couple. En cas de couple de même sexe, nous considérons comme « personne de référence » celle qui travaille ou à défaut la plus âgée.

CAS 2 : Ménage de plusieurs couples.

PR = Conjoint masculin, actif le plus âgé ou, à défaut d'actif, le plus âgé.

CAS 3 : Ménage de personnes sans lien familial : pas de couple, ni d'enfant.

Toutes les personnes sont considérées comme des personnes isolées.

PR = Personne active la plus âgée ou, à défaut d'actif, la personne la plus âgée.

CAS 4 : Famille monoparentale simple, l'Assuré Principal (AP) étant le parent.

PR = AP : le père ou la mère des enfants.

CAS 5 : Famille monoparentale simple, l'Assuré Principal (AP) est un enfant.  
PR = Parent de l'AP : le père ou la mère des enfants

CAS 6 : Trois générations et monoparentalité : L'AP a + de 25 ans<sup>1</sup>, vit avec ses enfants, ses frères et sœurs et l'un de ses parents  
Il y a alors deux familles monoparentales, celle de l'AP et celle du parent de l'AP.  
PR = Parmi les deux parents de familles monoparentales, personne active la plus âgée ou à défaut d'actif la personne la plus âgée.

CAS 7 : Trois générations et monoparentalité : L'AP a - de 25 ans, vit avec ses enfants, éventuellement ses frères et sœurs et l'un de ses parents  
Il y a alors deux familles monoparentales, celle de l'AP et celle du parent de l'AP.  
PR = Parmi les deux parents de familles monoparentales, personne active la plus âgée ou à défaut d'actif la personne la plus âgée.

CAS 8 : Trois générations et monoparentalité : L'AP a plus de 25 ans et vit seul avec ses enfants et l'un de ses parents (considéré comme personne isolée pour la détermination de la personne de référence). Il reste donc une seule famille : AP avec ses enfants  
PR = AP : le père ou la mère des enfants

CAS 9 : Trois générations et monoparentalité : AP seul avec ses enfants + un petit-enfant  
Le petit-enfant dont nous ne savons pas qui est le parent est considéré comme une personne isolée pour la détermination de la personne de référence qui est donc soit l'AP, soit un de ses enfants.  
PR = AP : le père ou la mère des enfants sauf si l'AP est retraité et au moins un des enfants est actif. Dans ce cas, la PR est l'actif le plus âgé.

CAS 10 : AP vit seul avec un enfant et plusieurs petits enfants.  
L'AP est alors considéré comme personne isolée, pour la détermination de la personne de référence qui est alors parmi les descendants de l'AP.  
PR = l'enfant de l'AP

CAS 11 : AP vit seul avec un de ses parents et d'autres personnes isolées.  
Pour la détermination de la personne de référence, nous considérons que tous les membres de ce ménage sont des personnes isolées.  
PR = personne active la plus âgée ou à défaut d'actif la personne la plus âgée

CAS 12 : AP vit seul avec ses enfants et d'autres personnes isolées  
PR = AP : le père ou la mère des enfants

▪ **Le type de ménage (TYPMEN) :**

Nous ne tenons pas compte de l'âge pour définir la notion d'enfant. Ainsi, le lien familial direct (parent/enfant) suffit pour déterminer les familles monoparentales ainsi que les couples avec enfants.

Les ménages présents dans la modalité « 7 : Autres » sont dits « complexes ».

---

<sup>1</sup> Seuls les célibataires de moins de 25 ans peuvent être considérés comme des enfants.



Ils sont définis négativement, c'est-à-dire qu'ils ne sont classables dans aucune des 6 autres catégories (homme seule, femme seule, homme seul avec enfant(s), femme seule avec enfant(s), couple avec ou sans enfants).

Les ménages complexes comprennent au moins une famille principale (parent(s)/enfant(s)) à laquelle s'ajoute un ou plusieurs individu(s) isolé(s) (sans liens familiaux directs) ou ils peuvent aussi regrouper des personnes qui n'ont pas de liens familiaux directs (colocataires par exemple).

▪ **Les variables sur le revenu du ménage :**

L'enquête peut déclarer de deux façons différentes son revenu, soit en continu, soit en tranches. Il peut le renseigner soit en francs, soit en euros et soit par mois, soit par année. Au final, les informations dans les tables ESPS sont toutes converties en euros par mois.

