

Université Victor Segalen Bordeaux 2

Année 2009

Thèse n°1676

THÈSE

pour le

DOCTORAT DE L'UNIVERSITÉ BORDEAUX 2

Mention : Sciences, Technologie, Santé

Option : Épidémiologie et Santé Publique

Présentée et soutenue publiquement

Le 15 décembre 2009

Par Pernelle NOIZE

Née le 20 novembre 1979 à Dijon

**Mesure de l'exposition médicamenteuse en pharmaco-épidémiologie :
étude comparative de données issues des bases de remboursement
de l'Assurance Maladie française et de données déclaratives**

Membres du Jury

Mr le Pr Bernard BÉGAUD	Président
Mme le Dr Joëlle MICALLEF-ROLL	Rapporteur
Mr le Pr Jean-Louis MONTASTRUC	Rapporteur
Mr le Dr Alain WEILL.....	Membre du jury
Mme le Dr Annie FOURRIER-RÉGLAT	Membre du jury
Mr le Pr Nicholas MOORE	Directeur de thèse

À la mémoire de Pépé Georges, Mémé Suz et Mémé Marcelle

REMERCIEMENTS

À Mr le Professeur Bernard Bégaud pour m'avoir accueillie comme doctorante au sein de l'unité Inserm U657 et avoir accepté la présidence de cette thèse. Veuillez trouver ici l'expression de mon plus grand respect.

À Mme le Docteur Joëlle Micallef-Roll envers qui je suis extrêmement reconnaissante d'avoir accepté de juger ce travail. Avec tout mon respect et ma gratitude.

À Mr le Professeur Jean-Louis Montastruc pour avoir également accepté de juger ce travail. Qu'il soit l'occasion de vous témoigner à nouveau mon plus grand respect.

À Mr le Docteur Alain Weill envers qui je suis extrêmement reconnaissante d'avoir accepté de participer à l'évaluation de ce travail. Avec tout mon respect et ma gratitude.

À Mme le Docteur Annie Fourier-Réglat pour me faire l'honneur d'être présente dans le jury de cette thèse comme elle l'a toujours été tout au long de mon parcours de doctorante. Avec mon plus profond respect et mon extrême reconnaissance.

À Mr le Professeur Nicholas Moore pour m'avoir confié et avoir dirigé ce travail, mais aussi et surtout, pour m'avoir ouvert les portes de la pharmaco-épidémiologie dès le début de mon internat et m'avoir depuis accordé sa confiance. Veuillez trouver ici l'expression de mon plus profond respect et de mon extrême reconnaissance.

À Fabienne Diaz pour son aide précieuse et son soutien de chaque instant et Antoine Pariente pour l'intérêt qu'il a porté à ce travail et ses conseils avisés. Avec toute mon amitié.

À toutes les personnes du département de pharmacologie, de l'unité Inserm U657 ou d'ailleurs avec lesquelles j'ai eu grand plaisir à travailler durant ces années de doctorat et que je ne peux pas toutes citer ici...

À mes amis qui vont enfin pouvoir arrêter de me demander si j'ai fini mes études !

À Pépé Paul (j'espère n'avoir oublié aucun accent sur les lettres majuscules !), Maman, Papa, Fred, Nico et toute ma famille. Quel bonheur immense d'être née parmi vous...

À Blanche, quel bonheur immense de t'avoir vu naître et quelle fierté d'être ta marraine !

Enfin, à Yannick pour les collines du pays basque, les paris à 100 euros, le temple d'Abou Simbel, les dîners au Café du Port, les parties de Time's up, les pique-niques à Central Park, les journées de ligue 1 (et du top 14 !) et tout ce que nous avons encore à partager. Avec tout mon amour.

RÉSUMÉ

En pharmaco-épidémiologie, la mesure de l'exposition médicamenteuse est fondamentale. Des données déclaratives recueillies par interrogatoire des sujets ou des données extraites des bases de remboursement de l'Assurance Maladie peuvent être utilisées. Les objectifs de ce travail étaient de comparer l'exposition mesurée à partir de ces deux sources de données et d'évaluer l'impact du choix de l'une ou l'autre source sur les mesures de risque dans les études étiologiques. Les travaux ont été conduits au sein de l'étude des Trois-Cités, cohorte de sujets âgés pour laquelle étaient disponibles des données déclaratives et de remboursement. La concordance entre les expositions issues des deux sources ou la validité de l'exposition issue de l'une par rapport à celle issue de l'autre ont été évaluées dans différents contextes. La survenue d'événements a été simulée dans la population et des études cas-témoins nichées ont été conduites pour évaluer l'association entre l'exposition issue de chaque source et chaque événement. Pour les médicaments cardiovasculaires, les mesures d'exposition étaient peu différentes entre les deux sources. L'impact du choix de l'une ou l'autre source sur les mesures d'association entre l'exposition et un événement était faible. Pour les benzodiazépines ou les anti-inflammatoires non stéroïdiens, des différences étaient observées entre les expositions issues des deux sources. Pour ces médicaments, l'association avec un événement pouvait varier selon la source utilisée. Pour des médicaments pris de manière irrégulière ou intermittente, le choix de la source de données pour la mesure de l'exposition peut donc être un élément déterminant.

MOTS-CLÉS

Pharmaco-épidémiologie

Exposition

Médicaments

Interrogatoire

Bases de données de remboursement

INTITULÉ ET ADRESSE DU LABORATOIRE

Inserm U657 « Pharmaco-épidémiologie et évaluation de l'impact des produits de santé sur les populations » - Département de pharmacologie

Université Victor Segalen Bordeaux 2

Bât 1A, Zone Nord, Site Carreire, Case 36

146, rue Léo Saignat 33076 Bordeaux Cedex

TITLE AND ABSTRACT

Assessment of drug exposure in pharmacoepidemiology : comparison of interview data and reimbursement claims data from the French national healthcare insurance system

In pharmacoepidemiology, assessment of drug exposure is fundamental. It can rely on data collected through patient interviews or extracted from healthcare insurance system databases recording reimbursement claims. This work aimed to compare drug exposure measured from these two data sources and to evaluate the impact of choosing one source or the other on risk estimates in etiological studies. It was conducted as part of the Three-City Study, a cohort of French elderly persons for which both interview and reimbursement data were available. Agreement between exposures measured from both sources or validity of exposure measured from one source with reference to that measured from the other were evaluated in different backgrounds. Simulated outcomes were generated in the study population and nested case-control studies were conducted in order to estimate the association between the drug exposure measured from each source and each simulated outcome. For cardiovascular system drugs, exposure measured from interview data was close to that measured from reimbursement data. The choice of one source or the other had few impact on the estimated associations between the exposure and an outcome. For benzodiazepines or non-steroidal anti-inflammatory drugs, exposure measured from both sources could differ. For these drugs, the association between the exposure measured from each source and an outcome could vary. For drugs that can be used irregularly or intermittently, the choice of the source of data for drug exposure assessment could thus be of great importance.

KEYWORDS

Pharmacoepidemiology

Exposure

Medicines

Interview

Reimbursement databases

TABLE DES MATIÈRES

I- INTRODUCTION	10
I.1- CONTEXTE GÉNÉRAL.....	11
I.1.1- De l'importance de la mesure de l'exposition médicamenteuse en pharmaco-épidémiologie.....	11
I.1.2- Principales sources de données disponibles pour la mesure de l'exposition médicamenteuse en pharmaco-épidémiologie	12
I.1.2.1- Le patient comme source de données déclaratives.....	13
I.1.2.2- Les sources de données pré-enregistrées.....	15
I.1.2.3- Mesure de l'exposition médicamenteuse en pharmaco-épidémiologie : une revue de la littérature récente.....	18
I.1.3- État des lieux de la situation française	24
I.2- OBJECTIFS	26
II- DONNÉES DISPONIBLES ET POPULATION SOURCE POUR LES TRAVAUX.....	27
II.1- PRÉSENTATION GÉNÉRALE DE LA COHORTE DES TROIS-CITÉS (3C)	28
II.2- DESCRIPTION DES DONNÉES DÉCLARATIVES ET DE REMBOURSEMENT DISPONIBLES.....	29
II.3- DÉFINITION ET DESCRIPTION DE LA POPULATION SOURCE	31
III- PREMIÈRE PARTIE : DESCRIPTION ET COMPARAISON DE L'EXPOSITION MÉDICAMENTEUSE MESURÉE À PARTIR DE DONNÉES DÉCLARATIVES OU DE DONNÉES DE REMBOURSEMENT	34
III.1- SYNTHÈSE DE LA LITTÉRATURE.....	35
III.2- DESCRIPTION ET COMPARAISON DE L'EXPOSITION MÉDICAMENTEUSE MESURÉE À PARTIR DE DONNÉES DÉCLARATIVES OU DE REMBOURSEMENT UN JOUR DONNÉ (ARTICLE N°1).....	40
III.2.1- Présentation de l'article	40
III.2.1.1- Contexte et objectifs.....	40
III.2.1.2- Considérations méthodologiques.....	41
III.2.1.3- Principaux résultats.....	43
III.2.2- Article n°1 publié	43
III.3- DESCRIPTION ET COMPARAISON DE L'EXPOSITION MÉDICAMENTEUSE MESURÉE À PARTIR DE DONNÉES DÉCLARATIVES OU DE REMBOURSEMENT AU COURS D'UNE PÉRIODE DONNÉE (ARTICLES N°2 ET N°3).....	55
III.3.1- Présentation des articles	55
III.3.1.1- Contexte et objectifs.....	55
III.3.1.2- Considérations méthodologiques.....	56
III.3.1.3- Principaux résultats.....	59
III.3.2- Article n°2.....	59
III.3.3- Article n°3.....	81

IV- DEUXIÈME PARTIE : IMPACT DE LA SOURCE DE DONNÉES UTILISÉE POUR DÉTERMINER L'EXPOSITION MÉDICAMENTEUSE SUR LA MESURE D'UNE ASSOCIATION ENTRE UN MÉDICAMENT ET UN ÉVÉNEMENT.....	108
IV.1- CONTEXTE ET OBJECTIF.....	109
IV.2- MÉTHODES.....	111
IV.2.1- Échantillon d'étude	111
IV.2.2- Choix des classes médicamenteuses étudiées	111
IV.2.3- Définitions de l'exposition médicamenteuse.....	112
IV.2.3.1- Définition de l'exposition médicamenteuse selon les données déclaratives.....	112
IV.2.3.2- Définitions de l'exposition médicamenteuse selon les données de remboursement	112
IV.2.4- Simulation de la survenue d'un événement dans l'échantillon d'étude	115
IV.2.5- Type d'étude utilisé pour évaluer l'impact de la source de données : étude cas-témoins nichée	116
IV.2.6- Analyse complémentaire	117
IV.3- RÉSULTATS.....	119
IV.4- DISCUSSION.....	125
IV.4.1- Interprétation des résultats	125
IV.4.2- Données de la littérature.....	127
V- CONCLUSION GÉNÉRALE ET PERSPECTIVES	129
VI- RÉFÉRENCES BIBLIOGRAPHIQUES	133
VII- ANNEXES	148

INDEX DES FIGURES ET TABLEAUX

Figure 1- Représentation schématique des données déclaratives et de remboursement disponibles dans la cohorte des Trois-Cités (1999 – 2005).....	30
Figure 2. Illustration de la mesure transversale de l'exposition médicamenteuse d'après les données déclaratives et de remboursement disponibles pour la première visite de suivi de la cohorte 3C	41
Figure 3. Illustration de la mesure de l'exposition médicamenteuse entre les deux visites de suivi de la cohorte 3C d'après les données déclaratives et de remboursement.....	56
Figure 4. Illustration de la mesure de l'exposition médicamenteuse pour les études étiologiques dans une cohorte à suivis transversaux répétés : la cohorte 3C.....	110
Figure 5. Illustration des définitions de l'exposition médicamenteuse selon chaque source de données dans le cas d'un événement lié à une exposition « en cours » (transversale).....	114
Figure 6. Illustration des définitions de l'exposition médicamenteuse selon chaque source de données dans le cas d'un événement lié à une exposition chronique régulière au cours d'une période (« longitudinale »).....	115
Figure 7- Influence de la source de données utilisée pour déterminer l'exposition médicamenteuse sur la mesure d'une association entre un médicament et un événement : le cas des benzodiazépines	121
Figure 8- Influence de la source de données utilisée pour déterminer l'exposition médicamenteuse sur la mesure d'une association entre un médicament et un événement : le cas des anti-inflammatoires non stéroïdiens.....	122
Figure 9- Influence de la source de données utilisée pour déterminer l'exposition médicamenteuse sur la mesure d'une association entre un médicament et un événement : le cas des médicaments cardiologiques	123
Figure 10- Représentation graphique de la variation relative entre les rapports de cotes mesurant respectivement l'association entre l'exposition selon les données déclaratives ou selon les données de remboursement et l'événement simulé en fonction de la proportion de sujets discordants entre les deux sources de données vis-à-vis de l'exposition médicamenteuse	124
Tableau 1. Descriptif des études étiologiques publiées en 2008 dans la revue <i>Pharmacoepidemiology and Drug Safety</i> (n = 39)	21
Tableau 2. Comparaison des sujets de la population source et des autres sujets inclus dans la cohorte 3C (n = 9 294).....	32
Tableau 3. Descriptif des études publiées de comparaison de l'exposition médicamenteuse déclarée par les sujets lors d'interrogatoires ou issue de bases de données informatisées (n = 24).....	37
Tableau 4. Concordance entre les données déclaratives recueillies à l'interrogatoire de suivi S1 et les données de remboursement dans les n (15 à 90) jours précédant ce même interrogatoire déterminée à l'aide du coefficient κ (n = 2985)	155

LISTE DES ABRÉVIATIONS

3C	Trois-Cités
AINS	Anti-Inflammatoire Non Stéroïdien
AMM	Autorisation de Mise sur le Marché
ATC	Anatomical Therapeutic Chemical
CADEUS	COX-2 et AINS classiques : DEscription des UtiliSateurs
CépiDC	Centre d'épidémiologie des causes de décès
CES-D	Center for Epidemiologic Studies - Depression Scale
CESIR-U	Combinaison d'Études sur la Santé et l'Insécurité Routière – Urgences
CHU	Centre Hospitalier Universitaire
CIP	Club Inter-Pharmaceutique
CNAM-TS	Caisse Nationale d'Assurance Maladie des Travailleurs Salariés
DDJ	Dose Définie Journalière
E3N	Étude Épidémiologique auprès de femmes de l'Éducation Nationale
ERASME	Extraction Recherche Analyse pour le Suivi Médico-Économique
EU-ADR	Exploring and Understanding Adverse Drug Reactions
GHC	Group Health Cooperative
GPRD	General Practice Research Database
HMO	Health Maintenance Organization
IC 95 %	Intervalle de confiance à 95 %
ISPE	International Society for Pharmacoepidemiology
MGEN	Mutuelle Générale de l'Éducation Nationale
MMSE	Mini Mental State Examination
MPR	Medication Possession Ratio
OTC	Over The Counter
Paquid	Personnes âgées quid ?
PHRC	Programme Hospitalier de Recherche Clinique
PLASTICO	PLAte-forme Scientifique et TechnIque pour l'aide à la gestion de COhortes et de grandes enquêtes épidémiologiques
PMSI	Programme de Médicalisation du Système d'Information
RAMQ	Régie d'Assurance Maladie du Québec
RC	Rapport de cotes
RR	Risque relatif
THIN	The Health Improvement Network

I- INTRODUCTION

I.1- CONTEXTE GÉNÉRAL

I.1.1- De l'importance de la mesure de l'exposition médicamenteuse en pharmaco-épidémiologie

Au cours des dernières décennies, l'évaluation du médicament a subi de profondes mutations pour s'inscrire de plus en plus dans une perspective de santé publique. En vue de l'obtention de l'Autorisation de Mise sur le Marché (AMM), des études expérimentales ont principalement pour objectifs de mettre en évidence l'efficacité du médicament et de fournir une première évaluation de sa sécurité d'emploi. Ces études, souvent conduites sur de courtes périodes, chez des sujets particulièrement sélectionnés et suivis dans des conditions strictes, posent des problèmes de transposabilité des résultats, aussi bien d'efficacité que de sécurité, à la population future des utilisateurs.^{1,2}

Face à ces limites, des systèmes de surveillance du médicament après commercialisation se sont progressivement organisés autour de plusieurs disciplines et, notamment, la pharmaco-épidémiologie. Cette discipline met en application les méthodes et/ou le raisonnement épidémiologique(s) pour évaluer, généralement sur de grandes populations, l'efficacité, le risque et l'usage des médicaments.³ Contrairement aux études expérimentales, les études pharmaco-épidémiologiques adoptent généralement une approche observationnelle en ce sens qu'elles tentent de modifier le moins possible le cours naturel des choses (par exemple, la relation entre un médecin et son patient déterminant le choix d'un médicament et les modalités de suivi du traitement). Si l'intérêt majeur des études observationnelles est de décrire le plus fidèlement possible la réalité du terrain, un de leurs principaux inconvénients est qu'elles sont davantage sujettes à différents types de biais. Les biais classiquement décrits en pharmaco-épidémiologie peuvent intervenir au niveau du choix des sujets recrutés pour l'étude (biais de sélection), du recueil des données nécessaires (biais d'information) ou encore de l'interprétation des résultats.

Vis-à-vis de la minimisation des biais d'information, la mesure de l'exposition médicamenteuse requiert une attention particulière au cours des études observationnelles en pharmaco-épidémiologie.⁴ Le recueil de données de qualité discutable concernant le(s) médicament(s) d'intérêt peut remettre en cause les résultats d'études décrivant leur utilisation en vie réelle. Néanmoins, l'impact des biais d'information peut être d'autant plus important sur les conclusions d'études étiologiques visant à déterminer le rôle, éventuellement causal, de l'exposition à un

médicament dans la survenue d'un événement. Dans cette situation, l'existence de biais d'information peut aboutir à une mauvaise classification des sujets en tant qu'exposés ou non au médicament étudié. On parle alors plus spécifiquement de biais de classification, ces derniers pouvant être plus ou moins gênants selon qu'ils sont ou non différentiels c'est-à-dire qu'ils affectent ou non dans les mêmes proportions les groupes comparés (sujets exposés et non exposés dans une étude de cohorte ou cas et témoins dans une étude cas-témoins).⁵

Il paraît peu envisageable, dans les conditions observationnelles, d'éviter totalement l'introduction de biais quant à la mesure de l'exposition médicamenteuse. Toutefois, en fonction de la source de données utilisée pour l'estimer, tous les moyens possibles doivent être mis en œuvre pour qu'elle reflète au mieux l'exposition réelle des sujets.

I.1.2- Principales sources de données disponibles pour la mesure de l'exposition médicamenteuse en pharmaco-épidémiologie

Une classification possible des principales sources de données utilisables en pharmaco-épidémiologie consiste à distinguer les sources de données déclaratives, des sources de données dites pré-enregistrées.⁶ Les sources de données déclaratives sont des personnes physiques interrogées spécifiquement pour une étude : patients, prescripteurs, pharmaciens ou tout autre acteur du système de santé. Elles sont nécessairement rétrospectives, en ce qu'elles interrogent sur un comportement passé et sont donc sujettes à des biais de mémorisation. Les sources de données pré-enregistrées sont déjà constituées au moment de l'étude. Elles se présentent généralement sous la forme de bases de données ou fichiers électroniques renfermant des informations sur les médicaments prescrits, dispensés ou remboursés à une certaine population de sujets. Ces données sont recensées de façon prospective, au moment de l'acte (prescription, délivrance, remboursement) et ne sont donc pas sujettes aux biais de mémorisation. Dans le domaine spécifique du médicament, elles n'informent que sur une intention d'exposition et pas forcément sur la réalité de celle-ci. Le choix d'une ou plusieurs source(s) de données est un élément important lors de la conception d'une étude observationnelle en pharmaco-épidémiologie et il doit tenir compte des avantages et limites respectifs de chaque type de source de données.

I.1.2.1- Le patient comme source de données déclaratives

Parmi les sources potentielles de données déclaratives, le patient est plus particulièrement intéressant puisque son interrogatoire constitue *a priori* la seule et unique façon de mesurer de manière exhaustive son exposition médicamenteuse réelle.⁷ En effet, les autres sources de données déclaratives permettent seulement d'estimer cette exposition au travers des prescriptions (médecins) ou des dispensations (pharmaciens) de médicaments. Ainsi, le recueil de données déclaratives auprès des patients a fait l'objet de nombreuses études ayant montré que la qualité de ces données dépend de nombreux facteurs liés aux modalités de l'interrogatoire, aux patients interrogés ou encore à l'exposition médicamenteuse d'intérêt.

I.1.2.1.1- Facteurs liés aux modalités de l'interrogatoire

Un premier point à noter concernant les modalités de l'interrogatoire est l'existence possible d'un biais lié à l'enquêteur dans le cas où le patient est interrogé par une tierce personne. Il s'agit d'un biais d'information dans lequel un enquêteur interroge de manière systématiquement différente les sujets selon qu'ils appartiennent à l'un ou l'autre des groupes comparés³ (notamment, les cas et les témoins dans une étude cas-témoins). Il est donc important de dispenser une formation commune spécifique à l'ensemble des personnes susceptibles d'interroger les patients dans une étude et d'utiliser des outils de recueil standardisés. Les enquêteurs doivent aussi rester autant que possible en aveugle par rapport aux hypothèses testées ou au statut de cas ou de témoin des patients interrogés.^{8,9}

Que le sujet soit interrogé par un enquêteur ou qu'il remplisse un auto-questionnaire, la nature même du questionnaire est un facteur déterminant pour la qualité des données recueillies.¹⁰ En particulier, le type de question posée, selon qu'il s'agit de questions ouvertes ou orientées, constitue un élément essentiel.^{11,12} Parmi les travaux spécifiquement menés autour de cette problématique^{7,13-15}, citons l'étude de Moore *et al.*, dans laquelle l'effet d'une seule question ouverte (« Avez-vous pris des médicaments pendant la semaine dernière. Si oui, lesquels, quand et à quelle dose ? ») a été comparé à celui de questions orientées respectivement sur l'indication du médicament, sa forme galénique, un éventuel changement des modalités de traitement et le moment d'administration du médicament au cours de la journée.⁷ Dans cette étude, le nombre de médicaments déclarés à l'issue de la question ouverte doublait après la série des quatre questions

orientées. L'ordre dans lequel ces quatre questions étaient posées semblait également impliqué dans la qualité des réponses.

