



*Bundesamt für Sozialversicherung  
Office fédéral des assurances sociales  
Ufficio federale delle assicurazioni sociali  
Uffizi federal da las assicuranzas socialas*

Dossier technique

*Données administratives de  
l'assurance-maladie:  
Analyse de qualité, statistique  
Élémentaire et base pour les  
exploitations*

Rapport de recherche n° 9/98

**ASPECTS DE LA SECURITE SOCIALE**

**Rédaction:** Gabriel Sottas

**Avec la collaboration de:** Till Bandi, Mehdi Chikhi, Hans-Peter Ittig (BFI),  
Herbert Känzig, Peter Lutz, Stefan Schütz

**Renseignements:** Office fédéral des assurances sociales  
Section statistique  
3003 Berne  
Gabriel Sottas, tél. 031 / 324 06 90

**Diffusion:** Office central fédéral des imprimés et du matériel  
(OCFIM), 3000 Berne

**Copyright:** Office fédéral des assurances sociales  
CH – 3003 Berne, Novembre 1997

Reproduction d'extraits autorisés – excepté à des fins  
commerciales – avec mention de la source; copie à  
l'Office fédéral des assurances sociales.

**Numéro de commande:** 318.010.9/98 f 1.98 300

***Données administratives de  
l'assurance-maladie:  
Analyse de qualité, statistique  
élémentaire et base pour les  
exploitations***

***Editeur: Office fédéral des assurances sociales***

## Table des matières

<b>Table des matières</b> .....	<b>i</b>
<b>Liste des tables</b> .....	<b>iii</b>
<b>Liste des figures</b> .....	<b>v</b>
<b>1 Introduction</b> .....	<b>1</b>
<b>2 Données administratives et validation</b> .....	<b>3</b>
2.1 Description succincte de la structure des données .....	3
2.2 Concepts généraux de validation .....	4
2.3 Description sommaire de la phase 1 de la validation .....	4
2.4 Corrections spéciales .....	5
2.5 Intégration des données des HMOs avec celles des caisses-maladie .....	6
<b>3 Données administratives disponibles – Etat au 31 mars 1996</b> .....	<b>9</b>
3.1 Données administratives à fournir .....	9
3.2 Données administratives des caisses-maladie .....	9
3.3 Données administratives des HMOs.....	11
<b>4 Analyse de qualité des données</b> .....	<b>13</b>
4.1 Etat des livraisons .....	13
4.1.1 Données de type assuré.....	13
4.1.2 Données de type prestation.....	16
4.2 Correction du nombre de blocs d'informations .....	17
4.2.1 Correction du nombre de branches d'assurance.....	17
4.2.2 Correction du nombre de composantes de frais.....	20
4.2.3 Suppression de composantes de frais vides .....	20
4.3 Correction des identificateurs des assurés.....	21
4.3.1 Données de type assuré.....	22
4.3.2 Données de type prestation.....	22
4.4 Correction des branches d'assurance .....	23
4.4.1 Correction du genre d'assurance .....	23
4.4.2 Correction des valeurs des franchises .....	24

4.5	Regroupement des enregistrements de type assuré correspondant à une même personne .....	25
4.6	Intégration des données des HMOs.....	27
4.7	Validation de l'âge d'entrée pour les branches d'assurance .....	27
4.8	Validation des durées de risque .....	28
4.9	Validation des dates de début et de fin de prestation .....	29
4.10	Remarque générale.....	29
<b>5</b>	<b>Statistique élémentaire des données administratives.....</b>	<b>31</b>
5.1	Effectifs des assurés par types d'assurance .....	31
5.1.1	Effectifs des assurés en 1990.....	33
5.1.2	Effectifs des assurés en 1991 .....	33
5.1.3	Effectifs des assurés en 1992.....	35
5.1.4	Effectifs des assurés en 1993.....	36
5.1.5	Effectifs des assurés en 1994.....	36
5.2	Primes par types d'assurance .....	36
5.2.1	Répartition des primes en 1990.....	37
5.2.2	Répartition des primes en 1991.....	37
5.2.3	Répartition des primes en 1992.....	39
5.2.4	Répartition des primes en 1993 et 1994 .....	39
5.3	Répartition des coûts par branches d'assurance et par genres de cas .....	39
5.3.1	Analyse des coûts pour l'année 1990.....	43
5.3.2	Analyse des coûts pour l'année 1991 .....	43
5.3.3	Analyse des coûts pour l'année 1992.....	44
5.3.4	Analyse des coûts pour l'année 1993.....	44
5.3.5	Analyse des coûts pour l'année 1994.....	44
<b>6</b>	<b>Base pour les exploitations .....</b>	<b>47</b>
6.1	Choix de l'échantillon .....	47
6.1.1	Période 1991 – 1994 .....	47
6.1.2	Période 1989 – 1990 .....	48
6.2	Représentativité de l'échantillon par rapport à l'ensemble des 25 caisses participantes — période 1991–1994 .....	48
6.3	Comparaison des données de l'échantillon avec celles de la statistique de l'assurance-maladie 1994 .....	50
6.3.1	Effectifs des assurés par type d'assurance pour soins médico-pharmaceutiques .....	52
6.3.2	Primes par type d'assurance pour soins médico-pharmaceutiques.....	54
6.3.3	Prestations à la charge de l'assurance pour soins médico-pharmaceutiques .....	55

---

## Liste des tables

3.1 :	Données administratives des caisses.....	10
3.2 :	Données administratives de type prestation des HMOs.....	12
5.1 :	Les niveaux de franchise.....	32
5.2 :	Les niveaux de bonus.....	33
5.3 :	Effectifs des assurés par types d'assurance.....	34
5.4 :	Primes par types d'assurance.....	38
5.5 :	Statistique globale des coûts.....	41
5.6 :	Coûts par branches d'assurance.....	42
5.7 :	Coûts par genre de cas.....	43

## Liste des figures

4.1 :	Répartition annuelle relative des 25 caisses-maladie selon l'état des données de type assuré. ....	14
4.2 :	Répartition annuelle relative des assurés des 25 caisses-maladie selon l'état des données de type assuré. ....	15
4.3 :	Répartition annuelle relative des 25 caisses-maladie selon l'état des données de type prestation. ....	18
4.4 :	Répartition annuelle relative des factures des 25 caisses-maladie selon l'état des données de type prestation. ....	19
4.5 :	Part proportionnelle des enregistrements de type assuré ayant subi une correction du nombre de branches d'assurance. ....	20
4.6 :	Part proportionnelle des enregistrements de type prestation ayant subi une correction du nombre de composantes de frais. ....	21
4.7 :	Part proportionnelle des enregistrements de type prestation comportant des composantes de frais vides. ....	22
4.8 :	Part proportionnelle des enregistrements ayant subi une correction de l'identificateur de l'assuré. ....	23
4.9 :	Part proportionnelle des enregistrements de type assuré ayant subi des corrections du genre d'assurance. ....	24
4.10 :	Part proportionnelle des enregistrements de type assuré ayant subi des corrections des valeurs des franchises. ....	25
4.11 :	Part proportionnelle des assurés pour lesquels des enregistrements de type assuré ont été regroupés. ....	26
4.12 :	Taux de réussite pour le regroupement des enregistrements de type assuré appartenant à un même assuré. ....	26
4.13 :	Taux de réussite lors de la conversion des identificateurs d'assurés dans les données de type prestation des HMOs. ....	28
6.1 :	Proportion des assurés avec une assurance pour soins médico-pharmaceutiques par rapport à l'ensemble des membres. ....	49
6.2 :	Répartition par type d'assurance des assurés en soins médico-pharmaceutiques des 25 caisses. ....	50
6.3 :	Répartition par type d'assurance des assurés en soins médico-pharmaceutiques — comparaison de l'échantillon avec l'ensemble des caisses. ....	51
6.4 :	Différences entre les données de l'échantillon et celles de la statistique de l'assurance-maladie — Effectifs des assurés par branches pour 1994. ....	52
6.5 :	Différences entre les données de l'échantillon et celles de la statistique de l'assurance-maladie — Effectifs des assurés "franchise à option" pour 1994. ....	53
6.6 :	Différences entre les données de l'échantillon et celles de la statistique de l'assurance-maladie — Primes par branches pour 1994. ....	55
6.7 :	Différences entre les données de l'échantillon et celles de la statistique de l'assurance-maladie — Prestations par type de soins pour 1994. ....	56
6.8 :	Différences entre les données de l'échantillon et celles de la statistique de l'assurance-maladie — Prestations par type de soins pour 1994 et pour un sous-ensemble de l'échantillon constitué de 11 caisses. ....	57

---

# **DONNÉES ADMINISTRATIVES DE L'ASSURANCE-MALADIE :**

## **ANALYSE DE QUALITÉ, STATISTIQUE ÉLÉMENTAIRE ET BASE POUR LES EXPLOITATIONS**

### **1 Introduction**

Les données administratives de l'assurance-maladie sont composées d'un ensemble de données individuelles livrées par les caisses-maladie à l'OFAS. Ces données individuelles regroupent des informations relativement détaillées sur l'assuré, sa couverture d'assurance et les primes relatives à cette couverture ainsi que sur les prestations médicales pour lesquelles il a introduit une demande de remboursement auprès de son assureur. Précisons d'emblée que l'identité de l'assuré demeure anonyme (il n'est connu que par un identificateur numérique établi par le fournisseur des données) et qu'aucune information concernant les diagnostics n'est communiquée à l'OFAS. Toutes les données rassemblées ici proviennent des banques de données des caisses-maladie concernées. Ainsi, contrairement à d'autres études empiriques, une collecte spécifique des données n'a pas été nécessaire; seule une extraction de ces données et une mise en forme appropriées a été effectuée par les fournisseurs<sup>1</sup>.

Les données administratives doivent être livrées par chaque caisse participant à l'évaluation des nouvelles formes d'assurance-maladie et pour chacune des années comptables 1990 à 1995. Pour les cinq années antérieures à la période d'essai des nouvelles formes d'assurance (1985 à 1989), ces caisses devaient également fournir des données administratives, mais avec moins de détails. En raison de la structure organisationnelle particulière des HMOs, les coûts des traitements effectués dans ces institutions ne sont pas toujours transmis aux caisses-maladie. Dans un tel cas, pour pallier la perte d'information au niveau des caisses, les HMOs eux-mêmes doivent fournir directement à l'OFAS les données administratives les concernant.

Cette base de données administratives recouvre actuellement les cinq années 1990 à 1994. Pour chaque année, elle regroupe des informations concernant approximativement 5 millions d'assurés et,

---

<sup>1</sup> Ce qui représente déjà un travail non négligeable pour les assureurs.

par voie de conséquence, concernant approximativement 25 à 30 millions de factures de prestations médicales traitées par des caisses-maladie. La constitution d'une telle banque de donnée ne va pas sans poser de problèmes. En particulier, il est clair que les données administratives reçues des caisses et des HMOs peuvent comporter une multitude d'erreurs et d'imprécisions. De ce fait, il est indispensable de disposer d'un outil de validation automatique et systématique de ces données permettant de détecter ces problèmes. L'extrême complexité des données à traiter n'a pas encore permis de mettre complètement au point un tel outil. Cependant, les validations effectuées jusqu'à présent ont déjà mis en évidence un grand nombre d'erreurs, erreurs qui, dans la mesure du possible, ont été corrigées soit par les fournisseurs eux-mêmes, soit par nos soins. Ainsi, bien que ce processus de validation ne soit pas encore complètement opérationnel, les données sont déjà dans un état permettant de les exploiter, à condition de le faire avec certaines précautions.

L'objectif du présent rapport est de donner un aperçu de la qualité des données administratives reçues par le biais des résultats des validations opérées jusqu'ici. Observons aussi que des validations autres que celles rapportées ici ont été effectuées en fonctions des besoins spécifiques des diverses études statistiques en cours sur ces données administratives. Les résultats de ces validations ne seront pas intégrés dans ce compte-rendu.

Après une présentation succincte de la structure des données, du concept de validation systématique des données et de l'état actuel de la mise en oeuvre de ce concept (section 2), nous dresserons un panorama détaillé des données administratives actuellement en notre possession (section 3). L'analyse de la qualité des données (section 4) débutera par une étude de l'état des données disponibles incluant l'histoire des livraisons<sup>1</sup>. Cette analyse de qualité se poursuivra ensuite par l'examen des principales corrections qui ont été effectuées par nos soins sur ces données (spécialement de l'ampleur de ces corrections) ainsi que par une étude des problèmes posés par l'intégration des données des HMOs dans les données des caisses-maladie.

Afin de ne pas limiter notre horizon uniquement à des questions purement techniques de codage des données, quelques analyses simples du contenu des fichiers de données ont également été réalisées. Un résumé des résultats de ces analyses complètera cet aperçu qualitatif du contenu des données administratives (section 5).

Pour terminer, nous aborderons la question de la définition d'un échantillon de données (au sens d'un échantillon de caisses-maladie) utilisable pendant toute la période couverte par notre étude (ce qui aura également pour conséquence la définition d'une période d'étude plus restreinte) et permettant des comparaisons multi-annuelles (section 6)<sup>2</sup>. Une comparaison des données de cet échantillon avec les données de la statistique de l'assurance-maladie sera également présentée.

---

<sup>1</sup> Cet historique permet de différencier les données directement utilisables sans problèmes majeurs des données qui ont dû être refaites, éventuellement à plusieurs reprises, par les fournisseurs avant qu'elles ne soient utilisables.

<sup>2</sup> Cet échantillon est constitué des données des caisses-maladie pour lesquelles les données administratives sont disponibles et utilisables pour toute la période étudiée. Ainsi, l'ensemble des assurés considéré peut varier d'une année à l'autre. Il ne s'agit donc pas d'un échantillon d'assurés au sens statistique du terme.

## 2 Données administratives et validation

### 2.1 Description succincte de la structure des données

Les données administratives sont subdivisées en deux groupes bien distincts. Nous avons tout d'abord les données dites de "type assuré" qui regroupent l'ensemble des informations relatives à l'assuré et à sa couverture d'assurance. Dans ce cas, nous avons un seul enregistrement par assuré. Cet enregistrement est constitué tout d'abord d'une partie de longueur fixe contenant une description de l'assuré (notamment le sexe, l'âge et la région de domicile). A cela s'ajoute une description de la couverture d'assurance de l'assuré, sous la forme d'un bloc de donnée pour chaque branche d'assurance souscrite. Le nombre de tels blocs dépend de la couverture effective de l'assuré. Chaque bloc est constitué des mêmes informations, il contient en particulier le type de branche<sup>1</sup>, la durée de la couverture et la prime encaissée pour cette branche.

Le deuxième groupe de données réunit les informations sur les coûts des prestations médicales pour lesquelles une demande de prise en charge a été formulée auprès d'un assureur maladie. Il s'agit des données dites de "type prestation". Chaque facture donne lieu à la génération d'un enregistrement de données. Ainsi, dans ce cas, il peut y avoir plusieurs enregistrements relatifs au même assuré. Du point de vue technique, ces enregistrements sont également constitués d'une partie de longueur fixe à laquelle s'ajoute un nombre variable de blocs. Dans la partie de longueur fixe, nous trouvons l'identification de l'assuré concerné, des informations sur le type de fournisseur de prestations (médecin, pharmacie, hôpital, etc.), le montant de la facture et le montant de la participation de l'assuré. Selon les informations à disposition des caisses-maladie, le montant total de la facture est ventilé de façon plus ou moins détaillée en composantes de frais<sup>2</sup>. Chaque composante de frais (type et montant pris en charge) est décrite dans un des blocs de données ajoutés à la suite de la partie générale.

Chaque caisse-maladie participant à l'évaluation doit livrer, pour chacune des années 1990 à 1995, les données décrites ci-dessus, regroupées par type dans deux fichiers distincts. Dans le but de déceler d'éventuels changements de comportement vis à vis de la consommation de soins médicaux parmi les assurés qui ont opté pour une assurance de base de type assurance avec système de bonus ou avec choix limité du fournisseur de prestations, des données relatives aux cinq années précédente, soit 1985 à 1989, doivent également être fournies. Cependant, plutôt que des données détaillées sur les prestations reçues, seules des données globales relatives aux coûts ambulatoires et hospitaliers sont exigées dans ce cas. Ces données sont dites de "type coûts". De plus, pour cette

---

<sup>1</sup> Par exemple : assurance de base avec franchise à option, assurance d'indemnités journalières ou assurance complémentaire d'hospitalisation.

<sup>2</sup> Par exemple : prestations de l'assurance de base pour traitement hospitalier, médicaments, prestation de l'assurance combinée en cas d'hospitalisation.

période, aucune information sur l'assuré et, en particulier, sur sa couverture d'assurance n'est demandée.

Pour des raisons de simplification, nous nous limiterons dès à présent aux problèmes des données de type assuré et de type prestation. Ces données sont d'ailleurs celles qui, par leur complexité, présentent le plus de problèmes.

## **2.2 Concepts généraux de validation**

La validation est divisée en deux phases. Au cours de la première, les fichiers d'assurés et de prestations subissent des contrôles de base afin de voir si leur format est bien compatible avec les spécifications. Les enregistrements comportant des erreurs sont soit corrigés, soit éliminés car inutilisables pour les évaluations. Les données de type assuré et celles de type prestation sont traitées de manière indépendante. De ce fait, aucune validation croisée des données d'un fichier de prestations avec celles du fichier d'assurés correspondant n'est effectuée à ce stade de la validation.

La deuxième phase est une phase de regroupement et d'analyse de détail. Les données des champs sont alors examinées en détail afin de voir si elles sont présentes et si elles sont conformes aux spécifications de détail. De plus, dans certains cas, la compatibilité des données de divers champs est vérifiée. Les erreurs détectées sont corrigées dans la mesure du possible. Lorsqu'aucune correction ne peut être définie, un code d'erreur correspondant est défini.

Actuellement, seule la première phase est réalisée de manière systématique. La deuxième n'est réalisée que partiellement, en relation avec les besoins spécifiques des diverses exploitations des données en cours. Le panorama de la qualité des données que nous présentons dans la suite de ce chapitre se fonde principalement sur les résultats de cette première phase de la validation. Pour la compréhension de ce qui va suivre, il est maintenant nécessaire de présenter un peu plus en détail le travail effectué au cours de cette phase 1.

## **2.3 Description sommaire de la phase 1 de la validation**

Cette phase consiste tout d'abord à examiner si tous les champs numériques d'un enregistrement ne contiennent que des caractères numériques ou blancs. Comme rappelé ci-dessus, les deux types d'enregistrements se composent d'une partie générale complétée d'un ensemble de blocs de données dont le nombre varie. Ce nombre, qui est spécifié dans la partie générale, est comparé avec le nombre de parties effectivement présentes. Les blocs de données complètement vides sont également détectés puis éliminés. Tous les enregistrements pour lesquels aucune erreur n'est détectée ou pour lesquels tous les problèmes ont pu être corrigés sont regroupés dans un même fichier de sortie. Les enregistrements pour lesquels des erreurs n'ont pas pu être corrigées sont, pour leur part, augmentés d'un code d'erreur et groupés dans un autre fichier.

Les fichiers d'assurés ne devraient contenir qu'un enregistrement par membre, ce qui n'est malheureusement pas toujours le cas. Ainsi, nous avons été amenés à détecter les assurés pour lesquels

plusieurs enregistrements existent puis à analyser (électroniquement) ces enregistrements. Si le résultat de cette analyse montre que les informations des divers enregistrements sont cohérentes entre elles, le groupe d'enregistrements est remplacé par un enregistrement unique contenant une synthèse de ces informations. Si, au contraire, le groupe comporte des incohérences, il est supprimé des données (avec pour conséquence la suppression d'au moins un assuré). Ainsi, après ce traitement, les fichiers d'assurés ne contiennent plus qu'un seul enregistrement par assuré.

Une fois effectués les contrôles ci-dessus, une analyse sommaire du contenu des fichiers est encore exécutée. Les résultats de ces analyses permettent de détecter (manuellement) un certain nombre d'autres problèmes dans les données, problèmes auxquels il est ensuite remédié, dans la mesure du possible, à l'aide de corrections spéciales<sup>1</sup>. Les principaux problèmes ainsi corrigés sont décrits dans la section suivante.

## 2.4 Corrections spéciales

Le travail avec les fichiers de données peut faire ressortir des erreurs systématiques, erreurs ne faisant pas partie des problèmes testés au cours de la phase 1 de la validation. Certaines de ces erreurs font l'objet de procédures spéciales de correction. Il s'agit ici d'un travail spécifique à chacune des caisses traitées.

Le premier type de problème traité concerne l'identificateur assuré. Plusieurs variantes ont été observées. La première est d'ordre purement informatique : l'identificateur est cadré à droite de son champ pour les assurés et à gauche pour les prestations ou vice et versa. Ce cas est facile à traiter et ne nécessite que la correction de l'un ou l'autre type de fichier, selon la version de l'identificateur retenue. Il est à noter que cette correction est effectuée de sorte que la cohérence des valeurs soit préservée au cours des années. La deuxième variante observée est liée au processus de gestion interne des caisses. Dans certains cas, l'identificateur des assurés comporte une portion spécifique à l'assuré et une portion liée à la sous-caisse ou à la section à laquelle cet assuré est affilié. Cette façon de faire n'est pas conforme avec le principe d'unicité de l'identificateur et de l'assuré puisque lorsque l'assuré est transféré d'une sous-caisse dans une autre au cours de l'année, son identificateur change. Ceci a pour conséquence la production de deux enregistrements (avec identificateurs différents) concernant la même personne physique dans les données assurés. De plus, dans certains cas, la portion de l'identificateur spécifiant la sous-caisse n'est remplie que pour les enregistrements de type assuré (ce qui complique le regroupement des assurés et des prestations). La correction de ce problème consiste à annuler la portion de l'identificateur liée à la sous-caisse (puis à traiter le problème des éventuels membres avec plusieurs enregistrements de type assuré). Une troisième variante de problèmes affectant les identificateurs assurés a également été observée : un digit de cet identificateur change chaque année (en fait, il est égal au dernier digit de l'année). Ceci complique le suivi des assurés d'une année à l'autre et, par conséquent, a également été corrigé.