- Type de déclaration du revenu :

Par défaut, nous demandons le détail de toutes les sources de revenu du foyer (revenu salarié : EUM\_SALAIRE, pension de retraite : EUM\_RETRAITE, chômage : EUM\_CHOMAGE...).

Au cours de cette déclaration, la personne interrogée peut :

- Ne pas connaître un ou plusieurs éléments du revenu de son ménage.

Dans ce cas, nous l'interrogeons tout de même sur toutes ses sources de revenu afin de recueillir le maximum d'informations possibles. Ensuite, nous proposerons à l'enquête de déclarer son revenu global en tranches → TREVNSP

- Refuser d'indiquer un des éléments du revenu de son ménage.

Les sources de revenu déclarées auparavant sont alors conservées et nous proposons directement à l'enquête de déclarer son revenu global en tranches → TREVREFUS.

Si la personne interrogée a su déclarer les revenus du ménage en détail, nous lui rappelons le récapitulatif des sources de revenus nets du foyer et nous lui demandons de valider le montant total net revenus du ménage (nouveau de 2006).

Deux solutions s'offrent alors à l'enquête :

- Soit il valide ce montant → REVDETAIL
- Soit il ne le valide pas : dans ce cas, nous lui demandons d'indiquer ce qui n'est pas correct dans le récapitulatif précédent. (Variable temporaire : ERREUR)

- Création de la variable REVENU :

Quel que soit le mode de déclaration, les revenus sont convertis en tranches.

La variable REVENU prend la valeur de la variable renseignée entre REVDETAIL converti en tranches, TREVNSP et TREVREFUS sachant que ces dernières sont exclusives, c'est-à-dire dès que l'une d'entre elles est renseignée les autres ne le sont pas.

- Revenu de l'assuré principal (AP) :

En 2006, lorsque l'assuré principal est actif occupé parmi plusieurs actifs occupés dans le ménage (APACTIF « 1 : Oui »), nous demandons à l'enquête le montant net du revenu salarié et/ou d'activité de l'assuré principal que nous souhaitons isolé pour nos études →

EUM\_SALAIREAP. Si la personne interrogée ne connaît pas le montant exact, nous lui proposons alors des tranches de revenu → TSALAIREAP.

Au final, comme précédemment pour le revenu du ménage, ces variables sont toutes converties en euros par mois.

En face-à-face, des problèmes de compréhension de l'enquêteur sont survenus quant à définir si « l'assuré principal est actif occupé parmi plusieurs actifs occupés dans le ménage ». Et cela, malgré la présence de l'information sur le dossier récapitulatif ménage.

Cette erreur a conduit à relever à tort le revenu de l'assuré principal dans des ménages où il était le seul à percevoir un revenu d'activité ou bien à ne pas relever un revenu qui aurait dû l'être dans des ménages où il y a plusieurs actifs.

Au téléphone ce filtre était informatisé et n'a donc pas provoqué les mêmes erreurs.

Cette erreur a servi de contrôle de cohérence entre le revenu de l'assuré principal et les sources de revenu de ménage (cf. en haut de page 20 « Différents tests pour corriger le revenu »).

▪ **Les variables REVUCCMU, REVUCINSEE, REVUCOXFORD :**

Le revenu par unité de consommation est égal au revenu déclaré (en continu → REVDETAIL ou en tranches → TREVNSP ou TREVREFUS) rapporté au nombre d'unités de consommation dans le ménage.

Nous proposons trois modes des *unités de consommation*.

- coefficient CMU :

- Première personne du foyer : coef1 = 1
- Deuxième personne du foyer : coef2 = 0,5
- Troisième et quatrième personne du foyer : coef3 = 0,3
- Personne supplémentaire : coef4 = 0,4

- coefficient INSEE :

- Première personne du foyer : coef1 = 1
- Personne supplémentaire et si âge supérieur ou égal à 14 ans : coef2 = 0,5
- Personne supplémentaire et si âge inférieur à 14 ans : coef3 = 0,3

- coefficient OXFORD :

- Première personne du foyer : coef1 = 1
- Personne supplémentaire et si âge supérieur ou égal à 14 ans : coef2 = 0,7
- Personne supplémentaire et si âge inférieur à 14 ans : coef3 = 0,5

Dénominateur - Nombre d'unité de consommation :

Il est obtenu en sommant les coefficients de chaque individu du ménage.