D'une manière intuitive, le niveau de précision de la question influence également beaucoup la qualité des déclarations.^{16,17} À titre d'exemple, dans leur étude relative à l'utilisation d'Anti-Inflammatoires Non Stéroïdiens (AINS), West *et al.* ont recueilli une information de qualité décroissante selon que la question se rapportait à l'utilisation d'un AINS quel qu'il soit, au nom de l'AINS utilisé, à son nom et à son dosage voire à son nom, son dosage et ses dates d'utilisation.¹⁶

I.1.2.1.2- Facteurs liés aux patients interrogés

Certaines caractéristiques sociodémographiques des patients interrogés ont déjà été citées comme pouvant affecter le niveau de déclaration. Tel est le cas de l'âge plus élevé^{17,18}, du niveau d'éducation plus faible^{11,19} ou de l'origine ethnique et du statut marital.⁹

L'état de santé du patient pourrait aussi avoir une influence que ce soit la présence de certains antécédents médicaux⁹ ou l'existence d'un certain niveau de dépendance vis-à-vis des activités de la vie courante.¹⁸

I.1.2.1.3- Facteurs liés à l'exposition médicamenteuse d'intérêt

Un dernier groupe de facteurs susceptibles d'intervenir sur la qualité du recueil de données déclaratives concerne directement la nature de l'exposition médicamenteuse à mesurer. S'agissant de l'exposition globale (toutes classes confondues), il semble que plus le patient consomme un nombre important de médicaments, plus son aptitude à tous les déclarer diminue.^{13,17}

Par ailleurs, il a été montré que plus le médicament d'intérêt a été utilisé sur une période longue, plus la qualité des déclarations s'améliore.^{9,13} À ce titre, dans l'étude conduite par Goodman *et al.* afin d'estimer la concordance entre des données recueillies auprès des femmes incluses dans une étude cas-témoins sur le cancer du sein ou auprès de leurs médecins à propos de l'exposition aux oestrogènes, les femmes ayant oublié de déclarer la prise du médicament l'avaient utilisé sur des périodes significativement plus courtes.⁹ Une consommation répétée du médicament au cours du temps est aussi associée à de meilleures déclarations.^{16,19}

Enfin, l'intervalle de temps s'écoulant entre la prise du médicament et l'interrogatoire est également déterminant : le recueil est meilleur vis-à-vis de médicaments pris au moment de l'interrogatoire ou stoppés récemment avant ce dernier que vis-à-vis de médicaments plus anciens. Ceci remet en cause la validité du patient comme source de données pour la mesure d'une exposition médicamenteuse passée.^{11,13,16}

Devant ces apparentes limites, des réflexions sur la manière d'interroger les patients, comme de compléter le questionnaire par une inspection visuelle des boîtes de médicaments à domicile¹⁸ ou au centre d'étude²⁰ (« brown bag method »), et l'utilisation de divers outils pour stimuler leur mémoire, tels que des listes et illustrations des médicaments d'intérêt^{21,22} ou des calendriers visant à rappeler des événements de la vie²³, ont permis d'améliorer la qualité des données déclaratives. Néanmoins, elles restent trop souvent sujettes à des erreurs, des omissions.^{9,12,16}

I.1.2.2- Les sources de données pré-enregistrées

Outre l'utilisation du patient, du médecin ou du pharmacien comme sources de données déclaratives, l'enregistrement électronique des consommations de soins médicaux ambulatoires et hospitaliers, en constante progression, a donné aux chercheurs en pharmaco-épidémiologie de nouvelles opportunités en leur permettant d'accéder à des bases de données informatisées de prescription, dispensation ou remboursement de médicaments.²⁴

Il convient de distinguer deux catégories de bases de données, à savoir les bases de données médicales et les bases de données administratives.²⁵ Les bases de données dites médicales sont alimentées, en pratique courante, par des médecins (le plus souvent, généralistes) acceptant d'enregistrer dans un système informatique un ensemble de données cliniques inhérentes à leurs patients : antécédents médicaux, facteurs de risque (par exemple, consommation de tabac ou d'alcool), diagnostics médicaux, médicaments prescrits. Ces bases se développent parallèlement au déploiement des dossiers médicaux informatisés dont elles reprennent les données sous forme généralement anonymisée. Les bases de données dites administratives ont été initialement constituées à des fins « médico-tarifaires » dans le cadre d'un système de soins et non pas à des fins de recherche. Elles ne renferment généralement pas de données cliniques mais peuvent parfois être croisées avec des bases de données d'hospitalisation ou des registres de pathologies. Dans certains cas, le remboursement des soins est conditionnel au diagnostic qui, de ce fait, peut être renseigné.

L'exploitation des bases de données informatisées en pharmaco-épidémiologie a débuté dans les pays d'Amérique du Nord où le potentiel de bases de données administratives a rapidement été perçu : aux États-Unis, celles de certaines Health Maintenance Organizations (HMO) comme Group Health Cooperative (GHC) of Puget Sound^{26,27}, Medicaid^{28,29}, Kaiser Permanente³⁰ et, au Canada, celles de la province de Saskatchewan³¹ ou de la Régie d'Assurance Maladie du Québec (RAMQ).³² En Europe du Nord, les Pays-Bas³³ et les pays scandinaves³⁴ sont eux aussi dotés de systèmes informatisés intervenant principalement au niveau de la dispensation des médicaments. Enfin, le Royaume-Uni se distingue par ses bases de données médicales dont la General Practice Research Database (GPRD)³⁵, connue comme la plus importante à ce jour ou, plus récemment, The Health Improvement Network (THIN).³⁶ Le développement de ces bases de données est en plein essor et la situation évolue rapidement avec, en particulier la mise en place de réseaux multibases tels le projet européen EU-ADR (Exploring and Understanding Adverse Drug Reactions by integrative mining of clinical records and biomedical knowledge).

1.1.2.2.1- Des atouts incontestables

Un des principaux intérêts des bases de données pré-enregistrées est que la qualité des données qu'elles contiennent est indépendante des caractéristiques du patient. En outre, elles ne sont pas limitées, comme peuvent l'être des données déclaratives recueillies par interrogatoire des patients, par les biais de mémorisation.¹⁶ Renfermant souvent des informations longitudinales sur plusieurs années, elles peuvent être davantage adaptées que des données déclaratives pour la mesure d'une exposition médicamenteuse passée.^{16,24,35,37} Par ailleurs, les diverses informations qui y sont enregistrées offrent certaines possibilités : la quantité d'unités de médicaments dispensée ou remboursée, le dosage unitaire et la posologie peuvent permettre de calculer des durées théoriques d'utilisation, détails qui ne sont pas forcément bien déclarés par les patients.³⁷ L'ensemble des données démographiques, médicales, pharmaceutiques d'une même base de données, comme la GPRD, peut également être utilisé afin d'établir des profils de patients vis-à-vis de la consommation d'un médicament donné.³⁵

Parmi les autres atouts fréquemment cités pour ces bases de données, on retrouve le fait qu'elles peuvent concerner de larges populations ce qui est intéressant en termes de représentativité.²⁴ S'agissant à nouveau de la GPRD, García-Rodríguez *et al.* indiquent, par exemple, qu'elle peut couvrir jusqu'à 3 millions de sujets.³⁵ Enfin, les études conduites à partir de données extraites de

ces bases pourraient bénéficier d'une meilleure faisabilité en termes de coûts et surtout de délais.^{24,27}

I.1.2.2.2- Des limites à ne pas négliger

Bien que très attractives, ces bases de données présentent un certain nombre de limites et, principalement, qu'elles ne concernent que des médicaments prescrits. Contrairement au patient, qui s'il est convenablement interrogé, peut livrer l'exhaustivité des médicaments pris, ces bases de données n'incluent aucune information sur l'automédication que ce soit avec des médicaments disponibles sans ordonnance ou Over The Counter (OTC), des médicaments obtenus d'une tierce personne ou des médicaments restants de traitements antérieurs.^{7,16,35,37} La plupart d'entre-elles ne renseignent pas non plus sur les médicaments administrés lors de séjours hospitaliers.^{32,35,37}

Même si elles permettent de l'approcher, par la présence de renouvellements réguliers de prescriptions, dispensations ou remboursements, ces bases de données ne sauraient refléter l'observance du patient pour son traitement médicamenteux. En effet, ces trois étapes ne garantissent pas que le médicament a effectivement été consommé et, encore moins, qu'il a été consommé conformément à la prescription.^{7,16} Cette limite a été, en outre, soulignée concernant les calculs de durée de traitement qui présupposent un parfait respect de la posologie prescrite indiquée dans la base de données.^{13,32,37}

L'exhaustivité des données peut être affectée par certains comportements des patients comme de ne pas exécuter une prescription, de ne pas demander de remboursement, d'acheter le médicament dans une autre pharmacie que celle où ils sont habituellement enregistrés (dans le système néerlandais, plus spécifiquement), de changer d'affiliation ou encore de déménager.^{13,37} Leur représentativité fait aussi parfois l'objet de controverse : tel est le cas des bases de données de Medicaid aux États-Unis qui couvre des sujets âgés et de milieux sociaux défavorisés.^{28,29} Enfin, initialement prévues pour des finalités financières, les bases de données administratives peuvent être d'un accès difficile et contenir des données de qualité variable (erreurs de codage...)^{24,32,35}.

Même si l'engouement récent pour l'utilisation de ces bases de données à des fins de recherche en pharmaco-épidémiologie est totalement justifié, elle doit rester prudente et tenir compte de leurs possibles limites.³⁸

Cette description des principaux avantages et inconvénients des données déclaratives ou pré-enregistrées permet d'appréhender toute la complexité du choix d'une source de données pour la mesure de l'exposition médicamenteuse en pharmaco-épidémiologie. Mais qu'en est-il de l'utilisation respective de ces différentes sources de données dans les études observationnelles récemment conduites dans cette discipline ?

I.1.2.3- Mesure de l'exposition médicamenteuse en pharmaco-épidémiologie : une revue de la littérature récente

Afin de décrire les sources de données et définitions utilisées pour mesurer l'exposition médicamenteuse dans les études pharmaco-épidémiologiques récemment publiées, une synthèse de la littérature a été effectuée dans la revue *Pharmacoepidemiology and Drug Safety*, revue officielle de l'International Society for Pharmacoepidemiology (ISPE). Cette synthèse a porté sur l'ensemble des études étiologiques publiées, en 2008, dans cette revue. En effet, comme précédemment précisé dans ce manuscrit, les études étiologiques peuvent plus particulièrement être affectées par des erreurs de classification des sujets en tant qu'exposés ou non au(x) médicament(s) d'intérêt.

Au total, 139 références ont été consultées parmi lesquelles, 31 (22,3 %) ont été sélectionnées pour cette synthèse.³⁹⁻⁶⁹ Au travers de ces 31 références, 39 études étiologiques étaient présentées. Leurs principales caractéristiques ont été résumées dans le Tableau 1. Onze de ces études (repérées par les cellules grisées) avaient des objectifs d'ordre méthodologique comme, par exemple, de reproduire des essais cliniques analysés en « intention de traiter » dans les conditions observationnelles.^{54,57,58}

Ces 39 études étaient dans 71,8 % des cas des études de cohorte. Par ailleurs, parmi les études cas-témoins (n = 11), 81,8 % correspondaient à des études cas-témoins intra-cohortes. Quatorze études (35,9 %) étaient conduites en Amérique du Nord (États-Unis et Canada) et 25 (64,1 %) en Europe, dont 14 (56,0 %) au Royaume-Uni, 8 (32,0 %) dans les pays d'Europe du Nord (Pays-Bas, Danemark, pays scandinaves) et 3 (12,0 %) dans d'autres pays européens (France, Italie et Espagne).

Les médicaments du système cardiovasculaire étaient les plus fréquemment étudiés (25,6 %), principalement les statines, suivis par les médicaments du système musculosquelettique (20,5 %), principalement les AINS, puis par les psychotropes (17,9 %), principalement les antidépresseurs. Dans la grande majorité des études (84,6 %), la source de données utilisée pour la mesure de l'exposition médicamenteuse était une base de données pré-enregistrées. Les deux types de bases de données, administratives et médicales, étaient représentés dans des proportions relativement similaires (54,5 % *versus* 45,5 %). Cinq études (12,8 %) se basaient sur les déclarations des patients pour estimer l'exposition médicamenteuse et 1 étude (2,6 %) sur les données extraites de dossiers médicaux hospitaliers.

Hormis dans cette dernière étude⁴³, d'autres critères que la classe médicamenteuse et la source de données étaient précisés pour définir l'exposition. À noter que la compréhension de ces critères pouvait parfois être difficile. Ils ont été plus particulièrement examinés pour les 24 études utilisant des bases de données pré-enregistrées et n'ayant pas pour objectif de mimer des essais cliniques. Dans 62,5 % de ces études, l'exposition médicamenteuse était définie de manière « longitudinale » avant la survenue de l'événement. Différentes méthodes étaient employées et, principalement, le calcul de la durée théorique de chaque prescription afin de détecter des épisodes d'utilisation continue du médicament.^{39,45,48,49,50,53,61,67} D'autres auteurs calculaient la somme des Doses Définies Journalières (DDJ) dispensées au cours du suivi^{52,59} ou des indicateurs tels que le Medication Possession Ratio (MPR) défini comme le rapport entre le nombre total de jours de traitement couverts par les dispensations et le nombre total de jours de la période d'étude.⁴⁴ Dans 58,3 % des études, l'existence d'éventuels problèmes d'observance était prise en compte soit en autorisant des espaces (ou « gaps ») entre la fin théorique d'une prescription et le début de la suivante^{45,50,53}, soit en allongeant artificiellement la durée théorique de chaque prescription.^{48,49} Par ailleurs, dans l'étude de Perreault *et al.*, utilisant le MPR, un seuil de 80 % était considéré comme suffisant.⁴⁴ Enfin, dans un tiers des études, l'exposition médicamenteuse était catégorisée selon qu'elle était « en cours », récente ou passée par rapport à la survenue de l'événement.^{41,45,46,48,50,65,66,67,69} Les seuils utilisés pour définir chaque catégorie « temporelle » étaient très variables d'une étude à l'autre (de quelques mois à plusieurs années).

En conclusion, deux points semblent importants à souligner à l'issue de cette synthèse : d'une part, la « suprématie » des bases de données pré-enregistrées comme sources de données d'exposition médicamenteuse dans ces études et, d'autre part, la grande variabilité des définitions utilisées pour classer les sujets en tant qu'exposés ou non en fonction de la nature des données

disponibles et des classes médicamenteuses d'intérêt. S'il a été possible de définir de manière consensuelle au niveau international un grand nombre d'événements, il paraît peu envisageable d'en faire de même pour l'exposition à une classe médicamenteuse donnée. En effet, l'existence d'un risque lié à un médicament dépend souvent de ses conditions d'utilisation. Or, ces conditions d'utilisation « à risque » varient selon l'événement considéré et restent souvent mal connues.

Tableau 1. Descriptif des études étiologiques publiées en 2008 dans la revue *Pharmacoepidemiology and Drug Safety* (n = 39)

Premier auteur (référence n°)	Localisation	Effectif	Classes médicamenteuses	Événements	Source de données (exposition)
Friedman (39)	États-Unis	n = 361 659	Statines	Cancers	Base de données administratives
Sukel (40)	Pays-Bas	n = 648	Chimiothérapies anticancéreuses	Événements cardiovasculaires	Base de données administratives
Tamim (41)	Canada	n = 15 495	Antibiotiques	Cancer du sein	Base de données administratives
Holst (42)	Suède	n = 860 215	Phytothérapie ^a	Événements à la naissance	Patientes
Ibáñez (43)	Espagne	n = 135	G-CSF ^b	Agranulocytose	Dossiers médicaux hospitaliers
Perreault (44)	Canada	n = 20 064	Médicaments des désordres osseux ^c	Fractures	Base de données administratives
Hammad (45)	Royaume-Uni	n = 283 136	Anti-inflammatoires non stéroïdiens	Infarctus du myocarde	Base de données médicales
Kelly (46)	États-Unis	n = 8 135	Anti-inflammatoires non stéroïdiens	Cancer du poumon	Patients
Lewis (47)	États-Unis	n = 2 248	Antidépresseurs	Toxicité gastro-intestinale haute	Patients
Derijks (48)	Pays-Bas	n = 2 446	Antidépresseurs	Hypoglycémie	Base de données administratives
Said (49)	États-Unis	n = 5 862	Neuroleptiques	Blessures involontaires	Base de données administratives
Nielsen (50)	Danemark	-	Traitement hormonal de la ménopause	Infarctus du myocarde	Base de données administratives
Guiguet (51)	France	n = 539	Antirétroviraux	Infarctus du myocarde	Base de données médicales
Russo (52)	Italie	n = 64 114	Spironolactone	Hémorragie digestive haute	Base de données administratives
Meropol (53)	Royaume-Uni	n = 999 173	Antibiotiques	Effets indésirables graves ^d	Base de données médicales
Meropol (53)	États-Unis	n = 1 047 496	Antibiotiques	Effets indésirables graves ^d	Base de données administratives
Meropol (53)	États-Unis	n = 1 819 004	Antibiotiques	Effets indésirables graves ^d	Base de données administratives

^aÉvaluation de l'exposition médicamenteuse pendant la grossesse

^bGranulocyte Colony Stimulating Factor

^cBiphosphonates, raloxifène et calcitonine

^dEffets indésirables entraînant une hospitalisation

Tableau 1. Suite

Premier auteur (référence n°)	Localisation	Effectif	Classes médicamenteuses	Événements	Source de données (exposition)
Delaney (54)	Royaume-Uni	n = 21 532	Warfarine	Variations de pression artérielle	Base de données médicales
Delaney (54)	Royaume-Uni	n = 92 037	Ibuprofène	Variations de pression artérielle	Base de données médicales
Delaney (54)	Royaume-Uni	n = 153 695	Inhibiteurs de la pompe à protons	Variations de pression artérielle	Base de données médicales
Delaney (54)	Royaume-Uni	n = 118 704	Statines	Variations de pression artérielle	Base de données médicales
Delaney (54)	Royaume-Uni	n = 6 399	Rofécoxib	Variations de pression artérielle	Base de données médicales
Delaney (54)	Royaume-Uni	n = 6 217	Célécoxib	Variations de pression artérielle	Base de données médicales
Knol (55)	Pays-Bas	n = 313	Antidépresseurs	Variations de glycémie	Base de données administratives
Cunnington (56)	États-Unis	n = 80 826	Anti-inflammatoires non stéroïdiens	Événements cardiovasculaires	Base de données administratives
Weiner (57)	Royaume-Uni	n = 4 151	Statines	Plusieurs ^a	Base de données médicales
Tannen (58)	Royaume-Uni	n = 35 521	Inhibiteurs de l'enzyme de conversion	Événements cardiovasculaires	Base de données médicales
Tannen (58)	Royaume-Uni	n = 19 958	Inhibiteurs de l'enzyme de conversion	Événements cardiovasculaires	Base de données médicales
Haukka (59)	Finlande	n = 1 611	Psychotropes	Plusieurs ^b	Base de données administratives
Margolis (60)	Royaume-Uni	n = 63 579	Hypoglycémiant (insuline et oraux)	Événements cardiovasculaires	Base de données médicales
Walker (61)	États-Unis	n = 543 000	Hypoglycémiant oraux	Événements cardiovasculaires	Base de données administratives
Källén (62)	Suède	n = 831 324	Antidépresseurs ^c	Hypertension pulmonaire	Patientes
Yang (63)	Royaume-Uni	n = 48 724	Statines	Cancer colorectal	Base de données médicales
Skurtveit (64)	Norvège	n = 13 390	Benzodiazépines	Consommation d'opioïdes ^d	Patients

^aMortalité toutes causes et infarctus du myocarde

^bSuicide, tentative de suicide entraînant une hospitalisation et mortalité

^cÉvaluation de l'exposition médicamenteuse pendant la grossesse

^dConsommation d'antalgiques opioïdes vingt ans plus tard

Tableau 1. Suite

Premier auteur (référence n°)	Localisation	Effectif	Classes médicamenteuses	Événements	Source de données (exposition)
García-Rodriguez (65)	Royaume-Uni	n = 127 391	Statines	Plusieurs ^a	Base de données médicales
García-Rodriguez (66)	Canada	n = 25 238	Statines	Plusieurs ^a	Base de données administratives
Patterson (67)	États-Unis	n = 8 677	Anti-inflammatoires non stéroïdiens	Hémorragie digestive haute	Base de données administratives
Koro (68)	États-Unis	n = 39 480	Hypoglycémiants (insuline et oraux)	Infarctus du myocarde	Base de données administratives
Assimes (69)	Canada	n = 128 667	Antihypertenseurs	Cancers	Base de données administratives

^aMortalité et insuffisance rénale aiguë

1.1.3- État des lieux de la situation française

Si, depuis plus d'une vingtaine d'années, la pharmaco-épidémiologie est une discipline bien implantée en Amérique ou Europe du Nord, elle est encore en cours de développement en France où elle présente, par conséquent, un certain retard.^{1,70}

La plupart des études observationnelles françaises consistent en des études *ad hoc* auprès des patients, médecins ou pharmaciens. Par ailleurs, l'épidémiologie française s'est enrichie ces dernières années, et continue de s'enrichir, de larges cohortes populationnelles constituées pour des objectifs variés. Parmi elles, peuvent être citées :

- la cohorte Gazel⁷¹ mise en place afin de suivre l'évolution de certains indicateurs de santé chez des employés de l'entreprise Électricité de France – Gaz de France,
- la cohorte Paquid⁷² (Personnes âgées quid ?) étudiant le vieillissement cérébral pathologique et normal à partir de près de 4 000 sujets âgés de Gironde et Dordogne,
- la cohorte E3N⁷³ (Étude Épidémiologique auprès de femmes de l'Éducation Nationale) étudiant les facteurs de risque de cancer et d'autres maladies, notamment les maladies cardiovasculaires, chez 98 995 femmes, nées entre 1925 et 1950, et adhérentes à la Mutuelle Générale de l'Éducation Nationale (MGEN) ou,
- la cohorte des Trois-Cités⁷⁴ (3C) visant à estimer le risque de démence attribuable aux facteurs vasculaires à partir de plus de 9 000 sujets âgés vivant dans trois régions françaises.