---

<sup>1</sup> Par corrections spéciales, on entend ici les corrections d'erreurs ou de problèmes observés dans les données de type assuré ou prestation (corrigeables au niveau de l'OFAS!), à l'exclusion des erreurs détectées automatiquement avant la procédure d'analyse sommaire des données.

Un deuxième type de problèmes systématiques observé et faisant l'objet de traitement ici affecte les données de type assuré. Le premier champ de chacun des blocs de données décrivant la couverture d'assurance de l'assuré sert à identifier la branche d'assurance concernée par le bloc auquel il appartient. Dans certaines caisses, un nombre important de tels champs demeurés vides a été observé. Après discussion avec ces caisses, il est apparu que ces blocs devaient être soit attribués à un type précis d'assurance de soins médico-pharmaceutiques (par affectation systématique de la valeur correspondante au champs concernés), soit purement et simplement supprimés car ces blocs reflètent des branches d'assurance suspendues. Une autre classe d'erreurs liée à ces champs est l'erreur d'affectation comme par exemple la présence d'assurés bonus en 1990 alors que la caisse n'a introduit ce mode d'assurance qu'en 1991. Ce type de problème fait également l'objet d'un traitement approprié.

Une autre source de problèmes provient des valeurs des franchises. Une caisse a, par exemple, établi les données assurés en affectant des franchises de type adulte également aux jeunes entre 15 et 20 ans<sup>1</sup>. Comme la combinaison âge et montant de la franchise sert à distinguer les assurés en franchise à option des assurés en mode traditionnel, cette erreur nécessite également une correction. D'autre part, dans le cas de quelques caisses, l'analyse des assurés révèle la présence d'un trop grand nombre d'enfants avec une franchise à option de niveau 2 (plusieurs dizaines de milliers en lieu et place de quelques centaines par exemple)<sup>2</sup>. Dans un tel cas, il a été jugé préférable de transférer de franchise à option niveau 2 en traditionnelle (par annulation des franchises) l'ensemble de ces enfants. Cette correction n'est évidemment pas 100 % correcte mais elle permet d'avoir un nombre beaucoup moins élevé d'assurés mal classés.

L'énumération ci-dessus n'est bien évidemment pas exhaustive. Elle se limite à présenter les principaux problèmes observés et corrigés ici.

## 2.5 Intégration des données des HMOs avec celles des caisses-maladie

Une des particularités des HMOs est de ne pas obligatoirement adresser de facture aux caisses-maladie pour les soins dispensés au sein de l'organisation elle-même. Cependant, les informations concernant ces prestations sont également nécessaires pour le bon déroulement de l'étude en cours. Par conséquent, elles doivent être mise à disposition de l'OFAS directement par les HMOs.

Ces données de type prestation subissent également le même processus de validation (phase 1 décrite ci-dessus) que les données provenant des caisses-maladie. Lorsque ce processus est terminé, les données qui en résultent sont réparties en sous-groupes selon les caisses-maladie auxquelles sont affiliés les membres du HMO. Ces sous-groupes sont alors intégrés aux données de type prestation des caisses avant que ne débute la phase 2 de la validation<sup>3</sup>.

---

<sup>1</sup> En principe, le passage des franchises enfant aux franchises adultes intervient au cours du mois ou au début de l'année qui suit l'accomplissement de la 20e année.

<sup>2</sup> Par comparaison avec les données reçues par l'OFAS en vue de l'établissement de la statistique de l'assurance-maladie.

<sup>3</sup> Mais une fois que la phase 1 de la validation des données des caisses concernées est terminée.

Pour que ces données puissent être intégrées aux données de type prestations reçues des caisses-maladie, il fallait que les HMOs identifient leurs patients à l'aide des identificateurs (anonymisés) définis par les caisses-maladie auxquelles ceux-ci sont affiliés. Malheureusement, ceci n'a pas été le cas, les HMOs ayant défini leurs propres identificateurs. Pour nous permettre de résoudre cette difficulté, les HMOs nous ont alors livré une table de conversion. Ces tables ont permis de remplacer les identificateurs définis par les HMOs par ceux définis par les caisses dans les enregistrements de type prestation reçus directement des HMOs. Cependant, pour que cette conversion soit réalisable à l'aide d'un processus informatique, il était nécessaire que les codes soient formatés de manière cohérente au sein de chaque table. Ceci n'étant, et de loin, pas le cas, il fallut également remédier à cette difficulté par un long et fastidieux retraitement manuel.

Au cours du processus d'intégration des données des HMOs aux données des caisses, nous avons pris la précaution de contrôler que les assurés concernés avaient bien une assurance de soins médico-pharmaceutiques de type HMO. Lorsque tel n'était pas le cas, deux situations pouvaient se présenter. Si l'assuré possède une assurance de soins d'un autre type et est domicilié dans le bassin de recrutement naturel du HMO, le type d'assurance de cet assuré est corrigé en HMO<sup>1</sup>. Dans le cas contraire<sup>2</sup>, les données d'assurance de cet assuré ne sont pas modifiées. Dans tous les cas, les données de type prestation fournies par le HMO ont été intégrées aux données de la caisse-maladie.

---

<sup>1</sup> Dans les données de type assuré reçues de la caisse.

<sup>2</sup> Dans ce cas, soit l'assuré n'a pas d'assurance pour soins médico-pharmaceutiques (décrite dans les données), soit son domicile ne se trouve pas dans la région du HMO.

### **3 Données administratives disponibles – Etat au 31 mars 1996**

#### **3.1 Données administratives à fournir**

Comme nous l'avons déjà mentionné précédemment, chaque caisse participant à l'évaluation doit fournir des données administratives de type assuré et de type prestation, et ce pour chacune des années comptables 1990 à 1995. Ces informations sont fournies pour chaque caisse et chaque année dans deux fichiers distincts. Pour les cinq années antérieures à la période d'essai des nouvelles formes d'assurance, les caisses devaient également fournir des données administratives, mais cependant moins détaillées. Ainsi, pour cette période, les caisses avaient le choix entre la fourniture de données de type prestation pour chacune des années ou de type coûts<sup>1</sup>. Les données de type assuré n'étaient pas demandées pour ces cinq années.

D'autre part, les HMOs pour lesquels les coûts des traitements effectués en leur sein ne sont pas transmis aux caisses-maladie doivent fournir directement à l'OFAS les données administratives de type prestation.

#### **3.2 Données administratives des caisses-maladie**

17 caisses-maladie participent à l'expérimentation des nouvelles formes d'assurance-maladie (HMO et assurances avec système de bonus) et sont donc soumises à l'obligation de fournir à l'OFAS des données administratives concernant l'ensemble de leurs assurés<sup>2</sup>. Il s'agit des assureurs suivant : CSS Versicherung, Krankenkasse Helvetia, Schweizerische Grütli, Krankenkasse des Schweizerischen Kaufmännischen Verbandes (KK SKV), OSKA Kranken- und Unfallversicherung, KONKORDIA Schweizerische Kranken- und Unfallversicherung, l'AVENIR Assurances, OeKK Luzern, Caisse-maladie HERMES, OeKK Basel, KFW Winterthur Schweizerische Kranken- und Unfallversicherung, Schweizerische Kranken- und Unfallkasse ZOKU, OeKK Winterthur, Mutuelle Valaisanne Caisse-maladie, Krankenkasse KBV, Caisse-maladie FUTURA et Caisse-maladie UNIVERSA.

Du fait de l'extension du projet à l'évaluation des assurances avec franchise annuelle à option, 8 autres caisses participent également, à titre bénévole, à l'enquête et mettent à disposition de l'OFAS leurs données administratives. Dans ce cas, nous avons : SUPRA Caisse-maladie, Krankenkasse KKB, La Caisse Vaudoise Assurance en cas de maladie et d'accidents, INTRAS Caisse-maladie,

---

<sup>1</sup> Voir la section 2.1 pour une description succincte des divers types de données.

<sup>2</sup> Cette obligation résulte de l'Ordonnance V sur l'assurance-maladie (article 23<sup>quinq</sup>) et de l'Ordonnance 14 du DFI sur l'assurance-maladie.

Table 3.1 : Données administratives des caisses.

Dans les en-têtes des colonnes, M désigne les données de type assuré et S celles de type prestation. Les lignes ombrées indiquent les caisses soumises à l'obligation de fournir les données administratives.

Le symbole "✓" désigne les fichiers reçus et validés tandis que "#" désigne les fichiers reçus et non validés. "X" désigne les fichiers devant être refaits par le fournisseur en raison d'une qualité des données insuffisante et "†" marque les fichiers que les fournisseurs ne sont plus en mesure de produire ou de corriger.

Caisse	Nr. OFAS	'89	1990		1991		1992		1993		1994	
		S	M	S	M	S	M	S	M	S	M	S
CSS	0008	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Helvetia	0058	#	†	†	✓	✓	✓	✓	✓	✓	✓	✓
SUPRA	0062			✓								
KKB	0063	†	†	†	✓	✓	✓	✓	✓	✓	✓	✓
Grütli	0124	✓ <sup>1</sup>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
KK SKV	0131	✓ <sup>1</sup>	X	✓ <sup>1</sup>	X	X	✓	X	X	X		
OSKA	0245	†	†	†	†	†	✓	✓	✓	✓	✓	✓
KONKORDIA	0290	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
L'AVENIR	0343		X	X	X	X	X	X	X	X	X	X
OeKK Luzern	0411	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
HERMES	0445	# <sup>1</sup>	✓		✓		✓		✓		✓	
OeKK Basel	0509	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
La Caisse Vaudoise	0749	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
KFW	1060	✓ <sup>1</sup>	†	✓ <sup>1</sup>	✓	✓	✓	✓	✓	✓	✓	✓
ZOKU	1089	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
OeKK Winterthur	1305	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Mutuelle Valaisanne	1479	# <sup>1</sup>	✓		✓		✓		✓		✓	
KBV	1519		X	X	X	X	X	X	X	X	X	X
INTRAS	1529	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
FRV	1531	✓	†	✓	†	✓	†	✓	†	✓	✓	✓
ASSURA	1542	✓ <sup>1</sup>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
FUTURA	1549	# <sup>1</sup>	✓		✓		✓		✓		✓	
UNIVERSA	1551	# <sup>1</sup>	✓		✓		✓		✓		✓	
FAMA	1553			✓								
EVIDENZIA	1559	†	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

<sup>1</sup> Données de type coût

Fondation de l'assurance rurale FRV, ASSURA Assurance maladie et accidents, FAMA Caisse-maladie et accidents et EVIDENZIA Schweizerische Kranken- und Unfallversicherung.

Toutes les caisses participant à l'expérimentation n'ont pas mis le même empressement à livrer leurs données à l'OFAS. Ainsi, certaines caisses ont livré années après années leurs données en respectant les délais imposés alors qu'à l'opposé d'autres caisses n'ont encore quasiment rien livré du tout. Entre ces deux extrêmes, nous trouvons des caisses qui n'avaient rien livré jusqu'au milieu de l'année 1995, moment où elles ont produit les données pour l'ensemble de la période 1985 à 1994. Cet état de fait a pour conséquence un retard non négligeable dans les travaux de validation et d'exploitation de ces données. La Table 3.1 présente une vue globale des données disponibles pour la période 1989 à 1994. Les données pour les années 1985 à 1988 ne sont pas prises en compte ici; nous mentionnerons simplement le fait que plus l'on remonte dans le temps, moins nous disposons de données. Quant aux données 1995, elles ne seront livrées qu'au cours du deuxième trimestre 1996.

Les validations et les analyses effectuées sur ces données ont permis de mettre en évidence un nombre relativement important d'erreurs dans ces données administratives. Certaines de ces erreurs peuvent être, plus ou moins facilement, corrigées par nos soins (éventuellement après discussion avec le fournisseur). D'autres fois, les données sont inutilisables pour nos évaluations et nous ne pouvons pas les corriger (en général par manque des informations individuelles nécessaires). Dans ces derniers cas, nous sommes obligés de prier le fournisseur de refaire les données.

La Table 3.1 présente la situation au 31 mars 1996, elle ne tient pas compte de l'histoire des livraisons. Ainsi, parmi les fichiers déclarés reçus et validés (marqués par le symbole ✓), il en est qui ont été refaits (éventuellement à plusieurs reprises) par le fournisseur. Les fichiers marqués par le symbole ✗ sont des fichiers qui, en raison d'une qualité insuffisante de leur contenu doivent être refaits par le fournisseur et pour lesquels aucune version acceptable ne nous est encore parvenue. Pour sa part, le symbole † désigne des fichiers qui, pour des raisons techniques, ne sont plus livrables par les caisses. Dans ce cas, il peut s'agir soit de fichiers qui n'ont jamais été livrés, soit de fichiers qui auraient dû être refaits. Finalement, nous avons désigné à l'aide du symbole # les fichiers dont nous disposons mais que nous n'avons pas validés car ils ne peuvent de toute manière pas être inclus dans les exploitations en raison de l'absence d'autres données de la même caisse.

Nous observons donc que sur les 17 caisses soumises à l'obligation de fournir les données administratives, seules 7 d'entre elles (soit 41 %) ont livré des données acceptables pour l'ensemble des années 1989 à 1994. D'autre part, sur les 8 caisses participant à titre bénévole, 3 ont été en mesure de livrer la totalité des données dans une qualité acceptable et, à l'opposé, deux n'ont quasiment rien livré.

### 3.3 Données administratives des HMOs

Trois HMOs (les trois plus anciens) participent à l'évaluation des nouvelles formes d'assurance. Il s'agit du HMO-Gruppenpraxis Zürich-Wiedikon (ancien IGAK Zürich), du Medizinisches Zentrum Helvetia (MZH), Zürich-Stadelhofen et du Gesundheitsplan HMO Basel (ancien IGAK Basel). Le MZH

est géré directement par la caisse-maladie Helvetia. Ainsi les données concernant les prestations fournies par le MZH ont été transmises directement par la caisse partenaire et seuls les deux autres HMO ont livrés leurs données de manière indépendante des caisses-maladie.

La Table 3.2 présente la situation au 31 mars 1996. Deux des HMOs n'ont été ouverts que dans le courant de 1991, ce qui explique qu'ils n'aient pas fournis de données pour la première ou les deux premières années de l'étude. Nous sommes ainsi en mesure de dire que les HMOs ont livré les données qui devaient l'être.

*Table 3.2 : Données administratives de type prestation des HMOs.  
Les cases ombrées marquent les années pour lesquelles des données doivent être fournies. Le symbole "✓" désigne les fichiers reçus et validés.*

HMO	1990	1991	1992	1993	1994
HMO Basel			✓	✓	✓
HMO ZH-Wiedikon	✓	✓	✓	✓	✓
MZH ZH-Stadelhofen		✓ <sup>1</sup>	✓ <sup>1</sup>	✓ <sup>1</sup>	✓ <sup>1</sup>

<sup>1</sup> Données livrées directement par la caisse-maladie Helvetia

## 4 Analyse de qualité des données

Dans cette section, nous présentons l'état qualitatif des données tel qu'il ressort des validations et corrections systématiques effectuées (cf. section 2).

### 4.1 Etat des livraisons

Une première indication sur la qualité des données est obtenue en analysant l'état des livraisons des données des caisses présenté dans la section précédente. Cette analyse est faite pour chacun des deux types de données.

#### 4.1.1 Données de type assuré

Pour chacune des années, nous pouvons répartir les caisses en 5 catégories selon l'état des données de type assuré dont nous disposons : "utilisable, version 1", "utilisable, version > 1", "à corriger", "aucune livraison" et "livraison impossible". Dans la catégorie "utilisable, version 1", nous classons toutes les caisses qui ont livré les données de type assuré pour l'année considérée et dont les données ont été qualifiées d'utilisables à la fin du processus de validation sans que le fournisseur aie eu besoin de les corriger<sup>1</sup>. Les autres caisses pour lesquelles nous disposons de données utilisables<sup>2</sup> constituent la classe "utilisable, version > 1". Les caisses pour lesquelles nous sommes dans l'attente d'une nouvelle version des données de type assuré pour l'année considérée (en raison d'une qualité insuffisante des données reçues précédemment) sont regroupées dans la classe "à corriger". Finalement, les caisses pour lesquelles nous ne disposons pas de données de type assuré pour l'année considérée sont réparties dans les deux catégories restantes, selon que nous n'ayons rien reçu sans raison apparente ("aucune livraison") ou que la caisse ne dispose plus des éléments nécessaires permettant de produire ces données ("livraison impossible"). Les caisses pour lesquelles nous avons reçu des données de qualité insuffisante et qui ne sont plus en mesure de produire une version corrigée de ces données sont également attribuées à la classe "livraison impossible".

L'importance relative de chacune des classes est illustrée par la Figure 4.1 (100 % = 25 caisses). Nous observons que, si la proportion des caisses qui ont été en mesure de livrer des données utilisables croît de 1990 à 1992, cette même proportion stagne aux environs de 80 % depuis 1992. Par contre, la part de la classe "utilisable, version 1" augmente au détriment de celle de la classe "utilisable, version > 1", ce qui est tout de même réjouissant. L'importance de la classe "livraison impossible" en 1990 provient essentiellement de deux facteurs. D'une part, certaines caisses ont pris

<sup>1</sup> Cependant, il se peut que ces données aient été corrigées par nos soins.

<sup>2</sup> C'est-à-dire les caisses pour lesquelles les données de type assuré de l'année considérée ont été refaites à une ou plusieurs reprises par le fournisseur avant qu'elles ne soient utilisables.