Numérateur - Revenu déclaré :

Plusieurs situations possibles selon sa déclaration :

- en continu : nous prenons la valeur contenue dans REVDETAIL
- en tranches (revenu issu de TREVNSP ou TREVREFUS), la valeur retenue dépend alors de la classe de revenu :

- Classe inférieure « 01 : Moins de 400 euros » → nous prenons la borne supérieure soit 400 €.
- Classe supérieure « 13 : Plus de 4 600 euros » → nous prenons la médiane des revenus déclarés en continu (REVDETAIL) > 4 600 €.
- Autres classes « XX : de min à moins de max euros » → nous prenons la moyenne des deux bornes.

#### ▪ La variable QUINTUC

Cette variable correspond à la répartition en quintile du revenu mensuel par unité de consommation appliquant le coefficient INSEE (REVUCINSEE).

Cette variable est générée automatiquement ainsi d'éventuelles corrections sur le revenu peuvent modifier les bornes des modalités de cette variable.

Le premier quintile correspond au REVUCINSEE inférieur à 800 €. Le second au REVUCINSEE compris entre 800 € inclus et 1 100 € exclus. Le troisième entre 1 100 € inclus et 1 400 € exclus. Le quatrième entre 1 400 € inclus et 1 866,67 € exclus. Le dernier quintile étant tout REVUCINSEE supérieur ou égal à 1 866,67 €.

Le revenu en continu est souvent déclaré de façon très arrondi par les enquêtés. Il y a donc une concentration importante des valeurs sur des données arrondies et donc une probabilité plus importante de tomber un chiffre rond quand on construit des quintiles.

#### **Apurement réalisé sur le revenu :**

##### - Variable ERREUR

Rappel : cette variable contient la déclaration en clair par l'enquêté de ce qui n'est pas correct dans le récapitulatif des sources de revenus nets du foyer.

Ces informations sont utilisées pour corriger le revenu lorsqu'elles sont interprétables.

La principale erreur est liée au fait que certains enquêtés avaient déclaré les revenus totaux du ménage dans les revenus salariés. Les revenus totaux calculés dans ESPS s'en trouvent donc surévalués puisque les revenus non salariés étaient comptabilisés deux fois.

Lorsque les informations contenues dans ERREUR ne sont interprétables nous avons mis à vide le revenu et le revenu par unité de consommation.

##### - Revenu = « 98 : Non renseigné suite au codage erroné de l'enquêteur »

La création de cette modalité fait suite à une rectification sur certains ménages où l'enquêteur a mal codé les composantes du revenu (refus et ne sait pas).

L'enquêteur avait pour consigne de coder :

- 888888 pour un « refus » alors nous basculions directement à la question en tranches.
- 999999 pour un « ne sait pas » alors nous continuions le questionnaire et nous posions la question du revenu en tranche à la fin.

Cette consigne n'a pas toujours été bien respectée. Certains enquêteurs ont codé 88888 au lieu de 888888, ainsi le questionnaire sur les autres sources de revenu a continué à tort.

Le logiciel ne basculant pas automatiquement dans le questionnaire par tranches, cette erreur a entraîné :

- Soit que le calcul du montant total du ménage est aberrant (calcul tout de même validé en grande majorité par l'enquêté ! (enquêteur ?)).

- Soit cela n'a eu aucune conséquence car il y a un autre refus déclaré ce qui a bien fait basculer vers le questionnement en tranches.

- Différents tests pour corriger le revenu :

- Lorsque l'enquêté a déclaré un revenu total en tranches (suite à un « ne sait pas » ou à un refus à l'une des sources de revenu) inférieur à la somme des éléments déclarés en détail → nous vidons REVENU.

Nous pensons que certains revenus déclarés en tranches (les plus faibles) correspondent, en fait, à la valeur non déclarée des prestations familiales. Il a été décidé d'attendre l'étude sur l'imputation du revenu qui permettra normalement de corriger cette erreur.