Les données déclaratives, régulièrement recueillies auprès des sujets inclus dans ces cohortes, peuvent faire l'objet d'études à visée pharmaco-épidémiologique. Néanmoins, les données concernant la consommation médicamenteuse ne sont pas toujours une priorité et peuvent souffrir d'un manque de fiabilité.

Pour ce qui est de l'existence de bases de données pré-enregistrées, le système d'Assurance Maladie français s'est progressivement informatisé générant ainsi des bases de données potentiellement utilisables en pharmaco-épidémiologie. Une des principales est la base de données ERASME, pour Extraction Recherche Analyse pour le Suivi Médico-Économique, de la Caisse Nationale d'Assurance Maladie des Travailleurs Salariés (CNAM-TS). Unique en termes de nombre de sujets, la CNAM-TS couvrant près de 80 % de la population générale française, cette base de données contient des informations relatives aux soins remboursés et, en particulier, aux médicaments prescrits. À noter que ni l'indication des médicaments, ni les posologies et durées de traitement prescrites ne figurent dans cette base de données. Leur potentiel a déjà été exploité

pour évaluer les modalités de suivi des malades et actualiser les données épidémiologiques disponibles à propos de certaines pathologies chroniques, comme le diabète⁷⁵, ou pour analyser les pratiques de prescription et de consommation, y compris dans des domaines spécialisés tels que la pharmacodépendance.^{76,77} Néanmoins, l'utilisation de ces bases de données à des fins de recherche est un concept relativement récent, sur lequel le recul est limité, et principalement dédié à l'épidémiologie descriptive.⁷⁸ Comparativement à ce qui a pu être mis en évidence dans la précédente revue de la littérature, il existe, en effet, que peu d'études étiologiques conduites à partir des bases de données de remboursement françaises.^{79,80}

Actuellement, très peu de liens existent entre les différentes sources de données épidémiologiques françaises. Un des rares exemples en la matière consiste en la disponibilité, au sein de la cohorte des Trois-Cités évoquée précédemment, de données déclaratives recueillies par interrogatoire des sujets inclus et de données de remboursement de médicaments extraites des bases de données de la CNAM-TS. Toutefois, il semble désormais y avoir une volonté politique de mutualiser les ressources existantes dans le but de faire progresser l'épidémiologie et la santé publique. Le futur projet PLASTICO (PLAte-forme Scientifique et TechnIque pour l'aide à la gestion de COhortes et de grandes enquêtes épidémiologiques) qui prévoit, entre autres, un accès aux données nationales de l'Assurance Maladie, aux données d'hospitalisation du PMSI (Programme de Médicalisation du Système d'Information) ou aux données de mortalité du CépiDC (Centre d'épidémiologie des causes de décès) en est la parfaite illustration.⁸¹ Quelques exemples d'études françaises récemment publiées, utilisant un croisement de plusieurs bases de données, vont également dans ce sens.⁸²⁻⁸⁴

I.2- OBJECTIFS

Dans ce contexte général, il paraissait intéressant d'explorer le potentiel représenté par les données de remboursement de l'Assurance Maladie française pour la mesure de l'exposition médicamenteuse dans les études pharmaco-épidémiologiques mais aussi de préciser celui représenté par les données déclaratives recueillies auprès des sujets inclus dans de grandes enquêtes épidémiologiques nationales. C'est pourquoi les données disponibles dans le cadre de la cohorte des Trois-Cités ont été utilisées afin de répondre à un premier objectif pour ce travail de thèse :

- décrire et comparer l'exposition médicamenteuse mesurée à partir de données déclaratives ou de données de remboursement.

Par ailleurs, la recherche et la quantification des biais (notamment de classification) dans les études étiologiques restant une préoccupation majeure en pharmaco-épidémiologie, les données de la cohorte des Trois-Cités ont également été exploitées afin de répondre à un second objectif pour ce travail de thèse :

- évaluer l'impact de la source de données utilisée pour déterminer l'exposition médicamenteuse sur la mesure d'une association entre un médicament et un événement.

Après une brève présentation de la cohorte des Trois-Cités, des données déclaratives et de remboursement disponibles ainsi que de la population source, ce manuscrit comprendra deux grandes parties organisées autour des deux objectifs définis ci-dessus puis une conclusion générale où seront abordées les perspectives des travaux de thèse. Outre la rédaction de plusieurs articles scientifiques présentés dans la suite du manuscrit, ces derniers ont également fait l'objet de communications orales et affichées à audience nationale et internationale listées en Annexe 1.

II- DONNÉES DISPONIBLES ET POPULATION SOURCE POUR LES TRAVAUX

II.1- PRÉSENTATION GÉNÉRALE DE LA COHORTE DES TROIS-CITÉS (3C)

La cohorte 3C a été constituée entre mars 1999 et mars 2001 avec, pour objectif général, d'estimer le risque de démence et de détérioration cognitive sévère attribuable aux facteurs de risque et pathologies vasculaires. Sa méthodologie a été exposée en détails dans une publication.⁷⁴

L'échantillonnage a été réalisé à partir de l'ensemble des sujets âgés de 65 ans et plus, non institutionnalisés et inscrits sur les listes électorales des villes de Bordeaux, Dijon et Montpellier. Les sujets ainsi identifiés (n = 34 922) ont été informés des objectifs et du protocole de l'étude par courrier puis par téléphone et, finalement, 9 294 sujets ont été inclus à raison de 2 104 participants à Bordeaux, 4 931 à Dijon et 2 259 à Montpellier. Les sujets ainsi sélectionnés ont été suivis sur une période de quatre ans avec deux visites prévues :

- la première, deux ans après inclusion soit entre 2001 et 2003 (suivi S1) et,
- la seconde, quatre ans après inclusion soit entre 2003 et 2005 (suivi S2).

À l'inclusion comme à chaque visite de suivi, les sujets ont été interrogés par des enquêtrices psychologues au moyen d'un questionnaire standardisé portant, d'une part, sur les principales caractéristiques sociodémographiques, d'autre part, sur l'état de santé dont les antécédents médicaux et, plus particulièrement, de pathologies et facteurs de risque cardiovasculaires, la symptomatologie dépressive, le niveau de dépendance et les performances cognitives et, enfin, sur la consommation médicamenteuse.

La symptomatologie dépressive était mesurée à l'aide de la version française de l'échelle CES-D (Center for Epidemiologic Studies - Depression Scale). Cette échelle est composée de vingt items cotés de 0 à 3 selon la fréquence des symptômes au cours de la semaine écoulée. La présence d'une symptomatologie dépressive a été définie à partir de seuils validés : score supérieur à 22 chez les femmes et à 16 chez les hommes.⁸⁵ Le niveau de dépendance était évalué au moyen de l'échelle d'activités de Lawton⁸⁶ et les performances cognitives par une batterie de tests neuropsychologiques dont le MMSE (Mini Mental State Examination). Il s'agit d'une échelle composite explorant l'orientation dans le temps et l'espace, la mémoire immédiate et différée, le calcul mental, le langage et les capacités visuoconstructives. Le score total varie de 0 à 30, le score maximal indiquant de bonnes performances.⁸⁷

II.2- DESCRIPTION DES DONNÉES DÉCLARATIVES ET DE REMBOURSEMENT DISPONIBLES

À chacun des interrogatoires de la cohorte 3C (inclusion, suivi S1 et suivi S2), les sujets devaient établir la liste de tous les médicaments, qu'ils aient été prescrits ou non, pris régulièrement (soit plusieurs fois par semaine) au cours du mois précédent. Cette liste, comprenant au maximum 15 médicaments, était validée par inspection des ordonnances et conditionnements que le sujet devait apporter à l'occasion de la visite. Les médicaments étaient ensuite enregistrés sous leur nom de spécialité puis codés selon la classification ATC (Anatomical Therapeutic Chemical) qui attribue un code unique à chaque principe actif.⁸⁸

Parallèlement, dès l'élaboration du protocole de l'étude 3C, l'accès à des données informatisées de remboursement de la CNAM-TS avait été prévu. À partir de la liste de tous les sujets inclus dans la cohorte 3C (n = 9 294), la CNAM-TS a, ainsi, identifié les sujets affiliés et extrait de ses bases de données l'ensemble de leurs remboursements de médicaments entre 2001 et 2005 pour les sujets de Bordeaux et Dijon et entre 2001 et 2003 pour ceux de Montpellier. Pour chaque médicament remboursé, la date de remboursement, le code ATC et le code CIP (Club Inter-Pharmaceutique) ont été fournis par la CNAM-TS. Le code CIP représente l'identifiant unique de chaque spécialité pharmaceutique et, contrairement au code ATC qui ne renseigne que sur la nature du principe actif du médicament, il prend également en compte la forme galénique, le dosage, le nombre d'unités par conditionnement...

Parmi l'ensemble des données décrites ci-dessus et comme l'illustre la Figure 1, les travaux de comparaison des données déclaratives et des données de remboursement vis-à-vis de la mesure de l'exposition médicamenteuse ont porté sur :

- des données déclaratives recueillies entre mars 2001 et mars 2005 aux deux visites de suivi S1 et S2 de la cohorte 3C,
- des données de remboursement couvrant entre trois et cinq années consécutives à partir du 1^{er} janvier 2001.

Préalablement à ces travaux, plusieurs semaines ont été consacrées à la préparation des données de remboursement. Une attention particulière a été portée sur les codes ATC des médicaments, ces derniers servant de base pour la comparaison des données. Plusieurs étapes, brièvement décrites en Annexe 2, ont donc été effectuées afin de garantir, autant que possible, l'exhaustivité

du codage ATC dans les données de remboursement et sa parfaite correspondance avec celui utilisé dans les données déclaratives de la cohorte 3C.

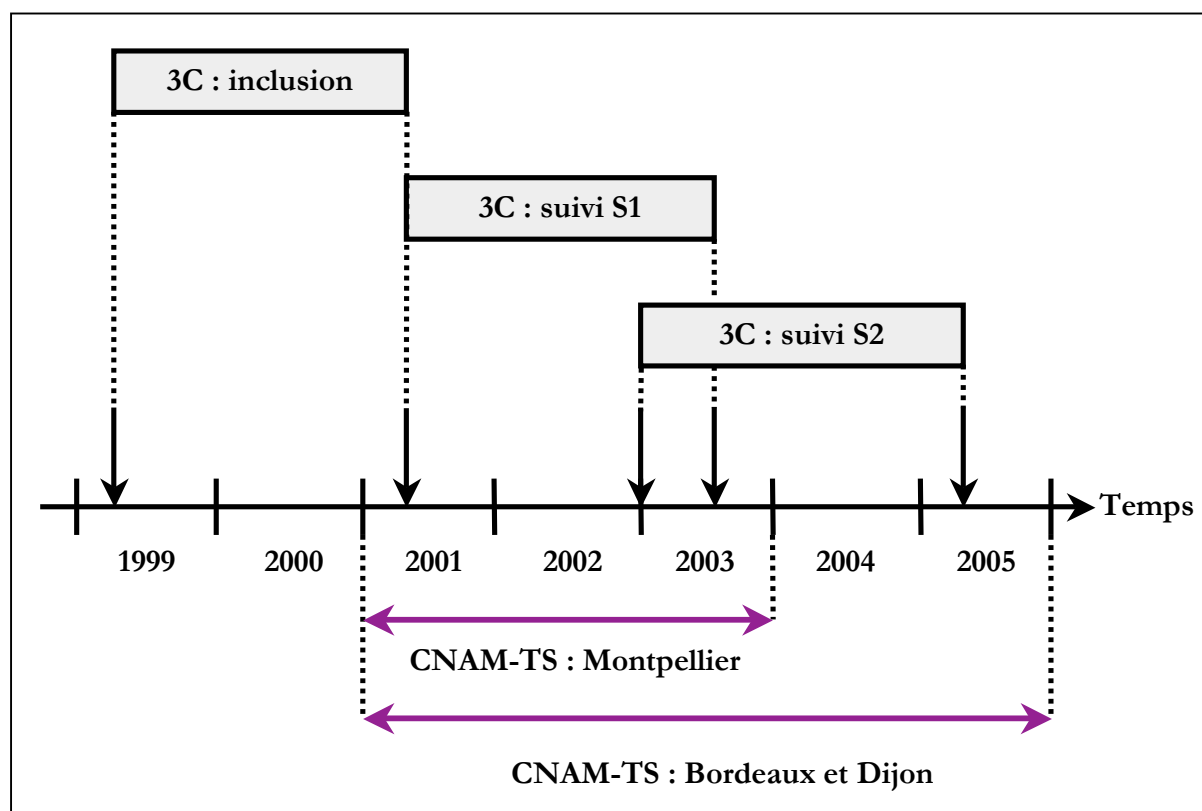


Figure 1- Représentation schématique des données déclaratives et de remboursement disponibles dans la cohorte des Trois-Cités (1999 – 2005)

II.3- DÉFINITION ET DESCRIPTION DE LA POPULATION SOURCE

La population source commune à l'ensemble des travaux de thèse était constituée de tous les sujets inclus dans la cohorte 3C ayant au moins une donnée de remboursement disponible au cours du suivi de la cohorte.

Les principales caractéristiques des sujets de cette population source ont été décrites et comparées à celles des autres sujets inclus dans la cohorte 3C (sujets n'ayant aucune donnée de remboursement sur la totalité de la période de suivi) afin d'identifier d'éventuels biais de sélection. Cette comparaison a porté sur des caractéristiques sociodémographiques et des variables relatives à l'état de santé et à la consommation médicamenteuse renseignées à l'inclusion dans la cohorte. Les variables quantitatives ont été décrites par la médiane, le minimum et le maximum. Ces variables n'étant pas distribuées selon une loi normale, elles ont été comparées entre les deux groupes par un test non paramétrique de Wilcoxon. Pour les variables qualitatives, l'effectif et la fréquence (en %) dans chaque classe ont été déterminés et le test du Khi-deux de Pearson a été utilisé pour la comparaison entre les deux groupes. Pour tous les tests statistiques, une valeur p égale à 5 % a été choisie comme seuil de significativité.

Parmi les 9 294 sujets inclus dans 3C, 4 962 (53,4 %) présentaient au moins une donnée de remboursement disponible au cours du suivi de la cohorte. L'âge médian des sujets de la population source était de 73 ans (étendue : 65 – 100 ans) et 64,7 % étaient des femmes.

Quoique parfois statistiquement significatives, les différences observées entre les sujets de la population source et les autres sujets inclus dans la cohorte 3C étaient faibles pour la majorité des variables, la significativité des tests ayant été favorisée par la grande taille des effectifs dans chacun des groupes (Tableau 2). Cependant, une différence notable était observée concernant le niveau d'étude, ce dernier étant inférieur chez les sujets de la population source. En effet, 30,4 % des sujets de cette population avaient un niveau d'étude au moins égal au baccalauréat contre 47,4 % des autres sujets inclus dans 3C ($p < 10^{-4}$).

Tableau 2. Comparaison des sujets de la population source et des autres sujets inclus dans la cohorte 3C (n = 9 294)

Caractéristiques, n (%)	Population source n = 4 962	Autres sujets n = 4 332	p
Âge (ans)			< 0,0001
65 - 69	1190 (24,0)	961 (22,2)	
70 - 74	1664 (33,5)	1324 (30,6)	
75 - 79	1314 (26,5)	1225 (28,3)	
≥ 80	794 (16,0)	822 (19,0)	
Sexe			< 0,0001
Hommes	1750 (35,3)	1900 (43,9)	
Femmes	3212 (64,7)	2432 (56,1)	
Niveau d'étude			< 0,0001
≤ Études primaires	1558 (31,5)	872 (20,2)	
1 ^{er} cycle	1884 (38,1)	1401 (32,4)	
2 ^{ème} cycle	838 (16,9)	1003 (23,2)	
Études supérieures	669 (13,5)	1048 (24,2)	
Mode de vie			0,0003
Seul(e)	2859 (37,5)	1464 (33,9)	
Couple / cohabitation	3092 (62,5)	1856 (66,1)	
Pathologies ou facteurs de risque cardiovasculaires ^a			0,01
Aucun	1443 (29,6)	1353 (31,9)	
Au moins un	3433 (70,4)	2883 (68,1)	
Fréquentation médicale ^b			0,05
Jamais	373 (7,5)	369 (8,6)	
Faible ^c	3839 (77,6)	3253 (75,6)	
Forte ^d	733 (14,8)	682 (15,8)	
Santé subjective			0,008
Bonne / très bonne	2834 (57,4)	2598 (60,5)	
Moyenne	1836 (37,2)	1469 (34,2)	
Mauvaise / très mauvaise	269 (5,4)	227 (5,3)	

^aHypertension artérielle, diabète, hypercholestérolémie, angor, troubles du rythme cardiaque, infarctus du myocarde et accident vasculaire cérébral

^bConsultation du médecin traitant en dehors de tout problème aigu de santé

^cMoins d'une fois par an ou une à deux fois par an

^dTrois à quatre fois par an ou au moins une fois tous les deux mois

Tableau 2. Suite

Caractéristiques, n (%)	Population source n = 4 962	Autres sujets n = 4 332	p
Affection de longue durée			0,3
Oui	1331 (27,2)	1123 (26,3)	
Non	3569 (72,8)	3147 (73,7)	
Dépendance (échelle d'activités de Lawton) ^a			0,03
Aucun	4312 (87,4)	3763 (87,2)	
1 ou 2	553 (11,2)	463 (10,7)	
3 ou 4	67 (1,4)	88 (2,0)	
Score au MMSE			< 0,0001
Médiane	28	28	
Minimum	6	7	
Maximum	30	30	
Symptomatologie dépressive (CES-D)			0,5
Oui	678 (13,9)	569 (13,4)	
Non	4196 (86,1)	3675 (86,6)	
Nombre de médicaments ^b			0,002
Aucun	385 (7,8)	404 (9,3)	
1 - 4	2275 (45,8)	2053 (47,4)	
Au moins 5	2302 (46,4)	1875 (43,3)	

^aNombre d'items affectés parmi les suivants : utilisation du téléphone, moyens de transport, prise de médicaments et gestion d'un budget

^bMédicaments déclarés être consommés à l'inclusion dans la cohorte

III- PREMIÈRE PARTIE
DESCRIPTION ET COMPARAISON DE L'EXPOSITION
MÉDICAMENTEUSE MESURÉE À PARTIR DE DONNÉES
DÉCLARATIVES OU DE DONNÉES DE REMBOURSEMENT

III.1- SYNTHÈSE DE LA LITTÉRATURE

Parmi les études de comparaison de plusieurs sources de données pour la mesure d'une exposition médicamenteuse, cette synthèse porte plus spécifiquement sur celles ayant confronté des données obtenues auprès des sujets lors d'interrogatoires ou extraites de bases de données informatisées de prescription, dispensation ou remboursement de médicaments.^{13,16,20,23,37,89-107} Majoritairement publiées dans les années 1990 et 2000, ces études ont été conduites aux États-Unis ou en Europe du Nord sur des populations de nature et d'effectif très variés. Leurs principales caractéristiques sont présentées dans le Tableau 3 et, notamment, les modalités de l'interrogatoire, le type de base de données et les classes médicamenteuses étudiées.

La diversité des méthodes utilisées pour comparer les expositions médicamenteuses mesurées à partir de chaque source de données mérite d'être soulignée. Dans certaines études, la concordance entre les expositions était mesurée au moyen de pourcentages de concordance, de coefficients kappa ou de coefficients de corrélation^{20,23,37,89,92,95-97,104,106,108} alors que, dans d'autres, la validité de l'exposition issue d'une source par rapport à celle issue de la seconde était exprimée par des paramètres comme la sensibilité, la spécificité ou la valeur prédictive positive.^{13,16,37,90,91,93,94,98,99,101-104} Comme souvent dans les conditions observationnelles, le choix de la méthode dépendait principalement du contexte de l'étude et, plus précisément ici, de la conviction ou non des auteurs dans chaque contexte spécifique que l'exposition issue d'une des deux sources reflétait davantage l'exposition réelle des sujets. Citons, pour illustration, trois de ces études dont celle de Klungel *et al.* dans laquelle la comparaison portait sur l'exposition aux médicaments antihypertenseurs chez des sujets adultes inclus dans une étude de cohorte.⁹⁷ Dans ce cas, les auteurs expliquaient ne pas avoir dégagé d'arguments particuliers pour considérer plutôt l'interrogatoire bien conduit des sujets (auto-questionnaire rempli à domicile puis vérifié au centre d'étude) que l'extraction, à partir d'une base de données, de l'historique de leurs dispensations pharmaceutiques, comme la méthode la plus à même d'approcher leur véritable exposition. En revanche, s'agissant de l'exposition médicamenteuse au cours de la grossesse, Olesen *et al.* ont considéré que l'interrogatoire des femmes enceintes en donnait une image plus réelle par rapport à leurs données personnelles de remboursement. Selon eux, certains médicaments prescrits, délivrés puis remboursés peuvent ne pas avoir été effectivement consommés par la femme enceinte préoccupée par les éventuels effets toxiques pour l'enfant.⁹⁸ À l'inverse, dans le cadre d'une étude cas-témoins sur le cancer du sein menée auprès de femmes âgées et nécessitant de mesurer leur exposition médicamenteuse au cours de longues périodes

dans le passé, Boudreau *et al.* avaient davantage d'arguments en faveur de l'utilisation de données informatisées de dispensation par rapport à des données déclaratives supposées être moins fiables du fait d'un biais de mémorisation potentiellement important.¹⁰²

Malgré l'évidente hétérogénéité de contexte, population ou méthodologie entre toutes ces études, il est intéressant de constater un certain nombre de similitudes dans leurs résultats. À titre d'exemple, la plupart de celles portant sur l'exposition médicamenteuse « toutes classes confondues » montraient, d'une part, de bons résultats (de concordance entre les expositions issues des deux sources de données ou de validité de l'exposition issue d'une des sources par rapport à celle issue de la seconde) pour les médicaments indiqués dans le traitement des pathologies et facteurs de risque cardiovasculaires ou du diabète ou encore de l'hypothyroïdie et, d'autre part, des résultats plus mitigés pour les anti-infectieux ou les antalgiques ou encore les médicaments dermatologiques.^{13,20,91,93,94,98,104,107}

À l'instar de ces études conduites aux États-Unis ou dans d'autres pays européens, les données disponibles de la cohorte 3C ont permis de comparer, dans une population âgée française, l'exposition médicamenteuse mesurée selon les déclarations des sujets ou selon leurs remboursements de médicaments extraits des bases de données de l'Assurance Maladie. La richesse de ces données a été exploitée afin d'envisager, à travers ces travaux de comparaison, différents contextes d'études rencontrés en pharmaco-épidémiologie depuis les études transversales descriptives jusqu'aux études étiologiques évaluant l'association entre la prise d'un médicament et la survenue d'un événement.