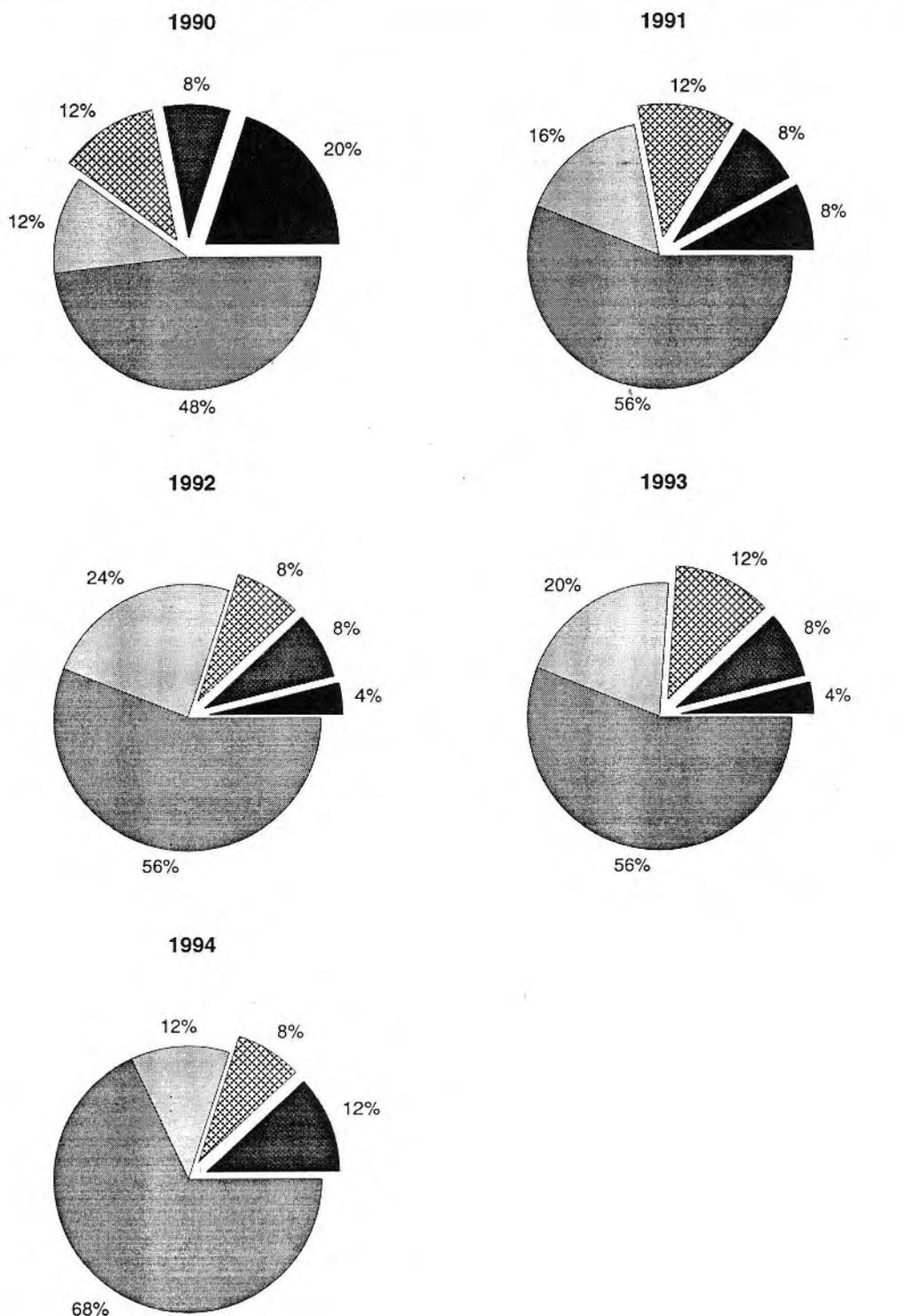


Figure 4.1 : Répartition annuelle relative des 25 caisses-maladie selon l'état des données de type assuré.

utilisable, version 1    
  utilisable, version > 1    
  à corriger    
  aucune livraison    
  livraison impossible

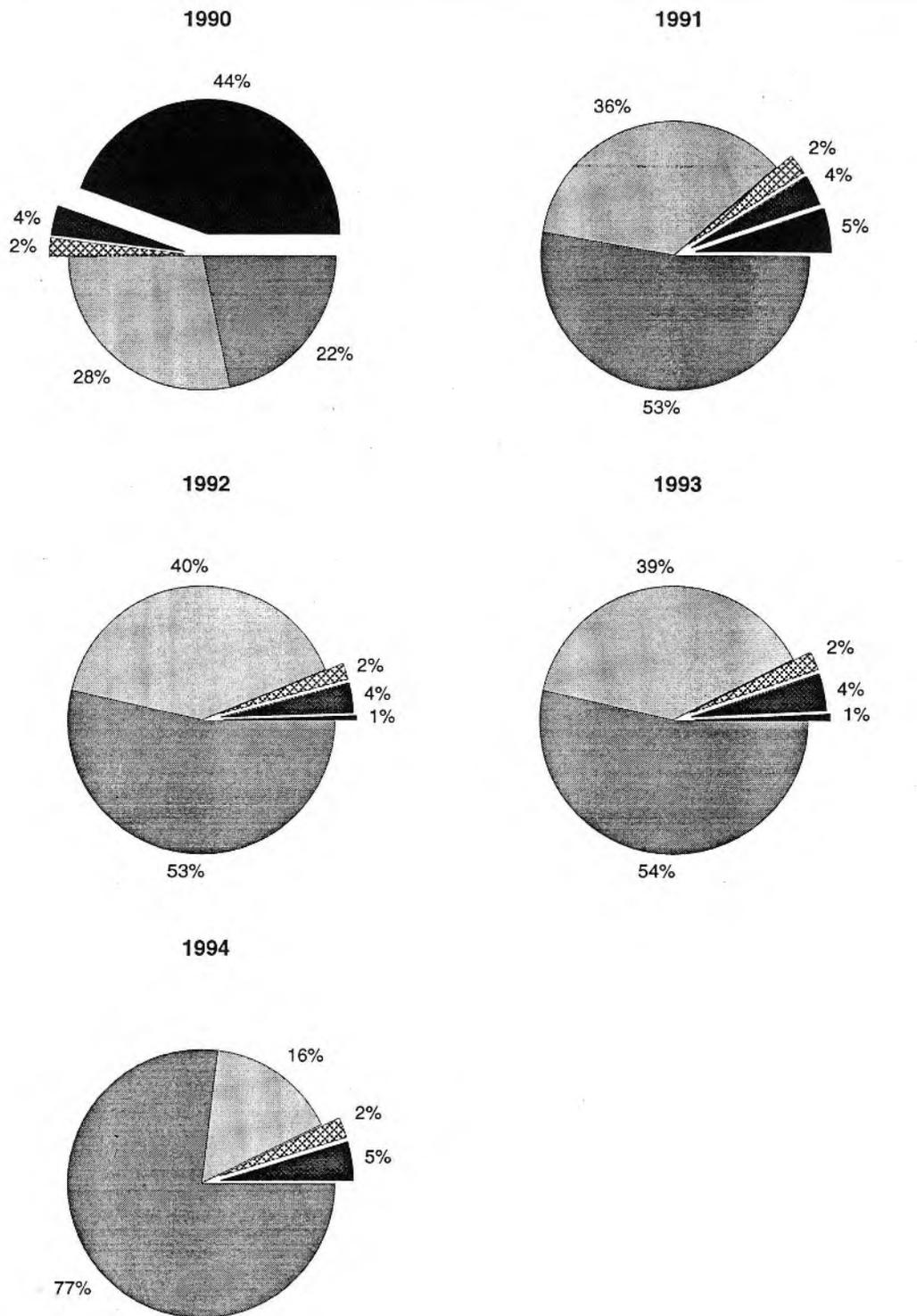


Figure 4.2 : Répartition annuelle relative des assurés des 25 caisses-maladie selon l'état des données de type assuré.

utilisable, version 1    
  utilisable, version > 1    
  à corriger    
  aucune livraison    
  livraison impossible

du retard dans l'établissement et la livraison des données administratives, avec pour conséquence l'impossibilité de livrer les données les plus anciennes. D'autre part, le temps nécessaire au développement du processus de validation des données a également eu pour conséquence des demandes tardives de correction des données auprès de certaines caisses, demandes que les caisses n'ont plus été en mesure de satisfaire.

La figure 4.1 illustre la situation lorsque l'on prend comme unité l'assureur. Cependant, les caisses-maladie ayant des tailles très diverses, leur poids en nombre d'assurés est extrêmement variable. Ainsi, le fait qu'une petite caisse de 10'000 assurés n'ait pas été en mesure de livrer des données n'a pas les mêmes conséquences sur les résultats des analyses que le fait qu'une caisse de 1'000'000 d'assurés fasse défaut. Lorsque nous remplaçons chaque caisse par son poids respectif en terme de nombre d'assurés, nous pouvons mesurer l'importance relative, en nombre d'assurés, de chacune des classes (100 % = nombre total d'assurés de l'année; cf. Figure 4.2). Nous insistons ici sur le fait que ce n'est pas chaque assuré qui a été classifié mais bien les caisses dans leur globalité. Pour les cas où nous n'avons reçu aucune donnée, nous avons utilisé l'effectif de la caisse tel qu'il ressort des données communiquées dans le cadre de la statistique de l'assurance-maladie. Observons d'emblée que si, en 1990, la classe "livraison impossible" correspond à 44 % des assurés (Figure 4.2), cela résulte pratiquement pour moitié du fait que la plus grande des caisses participant à l'évaluation (plus d'un million de membres) n'a plus été en mesure d'établir des données corrigées pour cette année. Pour le reste, la part des deux classes "utilisable" passe globalement de 89 à 93 % entre 1991 et 1994 alors que celle de la classe "utilisable, version 1" passe de 53 à 77 % au cours de ce même laps de temps. Nous pouvons encore remarquer, en particulier, que si les deux caisses qui doivent encore corriger leurs données constituent le 8 % des caisses participantes, elles ne représentent que le 2 % des assurés.

#### 4.1.2 Données de type prestation

La Figure 4.3 illustre, pour les années 1990 à 1994, la répartition annuelle des 25 caisses-maladie selon l'état des données administratives de type prestation (100 % = 25 caisses). La définition des 5 classes est analogue à celle utilisée dans le cadre des données de type assuré. Nous observons, comme dans le cas des données de type assuré, que la proportion des caisses qui ont été en mesure de livrer des données de type prestation utilisable croît de 1990 à 1992 puis stagne à 64 %. Par contre, la part de la classe "utilisable, version 1" augmente au détriment de celle de la classe "utilisable, version > 1" (de 1992 à 1994), ce qui est également réjouissant. Ce qui est étonnant ici, c'est le niveau de stagnation (64 %) comparativement aux 80 % de caisses qui ont été en mesure de livrer des données de type assuré utilisables pour la période 1992–1994. En réalité, ceci s'explique simplement par le fait qu'un groupe de 4 caisses n'a livré, à ce jour, que les données de type assuré, ce que confirme le fait que la part de la classe "aucune livraison" pour les données de type prestation est de 16 % supérieure à celle de la même classe pour les données de type assuré.

En remplaçant, maintenant, chaque caisse par son poids respectif en terme de nombre de factures, nous serons en mesure d'illustrer l'importance relative, en nombre de factures, de chaque classes (100 % = nombre total de factures de l'année; cf. Figure 4.4). A nouveau, nous relevons ici le fait que

ce n'est pas chaque facture qui est classifiée, mais bien la caisse dans son ensemble. Pour les cas où nous n'avons reçu aucune donnée, nous avons utilisé l'effectif d'assurés de la caisse<sup>1</sup> multiplié par un facteur cinq<sup>2</sup>. Pour 1990, comme pour les données de type assuré, la part des données utilisables dépasse à peine les 50 % (Figure 4.4) et l'importance de la classe "livraison impossible" résulte pratiquement pour moitié du fait que la plus grande des caisses participant à l'évaluation n'a plus été en mesure d'établir des données corrigées pour cette année. Pour le reste, la part des données utilisables avoisine les 90 % pour les années 1991 à 1994. Ainsi, si les caisses qui n'ont pas encore livré leurs données constituent en général le 24 % des caisses participantes (pour 1991–1994), elles ne représentent qu'approximativement 7 % des factures.

## 4.2 Correction du nombre de blocs d'informations

Comme nous l'avons rappelé précédemment, les enregistrements de données sont, quel que soit leur type, constitués d'une partie dite générale complétée d'un certain nombre de blocs d'informations. Au cours de la validation, le nombre de blocs effectivement présents est comparé avec le nombre spécifié dans l'enregistrement. Lorsque ces deux nombres sont différents, le nombre spécifié dans l'enregistrement est corrigé. Cette erreur a été détectée aussi bien dans les données de type assuré que dans celles de type prestation.

Un second type de problème donne lieu à une correction du nombre de blocs : la présence de blocs complètement vides (et donc inutiles) intercalées entre des parties utiles. Dans un tel cas, ces blocs vides sont supprimés. Ce phénomène n'a été détecté que pour les données de type prestation.

### 4.2.1 Correction du nombre de branches d'assurance

Pour les données de type assuré, le nombre de blocs correspond au nombre de branches d'assurance. Des différences entre le nombre de branches spécifié dans un enregistrement et le nombre de branches que contient effectivement l'enregistrement ont été détectées pour les années 1991 à 1994 dans les données d'une seule caisse (la même pour chacune des années). L'erreur affecte, selon les années, entre 5.5 et 11 % des enregistrements de cette caisse, ce qu'illustre le graphe de gauche de la Figure 4.5<sup>3</sup>. Comme cette caisse représente environ 6 % de l'ensemble des assurés<sup>4</sup>, c'est tout de même approximativement un enregistrement de type assuré sur 200 qui a fait l'objet de cette correction (Figure 4.5, graphe de droite), ceci pour chacune des années 1991 à 1994.

---

<sup>1</sup> Cet effectif est repris des données administratives de type assuré où, à défaut, des données de la statistique de l'assurance-maladie.

<sup>2</sup> Ce facteur correspond à la moyenne du rapport entre nombre de factures et nombre d'assurés pour les caisses ayant livré l'ensemble des données demandées.

<sup>3</sup> Dans ce cas, le 100 % correspond à l'ensemble des assurés des caisses dans lesquelles l'erreur a été détectée.

<sup>4</sup> Par ensemble des assurés, nous désignons ici le nombre total des membres des caisses pour lesquelles nous possédons des données utilisables pour l'année considérée (cf. Table 3.1).

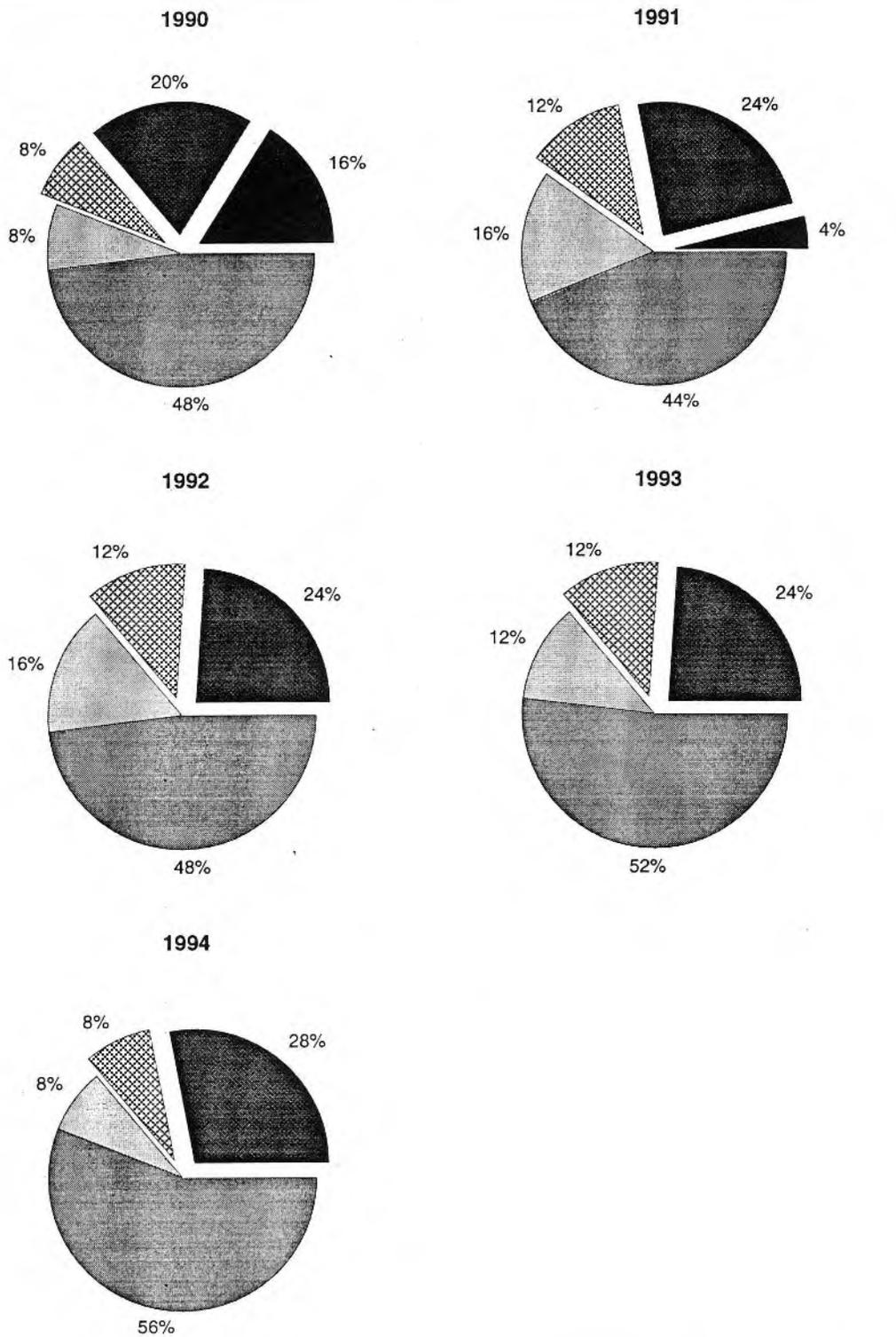


Figure 4.3 : Répartition annuelle relative des 25 caisses-maladie selon l'état des données de type prestation.

utilisable, version 1    
  utilisable, version > 1    
  à corriger    
  aucune livraison    
  livraison impossible

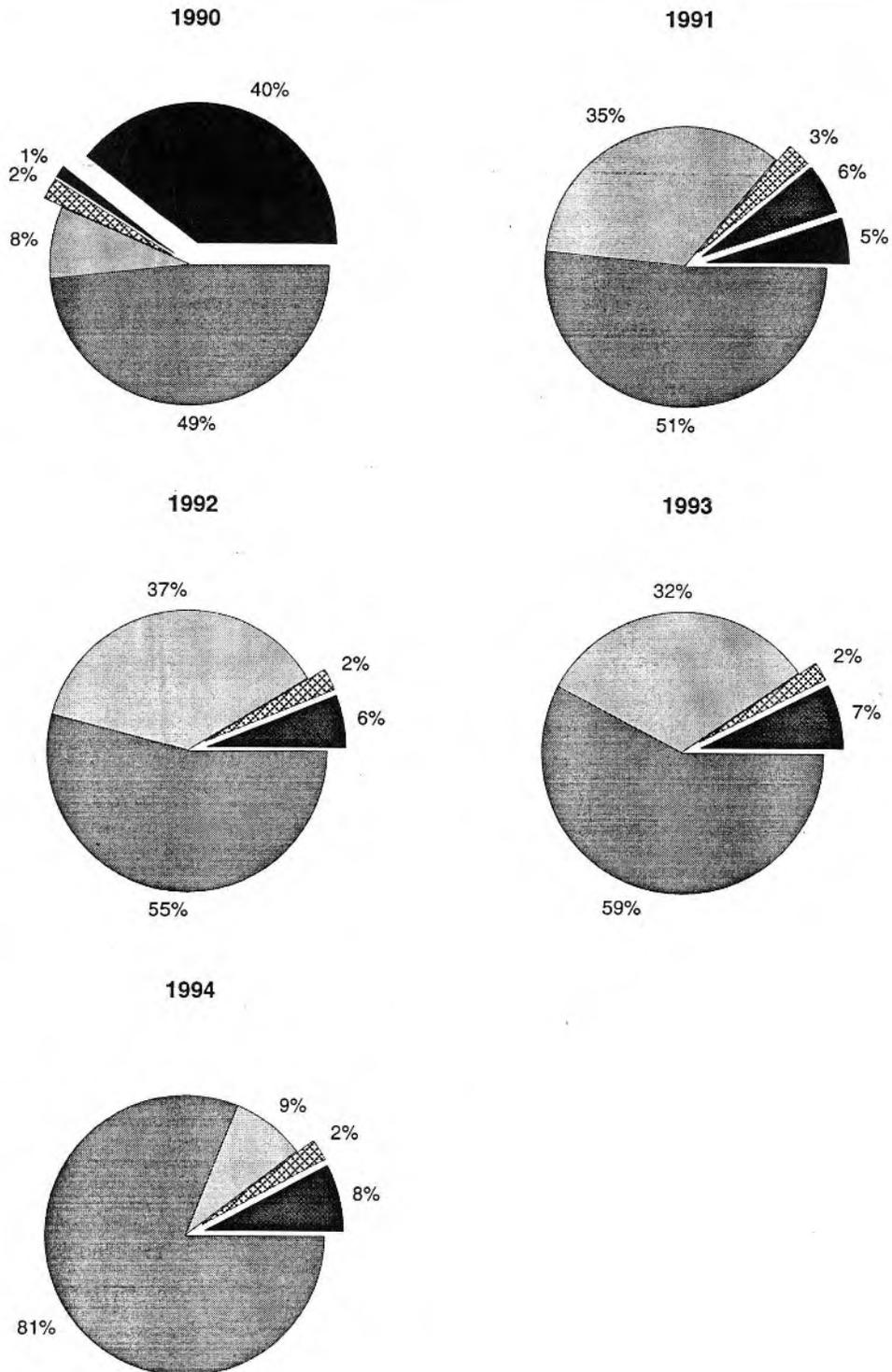
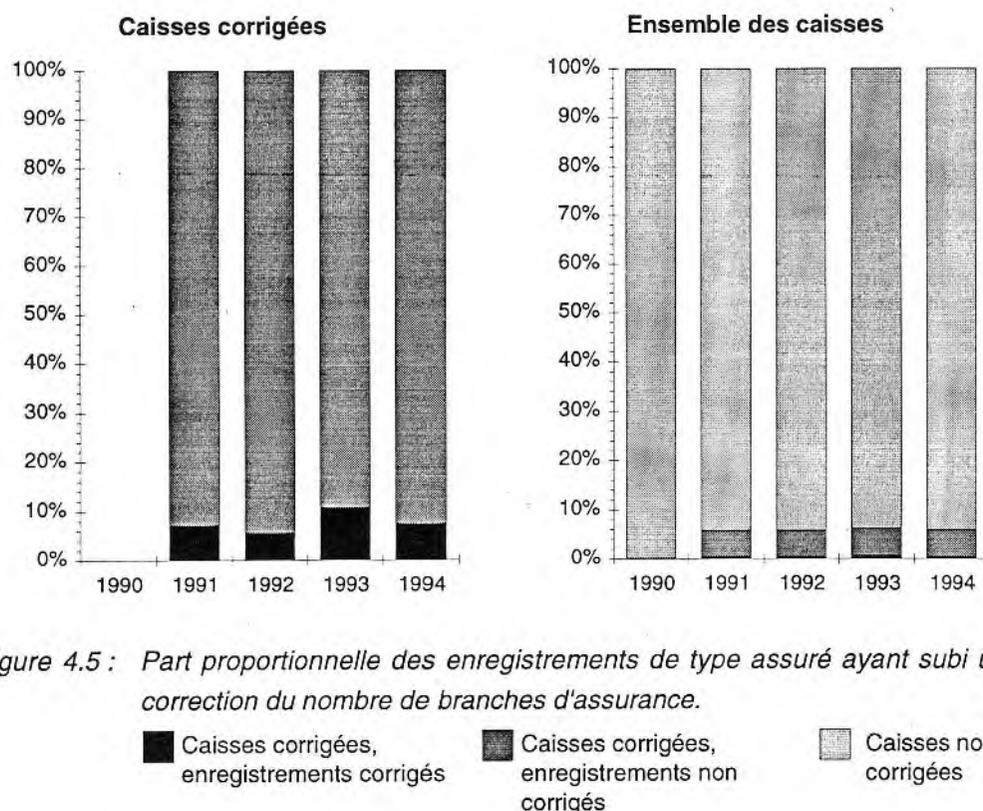


Figure 4.4 : Répartition annuelle relative des factures des 25 caisses-maladie selon l'état des données de type prestation.

utilisable, version 1    
  utilisable, version > 1    
  à corriger    
  aucune livraison    
  livraison impossible



#### 4.2.2 Correction du nombre de composantes de frais

Pour les données de type prestation, le nombre de blocs correspond au nombre de composantes de frais. Des différences entre le nombre de composantes spécifié dans un enregistrement et le nombre de composantes que contient effectivement l'enregistrement ont été détectées dans 2 caisses. Pour l'une d'elle, ce type d'erreur est présent pour chacune des années 1990 à 1994 tandis que pour la deuxième, ce n'est qu'en 1993 et 1994 que ce problème est apparu. Pour 1990, cette erreur n'affecte qu'un enregistrement sur 1'000 de la caisse concernée alors que cette proportion est 5 à 6 fois plus élevée en 1991 et 1992. En 1993 et 1994, c'est approximativement 8 enregistrements sur 1'000 qui ont dû être corrigés sur l'ensemble des deux caisses concernées (Figure 4.6, graphe de gauche). Comme, pour 1990 à 1992, l'unique caisse concernée ne représente qu'entre 0.8 et 1.1 % de l'ensemble des factures, c'est à peine un enregistrement sur 10'000 qui a dû être corrigé (Figure 4.6, graphe de droite). Pour 1993 et 1994, le fait que les deux caisses corrigées constituent le 1.4 % de l'ensemble des factures maintient la proportion de corrections à un enregistrement sur 10'000.