- Lorsque le revenu de l'AP est supérieur au revenu de l'ensemble du ménage (en continu ou en tranches)

Plusieurs corrections ont été faites :

- Correction suite à des problèmes d'unités (période de temps considérée – mois ou année, unité monétaire – euros ou francs)
- Lorsque l'AP est le seul actif salarié, que les revenus salariés ou d'activité du ménage n'ont pas été déclarés, mais que le revenu de l'AP est renseigné (alors qu'il n'aurait pas dû l'être).  
→ Nous reportons le revenu salarié de l'AP dans le revenu salarié du ménage et nous renseignons la composante correspondante (SALAIRE = « 1 : oui »). Même raisonnement pour le revenu agricole, industriel, non commercial (AGRICOM = « 1 : oui »)
- Lorsque le ménage est composé de plusieurs actifs, que le revenu de l'AP est renseigné et que le revenu salarié ou d'activité du ménage ne l'est pas ou bien semble incomplet.  
→ Nous vidons REVENU

- Problème des personnes seules avec le revenu de l'AP renseigné.

Cette information a été récupérée suite à l'erreur de déclaration de l'enquêté (vu précédemment). Cette situation a permis de vérifier la cohérence de la déclaration des différentes sources de revenus.

En effet, suite à cette vérification, des corrections ont été apportées sur les problèmes d'unités, ainsi que sur la valeur du revenu salarié (pour les salariés) ou du revenu agricole (pour les exploitants agricoles).

La correction sur le revenu salarié du ménage est soit :

- Un ajout suite à un oubli de déclaration du revenu principal.
- Une modification car l'enquêté avait déclaré le revenu total du ménage lorsque l'interview portait uniquement sur le montant des revenus salariés.

Une fois les différentes corrections effectuées, le revenu de l'AP est vidé sachant que la question n'aurait pas dû être posée et que sa présence n'a plus aucun sens.

- Etude des revenus par unité de consommation extrêmes (REVUCINSEE)

Remarques sur les revenus par UC les plus faibles (2 %), soit :

- Aucun signe apparent d'erreur (sera corrigé sinon par l'imputation). Nous y trouvons principalement des étudiants en location avec les études et le logement financés par les parents.
- Problème d'unités de déclaration (corrigé) : par année au lieu de par mois
- Problème de déclaration de composantes (à revoir après l'imputation) :
  - aucun revenu salarié alors que plusieurs salariés dans le ménage.
  - plusieurs retraités anciens actifs et aucun versement de pension de retraite.

Remarques sur les revenus par UC les plus élevés (1 %) soit :

- Le montant semble correct car nous sommes en présence de métiers rémunérateurs
  - Problème d'unités de déclaration (corrigé):
    - pour les actifs, erreur avec €/mois au lieu de €/a
    - pour les retraités, erreur avec €/mois au lieu de F/mois
- Revenu de l'AP a la même valeur avant conversion en €/mois qu'une des composantes du ménage mais l'unité (tps, monétaire) diffère.

Par exemple, l'AP actif salarié avait déclaré 1 200 F/mois pour les revenus salariés du ménage et 1 200 €/mois pour son revenu. Avec ces informations, nous pouvons donc corriger l'unité erronée. Une fois la correction effectuée, nous nous attardons sur le fait que le revenu de l'AP ne devait être donné que s'il était actif parmi plusieurs actifs dans le ménage (APACTIF = « 1 : Oui »). Ainsi, lorsque l'AP est la seule personne active : nous corrigeons APACTIF avec la valeur « 2 : non » ce qui conduit également à supprimer la partie sur le revenu de l'AP. En revanche, lorsque l'AP n'est pas le seul actif salarié et qu'il n'y a eu que son salaire déclaré alors REVENU est vidé du fait qu'il manque le revenu salarié des autres membres actifs du ménage.

- Revenu de l'AP a une valeur nettement inférieure à la composante principale à laquelle il devrait appartenir

Ex :  $EUM\_salaireap * 6 < EUM\_salaire$  (si l'AP est actif salarié), le facteur 6 sert à tenir compte de l'erreur mois au lieu d'année qui a pour ratio 12 et l'erreur Francs au lieu d'euros dont le ratio est de 6,56. Situation rencontrée :

- Problème d'unités de déclaration (valeur, tps, monétaire)
- Aucune erreur, le revenu total des autres membres est plus important que le revenu de l'AP
- L'AP n'est pas actif occupé → apactif = « 2 : non »
- Le revenu de l'AP semble faible et sera éventuellement corrigé lors de l'imputation du revenu