Tableau 3. Descriptif des études publiées de comparaison de l'exposition médicamenteuse déclarée par les sujets lors d'interrogatoires ou issue de bases de données informatisées (n = 24)

Premier auteur Année (référence n°)	Localisation	Population Effectif	Interrogatoire	Base de données	Classes médicamenteuses
Enlund 1981 (89)	Finlande	Sujets de 25 à 59 ans n = 310	Auto-questionnaire et inventaire	Dispensation	Antihypertenseurs
Van den Brandt 1991 (90)	Pays-Bas	Sujets âgés n = 207	Auto-questionnaire	Dispensation	Toutes classes
Johnson 1991 (91)	États-Unis	Sujets âgés n = 83	Questionnaire et inventaire à domicile	Dispensation (Kaiser Permanente)	Toutes classes
de Jong-van den Berg 1993 (92)	Pays-Bas	Femmes jeunes ^a n = 217	Questionnaire à domicile	Dispensation	Toutes classes
West 1995 (16)	États-Unis	Sujets d'au moins 50 ans n = 454	Questionnaire par téléphone	Dispensation (Group Health Cooperative)	Plusieurs classes ^b
Heerdink 1995 (93)	Pays-Bas	Sujets âgés n = 100	Questionnaire et inventaire à domicile	Dispensation	Toutes classes
Lau 1997 (94)	Pays-Bas	Sujets âgés n = 115	Questionnaire et inventaire à domicile	Dispensation	Toutes classes
Saunders 1998 (95)	États-Unis	Sujets de 20 à 80 ans n = 138	Questionnaire par téléphone	Dispensation (Group Health Cooperative)	Antidépresseurs

^aÉvaluation de l'exposition médicamenteuse pendant la grossesse

^bAnti-inflammatoires non stéroïdiens et oestrogènes (hors contraception orale)

Tableau 3. Suite

Premier auteur Année (référence n°)	Localisation	Population Effectif	Interrogatoire	Base de données	Classes médicamenteuses
Norell 1998 (23)	Suède	Femmes jeunes n = 427	Questionnaire à domicile	Dispensation	Contraceptifs oraux
Sjahid 1998 (96)	Pays-Bas	Sujets d'au moins 55 ans n = 1 682	Questionnaire et inventaire	Dispensation	Médicaments cardiovasculaires
Klungel 1999 (97)	Pays-Bas	Sujets de 20 à 59 ans n = 372	Auto-questionnaire à domicile	Dispensation (PHARMO)	Antihypertenseurs
Klungel 2000 (13)	Pays-Bas	Sujets de 20 à 59 ans n = 372	Auto-questionnaire à domicile	Dispensation (PHARMO)	Toutes classes
Olesen 2001 (98)	Danemark	Femmes jeunes ^a n = 2 041	Auto-questionnaire et questionnaire par téléphone	Remboursement (North Jutland Prescription Database)	Toutes classes
Monster 2002 (37)	Pays-Bas	Sujets de 28 à 75 ans n = 7 568	Auto-questionnaire	Dispensation	Plusieurs classes ^b
Metlay 2003 (99)	États-Unis	Sujets d'au moins 18 ans n = 103	Questionnaire par téléphone	Dispensation (National Veterans Affairs)	Plusieurs classes ^c
Al-Windi 2003 (100)	Suède	Sujets d'au moins 16 ans n = 1 436	Auto-questionnaire	Prescription	Toutes classes

^aÉvaluation de l'exposition médicamenteuse pendant la grossesse

^bAntihypertenseurs, hypolipémiants, hypoglycémisants oraux, dérivés nitrés, contraceptifs oraux, traitement hormonal de la ménopause, anti-inflammatoires non stéroïdiens, aspirine et paracétamol

^cAntibiotiques, antihypertenseurs et anti-inflammatoires non stéroïdiens

Tableau 3. Suite

Premier auteur Année (référence n°)	Localisation	Population Effectif	Interrogatoire	Base de données	Classes médicamenteuses
Strom 2004 (101)	États-Unis	Femmes de 50 à 79 ans n = 154	Questionnaire par téléphone	Dispensation	Traitement hormonal de la ménopause
Caskie 2004 (20)	États-Unis	Sujets âgés n = 294	Inventaire « Brown bag method »	Dispensation	Toutes classes
Boudreau 2004 (102)	États-Unis	Femmes âgées n = 403	Questionnaire et inventaire	Dispensation (Group Health Cooperative)	Plusieurs classes ^a
Løkkegaard 2004 (103)	Danemark	Femmes ménopausées ^b n = 2 694	Auto-questionnaire	Dispensation et remboursement	Traitement hormonal de la ménopause
Caskie 2006 (104)	États-Unis	Sujets de 23 à 97 ans n = 1 430	Inventaire « Brown bag method »	Dispensation (Group Health Cooperative)	Toutes classes
Glintborg 2007 (105)	Danemark	Sujets de 17 à 97 ans n = 500	Deux questionnaires dont un à domicile	Dispensation et remboursement	Toutes classes
Haukka 2007 (106)	Finlande	Sujets de 30 à 66 ans n = 905	Questionnaire et inventaire	Remboursement (Social Insurance Institution)	Médicaments psychotropes
Nielsen 2008 (107)	Danemark	Sujets d'au moins 16 ans n = 16 688	Questionnaire à domicile	Dispensation	Toutes classes

^aAntihypertenseurs, statines et antidépresseurs^bFemmes incluses dans une cohorte d'infirmières

III.2- DESCRIPTION ET COMPARAISON DE L'EXPOSITION MÉDICAMENTEUSE MESURÉE À PARTIR DE DONNÉES DÉCLARATIVES OU DE REMBOURSEMENT UN JOUR DONNÉ (ARTICLE N°1)

L'étude menée dans ce premier contexte de la mesure de l'exposition médicamenteuse un jour donné est présentée sous la forme d'un article scientifique publié début 2009 dans la revue *Pharmacoepidemiology and Drug Safety*.

III.2.1- Présentation de l'article

Noize P, Bazin F, Dufouil C, *et al.* Comparison of health insurance claims and patient interviews in assessing drug use: data from the Three-City (3C) Study. *Pharmacoepidemiol Drug Saf* 2009; 18: 310-9

III.2.1.1- Contexte et objectifs

En pharmaco-épidémiologie, la prise de décisions relatives aux risques et bénéfices des médicaments à l'échelle populationnelle nécessite de disposer de données actualisées concernant leur utilisation. Face à un problème de santé publique lié à un médicament, il est notamment indispensable de connaître le nombre de sujets qui y sont exposés ainsi que leurs principales caractéristiques (caractéristiques sociodémographiques, autres médicaments consommés...). Comparativement à la mise en place d'études de terrain de méthodologie lourde et souvent sujettes à des biais de représentativité, l'interrogation de bases de données informatisées s'avère, au moins en première approche, un moyen d'accéder rapidement à de telles informations et, de surcroît, sur de grands échantillons.¹⁰⁸ Ainsi, les bases de données de remboursement de l'Assurance Maladie ont progressivement été utilisées ces dernières années pour la conduite, à l'échelon local, régional ou national, d'études transversales descriptives d'utilisation médicamenteuse que ce soit par l'Assurance Maladie elle-même¹⁰⁹⁻¹²¹ ou par des équipes de recherche ayant demandé une extraction de données.¹²²⁻¹³³ Néanmoins, les sujets identifiés comme exposés à un médicament un jour donné sur la base de la soumission d'une prescription à remboursement peuvent ne pas l'avoir réellement consommé tel qu'ils l'auraient, par exemple, déclaré lors d'un interrogatoire sur leur consommation médicamenteuse actuelle.

Dans ce contexte, les objectifs de cette étude étaient de décrire et comparer, chez des sujets inclus dans la cohorte 3C, l'exposition à de nombreuses classes médicamenteuses mesurée, d'une part, d'après leurs déclarations de consommation à la première visite de suivi de la cohorte (S1) et, d'autre part, d'après leurs remboursements intervenus récemment avant cette visite (Figure 2).

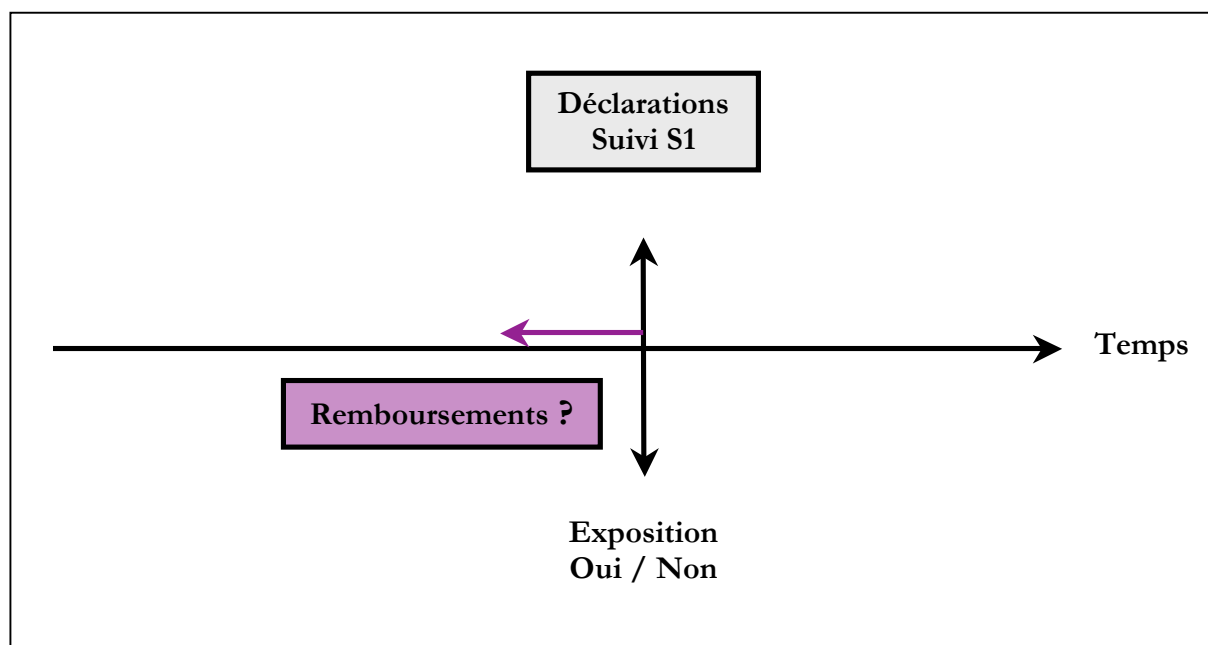


Figure 2. Illustration de la mesure transversale de l'exposition médicamenteuse d'après les données déclaratives et de remboursement disponibles pour la première visite de suivi de la cohorte 3C

III.2.1.2- Considérations méthodologiques

III.2.1.2.1- Échantillon d'étude

Compte-tenu des objectifs, la date à laquelle les sujets étaient vus à la première visite de suivi (S1) de la cohorte 3C a été choisie comme date de référence pour la comparaison de l'exposition médicamenteuse mesurée selon les données déclaratives ou selon les données de remboursement. Par ailleurs, il a été décidé de considérer les données de remboursement jusqu'à 60 jours avant cette date de référence (cf. § III.2.1.2.2). L'étude a donc été conduite sur un échantillon de 4 112 sujets (âge médian : 75 ans ; 64,7 % de femmes) de la population source interrogés au premier suivi de la cohorte et ayant des données de remboursement disponibles avec un recul de 60 jours par rapport à la date de cet interrogatoire.

III.2.1.2.2- Définitions de l'exposition médicamenteuse selon les données de remboursement

D'après les données déclaratives, les sujets ne pouvaient être classés comme exposés à une classe médicamenteuse que s'ils avaient déclaré consommer au moins un médicament de cette classe dans le mois précédant la date de l'interrogatoire (date de référence). En comparaison, plusieurs définitions de l'exposition ont été envisagées selon les données de remboursement. De manière intuitive, les sujets ont, tout d'abord, été classés comme exposés à une classe médicamenteuse s'ils avaient bénéficié du remboursement d'au moins un médicament de cette classe dans les 30 jours précédant la date de référence. Outre le fait que cette période correspondait à celle du recueil déclaratif, elle était également adaptée à la réglementation française selon laquelle la grande majorité des médicaments ne peuvent être délivrés par les pharmacies puis, le cas échéant, remboursés par l'Assurance Maladie que pour un mois de traitement.

Néanmoins, certains sujets n'observant pas correctement leur traitement peuvent progressivement accumuler certains médicaments et, à l'occasion d'un renouvellement d'ordonnance, ne pas en demander de délivrance (donc de remboursement). Il est, à l'inverse, fréquent que certains sujets renouvellent leur traitement avant la date prévue et, dans ce cas, le remboursement du médicament peut avoir été enregistré à une date antérieure à la période d'un mois considérée. En conséquence, les sujets ont secondairement été classés comme exposés à une classe médicamenteuse s'ils avaient bénéficié du remboursement d'au moins un médicament de cette classe dans les 60 jours précédant la date de référence.

III.2.1.2.3- Choix de la méthode de comparaison

Il est communément admis que, parmi les méthodes disponibles pour estimer l'exposition médicamenteuse actuelle d'un sujet, aucune ne constitue une véritable référence si ce n'est, dans certaines situations, la recherche du médicament ou de ses métabolites dans un prélèvement biologique.¹³⁴⁻¹³⁶ C'est pourquoi, dans un premier temps, les expositions médicamenteuses mesurées soit selon les données déclaratives, soit selon les données de remboursement ont été comparées au moyen du coefficient kappa.¹³⁷ Ce coefficient renseigne sur la concordance entre les deux types de données sans préjuger de leur capacité respective à déterminer l'exposition réelle des sujets.

Cependant, en fonction de chaque contexte et comme le montre la précédente synthèse de la littérature, une des méthodes à comparer peut apparaître supérieure aux autres et alors être considérée comme une potentielle référence (ou « alloyed gold standard »).^{4,138} Dans le cas présent, l'interrogatoire des sujets sur l'ensemble des médicaments consommés au cours d'une période courte et récente, complété par l'inspection des ordonnances et conditionnements, pouvait plus particulièrement représenter cette référence. Ainsi, dans un second temps, la validité de l'exposition médicamenteuse mesurée selon les données de remboursement a été évaluée en référence à celle issue des déclarations des sujets choisie comme « gold standard ».¹³⁹

III.2.1.3- Principaux résultats

L'exposition médicamenteuse mesurée selon les données de remboursement dans les 60 jours avant la date de référence reflétait beaucoup mieux l'exposition médicamenteuse déclarée par les sujets. Pour les médicaments utilisés dans le traitement du diabète, des maladies et facteurs de risque cardiovasculaires (troubles du rythme cardiaque, insuffisance cardiaque, hypertension artérielle, angor, hyperlipidémie), des maladies rhumatologiques (goutte, ostéoporose), des néoplasies, des maladies neurologiques (maladie de Parkinson, démence), des maladies psychiatriques (dépression, psychoses, anxiété, insomnie), des troubles de la coagulation, des maladies hormonales (hypo- ou hyperthyroïdie) et des maladies respiratoires (asthme, bronchopneumopathie chronique obstructive), les prévalences d'exposition mesurées à partir des données déclaratives ou des données de remboursement dans les 60 jours avant la date de référence étaient similaires. La concordance entre l'exposition mesurée selon chaque type de données ainsi que la validité de l'exposition mesurée à partir des données de remboursement dans les 60 jours avant la date de référence étaient bonnes. Parmi les classes médicamenteuses pour lesquelles elles étaient mauvaises, les vasodilatateurs périphériques, vasculoprotecteurs, psychostimulants, laxatifs, antispasmodiques, vitamines, médicaments du système génito-urinaire (chez les hommes seulement) et médicaments homéopathiques ou de phytothérapie présentaient une prévalence d'exposition nettement supérieure selon les données déclaratives. En revanche, cette prévalence était inférieure à celle mesurée selon les données de remboursement pour les antalgiques de palier 1 (paracétamol et aspirine), anti-infectieux, médicaments des organes sensoriels et médicaments dermatologiques.

III.2.2- Article n°1 publié

Comparison of health insurance claims and patient interviews in assessing drug use: data from the Three-City (3C) Study[†]

Pernelle Noize PharmD, MSc^{1,2,3}, Fabienne Bazin PhD^{1,2}, Carole Dufouil PhD^{4,5}, Nathalie Lechevallier-Michel PhD², Marie-Laure Ancelin PhD⁶, Jean-François Dartigues MD, PhD^{2,3,7}, Christophe Tzourio MD, PhD^{4,5}, Nicholas Moore MD, PhD^{1,2,3} and Annie Fourier-Réglat PharmD, PhD^{1,2,3*}

¹Inserm, Unit 657, Pharmacopépidémiologie et évaluation de l'impact des produits de santé sur les populations, Bordeaux, France

²Université Victor Segalen, Bordeaux, France

³Centre Hospitalier Universitaire, Bordeaux, France

⁴Inserm, Unit 708, Neuroépidémiologie, Paris, France

⁵Université Pierre et Marie Curie, Paris, France

⁶Inserm, Unit 888, Pathologies du système nerveux: recherche épidémiologique et clinique, Montpellier, France

⁷Inserm, Unit 897, Epidémiologie, santé publique, développement, Bordeaux, France

SUMMARY

Purpose Precise determination of drug exposure is fundamental in pharmacoepidemiology. Drug exposure is often presumed from health insurance claims but this may not correspond exactly to what subjects actually take. This study was designed to investigate French reimbursement databases in assessing drug use.

Methods Between 1999 and 2001, 9294 subjects were included in the Three-City (3C) Study, a French cohort studying the relationship between vascular risk factors and dementia. Of these, 4112 subjects had data available from both clinical interviews and the reimbursement databases of the French national health insurance system. Agreement between drugs reported as used at interview and drugs reimbursed during the previous 30 or 60 days was measured with κ coefficients. Using calculations of sensitivity (Se), specificity (Sp), positive predictive values (PPVs) and negative predictive values (NPVs), the validity of reimbursement data for the 30 or 60 days preceding the interview was investigated taking drugs reported at interview as the 'gold standard'.

Results Declared drug use at interview was less well predicted by 30-day than by 60-day reimbursement data. Agreement between reimbursement data and interviews as well as validity of reimbursement data with reference to interviews were substantial for drugs used in cardiovascular diseases, diabetes, rheumatic disorders or neuropsychiatric conditions and were poor for laxatives, vitamins, vasculoprotectives, first and second line analgesics, anti-infective products or dermatologicals.

Conclusions Reimbursement data with an appropriate time frame and interviews estimate exposure to chronically used drugs similarly. Self-medication was better described with interviews whereas reimbursement data seem more useful for drugs used topically or intermittently. Copyright © 2009 John Wiley & Sons, Ltd.

KEY WORDS—agreement; reimbursement databases; interview; pharmacoepidemiology; medicines; validity (epidemiology)

Received 12 June 2008; Revised 25 November 2008; Accepted 13 January 2009

INTRODUCTION

Pharmacoepidemiological studies often aim to assess the potential for a drug to be associated with adverse or

beneficial events in a particular population. In such investigations, the risk or benefit finally attributed to the drug of interest may differ according to the criteria used to classify subjects as users or non-users of this drug.^{1,2} Therefore, to perform such studies, valid and complete information on drug exposure is required.

Several sources provide useful data regarding drug use including patient interviews and reimbursement databases. Information collected by patient interviews

* Correspondence to: Dr A. Fourier-Réglat, Inserm U657, Département de Pharmacologie, Université Victor Segalen, 146 rue Léo Saignat, 33076 Bordeaux Cedex, France. E-mail: annie.fourrier@pharmaco.u-bordeaux2.fr
[†]The authors declared no conflict of interest.

is not limited which is particularly advantageous to estimate over the counter (OTC) drug consumption or self-medication patterns.³ However, these are subject to memory bias.^{4,5} The quality of the data collected can also be largely affected by the way subjects are interviewed (e.g. structured questionnaires, open-ended or precisely oriented questions, visual aids) as well as the nature of the drug exposure to assess (e.g. current or past use, number and type of drugs studied).^{3,6–11} Reimbursement databases are not subject to memory bias but present certain limits, primarily that reimbursement does not necessarily mean use. They also lack data regarding OTC drugs or self-medication.¹ Moreover, there may be a delay between the moment a drug is reimbursed and when it is actually used and the duration of use is not always predicted from pack size or theoretical daily dose.

In France, approximately 85% of the general population and 69% of those aged 65 years and over is covered by the national health insurance system, Caisse Nationale d'Assurance Maladie des Travailleurs Salariés (CNAM-TS), which reimburses prescribed drugs and other medical expenses to salaried persons and their families. The computerized databases of the CNAM-TS, in which all claims for reimbursement are recorded, could provide interesting data on drug exposure, though these data are currently rarely used.¹² It is recognized that when using health care databases, the validity of the data must be investigated. Several authors have done so and include Christensen *et al.*¹³, McKenzie *et al.*¹⁴ and Johnson *et al.*¹⁵, who respectively studied the Group Health Cooperative, the Medicaid and the Kaiser Permanente databases in the United States of America or Monstere *et al.*¹⁶ and Lau *et al.*¹⁷, who studied the Dutch pharmacy databases. The present study was designed to investigate the CNAM-TS databases by comparing data collected by face-to-face interviews of elderly subjects participating in the French Three-City (3C) cohort with the recordings in the CNAM-TS databases of their drug-related expenditures using different database time frames.