#### 4.2.3 Suppression de composantes de frais vides

Les composantes de frais complètement vides, intercalées entre d'autres composantes, sont supprimées au cours de la validation. De telles composantes ont été détectées dans les données de

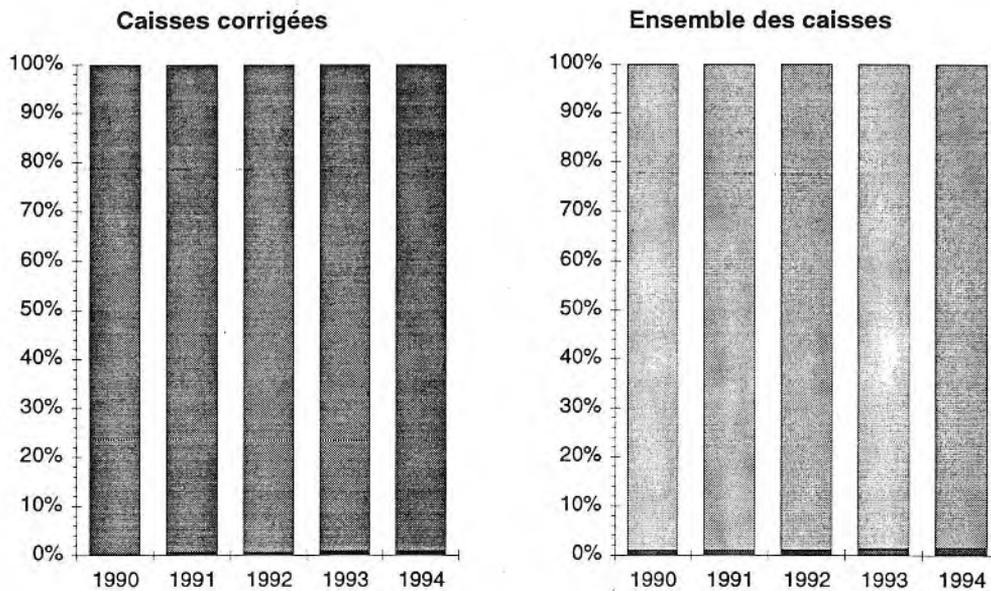


Figure 4.6 : Part proportionnelle des enregistrements de type prestation ayant subi une correction du nombre de composantes de frais.

Caisses corrigées, enregistrements corrigés
  Caisses corrigées, enregistrements non corrigés
  Caisses non corrigées

5 caisses. Pour l'une d'elle, ce type d'erreur n'est présent que pour les années 1992 à 1994 tandis que pour les quatre autres, ce problème apparaît chaque année. Comme le montre le graphe de gauche de la Figure 4.7, c'est à chaque fois plus de la moitié de l'ensemble des enregistrements de type prestation des caisses concernées qui a dû être corrigée. L'ensemble des factures des caisses concernées représente approximativement 9.5 %, respectivement 13.5 %, de la totalité des enregistrements pour 1990 et 1991, respectivement pour 1992 à 1994. Ainsi, des composantes de frais vides ont été éliminées dans 5 à 8 % des enregistrements (Figure 4.7, graphe de droite). En 1994, c'est un peu plus de 2 millions de composantes de frais inutiles qui ont ainsi été éliminées des données, ce qui équivaut à un gain de place de 27 MBytes.

### 4.3 Correction des identificateurs des assurés

Au cours du traitement des données administratives, il s'est avéré nécessaire de modifier les identificateurs des assurés de certaines caisses, soit uniquement pour l'un des types de données (assuré ou prestation), soit pour les deux. Les problèmes rencontrés sont décrits dans la section 2.4.

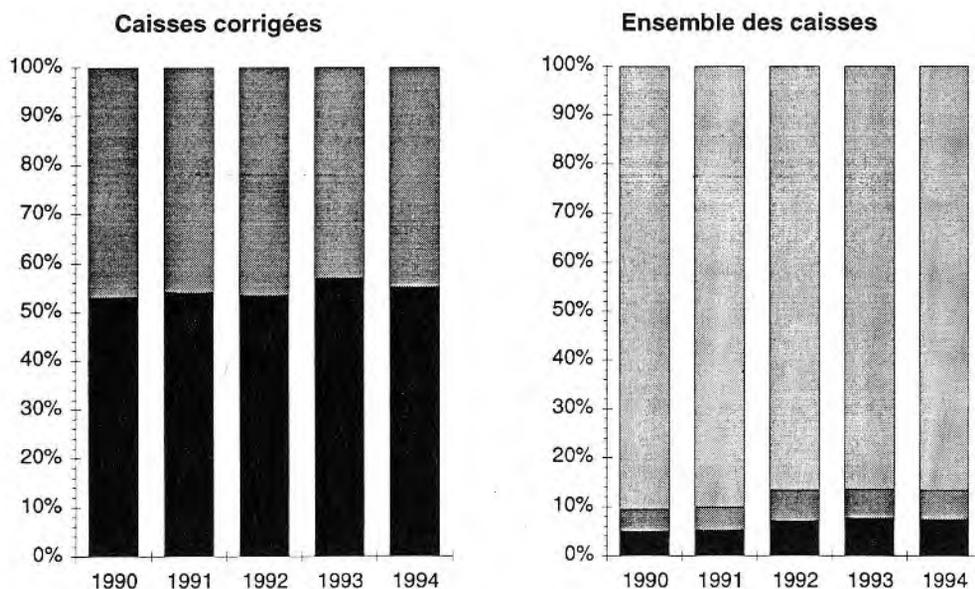


Figure 4.7 : Part proportionnelle des enregistrements de type prestation comportant des composantes de frais vides.

■ Caisses corrigées, enregistrements corrigés    ■ Caisses corrigées, enregistrements non corrigés    ■ Caisses non corrigées

#### 4.3.1 Données de type assuré

Pour les données de type assuré, 6 caisses ont vu l'ensemble des identificateurs des assurés modifiés pour les diverses raisons évoquées plus haut, sauf en 1990 où seuls les fichiers de 4 caisses ont dû être traités pour ce problème. Ceci équivaut à un peu plus de 45 % des assurés pour les années 1991 à 1994 contre près de 15 % en 1990 (Figure 4.8). Le bon résultat obtenu en 1990 (comparativement aux années postérieures) ne reflète cependant pas une meilleure qualité des données. Il résulte uniquement du fait que, pour cette année, les deux caisses qui ont également été corrigées par la suite n'ont pas livré de données de type assuré utilisables.

#### 4.3.2 Données de type prestation

Pour les données de type prestation, le nombre de caisses pour lesquelles les identificateurs des assurés ont été corrigés, oscille entre 3 et 4, ce qui équivaut à 15 à 30 % des factures (Figure 4.8). La chute de près de 19 % du nombre de corrections entre 1992 et 1993 s'explique simplement par le fait que l'une des 4 caisses corrigées en 1992 a remédié elle-même au problème dès 1993. Comme la caisse en question est l'une des plus importantes de Suisse, la diminution des corrections nécessaires est également très importante. Il s'agit donc bien ici d'une amélioration de la qualité des données.

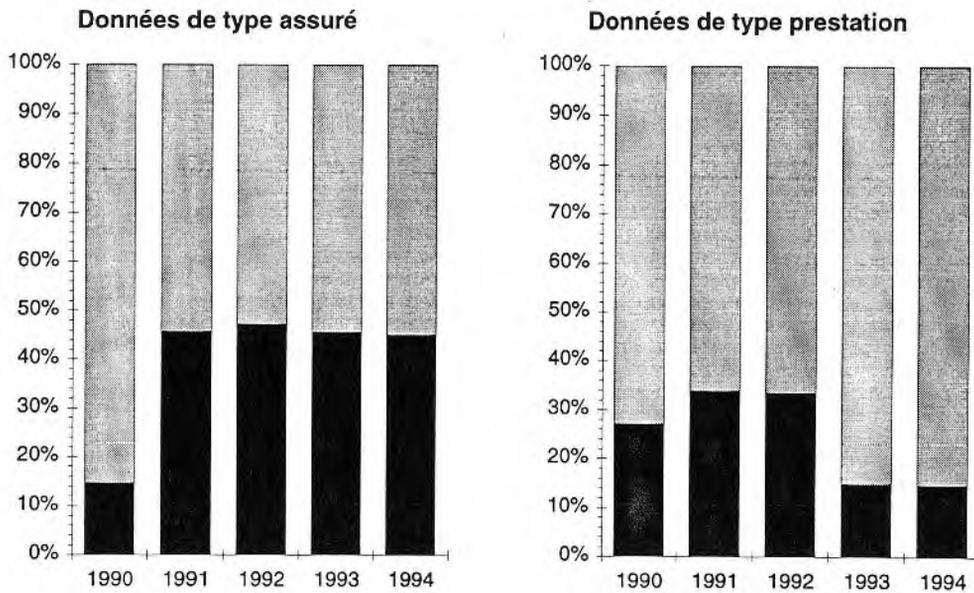


Figure 4.8 : Part proportionnelle des enregistrements ayant subi une correction de l'identificateur de l'assuré.

■ Enregistrements corrigés    □ Enregistrements non corrigés

## 4.4 Correction des branches d'assurance

Pour les enregistrements de type assuré, des corrections au niveau des branches d'assurances ont également été nécessaires. Parmi celles-ci, nous ne retiendrons ici que les deux types de corrections les plus significatifs : les corrections du genre d'assurance et celles des valeurs des franchises.

### 4.4.1 Correction du genre d'assurance

Nous considérons ici la correction des problèmes qui affectent la définition du genre d'assurance. Le travail consiste, selon les cas, soit à corriger le genre d'assurance, soit à supprimer la totalité du bloc de données correspondant à la branche d'assurance (cf. section 2.4). Dans la suite de cette section, une suppression d'un bloc est traitée comme une correction.

Le nombre de caisses pour lesquelles de telles corrections ont été nécessaires est passé de 2 à 6 de 1990 à 1994. Il faut malheureusement noter qu'en général, les données des caisses pour lesquelles ce type de problème se présente au cours d'une année doivent également être corrigées pour les années suivantes. Pour chaque année, il y a en plus entre 3 et 4 autres caisses qui ont subi ce type de correction, mais dans une proportion de moins de 1 % de leurs enregistrements. Une telle

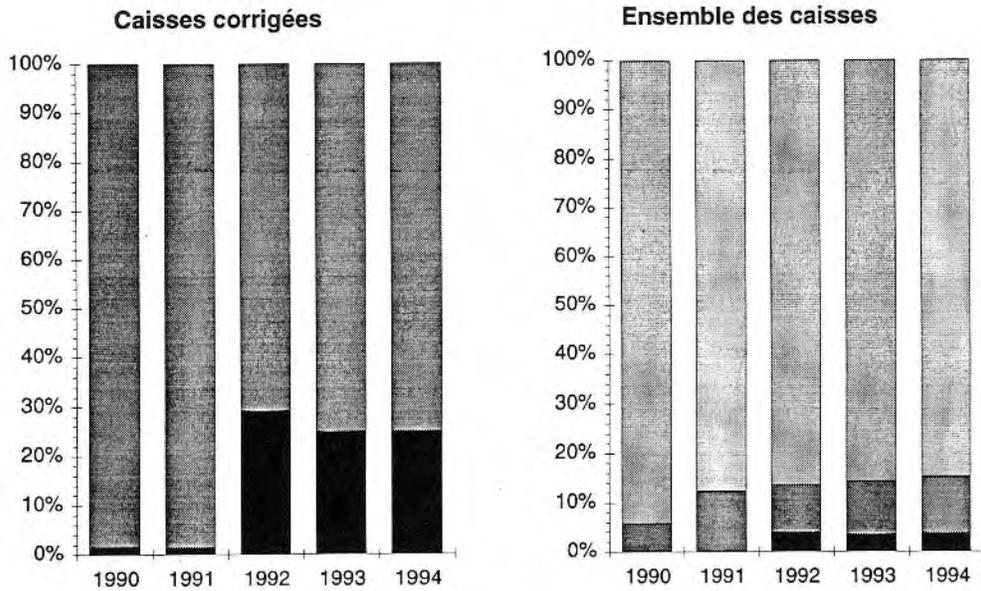


Figure 4.9 : Part proportionnelle des enregistrements de type assuré ayant subi des corrections du genre d'assurance.

Caisses corrigées, enregistrements corrigés
  Caisses corrigées, enregistrements non corrigés
  Caisses non corrigées

proportion ne nous a pas paru significative; par conséquent, nous n'avons pas tenu compte de ces caisses dans les résultats présentés dans la Figure 4.9.

Sur l'ensemble des caisses corrigées, la proportion d'enregistrements corrigés peut s'élever jusqu'à 30 % (Figure 4.9, graphe de gauche). L'ensemble des assurés des caisses concernées représente en général entre 10 et 15 % de la totalité des assurés et les assurés corrigés peuvent représenter jusqu'à 4 % de l'ensemble (Figure 4.9, graphe de droite).

Un assuré peut avoir plusieurs branches d'assurance. Ainsi, il arrive que ce soit plusieurs branches qui doivent être corrigées pour un même assuré. Le nombre moyen de corrections par assuré corrigé varie ainsi de 1.07 à 1.14 selon l'année.

#### 4.4.2 Correction des valeurs des franchises

Les divers problèmes affectant les valeurs des franchises sont décrits dans la section 2.4. Le nombre de caisses pour lesquelles des corrections des franchises ont été effectuées est en général égal à 5 ou 6. Sur l'ensemble des caisses corrigées, la proportion d'enregistrements corrigés varie entre 5.1 et 7.7 % (Figure 4.10, graphe de gauche). L'ensemble des caisses concernées représente entre 20 et 45 % de la totalité des assurés, ce qui a pour conséquence que la proportion d'assurés corrigés s'élève entre 1.5 et 2.5 % de la totalité (Figure 4.10, graphe de droite).

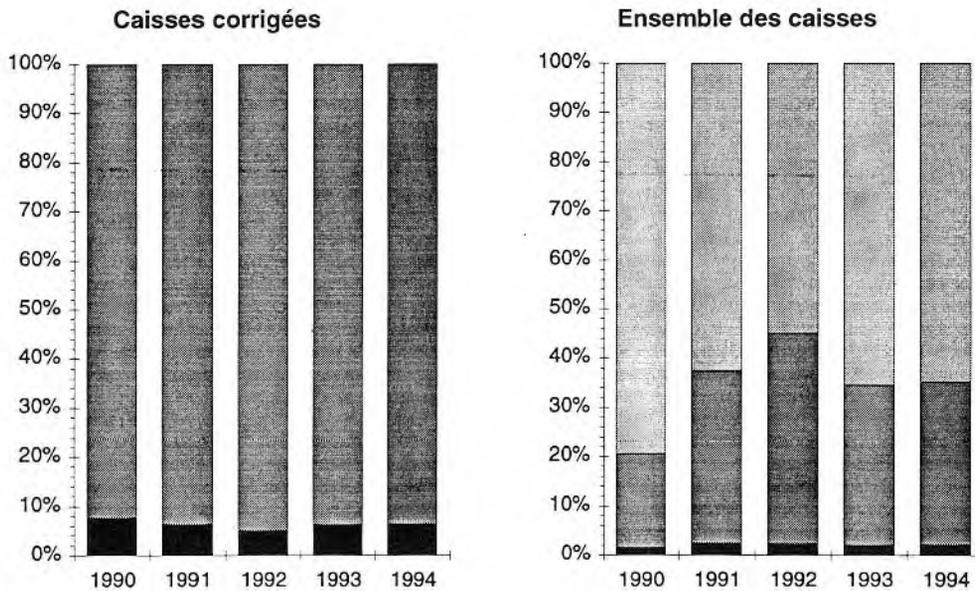
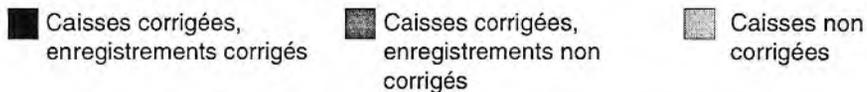


Figure 4.10 : Part proportionnelle des enregistrements de type assuré ayant subi des corrections des valeurs des franchises.



Comme la valeur de la franchise peut apparaître dans plusieurs blocs de données, il se peut que plusieurs corrections soient effectuées dans un même enregistrement. Le nombre moyen de corrections par assuré corrigé varie ainsi de 1.09 à 1.14 selon l'année.

#### 4.5 Regroupement des enregistrements de type assuré correspondant à une même personne

Selon les années, 3 ou 4 caisses ont livré des données de type assuré dans lesquelles nous avons détecté des membres pour lesquels plusieurs enregistrements avaient été créés. Comme cela est déjà le cas pour la plupart des problèmes présentés jusqu'ici, il s'agit d'une erreur qui persiste au fil des années et qui se retrouve toujours dans les mêmes caisses.

Lorsque les informations des divers enregistrements d'un même assuré sont cohérentes entre elles, le groupe d'enregistrements est fusionné en un seul enregistrement contenant une synthèse de ces données. Parmi l'ensemble des assurés des caisses affectées par ce type de problème, la part des assurés pour lesquels une telle fusion a été pratiquée avoisine chaque année les 2 % (Figure 4.11, graphe de gauche). La même proportion, mais considérée sur l'ensemble des 25 caisses, varie de 0.15 à 0.30 % selon les années et le poids des caisses corrigées passe quant à lui de 7.2 à 13.2 % de 1990 à 1994 (Figure 4.11, graphe de droite).

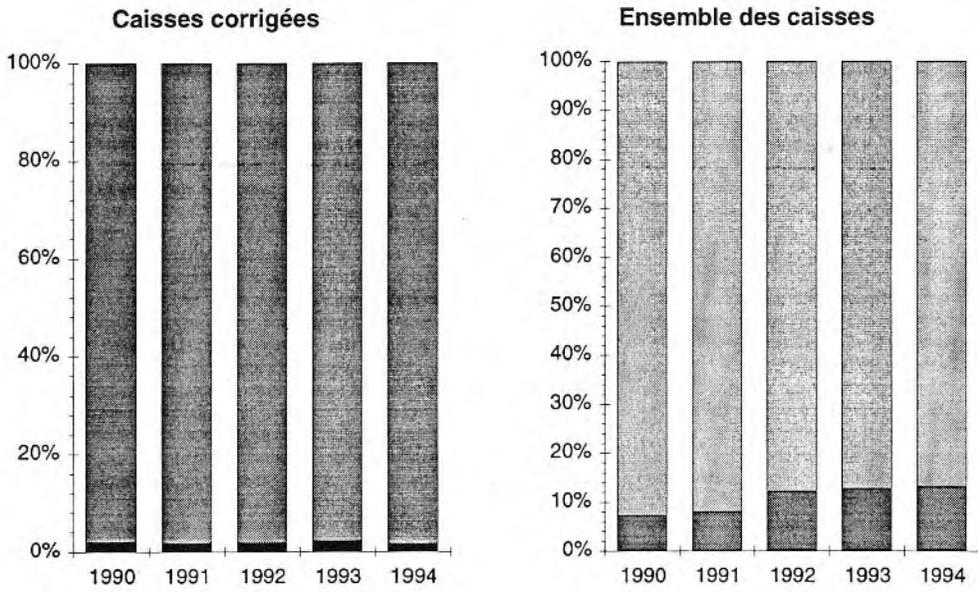


Figure 4.11 : Part proportionnelle des assurés pour lesquels des enregistrements de type assuré ont été regroupés.