METHODS

Study population

The 3C Study is a prospective cohort designed to assess the risk of dementia and cognitive decline due to vascular risk factors. The protocol of the 3C study was approved by the Ethics Committee of the Kremlin-Bicêtre University Hospital in Paris. Written informed consent was obtained from each participant. The methodology has been described in detail elsewhere.¹⁸

Briefly, between 1999 and 2001, 9294 subjects were included in the cohort according to the following criteria: (i) to be at least 65 years old, (ii) to live in the community and (iii) to be registered on the electoral rolls of one of three French cities: Bordeaux, Dijon or Montpellier. Participants were interviewed at inclusion, 2 years later (2001–2003) and 4 years later (2003–2005) with regard to socio-demographic characteristics, health status and drug use. In addition to interview data, reimbursement data were sought from the national health insurance system databases (CNAM-TS).

For the present study, reimbursement data were available for the 2-year follow-up period (2001–2003). Thus, the date of interview 2 years after inclusion in the cohort was chosen as reference date for each subject. Since reimbursement data were considered for up to 2 months before the reference date, the first date of interview had to occur at least 2 months after the 1 January 2001. Therefore, all subjects who participated in the 3C 2-year follow-up after the 1 March 2001 and for whom reimbursement data were available from CNAM-TS were eligible for the present study.

Drug exposure data

In France, drugs are dispensed for a maximum of 30 days whatever the actual prescription (i.e. a drug prescribed for 90 days will be dispensed in three times). Prescription forms are conserved by subjects following dispensation as these contain directions for use. At each dispensation, the pharmacist automatically submits reimbursement claims electronically. Moreover, as in most European countries, drugs are dispensed in packages which are indivisible by the pharmacists.

Data for the 3C Study were collected in each study centre by trained psychologists using a standardized questionnaire. Concerning drug use, subjects were asked to list all medications (prescribed and OTC) used more than once a week over the preceding month by an open-ended question. Not only to extend data collection on drug use but also to minimize memory bias, subjects were asked to present drug packages and prescription forms. A maximum of 15 drugs could be collected. Drugs were registered using their brand name and were then coded according to the anatomical therapeutic chemical (ATC) classification.¹⁹ Subjects were considered as users when at least one declaration for a particular drug class or sub-class was made even if the use could not be confirmed by visual inspection of prescriptions and packages during the interview.

The CNAM-TS provided for each subject the list of all drugs reimbursed over the 2 months prior to the

interview. For each drug, the ATC code and the corresponding reimbursement dates were made available. Use was defined when at least one reimbursement for a particular drug class or sub-class was found.

Daily dose or treatment duration data were not collected during the interview and were also absent from the reimbursement databases.

Statistical analyses

All analyses were conducted using two periods for CNAM-TS reimbursement data: the drugs reported during 3C interviews at reference date were compared to the drugs reimbursed by the CNAM-TS respectively over 30 or 60 days before the reference date (i.e. the date of interview).

Drug use was described according to each source of data and each reimbursement period using the number of drugs and the prevalence of use within ATC classes (see Appendix).

Because of the lack of a true 'gold standard', agreement between 3C interviews and CNAM-TS reimbursement data was first analysed by means of the κ statistic: κ coefficients and their 95% confidence intervals were calculated for each drug class using the same two periods for CNAM-TS reimbursement data.²⁰ Nevertheless, to better answer to the study objective (i.e. to investigate French reimbursement databases in assessing drug use), the validity of reimbursement data for the 30 or 60 days preceding the interview was next assessed taking drugs reported as being used at the time of interview as the 'gold standard'.²¹ Indeed, as information on drug use was collected at interview in the 3C Study for a recent period with confirmation by drug packages and prescription forms, the methodology of this interview presents certain similarities to drug in-home assessment that is often used as 'gold standard' in validation studies of computerized prescription or reimbursement data.^{15,17} Thus, for each drug class, sensitivity (Se), specificity (Sp) and the positive and negative predictive values were calculated. Se measured the ability of the reimbursement data (over 30 or 60 days before the interview) showing use of the drug class given actual consumption as indicated by the interview. Sp measured the ability of the reimbursement data showing non-use of the drug class given that this drug class was not found to be used according to the interview. The positive predictive value (PPV) was defined as the probability that a drug class was actually used (according to the interview) given consumption as indicated by the reimbursement data and the negative predictive value (NPV) as the probability that a drug

class was actually not used given that this drug class was not found to be used according to reimbursement data.

Many drug categories included very different types of drugs (e.g. anti-hypertensive agents and ointments against haemorrhoids are in the same cardiovascular system drugs class). As agreement between the two sources and validity of reimbursement data may both depend on the type of drug, these categories were subdivided (see Appendix). All the parameters listed above were then calculated for all drug sub-classes in order to determine their individual implication in the overall class agreement or validity.

RESULTS

Study population

Among the 9294 subjects included in the 3C cohort, 8061 were interviewed 2 years after their inclusion. Of these, 4112 subjects whose electronic records were available from the CNAM-TS were included in the study. Reasons for not having drug reimbursement data might be that subjects were affiliated with another health care system or that their data were not found (name change, change of address...). There was no difference regarding age, gender, health status or drug use between subjects included or not in the study (i.e. subjects with and without reimbursement data). The two groups significantly differed only with regard to education level (31.6% of subjects included in the study had high school level or above vs. 47.6% of subjects not included in the study). The median age of the study population was 75 years (range: 67–98), 64.7% were women and 72.2% reported at least one of the following cardiovascular risk factors or diseases: hypertension, diabetes, hypercholesterolemia, angina, heart rhythm disorders, myocardial infarction or stroke.

Description of drug use

Drug use data according to both sources and both reimbursement periods are shown for all drugs in Table 1 and for drug classes in Figure 1. The 60-day reimbursement data were much closer to interview data than the 30-day reimbursement data.

Comparison of reimbursement data and interview data

Thirty-day time frame. Agreement between sources was poor when reimbursement data were restricted to 30 days before the interview. Values of κ ranged

Table 1. Number of drugs used according to interviews and reimbursement data corresponding to the 30 or 60 days preceding the interview ($n = 4112$)

Number of drugs	3C interviews	CNAM-TS 30 days	CNAM-TS 60 days
Median (min-max)	5 (0-15)	2 (0-18)	4 (0-24)
No drug, n (%)	219 (5.3)	1388 (33.8)	543 (13.2)
One to four drugs, n (%)	1669 (40.6)	2050 (49.9)	1559 (37.9)
Five or more drugs, n (%)	2224 (54.1)	674 (16.4)	2010 (48.9)

3C, Three-City; CNAM-TS, Caisse Nationale d'Assurance Maladie des Travailleurs Salariés.

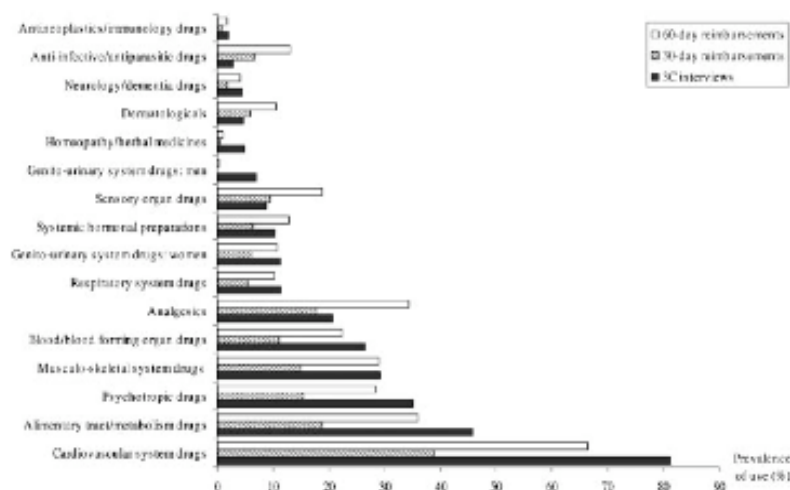


Figure 1. Prevalence of use according to interviews of elderly subjects participating in the French Three-City (3C) cohort study (2001–2003) and to their computerized reimbursement data from the French national health insurance system ($n = 4112$). For all drug classes, the proportion of users of at least one drug of the class was determined using 3C interviews, reimbursements corresponding to the 30 days preceding the interview and reimbursements corresponding to the 60 days preceding the interview (reference date)

between 0.03 for genito-urinary system drugs for men and 0.59 for antineoplastics and immunomodulating agents (Figure 2). Similarly, Se of reimbursement data limited to 30 days before the interview was low ranging from 2.4% for genito-urinary system drugs for men to 47.4% for anti-infective and antiparasitic drugs (Figure 3). Conversely, Sp was almost perfect for all drug categories (89.1–99.9%).

Sixty-day time frame. Agreement between sources was higher when the time frame used for reimbursement data was extended to 60 days before the interview and κ values increased considerably for almost all drug categories (Figure 2). However, agreement remained moderate for cardiovascular system drugs (κ value = 0.57) and alimentary tract and metabolism drugs (κ value = 0.56) and poor for other drug classes such as genito-urinary system drugs for men,

homeopathy and herbal medicines or anti-infective and antiparasitic drugs (κ values between 0.05 and 0.20). As the same, Se of reimbursement data corresponding to the 60 days before the interview was much higher (Figure 3). Values of at least 70% were observed for many drug classes but Se remained moderate for alimentary tract and metabolism drugs, dermatologicals, respiratory system drugs and anti-infective and antiparasitic drugs (50.3–66.7%). Homeopathy and herbal medicines and genito-urinary system drugs for men were characterized by a particularly low Se (10.8 and 3.5% respectively). Sp was always high ranging from 77.6 to 99.8%.

Since the study included a sufficient number of subjects, agreement between 60-day reimbursement data and interviews and validity of 60-day reimbursement data with reference to interviews could be assessed for sub-categories of drugs in some of these classes. Among cardiovascular system or alimentary

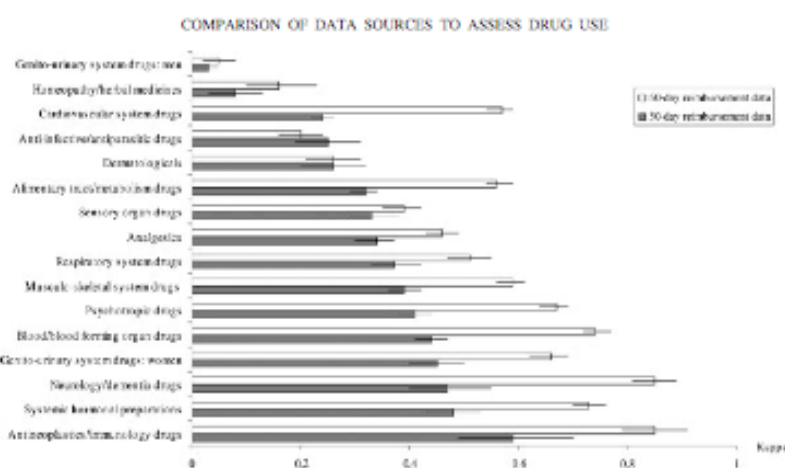


Figure 2. Agreement between reimbursement data and interview data of 4112 elderly subjects participating in the French Three-City (3C) cohort study (2001–2003). Coefficients κ and their 95% confidence intervals were calculated for each drug class using two reimbursement periods: 30 or 60 days before the date of interview (reference date)

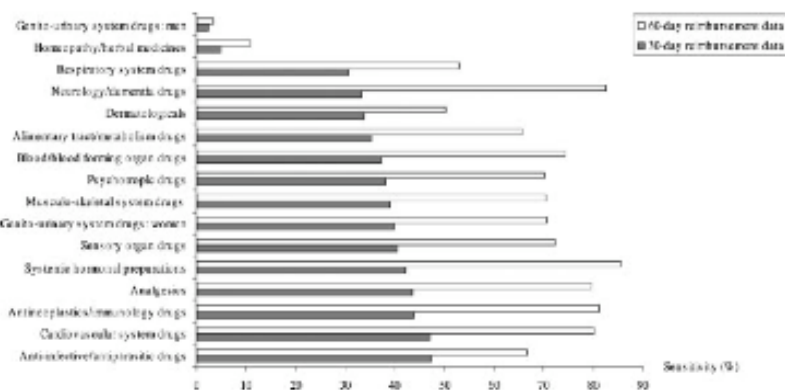


Figure 3. Sensitivity (Se) of reimbursement data in assessing drug use for 4112 elderly subjects participating in the French Three-City (3C) cohort study (2001–2003). Se of reimbursement data was calculated for each drug class using two reimbursement periods: 30 or 60 days before the date of interview (reference date), taking interview data collected as part of the 3C Study at the reference date as 'gold standard'

tract and metabolism drugs, substantial results were obtained regarding all parameters (κ coefficient, Se, Sp, PPV and NPV) for anti-arrhythmics, cardiac glycosides, vasodilators used in cardiac diseases, lipid lowering agents, antihypertensives, drugs used in diabetes, mineral supplements and drugs for acid related disorders (Table 2). Many other types of drugs were also concerned: antidepressants, antipsychotics, benzodiazepine derivatives, antitubercular preparations, some antirheumatic products, drugs for treatment of bone diseases, antineoplastics and immunomodulating agents, drugs used in neurological diseases and dementia, blood and blood forming organ drugs,

systemic hormonal preparations, drugs for obstructive airway diseases, genito-urinary system drugs for women and opioid analgesics (Table 2). However, agreement between both sources as well as validity of reimbursement data remained insufficient for a number of drug categories. Some of them showed a substantial Se associated to a poor PPV (e.g. non-steroidal anti-inflammatory drugs, first and second line analgesics, muscle relaxants, sensory organ drugs or anti-infective and antiparasitic drugs) (Table 2). Conversely, other drug categories such as peripheral vasodilators, laxatives, homeopathy and herbal medicines, drugs for functional gastrointestinal disorders, vitamins,

Table 2. Prevalence of use according to interviews (3C) and to 60-day reimbursement data (CNAM-TS), κ coefficients and sensitivity (Se), specificity (Sp), positive and negative predictive values (PPV and NPV) of 60-day reimbursement data with reference to interviews

Drug classes or sub-classes (n = 4112)	3C (%)	CNAM-TS (%)	κ [IC 95%]	Se (%)	Sp (%)	PPV (%)	NPV (%)
Drugs used in diabetes	7.4	7.1	0.93 [0.91–0.95]	91.1	99.7	95.9	99.3
Anti-arrhythmics	6.2	5.8	0.90 [0.87–0.93]	87.8	99.6	94.1	99.2
Antigout preparations	5.1	4.8	0.89 [0.85–0.92]	86.6	99.6	91.9	99.3
Cardiac glycosides	2.1	2.0	0.88 [0.83–0.94]	85.1	99.9	92.4	99.7
Antihypertensives	55.7	52.0	0.86 [0.84–0.87]	90.3	96.3	96.8	88.8
Vasodilators used in cardiac diseases	8.7	8.3	0.85 [0.82–0.88]	84.0	98.9	88.3	98.5
Lipid lowering agents	35.0	31.8	0.85 [0.84–0.87]	86.2	97.5	94.8	92.9
Antineoplastics/immunology drugs	2.0	1.8	0.85 [0.79–0.91]	81.3	99.8	89.3	99.6
Neurology/dementia drugs	4.5	4.2	0.85 [0.81–0.89]	82.5	99.5	88.4	99.2
Antidepressants	9.0	9.1	0.81 [0.77–0.84]	83.0	98.2	81.9	98.3
Antirheumatic products/chondroitin sulfate	4.1	3.4	0.78 [0.73–0.83]	72.5	99.5	87.2	98.8
Antipsychotics	1.8	1.5	0.76 [0.66–0.84]	69.9	99.8	83.8	99.4
Blood/blood forming organ drugs	26.5	22.3	0.74 [0.72–0.77]	74.3	96.4	88.1	91.2
Drugs for treatment of bone diseases	3.8	3.4	0.73 [0.67–0.79]	70.1	99.2	78.5	98.8
Systemic hormonal preparations	10.3	12.8	0.73 [0.70–0.76]	85.6	95.5	68.6	98.3
Drugs for obstructive airway diseases	4.8	5.4	0.72 [0.67–0.77]	78.1	98.3	97.7	83.1
Benzodiazepine derivatives	27.0	22.0	0.70 [0.67–0.73]	70.1	95.9	86.2	89.6
Mineral supplements	15.5	12.6	0.70 [0.67–0.74]	67.4	97.6	83.5	94.2
Drugs for acid related disorders	15.3	19.0	0.69 [0.66–0.72]	83.6	92.7	67.3	96.9
Genito-urinary system drugs: women (n = 2662)	17.3	16.6	0.67 [0.63–0.70]	70.7	94.7	73.7	93.9
Opioid analgesics	1.4	1.9	0.64 [0.55–0.74]	76.8	99.2	56.5	99.7
Non-steroidal anti-inflammatory drugs	12.8	18.8	0.57 [0.54–0.61]	78.7	90.0	53.5	96.6
Anilide and opioid analgesics in combination	8.3	13.1	0.54 [0.50–0.58]	75.6	92.6	48.0	97.7
Antihistamines for systemic use	3.8	5.4	0.52 [0.45–0.58]	65.0	97.0	45.8	98.6
Muscle relaxants	1.8	3.4	0.48 [0.40–0.56]	71.2	97.9	38.0	99.5
Peripheral vasodilators	32.1	11.2	0.39 [0.36–0.42]	33.1	99.1	94.8	75.8
Sensory organ drugs	8.7	18.8	0.39 [0.35–0.42]	72.5	86.3	33.6	97.1
Salicylates and anilide analgesics	16.8	27.0	0.38 [0.34–0.41]	66.0	80.9	41.0	92.2
Dermatologicals	4.7	10.6	0.26 [0.21–0.31]	50.3	91.3	22.3	97.4
Anti-infective/antiparasitic drugs	2.8	13.1	0.20 [0.16–0.24]	66.7	88.4	14.3	98.9
Laxatives	7.6	1.9	0.16 [0.11–0.21]	11.5	98.9	47.2	93.1
Homeopathy/herbal medicines	4.9	1.1	0.16 [0.10–0.23]	10.8	99.4	48.7	95.6
Drugs for functional gastrointestinal disorders	7.7	0.9	0.11 [0.06–0.15]	6.9	99.6	57.9	92.8
Vitamins	5.3	0.5	0.09 [0.04–0.14]	5.5	99.8	57.1	95.0
Genito-urinary system drugs: men (n = 1450)	19.9	1.0	0.05 [0.02–0.08]	3.5	99.6	66.6	80.6
Psychostimulants	2.4	<0.1	0.02 [–0.02–0.06]	1.0	100.0	49.9	97.6
Vasculoprotectives	22.0	<0.1	0.002 [–0.002–0.005]	0.1	100.0	100.0	78.0

psychostimulants or vasculoprotectives were characterized by a particularly low Se in comparison to the PPV (Table 2).

DISCUSSION

This study shows that drug exposure presumed from reimbursement databases compared to drug exposure collected from patient interviews varies according to drug classes and to the time frame considered for reimbursement data.

For this study, some limits can be addressed in terms of selection or information bias. As for all epidemiological studies, the representativity of the study population with regard to the general population is questionable. On the one hand, the 3C Study was conducted in three different geographical areas in France using electoral rolls to randomly select

participants. Such a methodology limits representativity biases but, on the other hand, the present study only concerned the 3C participants for whom reimbursement data were available from the CNAM-TS. Non-salaried subjects, who are affiliated to another health care systems (i.e. the specific health care systems for farmers or independent workers), were not included in this study. Even if there was no difference regarding age, gender, health status and drug use between these non-included subjects and the study sample, it remains difficult to know how the comparison of drug exposure would have changed if they had also been considered in the study. Another limit might be raised concerning reimbursement data which only concerned drugs both prescribed and approved for reimbursement. OTC drugs are not registered as well as non-reimbursed drugs which may appear in the interview data and, for technical reasons, could not be removed. At the time of the study (2001–2003), a large majority of the drugs

available in France were listed among those approved for reimbursement. Moreover, most of the drugs frequently involved in pharmacoepidemiological research (e.g. benzodiazepine derivatives, cardiovascular system drugs) are reimbursed by health care systems. The recent revision of the reimbursement list with an aim to limit health care expenditure has thus probably not affected the generalisability of the present study. Regarding the form of the questionnaire used in the 3C cohort to assess drug consumption, 15 drugs at the maximum could be listed while there was no limit in the number of drugs extracted from CNAM-TS reimbursement databases. However, only a handful of subjects had more than 15 drugs based on reimbursement data (respectively 0.02 and 1.1% for the 30-day and the 60-day time frame). The number of subjects whose declared drug use at interview could not be confirmed by inspection of drug packages and prescription forms is unfortunately unknown but assumed to be low. Elderly persons usually take an interest in their drug treatments and only a few subjects in this study population are likely to have lost or forgotten both drug packages and prescription forms for the interview. Theoretically, the method applied in the 3C cohort to validate the subjects' reporting of drug use can not totally avoid information biases. Nevertheless, given the specific context of this study, it seems appropriate to consider a carefully designed interview with only 30 days of recall as a reference for drug exposure assessment *versus* a list of reimbursements from a database that has never been really investigated for pharmacoepidemiological purposes. Furthermore, as discussed below, the results obtained with either comparison method (agreement between reimbursement data and interviews and validity of reimbursement data with reference to interviews) lead to equivalent interpretations.

Drugs involved in the treatment of chronic diseases of elderly subjects presented similar prevalences of use according to reimbursement data and interviews. For these drugs, agreement between both sources and validity of reimbursement data with reference to interviews were found to be substantial. These chronically used drugs are, on the one hand, easily remembered by the subjects and thus correctly reported at interview and, on the other hand, regularly reimbursed by health care systems therefore prone to be identified in reimbursement databases. However, it can be noticed that an almost perfect agreement according to the Landis and Koch criteria²⁰ (i.e. κ value greater than 0.80) was obtained for drugs requiring a particularly good compliance and often used daily (see Table 2). Indeed, in conditions such as diabetes, heart

rhythm disorders, gout, cardiac failure, angina, neoplasms or Parkinson's disease, compliance to drug treatments is needed to both improve clinical symptoms and limit disease worsening. This comment does not totally apply to risk factors such as hypertension and dyslipidemia though an almost perfect agreement was also observed for antihypertensives and lipid lowering agents. In the elderly, these drugs are often taken as part of a global cardiovascular treatment. Thus, compliance to antihypertensives and lipid lowering agents may be beneficially influenced by compliance to other cardiovascular system drugs. Moreover, certain drugs in the antihypertensives category are indicated for cardiac failure or heart rhythm disorders (e.g. certain angiotensin converting enzyme inhibitors or beta blocking agents).

Among the drug classes characterized by a poor agreement between both sources and also a poor validity of reimbursement data, two groups may be distinguished. Firstly, drugs that were less prevalent according to interviews than to reimbursement data especially with the 60-day time frame: non-steroidal anti-inflammatory drugs, anti-infective and antiparasitic drugs, first and second line analgesics, dermatologicals and sensory organ drugs. Secondly, drugs that were more prevalent according to interviews such as laxatives, vitamins, drugs for functional gastrointestinal disorders, vasculoprotectives, peripheral vasodilators, psychostimulants or homeopathy and herbal medicines. For the first category, the way the question on drug use was asked during the interview may explain part of the lower prevalence but this could also be due to subjects forgetting to report drugs used shortly or intermittently (i.e. treatment of acute conditions or on a 'take on demand' basis). With regard to drugs used topically, these are often not perceived as medicines³ and this may have contributed to under-reporting of dermatologicals and sensory organ drugs. For such drugs, Se of reimbursement data with reference to interviews was not so low. Poor validity was especially related to a poor PPV herself explained by the lower prevalence of use from the 'gold standard'. Regarding exposure to such drugs, the consistency of interview data as 'gold standard' can be discussed and reimbursement data could be more informative. However, it is also possible that they were bought then reimbursed but finally not used within the time frame considered in the interview which remains a typical limitation of reimbursement databases. For the second category, reimbursement data could not detect use of OTC drugs nor of drugs not reimbursed by the health insurance system. Additionally, detection failure may concern drugs used in

self-medication but bought before the 60 days considered in the reimbursement databases.

With regard to the time frame investigated in the reimbursement databases, the 30-day period was initially chosen because subjects were questioned about their drug consumption over the month preceding the interview. It was also adapted to French regulations since for most prescribed drugs only 1 month's treatment may be delivered and reimbursed. Our results show that the choice of this time frame also depends on the type of drug: for drugs with a chronic indication, a too short period may underestimate their use during the period of interest and conversely, for drugs used shortly or intermittently, a too long period could lead to an overestimation. For chronic use, the underestimation could be due to non-compliance where subjects stock medications thus delay refill and reimbursement but can also be related to the fact that subjects in France often seek a repeat prescription before finishing their monthly treatment leading to the possibility of use reported at the interview without reimbursement during the 30 preceding days. For drugs used intermittently, considering longer periods than the one of interest increases the probability of detecting reimbursements irrelevant to this period.

Our findings are consistent with earlier studies comparing drug exposure data from home inventories or questionnaires to pharmacy or reimbursement claim databases. When these studies concerned global drug exposure, agreement between both sources of data or Se of one source with reference to the other (chosen as 'gold standard') was always better for chronically used drugs especially cardiovascular system drugs than for drugs taken intermittently or as needed such as analgesics.¹⁵⁻¹⁷ Since agreement between both sources and validity of reimbursement data with reference to interviews could be tested for many drugs within a same class, our findings are also consistent with studies that have focused on a particular drug class.^{13,14,22,23} For example, among cardiovascular system drugs, agreement between pharmacy data and a home inventory was found almost perfect for beta blockers and very poor for ointments against haemorrhoids in a study from Sjahid *et al.*²³ The determination of the adequate time frame to consider data from computerized databases has also been already discussed. In the same way as in our study, Lau *et al.* determined the Se of pharmacy data in comparison to a home inventory using two fixed time windows of 30 and 90 days.¹⁷ Concerning antibiotics, they observed a certainly better Se with the longest period of 90 days but also significantly more false positive exposures leading to the same conclusion that this time window may be

KEY POINTS

- Conducting pharmacoepidemiological studies requires valid and complete information on drug exposure.
- Drug use as assessed from reimbursement data of the French national health insurance system was compared with drug use as declared during interviews of elderly subjects participating in the Three-City (3C) cohort.
- Using an appropriate time frame in the reimbursement databases, reimbursement data and interviews estimate exposure to chronically used drugs similarly. Self-medication seems to be better described with interviews whereas reimbursement data could be more useful for drugs used topically or intermittently.
- The choice between patient interviews and reimbursement databases to assess drug use for pharmacoepidemiological purposes depends on the type of drug studied.

too long for drugs used for a relatively short period like antibiotics. Van Staa *et al.* showed that the accuracy of the risk estimates attributed to a drug in pharmacoepidemiological studies can be influenced by the time window used to define drug exposure from automated databases.²⁴

To conclude, the choice between patient interviews and administrative databases to assess drug use for pharmacoepidemiological purposes must take into account the drug's main characteristics (e.g. prescription or OTC drug, reimbursement status, intermittent or chronic use, route of administration, compliance by subjects). From our findings, CNAM-TS reimbursement data could be used to measure drug exposure if attention is paid to the definition of an appropriate time window with regard to the drug studied. This study also confirmed the interest of patient interviews to assess exposure to drugs used chronically or by self-medication as well as their possible limits concerning drugs taken intermittently or topically.

Partnership and supports

The Three-City (3C) Study was conducted under a partnership agreement between the Institut National de la Santé et de la Recherche Médicale (INSERM), the Victor Segalen – Bordeaux 2 University and the Sanofi-Synthelabo Company. The Fondation pour la Recherche Médicale funded the preparation and initiation of the study. The 3C Study was also supported by the Caisse Nationale d'Assurance Maladie des Travailleurs Salariés, Direction Générale de la Santé, Conseils Régionaux of Aquitaine,

Languedoc-Roussillon and Bourgogne, Fondation de France, Ministry of Research-INSERM Programme 'Cohortes et collections de données biologiques', Mutuelle Générale de l'Éducation Nationale, Institut de la Longévité, Conseil Général de la Côte d'Or. The present study did not receive any specific funding.

ACKNOWLEDGEMENTS

The authors thank Annick Alperovitch and Philip Robinson for their valuable comments.

REFERENCES

- West SL, Strom BL, Poole C. Validity of pharmacoepidemiologic drug and diagnosis data. In *Pharmacoepidemiology*, Strom BL (ed.). John Wiley & Sons: Chichester, 2005; 709–765.
- Horwitz RJ, Feinstein AR, Stremlau JR. Alternative data sources and discrepant results in case-control studies of estrogens and endometrial cancer. *Am J Epidemiol* 1980; **111**: 389–394.
- Moore N, Masson H, Noblet C, Joannides R. What medicines do patients really take? A comparison of free form vs. oriented questionnaires. *Post Mark Surveill* 1993; **7**: 355–362.
- Van den Brandt PA, Petri H, Dorant E, Goldbohm RA, Van de Crommert S. Comparison of questionnaire information and pharmacy data on drug use. *Pharm Weekbl Sci* 1991; **13**: 91–96.
- West SL, Savitz DA, Koch G, Strom BL, Guess HA, Hartzema A. Recall accuracy for prescription medications: self-report compared with database information. *Am J Epidemiol* 1995; **142**: 1103–1112.
- Goodman MT, Nomura AM, Wilkens LR, Kolonel LN. Agreement between interview information and physician records on history of menopausal estrogen use. *Am J Epidemiol* 1990; **131**: 815–825.
- Klungel OH, de Boer A, Paes AH, Herings RM, Seidell JC, Bakker A. Influence of question structure on the recall of self-reported drug use. *J Clin Epidemiol* 2000; **53**: 273–277.
- Metlay JP, Hardy C, Strom BL. Agreement between patient self-report and a Veterans Affairs national pharmacy database for identifying recent exposures to antibiotics. *Pharmacoepidemiol Drug Saf* 2003; **12**: 9–15. DOI: 10.1002/pds.772
- Paganini-Hill A, Ross RK. Reliability of recall of drug usage and other health-related information. *Am J Epidemiol* 1982; **116**: 114–122.
- Harlow SD, Linet MS. Agreement between questionnaire data and medical records. The evidence for accuracy of recall. *Am J Epidemiol* 1989; **129**: 233–248.
- Moore N, Pierfite C, Pehourcq F, Lagnaoui R, Bégaud B. Comparison of patient questionnaires, medical records, and plasma assays in assessing exposure to benzodiazepines in elderly subjects. *Clin Pharmacol Ther* 2001; **69**: 445–450. DOI: 10.1067/mcp.2001.116147
- Depont F, Fourrier A, Merlière Y, et al. The CADEUS study: methods and logistics. *Pharmacoepidemiol Drug Saf* 2007; **16**: 571–580. DOI: 10.1002/pds.1348
- Christensen DB, Williams B, Goldberg HL, Martin DP, Engelberg R, LoGerfo JP. Comparison of prescription and medical records in reflecting patient antihypertensive drug therapy. *Ann Pharmacother* 1994; **28**: 99–104.
- McKenzie DA, Semradek J, McFarland BH, Mullooly JP, Mc Camant LE. The validity of Medicaid pharmacy claims for estimating drug use among elderly nursing home residents: The Oregon experience. *J Clin Epidemiol* 2000; **53**: 1248–1257.
- Johnson RE, Vollmer WM. Comparing sources of drug data about the elderly. *J Am Geriatr Soc* 1991; **39**: 1079–1084.
- Monster TB, Janssen WM, de Jong PE, de Jong-van den Berg LT. For the PREVENT Study Group. Pharmacy data in epidemiological studies: an easy to obtain and reliable tool. *Pharmacoepidemiol Drug Saf* 2002; **11**: 379–384. DOI: 10.1002/pds.722
- Lau HS, de Boer A, Beuning KS, Porsius A. Validation of pharmacy records in drug exposure assessment. *J Clin Epidemiol* 1997; **50**: 619–625.
- The 3C Study Group. Vascular factors and risk of dementia: design of the Three-City Study and baseline characteristics of the study population. *Neuroepidemiology* 2003; **22**: 316–325. DOI: 10.1159/000072920
- World Health Organization. Collaborating Centre for Drug Statistics Methodology. ATC Index with DDDs. World Health Organization, Oslo, 2000.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**: 159–174.
- Weiss NS. Studies of diagnostic and screening tests. In *Modern Epidemiology*, Rothman KJ, Greenland S (eds). Lippincott-Raven: Philadelphia, 1998; 520–524.
- Klungel OH, de Boer A, Paes AH, Herings RM, Seidell JC, Bakker A. Agreement between self-reported antihypertensive drug use and pharmacy records in a population-based study in The Netherlands. *Pharm World Sci* 1999; **21**: 217–220.
- Sjahid SI, Van der Linden PD, Stricker BH. Agreement between the pharmacy medication history and patient interview for cardiovascular drugs: the Rotterdam elderly study. *Br J Clin Pharmacol* 1998; **45**: 591–595.
- Van Staa TP, Abenhaim L, Leufkens H. A study of the effects of exposure misclassification due to the time-window design in pharmacoepidemiologic studies. *J Clin Epidemiol* 1994; **47**: 183–189.

APPENDIX - CLASSIFICATION OF DRUGS TESTED IN THE STUDY

Drug classes or sub-classes	Corresponding ATC codes
Alimentary tract/metabolism drugs	A
Drugs for acid related disorders	A02
Drugs for functional gastrointestinal disorders	A03
Laxatives	A06
Drugs used in diabetes	A10
Vitamins	A11
Mineral supplements	A12
Analgesics	N02
Opioid analgesics	N02A
Salicylates and anilide analgesics	N02BA/N02BE
Anilide and opioid analgesics in combination	N02AA59/N02AC54/N02BE51
Anti-infective/antiparasitic drugs	J/P
Antineoplastics/immunology drugs	L
Blood/blood forming organ drugs	B
Cardiovascular system drugs	C
Cardiac glycosides	C01A
Anti-arrhythmics	C01B
Vasodilators used in cardiac diseases	C01D
Antihypertensives	C02/C03/C07/C08/C09
Peripheral vasodilators	C04/C01EB15
Vasculoprotectives	C05
Lipid lowering agents	C10
Dermatologicals	D
Genito-urinary system drugs	G
Homeopathy/herbal medicines ^a	
Musculo-skeletal system drugs	M
Non-steroidal anti-inflammatory drugs	M01A
Antirheumatic products/chondroitin sulfate ^b	M01B/M01C/M09AX02
Muscle relaxants	M03
Antigout preparations	M04
Drugs for treatment of bone diseases	M05
Neurology/dementia drugs	N03/N04/N06D
Psychotropic drugs	N05/N06A/N06B/N06C
Antipsychotics	N05A
Benzodiazepine derivatives	N05BA/N05CD/N05CF
Antidepressants	N06A/N06CA
Psychostimulants	N06B/N06CB
Respiratory system drugs	R
Drugs for obstructive airway diseases	R03
Antihistamines for systemic use	R06
Sensory organ drugs	S
Systemic hormonal preparations	H

ATC, anatomical therapeutic chemical.

^aNo existing ATC code.^bAntirheumatic products other than non-steroidal anti-inflammatory drugs.

En plus de leur intérêt certain, mis en évidence à travers ce premier article, pour la mesure transversale de l'exposition à de nombreuses classes médicamenteuses, les données de remboursement de l'Assurance Maladie demeurent surtout avantageuses du fait de leur caractère continu au cours du temps. Par rapport aux données déclaratives, elles présentent donc un potentiel pour la mesure de l'exposition médicamenteuse au cours de périodes de temps plus longues.^{4,16}

III.3- DESCRIPTION ET COMPARAISON DE L'EXPOSITION MÉDICAMENTEUSE MESURÉE À PARTIR DE DONNÉES DÉCLARATIVES OU DE REMBOURSEMENT AU COURS D'UNE PÉRIODE DONNÉE (ARTICLES N°2 ET N°3)

Les travaux conduits dans ce second contexte de la mesure de l'exposition médicamenteuse au cours d'une période donnée sont présentés sous la forme de deux articles scientifiques actuellement en cours de soumission.

III.3.1- Présentation des articles

Noize P, Bazin F, Pariente A, *et al.* Description of drug use patterns from health insurance claims as part of a repeated cross-sectional cohort: the Three-City (3C) Study (article n°2)

Noize P, Bazin F, Pariente A, *et al.* Validity of drug exposure measured from repeated cross-sectional patient interviews: data from the Three-City (3C) Study (article n°3)

III.3.1.1- Contexte et objectifs

Pour une période de plusieurs mois à plusieurs années, l'obtention d'informations concernant la consommation médicamenteuse tout au long de la période via l'interrogatoire des sujets suscite de nombreux problèmes méthodologiques. Si la faisabilité d'un recueil prospectif « au jour, le jour » est compromise, la validité de données collectées rétrospectivement reste discutable du fait des biais de mémorisation.¹¹ Dans les cohortes telles que 3C ou encore Paquid⁷², le recueil prospectif de données à intervalles de temps réguliers représente une méthodologie « intermédiaire ». Les mesures transversales répétées de la consommation médicamenteuse peuvent, en effet, permettre d'étudier son évolution au cours du temps ainsi que les facteurs associés à cette évolution.¹⁴⁰⁻¹⁴³ Toutefois, compte-tenu de la longueur des intervalles de temps séparant deux interrogatoires (plusieurs mois à plusieurs années), ni les profils d'utilisation, ni l'exposition à proprement parler ne peuvent être mesurés précisément au cours de ces intervalles uniquement au moyen de données déclaratives recueillies au début et à la fin de l'intervalle.

Comme l'illustre la Figure 3, les objectifs étaient donc, chez des sujets inclus dans la cohorte 3C :

- de décrire, à partir des données de remboursement, les profils d'utilisation de plusieurs classes médicamenteuses au cours de la période entre la première visite de suivi (S1) et la seconde visite de suivi (S2) et ce, en fonction de la déclaration de consommation de ces classes à chacune des deux visites (article n°2) et,
- de comparer l'exposition à plusieurs classes médicamenteuses mesurée à partir des données déclaratives ou des données de remboursement pour la période comprise entre les deux visites de suivi (article n°3).

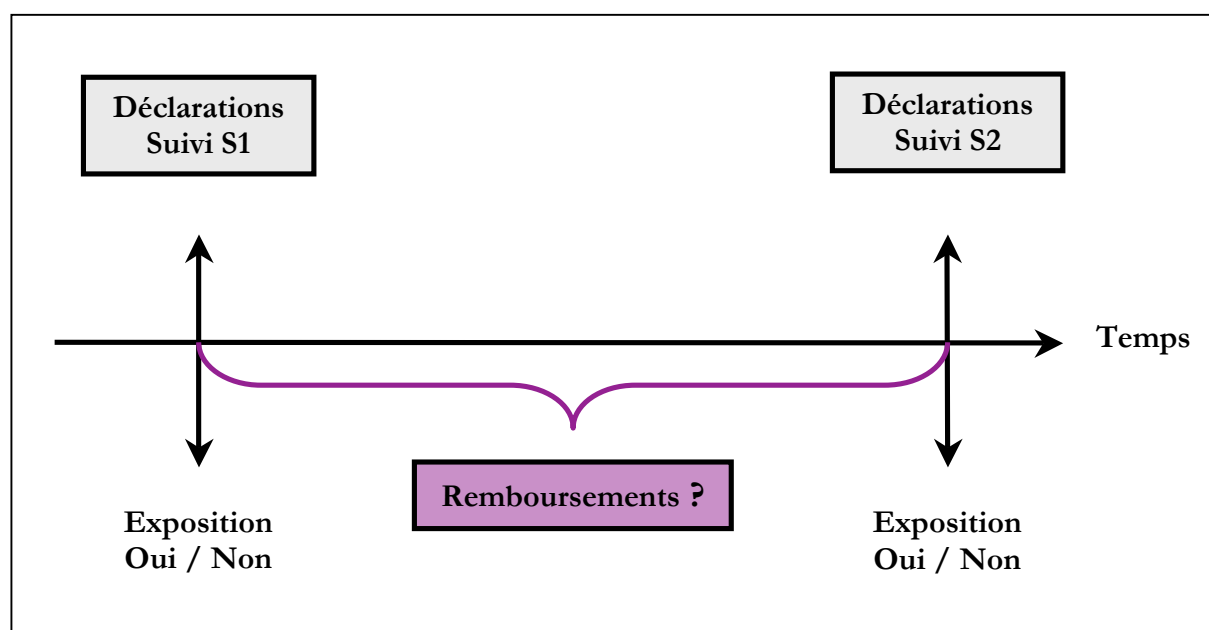


Figure 3. Illustration de la mesure de l'exposition médicamenteuse entre les deux visites de suivi de la cohorte 3C d'après les données déclaratives et de remboursement

III.3.1.2- Considérations méthodologiques

III.3.1.2.1- Échantillon d'étude

Les travaux présentés dans les articles n°2 et n°3 ont porté sur un échantillon de 2 985 sujets (âge médian à l'inclusion : 73 ans ; 65,7 % de femmes) de la population source interrogés aux premier et second suivis de la cohorte et ayant des données de remboursement disponibles de façon continue depuis 90 jours avant la date du premier interrogatoire (cf. § III.3.1.2.3) et jusqu'à la date du second interrogatoire.

III.3.1.2.2- Choix des classes médicamenteuses d'intérêt

Parmi l'ensemble des classes médicamenteuses étudiées dans l'article n°1, seules celles comptant une majorité de médicaments remboursables par l'Assurance Maladie française et pouvant faire l'objet d'une consommation chronique indépendamment de sa régularité dans le temps ont été sélectionnées pour les analyses ultérieures.

Les vasodilatateurs périphériques, laxatifs, vitamines, médicaments homéopathiques ou de phytothérapie, antispasmodiques, médicaments du système génito-urinaire chez les hommes, psychostimulants et vasculoprotecteurs dont la prévalence d'exposition était nettement inférieure selon les données de remboursement (y compris lorsqu'elles étaient considérées dans les 60 jours avant la date du recueil déclaratif) ont donc été éliminés ainsi que les antihistaminiques, médicaments dermatologiques et anti-infectieux supposés être consommés de manière très ponctuelle pour le traitement d'affections aiguës.

Certaines des classes médicamenteuses restantes (notamment celles utilisées dans le traitement du diabète, des maladies rhumatologiques, des néoplasies, des maladies neurologiques ou encore de l'asthme) ont été également exclues du fait d'un effectif trop faible de sujets en déclarant la consommation à au moins un des deux interrogatoires. Dans cette optique, certaines classes ont, par ailleurs, été regroupées comme les digitaliques, anti-arythmiques et dérivés nitrés.

III.3.1.2.3- Période et paramètres utilisés selon les données de remboursement

➤ Période utilisée

À chaque interrogatoire de la cohorte, les sujets devaient déclarer leur consommation de médicaments au cours du mois précédent. La période utilisée pour mesurer la consommation médicamenteuse entre les deux interrogatoires selon les données de remboursement devait donc inclure une certaine fenêtre de temps avant la date du premier de ces interrogatoires. Au vu des résultats présentés dans l'article n°1, cette fenêtre de temps dépend de la classe médicamenteuse. C'est pourquoi un travail préalable a été effectué afin de déterminer, pour chacune d'entre-elles, la fenêtre de temps la plus adaptée (entre 15 et 90 jours) pour considérer les données de remboursement avant la date de la première visite de suivi S1. La méthode et les résultats de ce travail sont décrits en Annexe 3.