Caisses corrigées, assurés corrigés
  Caisses corrigées, assurés non corrigés
  Caisses non corrigées

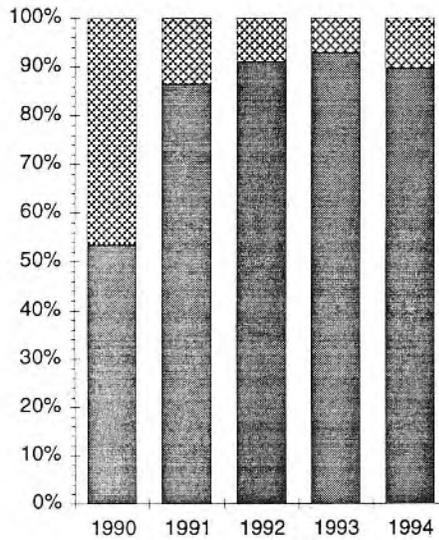


Figure 4.12 : Taux de réussite pour le regroupement des enregistrements de type assuré appartenant à un même assuré.

Assurés résolus
  Assurés supprimés

Lorsqu'un groupe d'enregistrements correspondant à un même assuré comporte des incohérences, il est totalement supprimé des données. L'ensemble des assurés ainsi supprimés représente entre 0.18 et 0.29 % de l'ensemble des caisses corrigées, sauf en 1990 où cette part s'élève à près de 1.9 %. Rapportée à l'ensemble des 25 caisses, cette part n'est plus que de 0.02 % (0.13 % en 1990); elle est donc négligeable.

Si, maintenant, nous comparons la part des assurés pour lesquels une fusion des multiples enregistrements a été possible par rapport à l'ensemble des assurés pour lesquels plus d'un enregistrement a été détecté, nous obtenons pour les années 1991 à 1994 un taux de réussite voisin de 90 % (Figure 4.12).

#### 4.6 Intégration des données des HMOs

Deux des trois HMOs intégrés dans le processus d'évaluation des nouvelles formes d'assurance-maladie nous ont transmis leurs données de type prestation de manière indépendante des caisses-maladie<sup>1</sup>. Au cours du processus d'intégration de ces données avec celles des caisses-maladie, nous avons dû faire face à plusieurs difficultés (voir également à ce sujet la section 2.5).

En particulier, sur la base de tables de conversion également transmises par les HMOs, il a été nécessaire d'effectuer une conversion des identificateurs des membres. Or, en raison de quelques manques dans ces tables, les identificateurs de certains membres des HMOs n'ont pas pu être convertis et les prestations correspondantes ont été éliminées. Malgré les problèmes pratiques rencontrés lors de ce travail, nous avons tout de même obtenu un taux de réussite remarquable de plus de 98 %, ce qui signifie que la part des enregistrements de type prestation de ces HMOs qui a été abandonnée est inférieure à 2 % (Figure 4.13).

#### 4.7 Validation de l'âge d'entrée pour les branches d'assurance

L'âge d'entrée dans une branche d'assurance est un des facteurs déterminants pour la fixation de la prime à payer et est donc nécessaire pour l'établissement d'une statistique des primes. Cette grandeur, répartie par groupes d'âge, devait figurer pour chaque branche d'assurance dans les données de type assuré. Or les contrôles effectués font apparaître que cette donnée n'est pas fiable. Parmi les divers problèmes observés nous noterons l'omission pure et simple de cette donnée, la spécification incomplète des groupes d'âge<sup>2</sup> et la spécification d'un groupe d'âge incompatible avec l'âge effectif de l'assuré. La plupart de ces erreurs n'ayant pu être corrigées, l'âge d'entrée est à utiliser avec précaution lors des exploitations des données.

<sup>1</sup> Le HMO-Gruppenpraxis Zürich-Wiedikon pour toutes les années et le Gesundheitsplan HMO Basel uniquement pour les années 1992 à 1994 (cf. section 3.3).

<sup>2</sup> Seul le début ou la fin du groupe d'âge est spécifié.

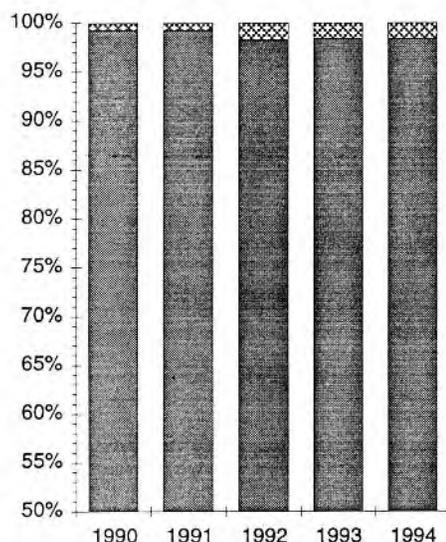


Figure 4.13 : Taux de réussite lors de la conversion des identificateurs d'assurés dans les données de type prestation des HMOs.

■ Enregistrements corrigés      ▨ Enregistrements supprimés

#### 4.8 Validation des durées de risque

La durée de risque fait également partie des grandeurs figurant dans les données de type assuré qui doivent être traitées avec une grande précaution. Il y a deux raisons principales à cet état de fait. La première est la présence de durées supérieures à une année dans les données de quelques assureurs. Ceci résulte d'erreurs commises par le fournisseur des données. Cette erreur est difficile à traiter à posteriori car, en particulier, nous ne savons pas si les primes associées sont les primes correspondant à cette durée (fausses!) ou s'il s'agit de la prime effectivement à la charge de l'assuré<sup>1</sup>.

La deuxième raison est plus complexe et trouve sa source dans la conception même de l'enregistrement de données de type assuré. Cet enregistrement doit contenir un bloc de données par branche d'assurance souscrite par l'assuré. Or l'examen des données a montré qu'il pouvait y avoir, dans un même enregistrement de type assuré, plusieurs blocs de données "branche d'assurance" parfaitement identiques à l'exception éventuelle des données de durée de risque et de primes. L'analyse du problème montre que cette situation peut se produire soit lorsque l'assuré possède deux couvertures distinctes pour l'assureur mais correspondant au même genre d'assurance dans les données administratives, soit lorsque l'assuré modifie la couverture d'assurance de l'une de ses

<sup>1</sup> En d'autres termes, nous ne savons pas si l'erreur commise affecte uniquement la durée du risque ou si elle affecte également la prime de la branche concernée.

branches au cours de l'année et que l'assureur génère, pour cette branche, un bloc de donnée pour chaque portion d'année même si la modification ne transparaît pas dans les éléments codés dans les données administratives. Pour traiter correctement les durées de risque, il est nécessaire d'être à même de distinguer entre ces deux cas de figure. Cependant, les informations à notre disposition dans les données administratives ne sont pas toujours suffisantes pour permettre d'effectuer cette distinction avec certitude.

#### **4.9 Validation des dates de début et de fin de prestation**

Chaque enregistrement de type prestation devait contenir les dates correspondant au début et à la fin des prestations englobées dans la facture qu'il représente. Ces deux données devaient, en particulier, servir à déterminer la durée des séjours hospitaliers. Force nous a été de constater que ces deux données font également partie des éléments auxquels nous ne pouvons pas nous fier. En effet, dans bien des cas, soit ces dates ne figurent pas dans les enregistrements, soit elles conduisent à des durées de traitement qui ne semblent pas plausibles. Et nous ne parlerons pas ici de tous les cas qui ont pu être corrigés comme par exemple les inversions entre le début et la fin du traitement !

#### **4.10 Remarque générale**

Le problème de la qualité dans un cadre aussi complexe que celui des données administratives de l'assurance-maladie est un problème extrêmement ardu dont la maîtrise nécessite beaucoup de temps et d'efforts. La liste des problèmes que nous venons de traiter (sections 4.1 à 4.9) n'est pas exhaustive mais illustre cependant parfaitement l'ampleur et la difficulté de la tâche. Ces défauts de qualité constatés, il a été nécessaire d'essayer d'y remédier. Malgré tous les efforts faits dans ce but, il n'a pas été possible d'éliminer tous les problèmes observés, ce qui a notamment pour conséquence d'interdire certaines analyses.

## 5 Statistique élémentaire des données administratives

Afin d'avoir un aperçu du contenu des données, quelques analyses simples ont été effectuées sur les enregistrements validés. Comme la typologie des informations contenues dans les fichiers de type assuré et dans ceux de type prestation est bien distincte, ces analyses sont spécifiques à chaque type d'enregistrements. Elles n'ont porté que sur les enregistrements des caisses qui ont subi avec succès toutes les validations et corrections présentées jusqu'ici (cf. Table 3.1). Cette statistique élémentaire ne sert pas uniquement à donner une idée du contenu des données mais elle a également permis de mettre en évidence certaines des erreurs mentionnées dans la section précédente, erreurs qui ont ainsi pu être corrigées.

### 5.1 Effectifs des assurés par types d'assurance

Pour les données de type assuré, deux études distinctes ont été effectuées. Il s'agit d'une part d'une classification des assurés en fonction de leur type d'assurance et, d'autre part, d'une comptabilisation des primes par types d'assurance. Dans cette section, nous abordons les résultats de la première de ces études alors que ceux de la deuxième seront présentés dans la section suivante.

Pour déterminer l'effectif des assurés par types d'assurance, ceux-ci sont tout d'abord répartis en 3 catégories : assurés possédant une assurance de base pour soins médico-pharmaceutiques, assurés sans assurance de base pour soins médico-pharmaceutiques et assurés avec assurance de base indéterminée.

La catégorie des assurés sans assurance de base pour soins médico-pharmaceutiques regroupe tous les assurés pour lesquels nous avons des informations concernant leur couverture d'assurance (blocs de données "branches d'assurance" présents dans les données de type assuré de ces personnes) et dont la couverture n'englobe pas l'assurance de base. De tels assurés ont en général des assurances d'indemnités journalières et/ou d'autres assurances; ils ont probablement également une assurance de base pour soins médico-pharmaceutiques, mais celle-ci est conclue auprès d'un assureur différent.

La catégorie des assurés avec assurance de base indéterminée est constituée de deux groupes d'assurés distincts. Il y a, d'une part, les personnes pour lesquelles nous avons des données de type assuré mais sans aucune indications sur leur couverture d'assurance<sup>1</sup> et, d'autre part, celles pour lesquelles il existe des données de type prestation<sup>2</sup> sans qu'aucune information de type assuré ne soit présente.

---

<sup>1</sup> Personnes pour lesquelles l'enregistrement de type assuré ne comporte aucun bloc de données "branches d'assurance".

<sup>2</sup> Personnes que nous qualifierons par le terme de malade par la suite.

La première de ces 3 catégories, celle des assurés avec une assurance de base pour soins médico-pharmaceutiques, est elle-même subdivisée en quatre sous-catégories selon le mode d'assurance : traditionnel, franchise à option, bonus, choix limité des fournisseurs de prestations. Lorsque l'assuré a modifié son assurance de base au cours de l'année, il se peut que son enregistrement comporte des mentions d'assurance de soins médico-pharmaceutiques de divers modes. Afin de classer chaque assuré dans un mode unique, nous avons adopté la convention suivante : un mode "choix restreint" prime sur tous les autres, le mode "bonus" prime sur les modes "franchise à option" et "traditionnel" et la franchise à option prime sur la franchise ordinaire (mode "traditionnel"). Comme les données administratives ne fournissent aucune indication sur l'ordre dans lequel ces changements de mode sont intervenus, cette manière de faire ne constitue pas l'unique façon de classer les assurés. D'autres définitions ont été appliquées dans certaines exploitations. Cependant, tous les résultats présentés dans le présent rapport sont établis selon la définition ci-dessus.

Les assurés avec franchise annuelle à option sont répartis selon le niveau de franchise choisi. L'attribution du niveau de franchise selon le montant de celle-ci est réglé par la Table 5.1. Lorsque le montant de franchise ne correspond pas à l'un des montants admissibles (et qu'il n'a pas été possible de le corriger), le niveau de franchise est considéré comme indéterminé. Rappelons que les montants des franchises ont été modifiés au 1er janvier 1991 avec faculté pour les caisses d'effectuer ces changements au 1er janvier 1992. Ainsi, le lien entre valeur et niveau de franchise dépend des caisses pour 1991. Lorsque le choix entre plusieurs niveaux se présente pour un assuré donné, c'est le niveau défini (1, 2 ou 3) le plus élevé qui est retenu.

Table 5.1 : Les niveaux de franchise.

Niveau	Franchises adulte		Franchises enfant	
	1990 – 1991	1991 – 1995	1990 – 1991	1991 – 1995
1	300	350	50	50
2	500	600	100	150
3	1000	1200	150	200
indéterminé	autre	autre	autre	autre

Pour les assurances avec bonus, nous présentons également la répartition des assurés par niveaux de bonus. Les niveaux admissibles sont définis par la Table 5.2. Lorsque le pourcentage de la prime initiale ne correspond pas à l'une des valeurs admissibles, le niveau de bonus est considéré comme indéterminé.

Finalement, nous observerons que les modèles d'assurances avec choix limité des fournisseurs de prestations se sont limités au modèle HMO jusqu'en 1994, année qui a vu l'introduction du modèle

HMO-IPA avec paiement à l'acte. La répartition des assurés entre ces deux modèles sera donc examinée dès 1994. Lorsque nous avons le choix entre ces deux possibilités pour un assuré donné, la convention retenue est que le modèle HMO prime sur l'autre.

Table 5.2 : Les niveaux de bonus.

Niveau	% de la prime initiale
100	100
85	85
75	75
65	65
55	55
indéterminé	autre > 0

Pour la période allant de 1990 à 1994, les répartitions des assurés sont présentées dans la Table 5.3. Dans ces résultats, seules sont prises en compte les données utilisables disponibles, selon l'état de la Table 3.1. Ainsi, les caisses pour lesquelles des données de type assuré utilisables sont disponibles fournissent une véritable répartition des assurés selon leur mode d'assurance tandis que celles pour lesquelles seules des données de type prestation utilisables existent ne contribuent qu'à la catégorie des assurés avec assurance de base indéterminée. Quelques commentaires sur ces résultats sont formulés ci-après, année par année.

#### 5.1.1 Effectifs des assurés en 1990

Pour 1990, les résultats se basent sur la distribution des assurés de 15 caisses, complétée par l'effectif des malades de 3 autres caisses (cf. Table 3.1). Ces trois caisses contribuent pour 77 % à l'effectif des assurés avec assurance de base indéterminée. Le reste de cette classe, soit approximativement 70'000 assurés, se répartit en 3 % d'assurés avec données de type assuré mais couverture d'assurance inconnue et 97 % de malades sans données de type assuré. Parmi les assurés avec une assurance de base de type franchise à option, il y en a 1 % dont le niveau est indéterminé; la raison de ce problème n'a pas été élucidée. Le mode bonus n'a été introduit par les caisses qu'en 1991; ainsi, il ne peut pas y avoir d'assurés bonus en 1990.

#### 5.1.2 Effectifs des assurés en 1991

Pour 1991, les résultats se basent sur la distribution des assurés de 18 caisses, complétée par l'effectif des malades d'une caisse supplémentaire (cf. Table 3.1). Cette caisse contribue pour 20% à

Table 5.3 : Effectifs des assurés par types d'assurance.

	1990	1991	1992
Soins médico-pharmaceutiques (base)	2 691 449	4 615 818	5 062 618
Traditionnelle	2 524 358	4 392 785	4 780 452
Franchise à option	165 096	210 344	257 801
niveau 1	90 847	109 262	126 708
niveau 2	52 923	68 137	85 729
niveau 3	19 725	32 720	45 168
niveau indéterminé	1 601	225	196
Bonus		7 841	14 066
niveau 100		7 841	4 802
niveau 85			9 264
niveau 75			
niveau 65			
niveau indéterminé			0
Choix restreint	1 995	4 848	10 299
HMO			
HMO-IPA (acte)	1 995	4 848	10 299
Sans assurance de base	105 932	162 413	162 861
Assurance de base indéterminée	304 182	191 707	212 908
Total assurés	3 101 563	4 969 938	5 438 387

	1993	1994
Soins médico-pharmaceutiques (base)	4 970 975	4 935 537
Traditionnelle	4 613 180	4 509 486
Franchise à option	330 630	388 073
niveau 1	163 155	200 275
niveau 2	102 906	112 313
niveau 3	64 293	75 174
niveau indéterminé	276	311
Bonus	13 093	15 545
niveau 100		4 557
niveau 85	2 051	2 132
niveau 75	3 307	2 498
niveau 65	7 733	6 358
niveau indéterminé	2	0
Choix restreint	14 072	22 433
HMO		
HMO-IPA (acte)	14 072	22 433
Sans assurance de base	150 677	142 570
Assurance de base indéterminée	260 783	200 108
Total assurés	5 382 435	5 278 215

l'effectif des assurés avec assurance de base indéterminée. Le reste de cette classe, soit 154'200 assurés, se réparti en 0.65 % d'assurés avec données de type assuré mais couverture d'assurance inconnue et 99.35 % de malades sans données de type assuré.

Comme nous l'avons mentionné plus haut, deux systèmes de franchises à option étaient en vigueur en 1991. Pour les caisses ayant appliqué pendant toute l'année exclusivement l'un ou l'autre des barèmes de franchises, la répartition en niveau est claire. Cependant, dans quelques cas, les deux systèmes ont été appliqués en parallèle au cours de cette année. Dans un tel cas, pour les enfants avec une franchise de 150.–, nous avons dû choisir entre niveau 2 ou niveau 3. Pour opérer ce choix, nous nous sommes basés sur deux critères. Premièrement, tous les enfants avec une franchise de 150.– ont été attribués au même niveau (soit le niveau 2 ou 3) et, deuxièmement, le niveau a été choisi de sorte que l'évolution des effectifs de 1990 à 1992 soit cohérente. Parmi les assurés avec une assurance de base de type franchise à option, il y en a 1 ‰ dont le niveau est indéterminé; il s'agit pour la plupart des cas d'enfants avec des montants de franchise erronés.

Le mode bonus a été introduit en cette année 1991 et, de ce fait, tous les assurés doivent être au niveau 100.

Pour des raisons techniques, une caisse-maladie n'a pas pu faire figurer les montants des franchises de ses assurés dans les données relatives à l'année 1991. Cependant, des données partielles relatives à cette même année nous avaient été remises précédemment, données qui comprenaient les montants des franchises. Sur cette base, nous avons attribué un montant de franchise à la plupart des assurés de cette caisse. Bien que ceci ne donne pas un reflet 100 % exact de la situation des assurés de cette caisse, cela permet tout de même d'inclure celle-ci dans l'ensemble des données utilisables.

### 5.1.3 Effectifs des assurés en 1992

Pour 1992, les résultats se basent sur la distribution des assurés de 20 caisses, complétée par l'effectif des malades d'une caisse supplémentaire (cf. Table 3.1). Cette caisse contribue pour 22% à l'effectif des assurés avec assurance de base indéterminée. Le reste de cette classe, soit approximativement 166'000 assurés, se réparti en 1.2 % d'assurés avec données de type assuré mais couverture d'assurance inconnue et 98.8 % de malades sans données de type assuré.

Parmi les assurés avec une assurance de base de type franchise à option, il y en a un peu moins de 1 ‰ dont le niveau est indéterminé; il s'agit d'enfants avec des montants de franchise erronés. 1992 étant la deuxième année d'application du mode bonus, les assurés ayant opté pour ce mode sont obligatoirement à l'un des niveaux 100 ou 85. Nous avons également observé qu'une caisse a totalement omis le mode bonus quand bien même elle devrait avoir des assurés de ce type. Cependant, en raison de la petitesse de l'effectif des assurés bonus de cette caisse, cette omission correspond à une perte d'environ 0.5 % de l'effectif global des assurés bonus.

Pour des raisons techniques, une caisse-maladie n'a pas pu faire figurer les montants des franchises de ses assurés dans les données relatives à l'année 1992. Ce problème s'étant déjà présenté pour les données 1991, il a été corrigé ici de manière analogue. Cependant, les données partielles servant de base à l'établissement de la correction sont des données pour 1991. Ceci a pour conséquence

que le résultat reflète moins bien la réalité de cette caisse qu'en 1991 mais cela permet tout de même d'inclure ces données dans l'ensemble des données utilisables.

#### **5.1.4 Effectifs des assurés en 1993**

Pour 1993, les résultats se basent sur la distribution des assurés de 19 caisses, complétée par l'effectif des malades d'une caisse supplémentaire (cf. Table 3.1). Cette caisse contribue pour 18% à l'effectif des assurés avec assurance de base indéterminée. Le reste de cette classe, soit approximativement 215'000 assurés, se répartit en 2.1 % d'assurés avec données de type assuré mais couverture d'assurance inconnue et 97.9 % de malades sans données de type assuré.

Parmi les assurés avec une assurance de base de type franchise à option, il y en a un peu moins de 1 % dont le niveau est indéterminé; il s'agit d'enfants avec des montants de franchise erronés. En 1993, la troisième année d'application du mode bonus, les assurés ayant opté pour ce mode ne peuvent pas être au niveau 65. Comme le montre la Table 5.3, il y a deux assurés mal classés. Nous avons également observé que deux caisses ont totalement omis le mode bonus quand bien même elles devraient avoir des assurés de ce type. Cependant, en raison de la relative petitesse des effectifs des assurés bonus de ces caisses, cette omission correspond à une perte d'environ 3 % de l'effectif global des assurés bonus.