➤ Paramètres utilisés

Les profils d'utilisation médicamenteuse au cours de la période comprise entre les deux visites de suivi ont été décrits (article n°2) puis l'exposition médicamenteuse au cours de cette même période définie (article n°3) à l'aide de deux paramètres construits à partir des informations disponibles pour chaque remboursement (date et code ATC du médicament). Il s'agit de la proportion de couverture et du délai entre deux remboursements successifs.

La proportion de couverture permet de refléter une quantité d'exposition au cours de la période. Pour une classe médicamenteuse donnée, elle correspond au rapport (exprimé en %) entre :

- le nombre de remboursements observés entre la date de la première visite de suivi S1 (moins la fenêtre de temps de 15 à 90 jours) et la date de la seconde visite de suivi S2 et,
- le nombre maximal de remboursements attendus pour la période entre les deux visites défini, compte-tenu du système d'Assurance Maladie français, comme le nombre de mois de la période.

Le délai entre deux remboursements successifs permet d'appréhender la régularité de l'exposition au cours de la période.

III.3.1.2.4- Choix de la méthode de comparaison (article n°3)

Contrairement au contexte de l'article n°1, il paraissait ici pertinent de considérer l'extraction de données informatisées des bases de remboursement de l'Assurance Maladie comme méthode de référence (« alloyed gold standard »)^{4,138} par rapport à l'interrogatoire des sujets. Ces données disponibles de manière continue au cours du temps semblaient plus à même de traduire l'exposition médicamenteuse tout au long de la période d'intérêt que des données déclaratives recueillies de manière transversale au début et à la fin de cette même période. Ainsi, pour chaque classe médicamenteuse, la validité de l'exposition définie au cours de la période comprise entre les deux visites de suivi par la déclaration d'une consommation à la première puis à la seconde de ces visites a été évaluée en considérant comme « gold standard » celle mesurée selon les données de remboursement en utilisant les paramètres décrits précédemment (proportion de couverture et délai entre deux remboursements successifs).¹³⁹

III.3.1.3- Principaux résultats

Pour l'ensemble des médicaments du système cardiovasculaire (antihypertenseurs, digitaliques, anti-arythmiques, dérivés nitrés et hypolipidémiants) et les médicaments hématologiques, la proportion de sujets déclarant une consommation aux deux interrogatoires était supérieure à la proportion de sujets ne déclarant une consommation qu'à un seul de ces interrogatoires. La proportion de couverture médiane calculée à partir des données de remboursement pour la période entre les deux interrogatoires excédait 100 % pour les sujets déclarant une consommation aux deux interrogatoires. Elle était comprise entre 34 et 61 % chez les sujets ne déclarant une consommation qu'à un seul des deux interrogatoires mais la médiane du délai moyen entre deux remboursements successifs restait proche de 30 jours chez ces sujets. Pour ces classes médicamenteuses, l'exposition définie au cours de la période entre les deux interrogatoires par la déclaration d'une consommation au premier puis au second de ces interrogatoires présentait une bonne validité en référence à celle mesurée selon les données de remboursement à l'aide de la proportion de couverture au cours de la même période et du délai maximal entre deux remboursements successifs.

À l'inverse, pour les antalgiques (aspirine et paracétamol associé ou non à un antalgique opiacé), suppléments minéraux, AINS, médicaments du système génito-urinaire chez les femmes et médicaments des organes sensoriels, la proportion de sujets déclarant une consommation aux deux interrogatoires était inférieure à la proportion de sujets ne déclarant une consommation qu'à un seul de ces interrogatoires. La proportion de couverture médiane calculée à partir des données de remboursement pour la période entre les deux interrogatoires était comprise entre 66 et 91 % pour les sujets déclarant une consommation aux deux interrogatoires et entre 25 et 38 % pour les sujets ne déclarant une consommation qu'à un seul des deux interrogatoires. Quelle que soit la catégorie de sujets, la médiane du délai moyen entre deux remboursements successifs était, le plus souvent, supérieure à 30 jours et pouvait atteindre 55 jours. Pour ces classes médicamenteuses, la validité de l'exposition définie au cours de la période entre les deux interrogatoires par la déclaration d'une consommation au premier puis au second de ces interrogatoires était relativement mauvaise.

III.3.2- Article n°2

Description of drug use patterns from health insurance claims as part of a repeated cross-sectional cohort: the Three-City (3C) Study

Pernelle Noize Pharm.D, MSc^{1,2,3}, Fabienne Bazin PhD^{1,2}, Antoine Pariente MD, PhD^{1,2,3}, Carole Dufouil PhD^{4,5}, Marie-Laure Ancelin PhD⁶, Jean-François Dartigues MD, PhD^{2,3,7}, Nicholas Moore MD, PhD^{1,2,3}, Annie Fourier-Réglat Pharm.D, PhD^{1,2,3}

¹Inserm U657, Pharmaco-épidémiologie et évaluation de l'impact des produits de santé sur les populations, Bordeaux, France

²Université de Bordeaux, France

³Centre Hospitalier Universitaire de Bordeaux, France

⁴Inserm U708, Neuroépidémiologie, Paris, France

⁵Université Pierre et Marie Curie, Paris, France

⁶Inserm U888, Pathologies du système nerveux: recherche épidémiologique et clinique, Montpellier, France

⁷Inserm U897, Épidémiologie, santé publique et développement, Bordeaux, France

Corresponding author:

Annie Fourier-Réglat

Inserm U657, Département de Pharmacologie, Université de Bordeaux

146 rue Léo Saignat, 33076 Bordeaux Cedex, France

Tel. +33-5-57571561

Fax. +33-5-57574660

E-mail: annie.fourrier@pharmaco.u-bordeaux2.fr

Conflict of interest statement:

The authors have no conflict of interest to declare.

Running head:

Description of drug use from reimbursement claims

Word count:

Abstract: 282 words

Text: 2628 words

ABSTRACT

Background: In repeated cross-sectional cohort studies, patients are interviewed regarding their current drug use at regular intervals, sometimes several years in duration. These interview data do not inform about real patterns of drug use during intervals that were therefore described in the present study using health insurance claims data.

Methods: The 3C Study is a cohort of 9,294 French elderly persons. For the follow-up period (2001-2005), information on drug use was obtained from two cross-sectional interviews conducted at a two-year interval and, for affiliated members, from reimbursement databases of the main French national healthcare insurance system. For each drug class studied, subjects who reported use during at least one follow-up interview were divided into two categories: use reported at both interviews or use reported at only one interview. Within each category, use of a given drug class between the two interviews was described from reimbursement data using the proportion of coverage and mean time between two claims.

Results: This study included 2,985 subjects. Between 65.3 and 86.7% of cardiovascular or hematological drug users reported use at both interviews. For these, median coverage measured from reimbursement data for the period between the two interviews ranged from 100 to 115% and median values of the mean time between claims from 30 to 33 days. For first line analgesics or non-steroidal anti-inflammatory drugs, between 67.0 and 73.7% of users reported use at only one interview, with median coverage ranging from 25 to 30%, and median values of the mean time between claims from 52 to 55 days.

Discussion: This study illustrates the usefulness of reimbursement data from the main French national healthcare insurance system to describe, between two cross-sectional interviews, specific patterns of drug use.

Keywords: Drug utilization; Pharmacoepidemiology; Reimbursement databases; Interview

INTRODUCTION

Studying the association between a drug exposure and the occurrence of an outcome remains a typical focus in pharmacoepidemiology.¹ Precise knowledge about specific drug utilization patterns are required to properly define exposure in such etiological studies. Several sources provide useful data regarding drug use including patient interviews or computerized databases recording prescription, dispensation or reimbursement claims. Information collected through well-conducted interviews may be exhaustive, especially with regard to Over The Counter (OTC) drug use or self-medication² but these are often subject to memory bias.^{3,4} On the other hand, information retrieved from computerized databases are restricted to prescription-only medicines and presuppose that prescribed, dispensed or reimbursed drugs are subsequently taken.⁵ Nevertheless, these are increasingly used, primarily because they totally avoid memory bias and allow an easier access to large datasets.⁶ Moreover, taking advantage of the longitudinal availability of claims for a given patient, these computerized databases may be of great interest, in comparison to interviews, to describe profiles of drug use over long periods of time. Indeed, recall accuracy of self-reported drug use over months or years may be questionable especially for specific information on drug dosage or use frequency^{4,7} and, to collect data prospectively on a day-by-day basis is hardly feasible for any length of time.

In repeated cross-sectional cohorts, such as the French “Three-City” (3C)⁸ or “Personnes âgées quid?” (Paquid)⁹ studies, prospective data collection by interviews conducted at regular intervals represents an intermediate method. For instance, such repeated cross-sectional drug use measurements may be useful to examine changes in drug use over time and factors associated to these changes.^{10,11} However, the length of intervals (i.e. several months or years) precludes precise determination of patterns of use between interviews. As part of the 3C Study, reimbursement claims data were available, in addition to interview data, for subjects affiliated to the main French healthcare insurance system (Caisse Nationale d’Assurance Maladie des Travailleurs Salariés, CNAM-TS). The CNAM-TS covers approximately 85% of the French general population and

69% of those aged 65 years and over for reimbursement of prescribed drugs and other medical expenses.¹² Thus, these reimbursement claims data were used in the present study in order to describe patterns of drug use over the period between two cross-sectional interviews of the 3C cohort.

MATERIALS AND METHODS

Methodology of the Three-City (3C) Study

The 3C Study is a prospective cohort designed to assess the risk of dementia and cognitive decline due to vascular risk factors. Between 1999 and 2001, 9,294 subjects were included in this cohort according to the following criteria: (i) to be at least 65 years old, (ii) to live in the community, (iii) to be registered on the electoral rolls of one of three French cities: Bordeaux, Dijon or Montpellier. Participants were followed for four years by means of two consecutive visits: two years after inclusion (S1: 2001-2003) and four years after inclusion (S2: 2003-2005). At inclusion and each follow-up visit, participants were interviewed by trained psychologists using a standardized questionnaire on socio-demographic characteristics, health status and drug use. The 3C Study protocol was approved by the Kremlin-Bicêtre university hospital ethics committee (Paris, France) and written informed consent was obtained from each participant. The methodology has been described in detail elsewhere.⁸

Drug use data

During 3C interviews, subjects were asked to list all medications, prescribed or OTC, used more than once a week over the preceding month by an open-ended question. A maximum of fifteen drugs could be collected. This list was validated by visual inspection of drug packages and prescription forms that subjects were asked to bring at interview. Drugs were registered using their brand name and coded according to the Anatomical Therapeutic Chemical (ATC) classification.¹³

In addition to interview data, reimbursement claims data were obtained from the main national healthcare insurance system (CNAM-TS) databases. All CNAM-TS affiliated members were identified among the 9,294 subjects included in the 3C cohort. For each identified subject, the list of all drugs reimbursed between 1 January 2001 and 31 December 2005 at the latest was extracted from computerized databases. The ATC code and date corresponding to each drug reimbursement claim were provided. The study timeline is illustrated in Figure 1.

Study population

The two years that elapsed between the first (S1) and the second (S2) 3C follow-up interviews were considered as the period of interest to describe patterns of drug use from health insurance claims data. Thus, the 3C participants interviewed at two and four years of follow-up and having reimbursement claims data available over the period between these two interviews were eligible for the present study.

Statistical analyses

The drug classes studied were chosen to include a majority of drugs reimbursed by the French healthcare insurance system and that can be chronically used whatever the regularity of use. Drug classes for which a minimum of 10% of subjects reported use during at least one of the two follow-up interviews were included in the analyses. The drug classes considered are listed in the Appendix in alphabetical order of ATC code.

Among the subjects who reported use during at least one interview (all users), two categories were identified: (i) subjects who reported use at both interviews and (ii) subjects who reported use at only one of the two interviews. For each drug class studied, the proportion of subjects in each category above was determined.

Patterns of drug use are expected to differ between the user categories defined by self-reported drug use at each interview. For each user category, reimbursement claims were used to

describe presumed patterns of drug use over the period between the two interviews. Coverage was calculated as the ratio between the number of observed claims and the maximal number of expected claims for the period between the two interviews. In France, most drugs are dispensed and reimbursed for a maximum of one month's treatment whatever the actual prescription. A treatment prescribed for six months will thus be dispensed six times and a reimbursement claim will be automatically submitted by the pharmacist at each of the six dispensations. The maximal number of expected claims for the period between the two interviews was therefore estimated as the number of months in the period. For each subject having at least two claims over the period, the regularity of use was assessed by calculating the mean time between two consecutive claims. Within each drug class and each user category, the proportion of coverage and mean time between two consecutive claims were described by the median and interquartile range of their distributions.

All analyses were performed using SAS[®] statistical software version 9.1 (SAS Institute Inc., Cary, North Carolina, USA).

RESULTS

Study population

Among the 9,294 subjects included in the 3C cohort, 4,962 (53.4%) had reimbursement claims data available from the main national healthcare insurance system (CNAM-TS). Reasons for not having reimbursement data might be affiliation with another healthcare insurance system or that claims data were not found (change of name or address...). There was no difference regarding age, gender, health status or drug use between subjects with or without reimbursement data. The two groups significantly differed only with regard to education level (30.4% of subjects with reimbursement data had high school level or above *vs.* 47.4% of subjects without reimbursement data).

Of these 4,962 subjects, 2,985 (60.2%) who were interviewed at two and four years of follow-up, and had reimbursement claims data over the period between these two interviews were considered in the study. At inclusion in the 3C cohort, median age of the study population was 73 years (range: 65 – 94), 65.7% were women and 71.6% reported at least one of the following cardiovascular risk factors or diseases: hypertension, diabetes, hypercholesterolemia, angina, heart rhythm disorders, myocardial infarction or stroke.

Distribution of subjects according to self-reported drug use

Overall, the proportion of subjects who reported use during at least one interview (all users) ranged between 12.6 and 63.6% depending on the drug class. The highest proportions were obtained for antihypertensives (63.6%), lipid lowering agents (41.0%), hematological drugs (34.5%) and benzodiazepines (32.4%). The lowest proportions were found for genito-urinary system drugs for women (14.7%), sensory system drugs (14.7%), second line analgesics (14.1%) and antidepressants (12.6%).

Among all users, the distribution of subjects into the two categories defined by self-reported use at each interview varied according to the drug class (Figure 2). There were between 60.7 and 86.7% of subjects who reported use at both interviews for systemic hormonal preparations (60.7%), benzodiazepines (61.3%), hematological drugs (65.3%), drugs for cardiac therapy (i.e. cardiac glycosides, anti-arrhythmics and vasodilators used in cardiac diseases such as nitrates, 69.4%), lipid lowering agents (74.0%) and antihypertensives (86.7%) while there were between 67.0 and 73.7% of subjects who reported use at only one of the two interviews for non-steroidal anti-inflammatory drugs (NSAIDs, 67.0%), first line analgesics (68.6%), sensory system drugs (70.6%) and second line analgesics (73.7%).

Description of drug use from reimbursement data between the two interviews

For subjects who reported use at both interviews, the median coverage calculated from reimbursement data for the period between the two interviews ranged from 66 to 115% depending on the drug class (Figure 3). It was below 80% only for benzodiazepines (79%), second line analgesics (78%), genito-urinary system drugs for women (74%), mineral supplements (73%) and first line analgesics (66%). As shown in Table 1, the interquartile range of the proportion of coverage was particularly narrow for antihypertensives (103 to 122%) and drugs used in cardiac therapy (92 to 112%) and the widest for first and second line analgesics (respectively 35 to 96% and 40 to 102%) and genito-urinary system drugs for women (35 to 101%). Concerning the regularity of use over the period, the median of the mean time between two consecutive claims varied between 30 and 42 days (Table 1). As expected from the distribution of coverage, the interquartile range of the mean time between claims was narrow for antihypertensives (28 to 32 days) and drugs used in cardiac therapy (29 to 35 days) and broader for first line analgesics (32 to 67 days), mineral supplements (34 to 64 days) and genito-urinary system drugs for women (32 to 62 days).

For subjects who reported use at only one of the two interviews, coverage ranged from 8 to 61% depending on the drug class (Figure 3). Only antihypertensives and lipid lowering agents showed more than 50% coverage between interviews (61 and 54% respectively). Whatever the drug class, the interquartile range of the proportion of coverage was broader than that observed for subjects who reported use at both interviews (Table 2). The median of the mean time between two consecutive claims ranged from 31 to 61 days (Table 2). For antihypertensives, drugs used in cardiac therapy, lipid lowering agents and hematological drugs, the interquartile range remained relatively narrow. It was the widest for benzodiazepines (41 to 121 days).

DISCUSSION

The results presented here show that patterns of drug use described from reimbursement claims data over the period between two cross-sectional interviews in function of self-reported drug use at each interview varies depending on the drug class considered.

Indeed, the results can be discussed according to two main categories of drugs. Firstly, drugs for which the proportion of subjects reporting use at both interviews was higher than the proportion of subjects reporting use at only one interview: cardiovascular drugs (i.e. antihypertensives, cardiac glycosides, anti-arrhythmics, nitrates and lipid lowering agents), hematological drugs, benzodiazepines and systemic hormonal preparations. Secondly, drugs for which the proportion of subjects reporting use at both interviews was lower than the proportion of subjects reporting use at only one interview: first and second line analgesics, mineral supplements, genito-urinary system drugs for women, sensory system drugs, NSAIDs and drugs for acid related disorders.

In the first category, except for benzodiazepines, median coverage was at least 100% for subjects reporting use at both interviews. This proportion was much lower for subjects reporting use at only one interview but the median of the mean time between claims was often near 30 days. Subjects might not be users over the whole period of interest (e.g. subjects initiating treatment between the two interviews) but when they were, use seem relatively regular. For these drugs involved in the treatment of chronic diseases in elderly subjects, to consider subjects reporting use at a cross-sectional interview as users during a long period of time around the interview seems appropriate. However, among these drug classes, an exception can be noted for systemic hormonal preparations: median coverage was high for subjects reporting use at both interviews and very low for subjects reporting use at only one interview. This could be explained by the presence in this large class of drugs taken chronically and regularly such as drugs used in thyroid therapy and of drugs that can be used for the treatment of acute conditions such as

corticosteroids. Regarding benzodiazepines, the proportion of subjects reporting use at both interviews may seem high (nearly 20%) but is not surprising as the elderly in France are known to often be long-term users of such drugs.¹⁴ Nevertheless, median values of both coverage and mean time between claims obtained from reimbursement data agreed with a chronic but irregular use, partly due to non-adherence as previously reported.¹⁵

In the second category, for subjects reporting use at both interviews, median proportions of coverage were lower than those obtained for drugs in the first category and interquartile ranges were also broader. This proportion was low for subjects reporting use at only one interview with median values of the mean time between claims up to 55 days. For analgesics, NSAIDs or hormone replacement therapy (HRT), these results confirm well-known patterns of use such as chronic but irregular use due to non-adherence or intermittent use in function of perceived symptoms.^{16,17} Without specific questions on adherence or frequency of use, cross-sectional interviews may not reflect well the use of such drugs over longer periods of time. Moreover, part of the higher proportion of subjects reporting use at only one interview can also be explained by subjects forgetting to report drugs used over short periods of time or topically⁵ that is another limitation of interviews based on an unique open-ended question. Nonetheless, another limit might be raised concerning reimbursement data which only include drugs both prescribed and approved for reimbursement. At interview, subjects could also report use of OTC drugs not registered in CNAM-TS databases. Since they were not asked to specify whether each drug was prescribed or not, OTC drugs could not be identified among drugs reported as used. However, in an elderly population often regularly followed by physicians (more than 70% of subjects in the study population presented with at least one cardiovascular disease or risk factor at inclusion in the cohort), self-medication with OTC drugs is not expected to occur frequently.¹⁸ This typical limitation of reimbursement databases especially affects drug classes such as NSAIDs or analgesics. For such classes, drug use described from reimbursement data over the period of interest may have been underestimated and should be completed by precise

data on self-medication patterns as reported by subjects. Among all the drug classes in this second category, drugs for acid related disorders are an exception: if median proportion of coverage was low for subjects reporting use at only one interview, it was very high for subjects reporting use at both interviews. As already explained for systemic hormonal preparations, drugs in this class may be used chronically (e.g. for the treatment of severe gastro-oesophageal reflux disease) or intermittently (e.g. for the treatment of acute dyspepsia symptoms or gastroprotection while using NSAIDs).

To conclude, this study illustrates the usefulness of reimbursement data from the French national healthcare insurance system to precisely describe patterns of drug use among this elderly population as well as the variability of such patterns according to each drug class studied.

ACKNOWLEDGEMENTS

The authors wish to thank all the persons who took part in the 3C Study and the CNAM-TS for providing reimbursement data. They also thank Philip Robinson for his valuable comments.

PARTNERSHIP AND SUPPORTS

The Three-City (3C) Study was conducted under a partnership agreement between the Institut National de la Santé et de la Recherche Médicale (INSERM), the Victor Segalen – Bordeaux 2 University and the Sanofi-Synthélabo Company. The Fondation pour la Recherche Médicale funded the preparation and initiation of the study. The 3C Study was also supported by the Caisse Nationale d'Assurance Maladie des Travailleurs Salariés, Direction Générale de la Santé, Conseils Régionaux of Aquitaine, Languedoc-Roussillon and Bourgogne, Fondation de France, Ministry of Research-INSERM Programme 'Cohortes et collections de données biologiques', Mutuelle Générale de l'Education Nationale, Institut de la Longévité, Conseil Général de la Côte d'Or. The present study did not receive any specific funding.