#### **5.1.5 Effectifs des assurés en 1994**

Pour 1994, les résultats se basent sur la distribution des assurés de 20 caisses (cf. Table 3.1). La classe des assurés avec assurance de base indéterminée se répartit en 1.6 % d'assurés avec données de type assuré mais couverture d'assurance inconnue et 98.4 % de malades sans données de type assuré.

Parmi les assurés avec une assurance de base de type franchise à option, il y en a un peu moins de 1 % dont le niveau est indéterminé; il s'agit à nouveau d'enfants avec des montants de franchise erronés. Au sujet du mode bonus, nous observons à nouveau qu'une caisse a totalement omis ce mode alors qu'elle devrait avoir des assurés de ce type. Cependant, en raison de la relative petitesse de l'effectif des assurés bonus de cette caisse, cette omission correspond à une perte d'environ 2.5 % de l'effectif global des assurés bonus.

## **5.2 Primes par types d'assurance**

Considérons maintenant la comptabilisation des primes par types d'assurance. Dans ce cas, nous avons tout d'abord réparti les assurances en 2 catégories : assurance de base pour soins médico-pharmaceutiques et autres assurances. Nous avons alors additionné toutes les primes payées par l'ensemble des assurés au titre d'une assurance de base pour soins médico-pharmaceutiques d'une part et toutes les autres primes d'autre part. Il est important de relever que le total des primes de la catégorie assurance de base ne correspond pas au total des primes payées par les assurés classés "assurés avec une assurance de base pour soins médico-pharmaceutiques" dans la section

précédente. En effet, les primes payées par ces assurés pour une assurance complémentaire ou pour une assurances d'indemnités journalières par exemple sont comptabilisées dans la catégorie "autres assurances".

La catégorie assurance de base comprend non seulement les primes pour l'assurance de base proprement dite<sup>1</sup>, mais également les primes pour l'assurance obligatoire d'indemnités journalières en cas d'hospitalisation. Les primes de l'assurance de base pour soins médico-pharmaceutiques sont alors elles-mêmes réparties en 5 groupes : traditionnel, franchise à option, bonus, choix restreint, indemnités journalières obligatoires en cas d'hospitalisation. Ici aussi, on notera que nous n'avons pas tenu compte de l'affectation de l'assuré définie dans la section précédente mais, bien de l'affectation des primes elles-mêmes. La mise en vigueur de l'arrêté fédéral urgent visant à réduire les prestations de l'assurance de base au minimum légal a pour conséquence que certaines prestations comprises dans l'assurance de base jusqu'en 1992 font partie, dès 1993, d'assurances complémentaires qualifiées de compléments statutaires. De ce fait, dès 1993, les primes correspondant à ces compléments statutaires sont enregistrées séparément. Cependant, pour des raisons de cohérence, ces primes sont également comptabilisées dans la catégorie assurance de base et constituent une catégorie supplémentaire dans notre analyse. Pour des raisons de simplification, la répartition des primes d'assurance soins médico-pharmaceutiques avec franchise à option en fonction des niveaux de franchise n'a pas été effectuée dans le cadre de cette statistique élémentaire. De même, les primes de l'assurance avec système de bonus n'ont pas été distribuées selon le niveau du bonus et celles des assurances avec choix limité des fournisseurs de prestation n'ont pas été ventilées selon les modèles d'assurance.

Pour la période allant de 1990 à 1994, les répartitions des primes sont présentées dans la Table 5.4. Dans ces résultats, seules sont prises en compte les répartitions des primes des caisses pour lesquelles des données de type assuré utilisables sont disponibles, selon l'état de la Table 3.1. Quelques commentaires sur ces résultats sont formulés ci-après, année par année.

### 5.2.1 Répartition des primes en 1990

Pour 1990, les résultats regroupent les primes annoncées par 15 caisses. Le mode bonus n'a été introduit par les caisses qu'en 1991; ainsi, il ne peut pas y avoir de prime bonus en 1990.

### 5.2.2 Répartition des primes en 1991

Pour 1991, les résultats regroupent les primes annoncées par 18 caisses. Pour une des caisses-maladie, la répartition des primes de l'assurance soins médico-pharmaceutiques en fonction du mode d'assurance est purement indicative. En effet, en raison des problèmes évoqués au cours de l'analyse de la répartition des assurés (section 5.1.2), les données de la caisse ont subi des modifications avant que cette répartition puisse être établie, modifications qui ne sont pas 100 % exactes. Malgré cela, la répartition des primes de cette caisse est prise en compte dans la Table 5.4.

---

<sup>1</sup> Assurance ordinaire (traditionnelle), assurance avec franchise à option, assurance avec bonus et assurances avec choix limité des fournisseurs de prestations.



### 5.2.3 Répartition des primes en 1992

Pour 1992, les résultats regroupent les primes annoncées par 20 caisses. De même qu'en 1991 et, pour les mêmes raisons, les données d'une des caisses ont dû subir des modifications afin de permettre de répartir les primes de l'assurance soins médico-pharmaceutiques selon le mode d'assurance. Comme les modifications ne sont pas absolument correctes, la répartition des primes obtenue pour cette caisse ne reflète pas exactement la réalité. Malgré cela, la répartition des primes de cette caisse est prise en compte dans la Table 5.4.

### 5.2.4 Répartition des primes en 1993 et 1994

Pour 1993, respectivement 1994, les résultats regroupent les primes annoncées par 19, respectivement 20 caisses.

## 5.3 Répartition des coûts par branches d'assurance et par genres de cas

L'analyse des données de type prestation porte ici uniquement sur les coûts; les aspects nombre de consultations, nombre d'hospitalisations, etc., sont volontairement omis afin de simplifier la présente étude. La répartition des coûts entre les différentes branches d'assurance<sup>1</sup> ainsi que la répartition entre les divers genres de cas<sup>2</sup> sont étudiées et le total des coûts des composantes de frais est comparé au total des factures.

Pour cette analyse, le premier pas consiste à répartir les enregistrements de type prestation en deux groupes : les enregistrements "conservés" et ceux qui sont "abandonnés". Les conditions qui entraînent un abandon d'un enregistrement sont soit l'absence de décomposition des coûts en composantes de frais<sup>3</sup>, soit une décomposition des coûts en composantes non cohérente avec le montant total de la facture<sup>4</sup>. La Table 5.5 présente les résultats généraux de cette analyse, pour les années 1990 à 1994.

La statistique des coûts recensés dans les enregistrements "conservés" est présentée dans la portion intitulée "Coûts analysés" de la Table 5.5. Le total des factures et le total des coûts des composantes de frais (total des frais répartis) sont comparés. Toutefois, une certaine différence entre ces deux totaux est admissible car les coûts non reconnus par les caisses-maladie sont inclus dans les totaux des factures alors qu'ils n'apparaissent pas dans la décomposition des frais. Ainsi, cet indicateur de qualité est à manipuler avec une certaine prudence. A ces éléments nous avons ajouté le total des participations des assurés aux frais de traitement médicaux. Observons que ce total est une valeur corrigée. En effet, lorsque le montant de la participation figurant sur un enregistrement est supérieur

---

<sup>1</sup> Assurance de base, assurances complémentaires, assurances d'indemnités journalières, ...

<sup>2</sup> Maladie – accident, maternité, autre.

<sup>3</sup> Dans ce cas, le montant total est en principe connu mais n'est pas réparti entre les différentes branches d'assurance.

<sup>4</sup> Le total des composantes de frais est supérieur au montant de la facture.

au total de la facture concernée, la participation prise en compte dans la statistique est plafonnée au total de la facture.

Pour le groupe des enregistrements "abandonnés", la même statistique est produite (portion "Coûts abandonnés" de la Table 5.5). Comme la cohérence des données des enregistrements concernés n'est pas garantie, ces chiffres sont donnés à titre purement indicatifs et le total des factures n'est pas comparé à celui des composantes de frais.

Un autre élément qu'il nous a également paru intéressant d'ajouter dans cette table est le nombre de malades concernés<sup>1</sup>.

Pour l'analyse des coûts des prestations en fonction des branches d'assurance, nous avons constitué les 5 catégories suivantes : assurance de base, assurances complémentaires, assurances d'indemnités journalières, autres assurances et indéfini. L'attribution des coûts dans ces catégories se fait sur la base des informations contenus dans les blocs de données décrivant les diverses composantes de frais des enregistrements "conservés". La catégorie assurance de base regroupe toutes les prestations à la charge de l'assurance de base pour soins médico-pharmaceutiques, les prestations prises en charge dans le cadre des compléments statutaires à l'assurance de base ainsi que les prestations de l'assurance obligatoire d'indemnités journalières en cas d'hospitalisation. Cette catégorie est également subdivisée en prestations de l'assurance de base en cas de séjours hospitaliers et prestations de l'assurance de base pour traitements ambulatoires. La catégorie des assurances complémentaires regroupe quant à elle toutes les prestations des assurances complémentaires<sup>2</sup>; la répartition entre compléments pour soins hospitaliers et compléments pour soins ambulatoires est aussi présentée. La catégorie des assurances d'indemnités journalières est constituée des prestations des assurances d'indemnités journalières pour pertes de gains et la catégorie autres assurances groupe les prestations de tous les types d'assurances ne faisant pas partie des trois catégories précédentes. Finalement, toutes les prestations portant un code dont la signification est inconnue sont regroupées dans la catégorie genre d'assurance indéfini. Pour les années 1990 à 1994, cette répartition des coûts par genre d'assurance est présentée dans la Table 5.6.

Pour terminer, une statistique des coûts par genre de cas est donnée dans la Table 5.7 (période allant de 1990 à 1994). Dans ce cas, les coûts sont répartis entre frais pour les cas de maladie ou accident, frais pour les cas de maternité et frais qui ne sont attribués à aucun des deux cas précédents. Comme pour la statistique des coûts par branches d'assurance, cette statistique se base sur l'ensemble des frais répartis des enregistrements "conservés".

Dans les Tables 5.5, 5.6 et 5.7, seuls sont pris en compte les coûts des caisses pour lesquelles des données de type prestation utilisables sont disponibles, selon l'état de la Table 3.1. Quelques commentaires sur ces résultats sont formulés ci-après, année par année.

---

<sup>1</sup> Un malade est un assuré pour lequel il existe au moins un enregistrement de type prestation pour l'année considérée.

<sup>2</sup> Par exemple, les assurances d'indemnités journalières en cas d'hospitalisation, les assurances de frais de traitement en cas d'hospitalisation, les assurances combinées en cas d'hospitalisation, les assurances pour frais de traitement dentaires. Par contre, les prestations des compléments statutaires ne sont pas inclus dans cette catégorie.

Table 5.5 : Statistique globale des coûts.

<b>Malades</b>					
	1990	1991	1992	1993	1994
Nombre de malades	2 336 835	3 699 329	4 028 430	4 035 433	3 965 828

<b>Coûts analysés</b>					
	1990	1991	1992	1993	1994
Total des frais répartis	4 260 926 906	7 565 363 929	8 928 250 565	9 764 398 583	10 113 650 114
Total des factures	4 489 111 051	7 996 945 015	9 472 256 994	10 220 587 910	10 560 346 664
Différence (factures - répartis)	228 184 145	431 581 086	544 006 429	456 189 327	446 696 550
Participation des assurés	564 611 396	994 642 014	1 343 325 123	1 195 735 685	1 171 788 107

<b>Coûts abandonnés</b>					
	1990	1991	1992	1993	1994
Total des frais répartis	1 095 761	6 794 586	60 765 920	78 325 898	73 762 898
Total des factures	6 689 452	9 315 636	55 108 252	80 314 423	82 995 649
Participation des assurés	6 201 708	7 379 067	10 203 202	3 282 654	4 016 496

Table 5.6 : Coûts par branches d'assurance.

	1990	1991	1992
Assurance de base	3 266 608 253	6 002 092 474	6 956 600 986
▲ Séjours hospitaliers	1 158 724 164	2 265 615 378	2 558 275 216
▲ Soins ambulatoires	2 107 884 089	3 736 477 096	4 398 325 770
Assurances complémentaires	685 447 583	1 233 863 042	1 507 174 841
▲ Soins hospitaliers	670 057 777	1 201 802 669	1 456 025 378
▲ Soins ambulatoires	15 389 806	32 060 373	51 149 463
Assurance indemnités journalières	197 648 765	278 651 067	405 207 061
Autres assurances	110 974 190	48 473 978	58 721 319
Genre d'assurance indéfini	248 115	2 283 368	546 358
Total des frais répartis	4 260 926 906	7 565 363 929	8 928 250 565

	1993	1994
Assurance de base	7 529 520 146	7 761 779 572
▲ Séjours hospitaliers	2 822 134 481	2 870 589 293
▲ Soins ambulatoires	4 707 385 665	4 891 190 279
Assurances complémentaires	1 685 100 249	1 749 556 168
▲ Soins hospitaliers	1 623 324 528	1 673 939 126
▲ Soins ambulatoires	61 775 721	75 617 042
Assurance indemnités journalières	476 214 125	483 714 106
Autres assurances	71 401 264	116 527 661
Genre d'assurance indéfini	2 162 799	2 072 607
Total des frais répartis	9 764 398 583	10 113 650 114

Table 5.7 : Coûts par genre de cas.

	1990	1991	1992
Maladie ou accident	3 958 801 149	7 042 307 043	8 229 492 063
Maternité	177 009 786	297 219 238	350 263 203
Autres ou non attribués	125 115 971	225 837 648	348 495 299
<b>Total des frais répartis</b>	<b>4 260 926 906</b>	<b>7 565 363 929</b>	<b>8 928 250 565</b>

	1993	1994
Maladie ou accident	9 033 939 544	9 392 727 375
Maternité	364 248 040	339 734 757
Autres ou non attribués	366 210 999	381 187 982
<b>Total des frais répartis</b>	<b>9 764 398 583</b>	<b>10 113 650 114</b>

### 5.3.1 Analyse des coûts pour l'année 1990

Pour 1990, les résultats regroupent les coûts annoncés par 14 caisses. Deux caisses-maladie ont des composantes de frais non attribuables; cependant le montant de cette catégorie ne constitue qu'une infime proportion de l'ensemble des factures et est bien inférieur au montant des factures qui n'est pas réparti. D'autre part, trois caisses-maladie présentent un montant de prestations non réparti supérieur à 10 % du total des factures (ce qui nous semble anormalement élevé) alors que cinq autres caisses présentent un montant de prestations non réparti compris entre 5 et 10 % du total des factures.

Deux caisses ont un total de participation des assurés inférieur à 1 % du montant total des factures, ce qui semble manifestement faux car pour toutes les autres caisses, cette participation dépasse le 7 % du total des coûts. A l'autre extrême, il y a deux caisses pour lesquelles la participation dépasse le 20 % du total des coûts, avec même une pointe à près de 50 % résultant fort probablement de données erronées.

### 5.3.2 Analyse des coûts pour l'année 1991

Pour 1991, les résultats regroupent les coûts annoncés par 15 caisses. Trois caisses-maladie comportent des composantes de frais non attribuables; cependant le montant de cette catégorie ne constitue qu'une infime proportion de l'ensemble des factures et est bien inférieur au montant des factures qui n'est pas réparti. D'autre part, une caisse-maladie présente un montant de prestations

non réparti supérieur à 10 % du total des factures et huit caisses présentent un montant de prestations non réparti compris entre 5 et 10 % du total des factures.

Deux caisses ont un total de participation des assurés inférieur à 1 % du montant total des factures, ce qui semble manifestement faux car pour toutes les autres caisses, cette participation dépasse le 7 % du total des coûts (avec même une pointe au dessus de 20 % dans un cas).

### **5.3.3 Analyse des coûts pour l'année 1992**

Pour 1992, les résultats regroupent les coûts annoncés par 16 caisses. Trois caisses-maladie comportent des composantes de frais non attribuables; cependant le montant de cette catégorie ne constitue qu'une infime proportion de l'ensemble des factures et est bien inférieur au montant des factures qui n'est pas réparti. D'autre part, aucune caisse-maladie présente un montant de prestations non réparti supérieur à 10 % du total des factures (contrairement au deux années précédentes) et huit caisses présentent un montant de prestations non réparti compris entre 5 et 10 % du total des factures.

Deux caisses ont un total de participation des assurés inférieur à 1 % du montant total des factures, ce qui semble manifestement faux car pour toutes les autres caisses, cette participation dépasse le 6 % du total des coûts. A l'autre extrême, il y a deux caisses pour lesquelles la participation dépasse le 20 % du total des coûts, avec même une pointe à près de 50 % résultant fort probablement de données erronées.

### **5.3.4 Analyse des coûts pour l'année 1993**

Pour 1993, les résultats regroupent les coûts annoncés par 16 caisses. Deux caisses-maladie comportent des composantes de frais non attribuables; cependant le montant de cette catégorie ne constitue qu'une infime proportion de l'ensemble des factures. Pour l'une d'elles, ce montant est bien inférieur au montant des factures qui n'est pas réparti tandis que pour l'autre, il atteint approximativement le 30 % du montant qui n'est pas réparti, ce qui peut être considéré comme important. D'autre part, une caisse-maladie présente un montant de prestations non réparti supérieur à 10 % du total des factures et cinq caisses présentent un montant de prestations non réparti compris entre 5 et 10 % du total des factures.

Deux caisses ont un total de participation des assurés inférieur à 1 % du montant total des factures, ce qui semble manifestement faux car pour toutes les autres caisses, cette participation dépasse le 6.5 % du total des coûts (avec même une pointe au dessus de 20 % dans deux cas).

### **5.3.5 Analyse des coûts pour l'année 1994**

Pour 1994, les résultats regroupent les coûts annoncés par 16 caisses. Quatre caisses-maladie comportent des composantes de frais non attribuables; cependant le montant de cette catégorie ne constitue qu'une infime proportion de l'ensemble des factures. Pour trois d'entre elles, ce montant est bien inférieur au montant des factures qui n'est pas réparti tandis que pour la quatrième, il atteint

approximativement le 16 % du montant qui n'est pas réparti, ce qui peut être considéré comme important. D'autre part, une caisse-maladie présente un montant de prestations non réparti supérieur à 10 % du total des factures et six caisses présentent un montant de prestations non réparti compris entre 5 et 10 % du total des factures.

Deux caisses ont un total de participation des assurés inférieur à 1 % du montant total des factures, ce qui semble manifestement faux car pour toutes les autres caisses, cette participation dépasse le 6 % du total des coûts (avec même une pointe au dessus de 20 % dans un cas).

## 6 Base pour les exploitations

Malheureusement, comme nous l'avons déjà remarqué précédemment, toutes les caisses n'ont pas été en mesure de livrer l'ensemble des données (et ce, pour diverses raisons). Selon la Table 3.1, seules 10 caisses (dont 7 soumises à l'obligation) nous ont fourni l'ensemble des données pour les années 1989 à 1994. Or, la répartition géographique des assurés d'une part et la répartition de ces assurés entre les divers types d'assurance de base pour soins médico-pharmaceutiques<sup>1</sup> d'autre part diffèrent fortement d'une caisse à l'autre. Ainsi, pour assurer la validité de nos travaux et pour permettre la comparaison des résultats annuels, nous sommes contraints de restreindre les exploitations à un ensemble de caisses ayant été en mesure de fournir des données complètes pour toute la période sur laquelle porte l'étude. Dans la suite de cette section, nous utiliserons le terme d'échantillon pour qualifier cet ensemble de caisses dont les données sont utilisées pour les analyses<sup>2</sup>.

### 6.1 Choix de l'échantillon

Pour ne pas perdre trop d'informations, nous avons tout d'abord limité la période considérée aux années 1989 à 1994 (ce qui justifie également à posteriori la restriction de la Table 3.1 à cette période). Des Figures 4.2 et 4.4, il ressort que, pour chaque type de données, l'ensemble des données utilisables est nettement plus petit en 1990 que pour les années 1991–1994. Cependant, ce qu'il faut considérer ici, c'est l'ensemble des caisses pour lesquelles à la fois les données de type assuré et celles de type prestation sont utilisables pour une année donnée. Dans cette optique, l'examen de la Table 3.1 confirme qu'il existe une différence importante entre les deux périodes 1989–1990 et 1991–1994. Ainsi, nous avons divisé la période 1989–1994 en deux : 1989–1990 et 1991–1994. Chacune de ces deux périodes sera considérée séparément et chacune aura son propre échantillon de caisses. Il est encore à relever ici que l'ouverture des premières caisses HMO s'est déroulée au cours des derniers mois de 1990 et que l'introduction de l'assurance bonus n'a eu lieu qu'en 1991. Ainsi, ce découpage reflète tout de même la distinction entre la phase d'essai des nouvelles formes d'assurance (1991–1994) et une phase antérieure (1989–1990). Cependant, la phase antérieure est nettement plus courte que celle prévue lors de la préparation du projet.

#### 6.1.1 Période 1991 – 1994

Pour cette période, l'ensemble retenu se compose de 15 caisses : CSS, Helvetia, KKB, Grütli, KONKORDIA, OeKK Luzern, OeKK Basel, La Caisse Vaudoise, KFW, ZOKU, OeKK Winterthur, INTRAS, FRV, ASSURA et EVIDENZIA.

<sup>1</sup> Caisses HMO, assurances bonus, assurances avec franchise annuelle à option et assurances traditionnelles.

<sup>2</sup> Il ne s'agit donc pas d'un véritable échantillon d'assurés au sens statistique du terme.

Selon la Table 3.1, la FRV n'a pas livré toutes les données nécessaires. Or, au cours de la phase préparatoire du projet, les caisses ont livré un fichier de données de type assuré représentant l'état au milieu de l'année 1991. Comme, de plus, l'effectif d'assurés de cette caisse est relativement petit<sup>1</sup>, nous avons estimé que, pour cette caisse, nous pouvions remplacer le fichier d'assuré de l'année 1991 par ce fichier intermédiaire et, les fichiers 1992 et 1993 par le fichier d'assurés 1994.

### 6.1.2 Période 1989 – 1990

Dans ce cas, l'ensemble des caisses retenues se réduit à 12 des 15 caisses de l'échantillon précédent. Les trois caisses de l'échantillon 1991–1994 qui ne font pas partie de l'échantillon 1989–1990 sont : Helvetia, KKB et EVIDENZIA. Ces trois caisses possèdent ensemble un peu plus de 1,5 millions d'assurés, ce qui justifie également le fait de travailler avec deux échantillons.

Selon la Table 3.1, la FRV et la KFW n'ont pas livré toutes les données nécessaires. Dans les deux cas, il est possible de remplacer le fichier d'assurés de l'année 1990 par le fichier de données intermédiaire (état des assurés au milieu de l'année 1991; cf. ci-dessus, période 1991–1994). Pour la FRV, ceci est cohérent avec ce qui est fait dans le cadre de l'échantillon pour la période 1991–1994. Pour la KFW, ce procédé ne peut pas être justifié aussi simplement car l'effectif de la caisse dépasse les 400 milles membres. Cependant, à fin 1991, plus de 97 % des assurés de cette caisse possédaient une assurance pour soins médico-pharmaceutiques de type traditionnel. De ce fait, l'emploi de données intermédiaire en lieu et place des données de type assuré de l'année 1990 n'introduit que peu d'erreurs dans les types de branches d'assurance. Seules les exploitations faisant appel aux durées de risques ou aux montants des primes doivent faire l'objet d'une certaine attention.

## 6.2 Représentativité de l'échantillon par rapport à l'ensemble des 25 caisses participantes — période 1991–1994

Comme nous l'avons relevé en introduction de cette section 6, la répartition des assurés entre les divers types d'assurance de base pour soins médico-pharmaceutiques diffère fortement d'une caisse à l'autre. La question se pose donc de savoir si le choix de l'échantillon introduit un biais dans cette répartition par rapport à celle que nous aurions eue s'il avait été possible de travailler avec l'ensemble des 25 caisses. Nous ne considérons ici que ce problème. Les questions telles que la représentativité de l'échantillon par rapport aux populations cantonales seront abordées ailleurs.

Pour traiter ce problème, nous avons besoin des effectifs d'assurés ventilés selon le type d'assurance pour chacune des caisses ne faisant pas partie de l'échantillon, ceci pour chacune des années 1991 à 1994. Dans le meilleurs des cas, ces données sont disponibles pour certaines années dans les données administratives<sup>2</sup> et ont été utilisées ici. Dans d'autres cas, des données de type assuré

---

<sup>1</sup> De l'ordre de 40'000, ce qui est peu à côté des 4 à 5 millions d'assurés que représente l'ensemble des caisses.

<sup>2</sup> Il s'agit, par exemple, de caisses qui n'ont pas été retenues dans l'échantillon en raison du manque de données de type prestation utilisables.

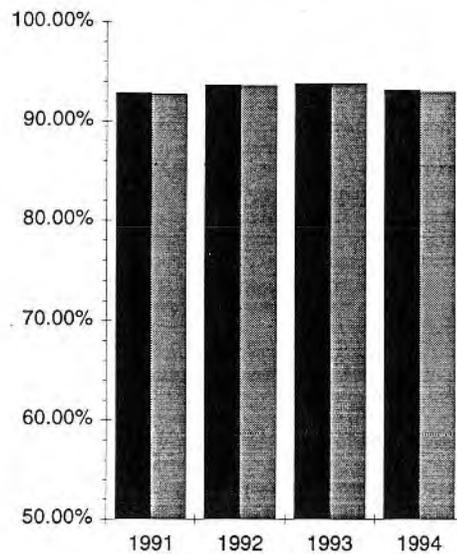
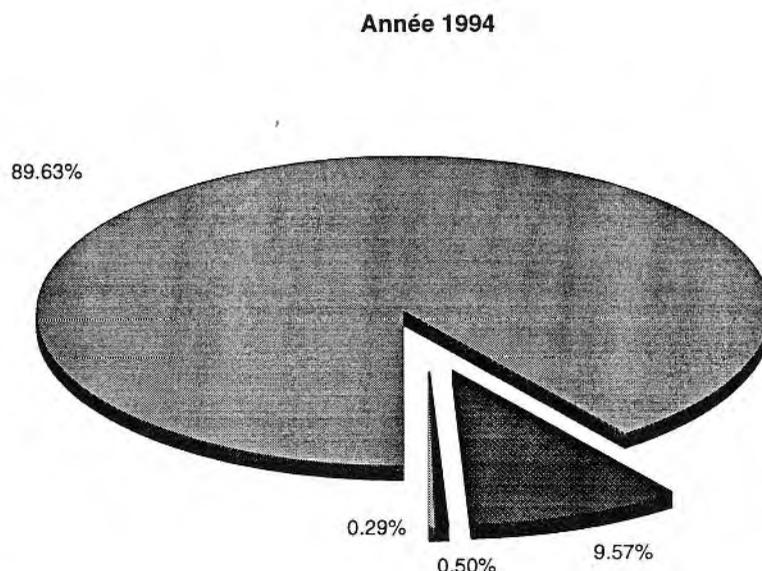


Figure 6.1 : Proportion des assurés avec une assurance pour soins médico-pharmaceutiques par rapport à l'ensemble des membres.

■ 25 caisses      ▨ 15 caisses

existent mais sont globalement inutilisables pour diverses raisons. Pour ce type de situation, nous avons procédé à une estimation en nous basant sur les observations faites dans ces données administratives et sur les données fournies par ces caisses dans le cadre du questionnaire pour la statistique de l'assurance-maladie. Finalement, dans le pire des cas, nous utilisons une estimation basée uniquement sur les données fournies par ces caisses dans le cadre du questionnaire pour la statistique de l'assurance-maladie.

La première question examinée est celle de la proportion des assurés avec une assurance pour soins médico-pharmaceutiques par rapport à l'ensemble des membres des caisses. Pour chacune des années 1991 à 1994, cette proportion diffère de moins d'un pour-cent entre l'échantillon (constitué de 15 caisses) et l'ensemble des 25 caisses (Figure 6.1). Il est donc clair que le choix de l'échantillon n'a pas eu d'influence à ce niveau. Le second point examiné est celui de la répartition des assurés avec assurance de soins médico-pharmaceutiques en fonction du type de leur assurance. La première observation qu'il convient de faire est que la proportion des assurés traditionnels par rapport à l'ensemble des assurés en soins varie de 95 à 90 % au cours de ces 4 années, celle des assurés en mode franchise à option varie de 4 à 10 % et celle des assurés en mode choix restreint, respectivement en mode bonus, croît continuellement pour atteindre 0.5 %, respectivement 0.3 %, en 1994. La Figure 6.2 illustre cette répartition pour l'année 1994 (ensemble des 25 caisses). La comparaison des répartitions par type d'assurance des assurés de l'échantillon et de ceux de l'ensemble des 25 caisses montre que les différences demeurent constantes au cours des années. Il apparaît ainsi que les assurés traditionnels sont légèrement sur-représentés dans l'échantillon alors que ceux en franchise à option sont légèrement sous-représentés. La proportion des assurés en



*Figure 6.2 : Répartition par type d'assurance des assurés en soins médico-pharmaceutiques des 25 caisses.*

■ traditionnelle   ■ franchise à option   ■ choix restreint   ■ bonus

---

mode choix restreint ainsi que celle des assurés en mode bonus reflètent de manière équitable la situation de l'ensemble des caisses (Figure 6.3).

### **6.3 Comparaison des données de l'échantillon avec celles de la statistique de l'assurance-maladie 1994**

Diverses données statistiques sont communiquées à l'OFAS par les caisses-maladie dans le cadre de la statistique de l'assurance-maladie. Nous y trouvons en particulier les effectifs d'assurés et les primes encaissées par branches pour l'assurance des soins médico-pharmaceutiques ainsi que les prestations à la charge de cette même assurance des soins. La comparaison de ces données avec les résultats tels qu'ils ressortent de nos analyses des données administratives fournit également des indications sur la qualité des données reçues des caisses-maladie. Nous n'avons effectué la comparaison que pour l'année 1994 car ce n'est que depuis cette année que nous disposons de données suffisamment détaillées dans la statistique de l'assurance-maladie.

Précisons d'emblée que, pour ces comparaisons, nous avons dû restreindre l'échantillon à 14 caisses. En effet, une caisse-maladie n'a livré les données administratives que pour une partie de ses membres tandis que la totalité de ceux-ci sont pris en compte dans les données pour la statistique de

l'assurance-maladie. Ce "déficit" dans la livraison des données administrative est motivé par des raisons de gestion informatique internes à cette caisse (il ne s'agit pas de données oubliées par erreur).

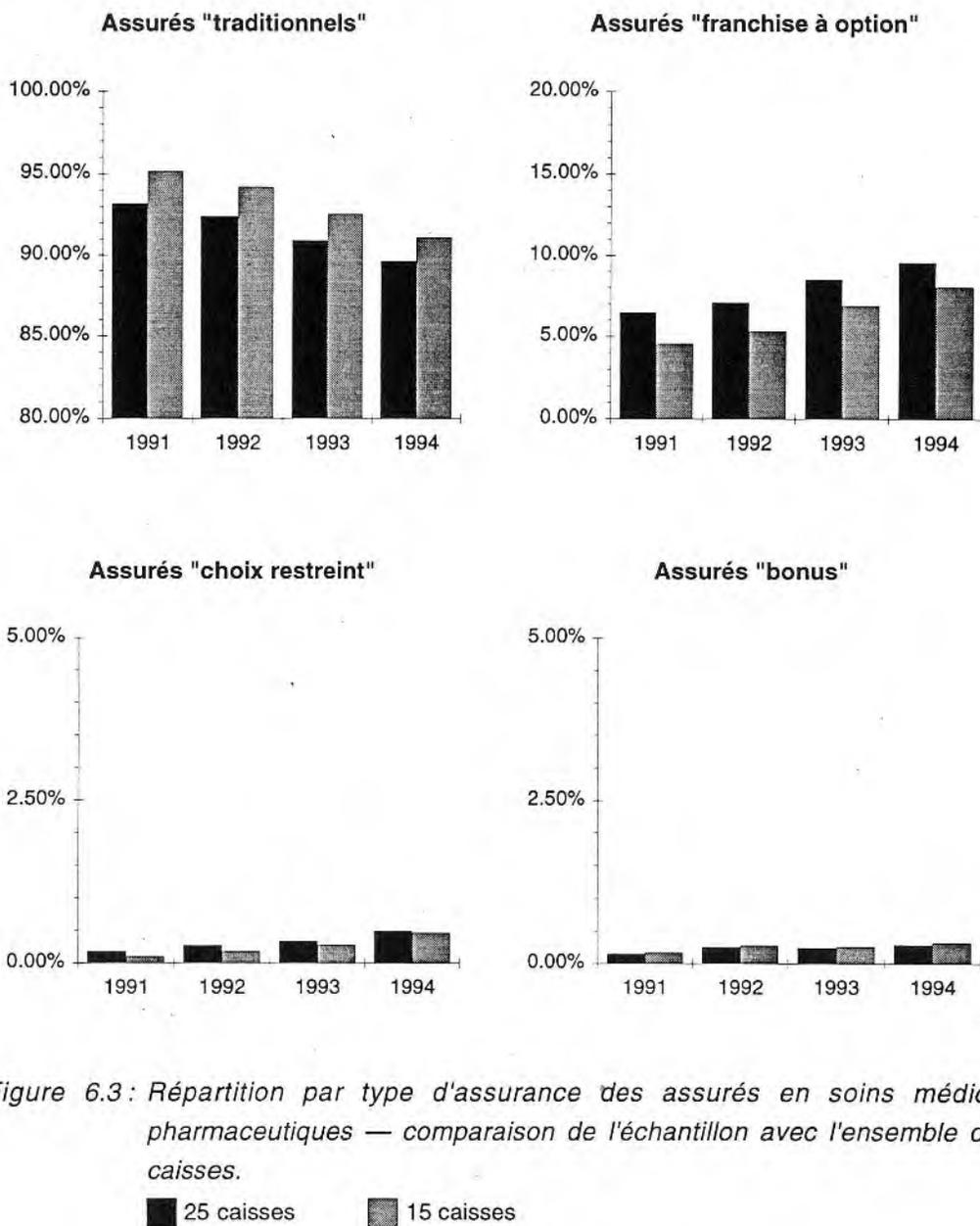


Figure 6.3: Répartition par type d'assurance des assurés en soins médico-pharmaceutiques — comparaison de l'échantillon avec l'ensemble des caisses.

■ 25 caisses    ■ 15 caisses

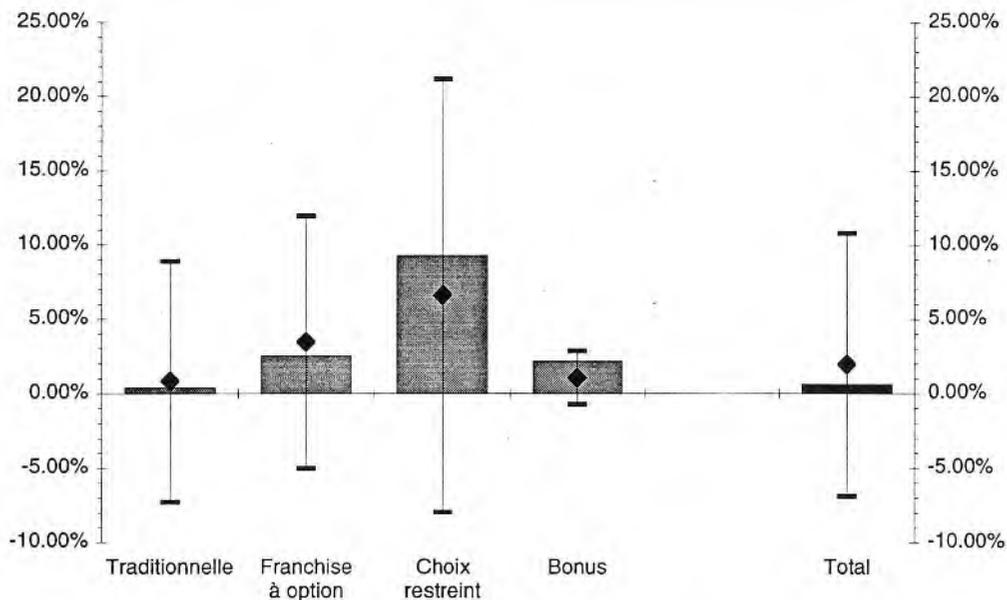


Figure 6.4 : Différences entre les données de l'échantillon et celles de la statistique de l'assurance-maladie — Effectifs des assurés par branches pour 1994.

■ Différence      ◆ Moyenne des différences

### 6.3.1 Effectifs des assurés par type d'assurance pour soins médico-pharmaceutiques

Considérons tout d'abord l'effectif des assurés par type d'assurance pour soins médico-pharmaceutiques. Pour cette comparaison, nous avons déterminé la différence relative des effectifs branche par branche. Par exemple, pour le type bonus, nous avons déterminé l'effectif total des assurés de ce type dans l'échantillon, soit à partir des données administratives, soit à partir des données pour la statistique. La différence relative résulte alors de la division de la différence de ces deux nombres par l'effectif résultant de la statistique. En effectuant ceci pour tous les types ainsi que pour l'ensemble des assurés en soins médico-pharmaceutiques ("Total"), nous obtenons l'histogramme représenté dans la Figure 6.4. Nous observons alors que cette différence est extrêmement faible pour l'ensemble des assurés en soins médico-pharmaceutiques ainsi que pour le type traditionnel et qu'elle est faible pour les groupes bonus et franchise à option alors qu'elle est sensiblement plus élevée pour le type choix restreint<sup>1</sup>. Cependant, pour tous les groupes, l'effectif est plus élevé dans les données administratives. Ceci peut s'expliquer par le fait que, dans ces données administratives, nous avons tous les assurés qui ont été membres d'une des caisses de l'échantillon

<sup>1</sup> Assurances avec choix limité des fournisseurs de prestations.

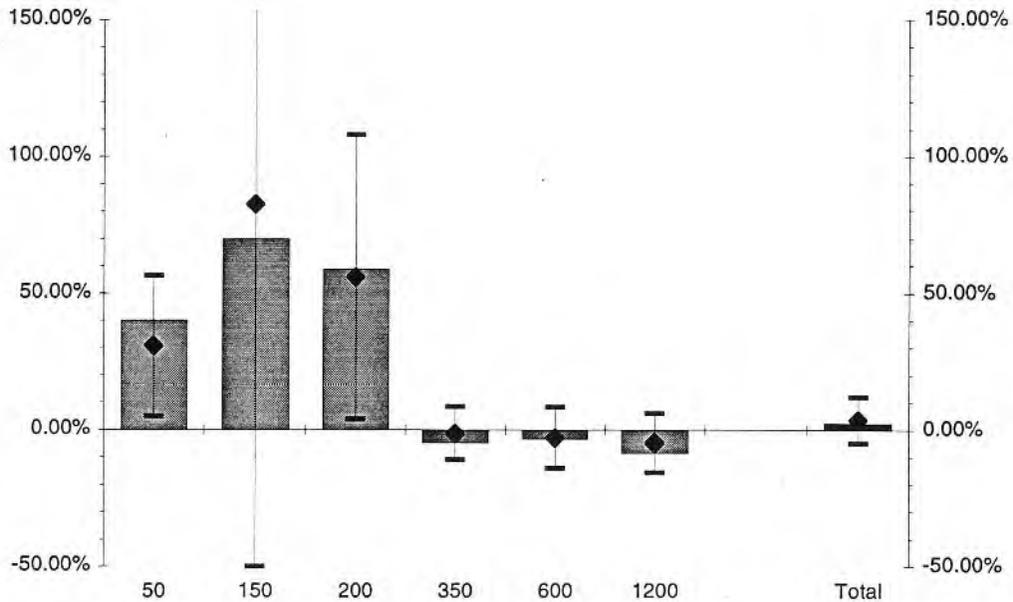


Figure 6.5 : Différences entre les données de l'échantillon et celles de la statistique de l'assurance-maladie — Effectifs des assurés "franchise à option" pour 1994.

■ Différence      ◆ Moyenne des différences

à un moment quelconque de l'année alors que les effectifs livrés pour la statistique de l'assurance-maladie décrivent la situation au 31 décembre 1994. Observons également que, comme la différence est relative à l'effectif du groupe considéré, le résultat global n'est pas égal à la somme des différences des quatre modes.

Cet histogramme ne fournit qu'une information globale pour chaque groupe de l'échantillon. Cependant, les résultats individuels des caisses sont également à considérer. Pour cela, nous avons déterminé les différences relatives pour chacune des 14 caisses retenues ici, puis nous avons calculé les moyennes de ces différences relatives, moyennes que nous avons superposé à l'historgramme de la Figure 6.4. De plus, nous avons également déterminé la variabilité des résultats des caisses autour de cette moyenne<sup>1</sup>, variabilité qui est également représentée dans la Figure 6.4. Une caisse présente une différence relative de -100 % pour le type bonus tandis qu'elle représente à peine le 2 % de l'effectif de ce groupe. Ainsi, pour ne pas fausser l'image présentée, la moyenne ainsi que la variabilité des différences reportées sur la Figure 6.4 pour le type bonus ne tiennent pas compte de cette caisse. En ce qui concerne les moyennes des différences, nous observons qu'elles peuvent

<sup>1</sup> Intervalle centré autour de la moyenne et de longueur égale à deux fois l'écart-type.

être supérieures aux différences observées sur l'échantillon (comme, par exemple, pour le type franchise à option). Ceci met simplement en évidence l'effet de compensation qui se produit entre les caisses à différence positive et celles à différence négative<sup>1</sup>. La variabilité est, quant à elle, relativement importante, sauf pour le type bonus.

Les données disponibles dans la statistique de l'assurance-maladie permettent de faire le même exercice pour le type franchise à option réparti par montant des franchises, ce qui aboutit aux résultats présentés dans la Figure 6.5. D'un point de vue global, nous observons que la différence relative est moyenne à faible pour les franchises adultes tandis qu'elle est extrêmement élevée pour les franchises enfants, ce qui permet de mettre en doute la fiabilité des informations concernant le type franchise à option pour les enfants. De plus, nous observons que cette différence est négative pour toutes les franchises adultes (ce qui traduit le fait surprenant que les effectifs issus des données administratives sont inférieurs à ceux déclarés par les caisses pour la statistique) tandis qu'elle est positive dans les cas des enfants. La variabilité des moyennes entre caisses est, quant à elle, moyennement importante dans le cas des franchises des adultes alors qu'elle est forte à très forte pour les franchises de 50 et 200 francs. Le cas de la franchise à 150 francs, pour lequel cette variabilité est encore plus élevée, constitue un cas à part, notamment en raison des problèmes déjà évoqués dans la section 2.4.

D'une manière générale, nous avons déjà remarqué que les effectifs issus des données administratives devraient toujours être supérieurs à ceux présentés dans la statistique. Or, pour plusieurs caisses, l'effectif total issu des données administratives est inférieur à celui donné dans la statistique. Pour de telles caisses, il est clair que l'un de ces deux effectifs est faux (sans que nous puissions déterminer lequel). Pour les effectifs par branches, il est plus difficile de tirer des conclusions car les différences peuvent également être dues à l'algorithme de classification des assurés que nous avons utilisé dans les données administratives.

### 6.3.2 Primes par type d'assurance pour soins médico-pharmaceutiques

Pour les données administratives, la répartition des primes d'assurance pour soins médico-pharmaceutiques selon le mode d'assurance est effectuée en fonction de l'affectation des primes elles-mêmes, sans tenir compte de la classification des assurés (cf. section 5.2). Cette façon de faire, purement comptable, devrait correspondre avec la pratique des caisses-maladie lors de l'établissement du questionnaire pour la statistique. La différence relative des primes par branches d'assurance, ainsi que la moyenne et la variabilité des différences par caisses sont déterminées de manière analogue à la procédure adoptée pour les effectifs d'assurés. Les résultats présentés dans la Figure 6.6 révèlent des différences relatives par type d'assurance nettement plus élevées que ce n'était le cas pour les effectifs (sauf pour la franchise à option et le bonus). Ceci est surprenant car, contrairement à ce qui était le cas pour les effectifs, nous n'avons pas de différence de principe sur l'établissement de ces montants. Si l'on observe de plus près les résultats caisse par caisse, on remarque que pour certaines, ces différences relatives ne dépassent pas le pour cent (pour chacun des

---

<sup>1</sup> Bien que nous ayons observé plus haut que l'effectif des assurés dans les données administratives devrait être supérieur à l'effectif communiqué pour la statistique, il existe des caisses-maladie pour lesquelles c'est le contraire qui se produit (d'où des différences négatives).

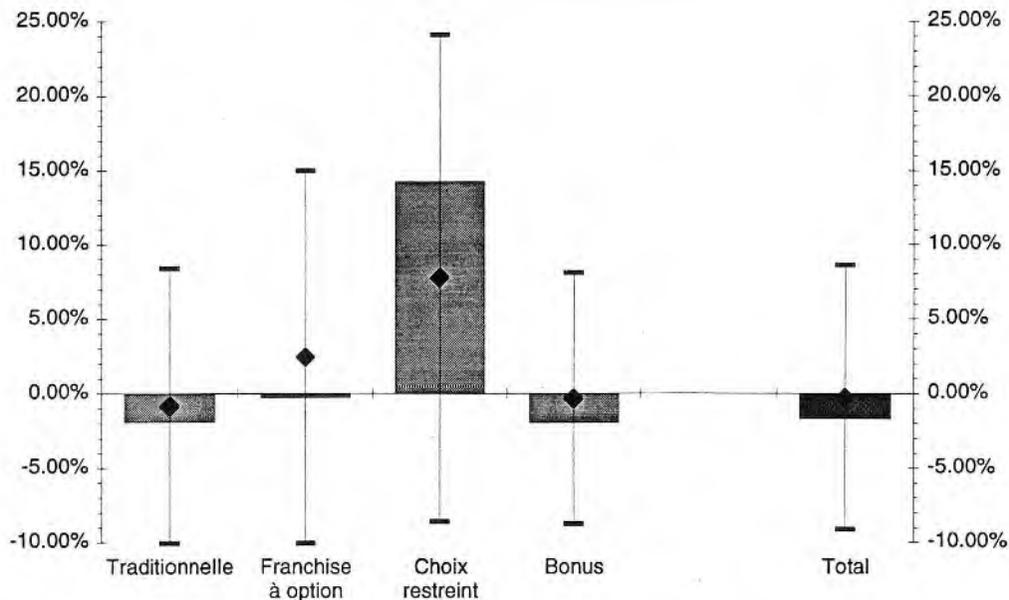


Figure 6.6 : Différences entre les données de l'échantillon et celles de la statistique de l'assurance-maladie — Primes par branches pour 1994.

■ Différence      ◆ Moyenne des différences

types) alors que d'autres présentent des différences relatives variant de 10 à 50 % selon la branche. Ceci se traduit par une variabilité importante des résultats des caisses autour de la moyenne.

Pour les divers types d'assurance de base, à l'exception du mode choix restreint, le total des primes obtenu à l'aide des données administratives est inférieur à celui que l'on obtient à l'aide des données de la statistique (Figure 6.6). Ce fait nous semble surprenant car, comme le montre la Figure 6.4 et comme nous l'avons relevé précédemment, les effectifs d'assurés sont plus élevés dans les données administratives que dans la statistique

### 6.3.3 Prestations à la charge de l'assurance pour soins médico-pharmaceutiques

Ces prestations se subdivisent en deux catégories : les prestations pour soins ambulatoires et celles pour soins hospitaliers. Pour les données administratives, nous avons déterminé les montants de ces prestations en nous basant uniquement sur la répartition des coûts des factures cohérentes<sup>1</sup>. Dans le cas présent, et contrairement à ce que nous avons fait lors de l'analyse élémentaire des données, les prestations dans le cadre des compléments statutaires à l'assurance de base ne font pas partie des

<sup>1</sup> Selon la discussion présentée dans la section 5.3, il s'agit ici des enregistrements "conservés".

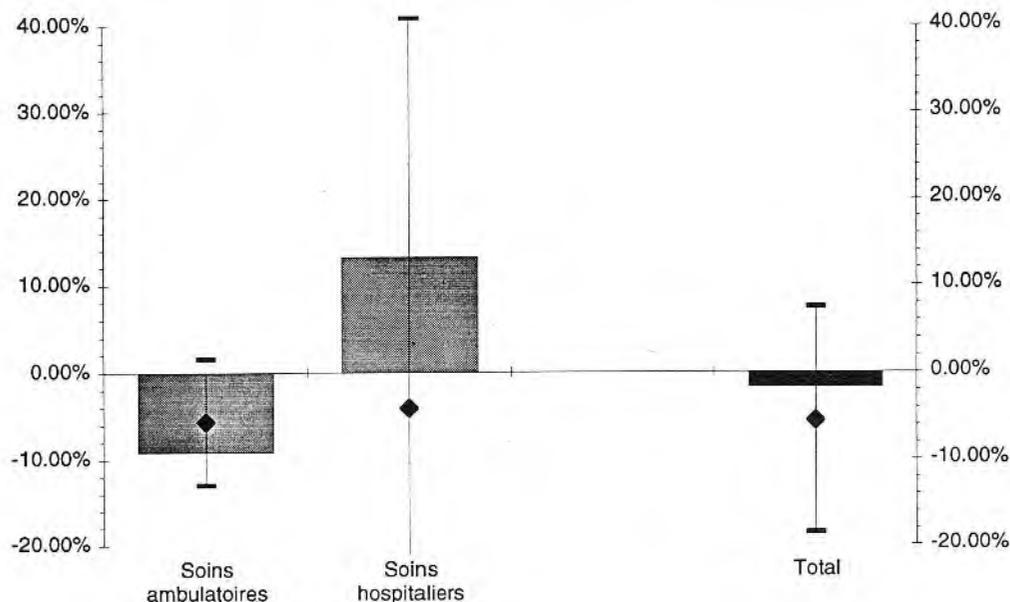


Figure 6.7 : Différences entre les données de l'échantillon et celles de la statistique de l'assurance-maladie — Prestations par type de soins pour 1994.

■ Différence      ◆ Moyenne des différences

prestations à la charge de l'assurance de base. Ceci assure une cohérence avec les données extraites de la statistique de l'assurance-maladie.

La différence relative des coûts par type de soins, ainsi que la moyenne et la variabilité des différences par caisses sont déterminées de manière analogue à la procédure adoptée pour les effectifs d'assurés. L'histogramme présenté dans la Figure 6.7 révèle une relativement bonne cohérence entre les données administratives et la statistique, du moins du point de vue de la globalité de l'échantillon. La différence globale ("Total") négative peut s'expliquer par le fait que nous avons été contraints d'ignorer un certain nombre d'enregistrements de type prestation présentant des incohérences. De cet histogramme ressort également clairement que la répartition des coûts entre soins ambulatoire et soins hospitaliers est sensiblement différente selon que l'on considère les données administratives ou la statistique. Ceci demeure pour l'instant difficilement explicable. Pour sa part, la variabilité des résultats des caisses autour de la moyenne demeure importante, voire même excessive dans le cas des coûts des soins hospitaliers.

Un examen plus détaillé des résultats caisse par caisse montre que, pour 3 des 14 caisses de l'échantillon prises en compte ici, les coûts issus des données administratives sont notoirement

inférieurs à ceux annoncés par la caisse pour la statistique<sup>1</sup>. Il est donc fort probable que les données de ces trois caisses ne soient pas correctes (soit les données administratives, soit celles pour la statistique). Si l'on considère l'échantillon sans les trois caisses incriminées, nous avons alors de meilleurs résultats, présentés dans la Figure 6.8. Pour ce sous-ensemble de l'échantillon, nous n'observons que de faibles différences au niveau global et la variabilité des résultats des caisses autour des moyennes est réduite, sauf pour le cas des soins hospitaliers.

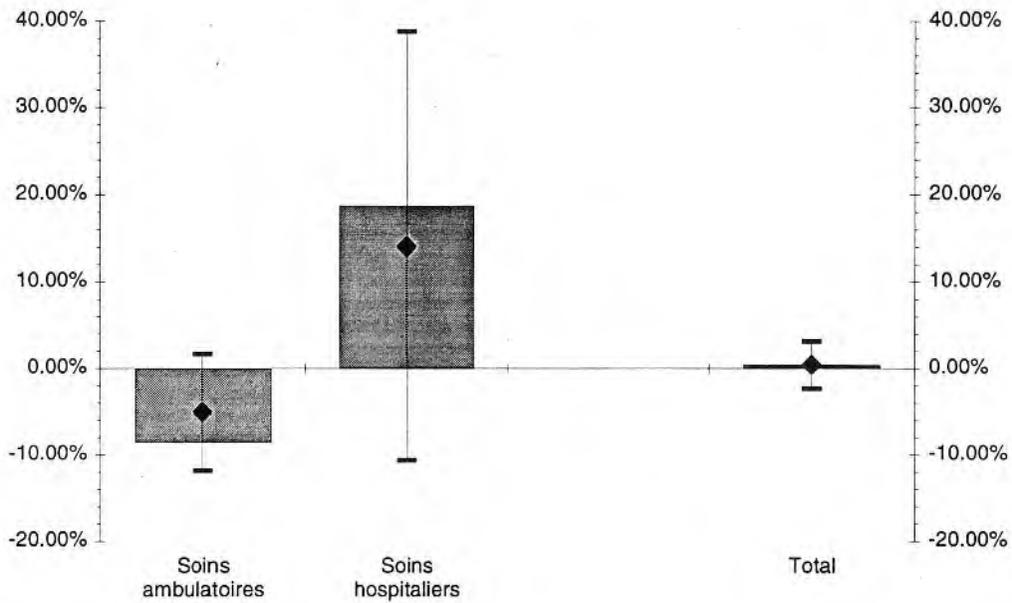


Figure 6.8 : Différences entre les données de l'échantillon et celles de la statistique de l'assurance-maladie — Prestations par type de soins pour 1994 et pour un sous-ensemble de l'échantillon constitué de 11 caisses.

■ Différence      ◆ Moyenne des différences

<sup>1</sup> La différence varie, pour ces trois caisses, d'environ 20 à plus de 40 %.

## Aspects de la sécurité sociale

Sous ce titre, l'Office fédéral des assurances sociales publie des rapports de recherche (signalés en gras) ainsi que d'autres contributions relevant de son champ d'activité. Ont déjà été publiés:

	Source N° de commande
<b>Rapport de recherche:</b> Wolfram Fischer, Possibilités de mesure des prestations hospitalières: considérations sur une réorganisation de la statistique hospitalière. N° 1/94	OCFIM* 318.010.1/94 f
<b>Rapport de recherche:</b> André Bender, M. Philippe Favarger, Dr. Martin Hoesli: Evaluation des biens immobiliers dans les institutions de prévoyance. N° 2/94	OCFIM* 318.010.2/94 f
<b>Forschungsbericht:</b> Hannes Wüest, Martin Hofer, Markus Schweizer: Wohneigentumsförderung – Bericht über die Auswirkungen der Wohneigentumsförderung mit den Mitteln der beruflichen Vorsorge. Nr. 3/94	EDMZ* 318.010.3/94 d
<b>Forschungsbericht:</b> Richard Cranovsky: Machbarkeitsstudie des Technologiebewertungsregister. Nr. 4/94	EDMZ* 318.010.4/94 d
<b>Rapport de recherche:</b> BRAINS: Inventaire du Spitex. N° 5/94	OCFIM* 318.010.5/94 f
<b>Forschungsbericht:</b> Jacob van Dam, Hans Schmid: Insolvenzversicherung in der beruflichen Vorsorge. Nr. 1/95	EDMZ* 318.010.1/95 d
<b>Forschungsbericht:</b> BASS: Tobias Bauer. Literaturrecherche: Modelle zu einem garantierten Mindesteinkommen. Nr. 2/95	EDMZ* 318.010.2/95 d
<b>Rapport de recherche:</b> IPSO: Peter Farago. Prévenir et combattre la pauvreté: forces et limites des mesures prises par l'Etat. N° 3/95	OCFIM* 318.010.3/95 f
Rapport du Département fédéral de l'intérieur concernant la structure actuelle et le développement futur de la conception helvétique des trois piliers de la prévoyance vieillesse, survivants et invalidité. Octobre 1995	OCFIM* 318.012.1/95 f/i
Universität Zürich, Interdisziplinäre Vorlesungsreihe 1995/96: Das neue KVG – Was ändert sich im Gesundheitswesen? Die Referate. Teil I	BSV** 96.217
Universität Zürich, Interdisziplinäre Vorlesungsreihe 1995/96: Das neue KVG – Was ändert sich im Gesundheitswesen? Die Referate. Teil II	BSV** 96.538
Groupe de travail interdépartemental "Perspectives de financement des assurances sociales" (IDA FiSo): Rapport sur les perspectives de financement des assurances sociales (eu égard en particulier à l'évolution démographique).	OCFIM* 318.012.1/96 f
<b>Rapport de recherche:</b> Laura Cardia-Vonèche et al.: Les familles monoparentales. N° 1/96	OCFIM* 318.010.1/96 f
Rapport du groupe de travail "Protection des données et liste des analyses / assurance-maladie". N° 2/96	OFAS** 96.568

\* EDMZ = Eidgenössische Drucksachen- und Materialzentrale, 3000 Bern

\* OCFIM = Office fédéral des imprimés et du matériel, 3000 Berne

\*\* BSV = Bundesamt für Sozialversicherung, Informationsdienst, 3003 Bern

\*\* OFAS = Office fédéral des assurances sociales, Service d'information, 3003 Berne

## Aspects de la sécurité sociale

Sous ce titre, l'Office fédéral des assurances sociales publie des rapports de recherche (signalés en gras) ainsi que d'autres contributions relevant de son champ d'activité. Ont déjà été publiés:

	Source N° de commande
Prévoyance professionnelle: Nouvelles prescriptions en matière d'établissement des comptes et de placements. Réglementation concernant l'utilisation des instruments financiers dérivés. Texte de l'ordonnance / commentaire / recommandations. N° 3/96	OCFIM* 318.010.3/96 f
<b>Forschungsbericht:</b> Martin Wechsler, Martin Savioz: Umverteilung zwischen den Generationen in der Sozialversicherung und im Gesundheitswesen. Nr. 4/96	EDMZ* 318.010.4/96 d
<b>Forschungsbericht:</b> Wolfram Fischer: Patientenklassifikationssysteme zur Bildung von Behandlungsfallgruppen im stationären Bereich. Nr. 1/97	EDMZ* 318.010.1/97 d
<b>Forschungsbericht:</b> Infrass: Festsetzung der Renten beim Altersrücktritt und ihre Anpassung an die wirtschaftliche Entwicklung. Überblick über die Regelungen in der EU. Nr. 2/97	EDMZ* 318.010.2/97 d
<b>Rapport de recherche:</b> Heinz Schmid: Procédure d'approbation des primes dans l'assurance-maladie. Expertise. N° 3/97	OCFIM* 318.010.3/97 f
<b>Forschungsbericht:</b> Eine Zusammenarbeit zwischen IPSO und Infrass: Perspektive der Erwerbs- und Lohnquote. Nr. 4/97	EDMZ* 318.010.4/97 d
<b>Forschungsbericht:</b> Stefan Spycher, BASS: Auswirkungen von Regelungen des AHV-Rentenalters auf die Sozialversicherungen, den Staatshaushalt und die Wirtschaft. Nr. 5/97	EDMZ* 318.010.5/97 d
<b>Forschungsbericht:</b> Günther Latzel, Christoph Andermatt, Rudolf Walther, BRAINS: Sicherung und Finanzierung von Pflege- und Betreuungsleistungen bei Pflegebedürftigkeit. Band I und II. Nr. 6/97	EDMZ* 318.010.6/97 d
Groupe de travail interdépartemental "Perspectives de financement des assurances sociales (IDA FiSo) 2": Analyse des prestations des assurances sociales; Concrétisation de modifications possibles en fonction de trois scénarios financiers.	OCFIM* 318.012.1/97 f

\* EDMZ = Eidgenössische Drucksachen- und Materialzentrale, 3000 Bern

\* OCFIM = Office fédéral des imprimés et du matériel, 3000 Berne

\*\* BSV = Bundesamt für Sozialversicherung, Informationsdienst, 3003 Bern

\*\* OFAS = Office fédéral des assurances sociales, Service d'information, 3003 Berne

## Aspects de la sécurité sociale

Sous ce titre, l'Office fédéral des assurances sociales publie des rapports de recherche (signalés en gras) ainsi que d'autres contributions relevant de son champ d'activité. Ont déjà été publiés:

	Source N° de commande
<b>Publications relatives à l'étude des nouvelles formes d'assurance-maladie</b>	
<b>Synthèse</b>	
<b>Rapport de recherche:</b> Rita Baur, Wolfgang Hunger, Klaus Kämpf, Johannes Stock (Prognos AG): Rapport de synthèse: Evaluation des nouveaux modèles d'assurance-maladie. N° 1/98	OCFIM* 318.010.1/98 f
<b>Dossiers techniques / Enquêtes</b>	
<b>Forschungsbericht:</b> Rita Baur, Doris Eyett (Prognos AG): Die Wahl der Versicherungsformen. Untersuchungsbericht 1. Nr. 2/98	EDMZ* 318.010.2/98 d
<b>Forschungsbericht:</b> Rita Baur, Doris Eyett (Prognos AG): Bewertung der ambulanten medizinischen Versorgung durch HMO-Versicherte und traditionell Versicherte. Untersuchungsbericht 2. Nr. 3/98	EDMZ* 318.010.3/98 d
<b>Forschungsbericht:</b> Rita Baur, Doris Eyett (Prognos AG): Selbstgetragene Gesundheitskosten. Untersuchungsbericht 3. Nr. 4/98	EDMZ* 318.010.4/98 d
<b>Forschungsbericht:</b> Rita Baur, Armin Ming, Johannes Stock, Peter Lang (Prognos AG): Struktur, Verfahren und Kosten der HMO-Praxen. Untersuchungsbericht 4. Nr. 5/98	EDMZ* 318.010.5/98 d
<b>Forschungsbericht:</b> Johannes Stock, Rita Baur, Peter Lang (Prognos AG); Prof. Dr. Dieter Conen: Hypertonie-Management. Ein Praxisvergleich zwischen traditionellen Praxen und HMOs. Nr. 6/98	EDMZ* 318.010.6/98 d
<b>Dossiers techniques / Etude des données administratives</b>	
<b>Forschungsbericht:</b> Stefan Schütz et al.: Neue Formen der Krankenversicherung: Versicherte, Leistungen, Prämien und Kosten. Ergebnisse der Administrativdatenuntersuchung, 1. Teil. Nr. 7/98	EDMZ* 318.010.7/98 d
<b>Forschungsbericht:</b> Herbert Känzig et al.: Neue Formen der Krankenversicherung: Alters- und Kostenverteilungen im Vergleich zu der traditionellen Versicherung. Ergebnisse der Administrativdatenuntersuchung, 2. Teil. Nr. 8/98	EDMZ* 318.010.8/98 d
<b>Rapport de recherche:</b> Gabriel Sottas et al.: Données administratives de l'assurance-maladie: Analyse de qualité, statistique élémentaire et base pour les exploitations. N° 9/98	OCFIM* 318.010.9/98 f
<b>Les questionnaires des enquêtes auprès des assurés (5 parties) peuvent être obtenus à l'adresse suivante:</b> <b>Office fédéral des assurances sociales, section Statistique,</b> <b>M. Herbert Känzig, 3003 Berne (Tél. 031 / 322 91 48)</b>	

\* EDMZ = Eidgenössische Drucksachen- und Materialzentrale, 3000 Bern

\* OCFIM = Office fédéral des imprimés et du matériel, 3000 Berne

\*\* BSV = Bundesamt für Sozialversicherung, Informationsdienst, 3003 Bern

\*\* OFAS = Office fédéral des assurances sociales, Service d'information, 3003 Berne