REFERENCES

- 1- Strom BL. *Pharmacoepidemiology*. Strom BL (eds). John Wiley & Sons: Chichester, 2005
- 2- Noize P, Bazin F, Dufouil C, *et al.* Comparison of health insurance claims and patient interviews in assessing drug use: data from the Three-City (3C) Study. *Pharmacoepidemiol Drug Saf* 2009; **18**: 310-319
- 3- Van den Brandt PA, Petri H, Dorant E, *et al.* Comparison of questionnaire information and pharmacy data on drug use. *Pharm Weekbl Sci* 1991; **13**: 91-96
- 4- West SL, Savitz DA, Koch G, *et al.* Recall accuracy for prescription medications: self-report compared with database information. *Am J Epidemiol* 1995; **142**: 1103-1112
- 5- Moore N, Masson H, Noblet C, *et al.* What medicines do patients really take? A comparison of free form *vs.* oriented questionnaires. *Post Marketing Surveillance* 1993; **7**: 355-362
- 6- Lau HS, de Boer A, Beuning KS, *et al.* Validation of pharmacy records in drug exposure assessment. *J Clin Epidemiol* 1997; **50**: 619-625
- 7- Paganini-Hill A, Ross RK. Reliability of recall of drug usage and other health-related information. *Am J Epidemiol* 1982; **116**: 114-122
- 8- 3C study group. Vascular factors and risk of dementia: design of the Three-City Study and baseline characteristics of the study population. *Neuroepidemiology* 2003; **22**: 316-325
- 9- Dartigues JF, Gagnon M, Barberger-Gateau P, *et al.* The Paquid epidemiological program on brain ageing. *Neuroepidemiology* 1992; **11**: 14-18
- 10- Fourrier A, Letenneur L, Dartigues JF, *et al.* Benzodiazepine use in an elderly community-dwelling population. Characteristics of users and factors associated with subsequent use. *Eur J Clin Pharmacol* 2001; **57**: 419-425
- 11- Soudry A, Dufouil C, Ritchie K, *et al.* Factors associated with changes in antidepressant use in a community-dwelling elderly cohort: the Three-City Study. *Eur J Clin Pharmacol* 2008; **64**: 51-59
- 12- Depont F, Fourrier A, Merlière Y, *et al.* The CADEUS study: methods and logistics. *Pharmacoepidemiol Drug Saf* 2007; **16**: 571-580

- 13- World Health Organization. Collaborating Centre for Drug Statistics Methodology. *ATC Index with DDDs*. World Health Organization, Oslo; 2000
- 14- Lagnaoui R, Depont F, Fourrier A, *et al.* Patterns and correlates of benzodiazepine use in the French general population. *Eur J Clin Pharmacol* 2004; **60**: 523-529
- 15- Lechevallier N, Fourrier A, Berr C. [Benzodiazepine use in the elderly: the EVA Study]. *Rev Epidemiol Sante Publique* 2003; **51**: 317-326
- 16- Moore N, Diris H, Martin K, *et al.* NSAID use profiles derived from reimbursement data in France. *Therapie* 2004; **59**: 541-546
- 17- Serfaty D, de Reilhac P, Eschwege E, *et al.* [Compliance with hormone replacement therapy in menopausal women: results of a two-year prospective French study comparing transdermal treatment with fixed oral combination therapy]. *Gynecol Obstet Fertil* 2003; **31**: 525-533
- 18- Colomes M, Rispail Y, Berlan M, *et al.* [Drug consumption in a retired population]. *Therapie* 1990; **45**: 321-324

FIGURE LEGENDS

Figure 1. Representation of drug use data available from the French national healthcare insurance system (Caisse Nationale d'Assurance Maladie des Travailleurs Salariés, CNAM-TS) and from patient interviews as part of the Three-City (3C) Study (1999-2005). Participants in the 3C cohort were interviewed two years (2001-2003) and four years (2003-2005) after inclusion (1999-2001). In parallel, for the CNAM-TS affiliated members, reimbursement claims data were extracted from computerized databases for the period between 1 January 2001 and 31 December 2003 or 2005 according to the study centre (Bordeaux, Dijon or Montpellier).

Figure 2. Proportion of subjects in the two categories defined by self-reported use of the drug class at each of the 3C follow-up interviews (S1 and S2): reported use at both interviews (S1 + / S2 +) and reported use at only one interview (S1 + / S2 - and S1 - / S2 +).

Figure 3. Description of drug use from reimbursement claims data over the period between the two 3C follow-up interviews (S1 and S2) within the two categories of subjects defined according to self-reported use at each interview: reported use at both interviews (S1 + / S2 +) and reported use at only one interview (S1 + / S2 - and S1 - / S2 +). For each drug class and each category of subjects, the proportion of coverage was calculated as the ratio between the number of observed claims and the maximal number of expected claims for the period between the two interviews (i.e. the number of months in the period).

Figure 1

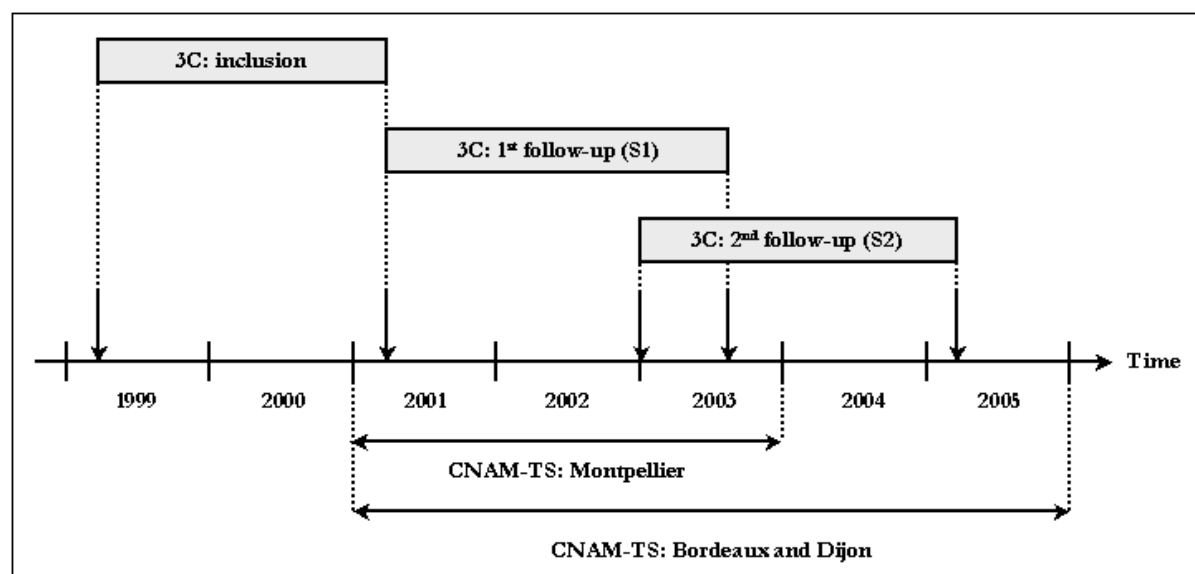


Figure 2

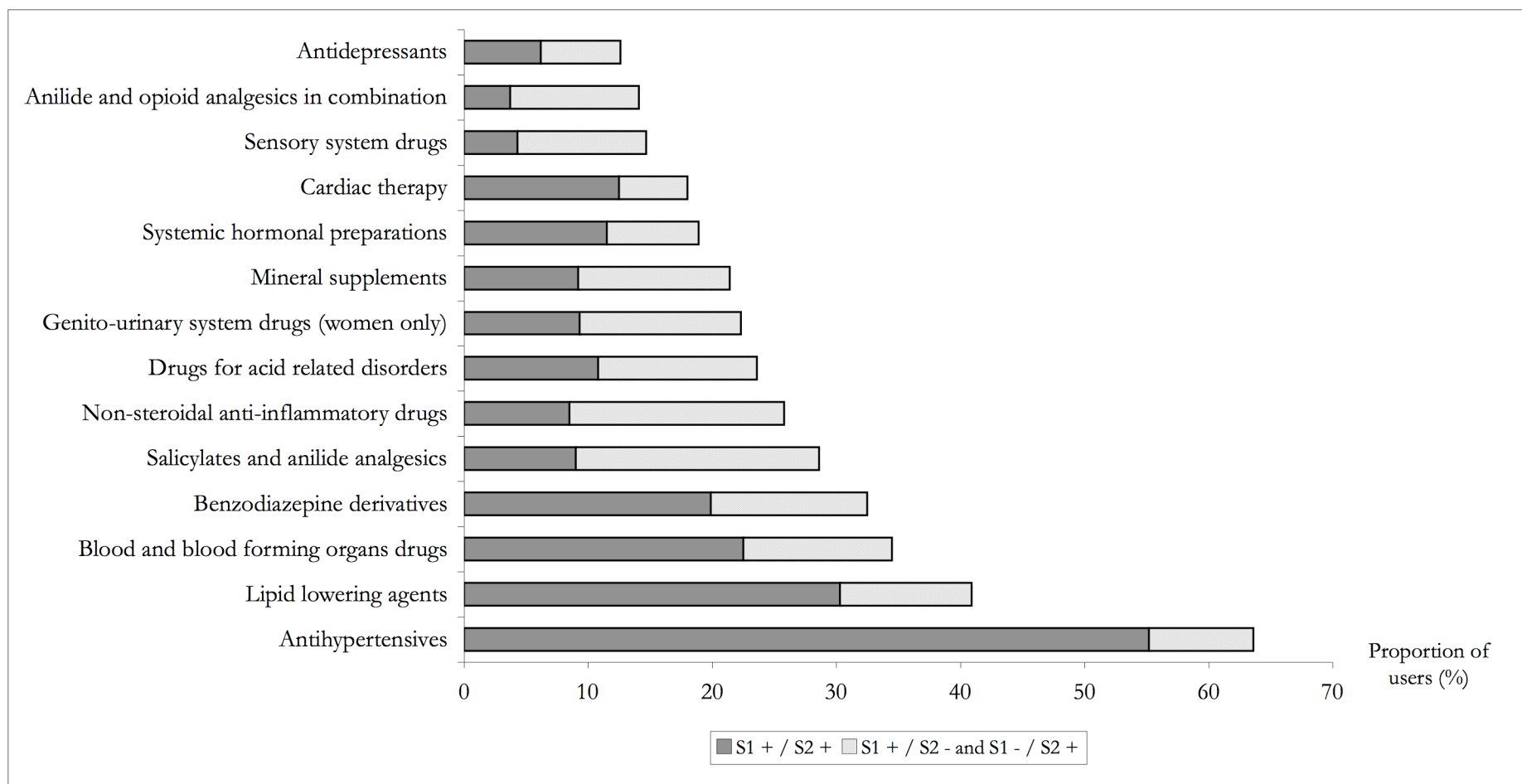


Figure 3

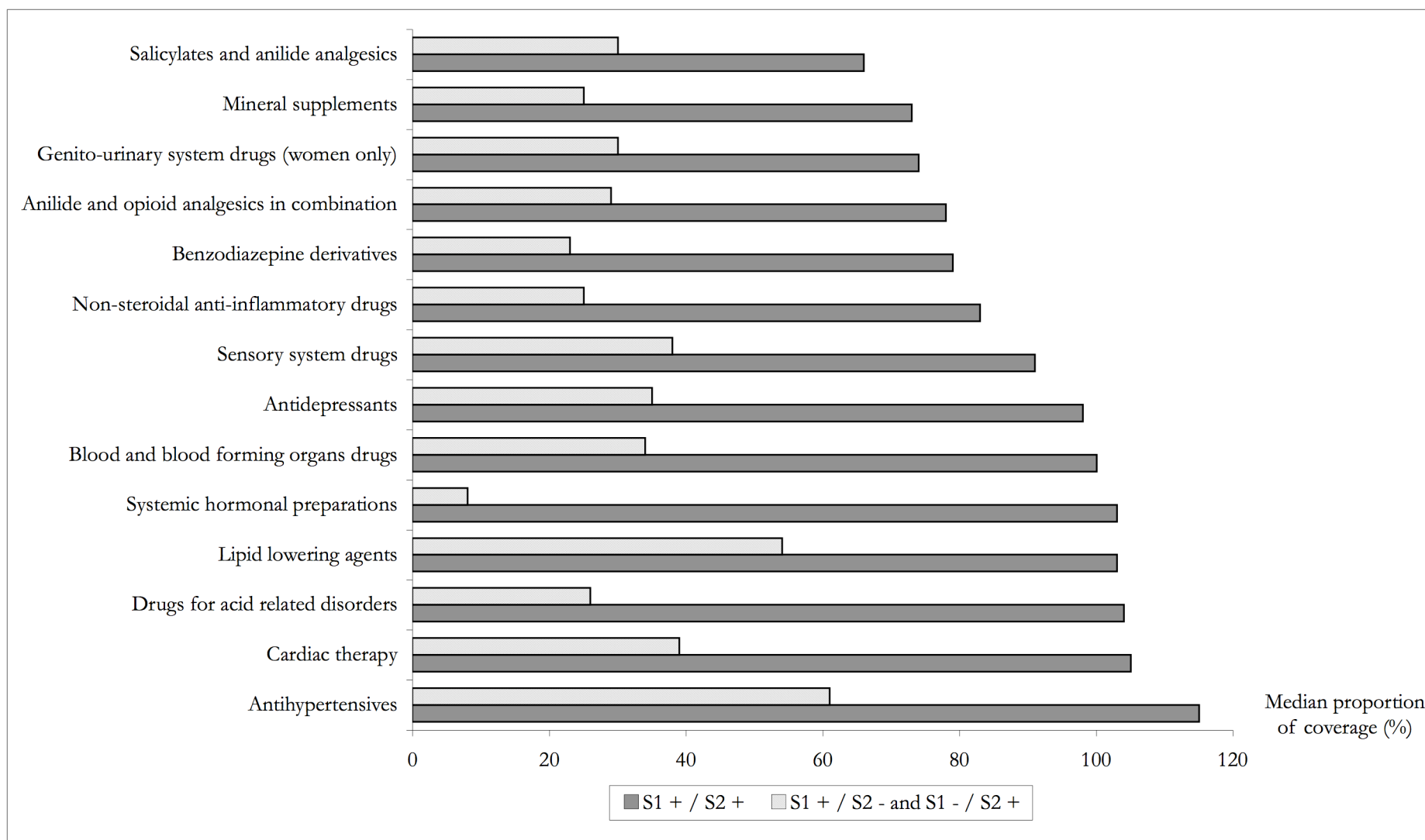


Table 1. Description of drug use from reimbursement claims data over the period between the two 3C follow-up interviews for subjects who reported use at both interviews: proportion of coverage (%) and mean time between two consecutive claims (days)

Drug classes (n = 2 985)	Proportion of coverage (%)	Mean time between claims (days)*
	median [interquartile range]	median [interquartile range]
Antihypertensives	115 [103 ; 122]	30 [28 ; 32]
Cardiac therapy	105 [92 ; 112]	31 [29 ; 35]
Drugs for acid related disorders	104 [87 ; 112]	30 [28 ; 35]
Systemic hormonal preparations	103 [61 ; 114]	30 [28 ; 35]
Lipid lowering agents	103 [75 ; 114]	33 [30 ; 43]
Blood and blood forming organs drugs	100 [67 ; 110]	32 [30 ; 39]
Antidepressants	98 [76 ; 108]	33 [30 ; 41]
Sensory system drugs	91 [51 ; 107]	32 [29 ; 42]
Non-steroidal anti-inflammatory drugs	83 [52 ; 101]	36 [31 ; 50]
Benzodiazepine derivatives	79 [52 ; 104]	41 [32 ; 59]
Anilide and opioid analgesics in combination	78 [40 ; 102]	38 [31 ; 56]
Genito-urinary system drugs (women only) [†]	74 [35 ; 101]	40 [32 ; 62]
Mineral supplements	73 [44 ; 98]	42 [34 ; 64]
Salicylates and anilide analgesics	66 [35 ; 96]	41 [32 ; 67]

*Calculated for the subjects who had at least two claims over the period

[†]n = 1 962

Table 2. Description of drug use from reimbursement claims data over the period between the two 3C follow-up interviews for subjects who reported use at only one interview: proportion of coverage (%) and mean time between two consecutive claims (days)

Drug classes (n = 2 985)	Proportion of coverage (%)	Mean time between claims (days)*
	median [interquartile range]	median [interquartile range]
Antihypertensives	61 [28 ; 91]	31 [28 ; 41]
Cardiac therapy	39 [13 ; 73]	32 [28 ; 47]
Drugs for acid related disorders	26 [9 ; 49]	51 [32 ; 92]
Systemic hormonal preparations	8 [0 ; 28]	50 [31 ; 102]
Lipid lowering agents	54 [26 ; 84]	34 [30 ; 54]
Blood and blood forming organs drugs	34 [11 ; 71]	34 [29 ; 48]
Antidepressants	35 [14 ; 67]	37 [29 ; 56]
Sensory system drugs	38 [12 ; 84]	38 [30 ; 65]
Non-steroidal anti-inflammatory drugs	25 [10 ; 51]	55 [35 ; 107]
Benzodiazepine derivatives	23 [9 ; 54]	61 [41 ; 121]
Anilide and opioid analgesics in combination	29 [11 ; 53]	52 [34 ; 88]
Genito-urinary system drugs (women only) [†]	30 [12 ; 57]	42 [31 ; 67]
Mineral supplements	25 [9 ; 54]	44 [33 ; 64]
Salicylates and anilide analgesics	30 [10 ; 57]	55 [36 ; 98]

*Calculated for the subjects who had at least two claims over the period

[†]n = 1 962

APPENDIX – List of the drug classes considered in the study

Drug classes	Corresponding ATC codes
Drugs for acid related disorders	A02
Mineral supplements	A12
Blood and blood forming organs drugs	B
Cardiac therapy	C01A / C01B / C01D
Antihypertensives	C02 / C03 / C07 / C08 / C09
Lipid lowering agents	C10
Genito-urinary system drugs (women only)	G
Systemic hormonal preparations	H
Non-steroidal anti-inflammatory drugs	M01A
Anilide and opioid analgesics in combination	N02AA59 / N02AC54 / N02BE51 / N02AX52
Salicylates and anilide analgesics	N02BA / N02BE
Benzodiazepine derivatives	N05BA / N05CD / N05CF
Antidepressants	N06A / N06CA
Sensory system drugs	S

ATC, Anatomical Therapeutic Chemical

III.3.3- Article n°3

**Validity of drug exposure measured from repeated cross-sectional patient interviews:
data from the Three-City (3C) Study**

Pernelle Noize Pharm.D, MSc^{1,2,3}, Fabienne Bazin PhD^{1,2}, Antoine Pariente MD, PhD^{1,2,3}, Carole Dufouil PhD^{4,5}, Marie-Laure Ancelin PhD⁶, Jean-François Dartigues MD, PhD^{2,3,7}, Nicholas Moore MD, PhD^{1,2,3}, Annie Fourier-Réglat Pharm.D, PhD^{1,2,3}

¹Inserm U657, Pharmaco-épidémiologie et évaluation de l'impact des produits de santé sur les populations, Bordeaux, France

²Université de Bordeaux, France

³Centre Hospitalier Universitaire de Bordeaux, France

⁴Inserm U708, Neuroépidémiologie, Paris, France

⁵Université Pierre et Marie Curie, Paris, France

⁶Inserm U888, Pathologies du système nerveux: recherche épidémiologique et clinique, Montpellier, France

⁷Inserm U897, Épidémiologie, santé publique et développement, Bordeaux, France

Corresponding author:

Annie Fourier-Réglat

Inserm U657, Département de Pharmacologie, Université de Bordeaux

146 rue Léo Saignat, 33076 Bordeaux Cedex, France

Tel. +33-5-57571561

Fax. +33-5-57574660

E-mail: annie.fourrier@pharmaco.u-bordeaux2.fr

Conflict of interest statement:

The authors have no conflict of interest to declare.

Running head:

Validity of drug exposure measured from interviews

Word count:

Abstract: 284 words

Text: 3890 words

ABSTRACT

Background: In certain cohort studies, data on current drug use are collected by regular interviews sometimes several years apart. Drug exposure between interviews can only be presumed. This study was designed to evaluate the validity of drug exposure presumed from two cross-sectional interviews taking continuous reimbursement data as reference.

Methods: The 3C Study is a cohort of 9,294 French elderly persons. Drug use data were collected by interviews at inclusion, two and four years later. For affiliated members, reimbursement data were obtained from the main national healthcare insurance system. Validity (sensitivity, Se; specificity, Sp; positive predictive value, PPV; negative predictive value, NPV) of exposure presumed from the two follow-up interviews was investigated for different drug classes taking reimbursement data as “gold standard” with several exposure definitions for the period between the two interviews: at least 50 or 80% coverage with or without a maximal time between claims of respectively 90 or 60 days.

Results: A total of 2,985 subjects was considered. For cardiovascular drugs, validity of exposure presumed from both interviews was good: Se 82.1 to 94.1%, Sp 81.2 to 99.3%, PPV 66.9 to 96.5% and NPV 88.1 to 98.7% except, to a certain extent, when the “gold standard” from reimbursement data corresponded to 80% coverage with a maximum of 60 days between claims. For benzodiazepines or analgesics, except when the “gold standard” consisted of a 50% coverage only, validity of exposure presumed from both interviews was questionable, especially regarding PPV (18.2 to 48.3%).

Discussion: Using reported use at two cross-sectional interviews as a proxy for continuous exposure over the period between interviews was found valid for chronically and regularly used drugs but highly approximative for drugs used more irregularly or intermittently.

Keywords: Medicines; Pharmacoepidemiology; Reimbursement databases; Interview; Validity (epidemiology)

INTRODUCTION

In pharmacoepidemiology, determining drug exposure is fundamental, particularly in the framework of etiological studies designed to evaluate the association between a drug exposure and an outcome occurrence.¹ In such studies, errors on exposure measurement may affect point estimates of associations by introducing misclassification bias that may potentially lead to erroneous conclusions.² There are several sources of data available for drug exposure assessment in pharmacoepidemiological research including patient interviews and computerized databases recording reimbursement claims but all are limited in a certain way in their ability to reflect true patient exposure. Whereas databases remain attractive because they totally avoid memory bias and allow an easier accessibility to large datasets, they suffer from a lack of data regarding Over The Counter (OTC) drugs and also that reimbursement does not necessarily mean use.^{3,4} Information obtained through well-conducted patient interviews may be exhaustive with regard to use of OTC drugs or self-medication⁵, yet when collecting data about drug exposure over long periods of time (i.e. several months or years), these are hampered by recall bias.⁶⁻⁸ In this respect, the method of repeated cross-sectional interviews employed in certain cohort studies^{9,10} may be of interest. From several cross-sectional measurements of current drug use (often including a short period preceding interview), hypotheses can be made concerning exposure during the periods between interviews.¹¹ This is especially the case for drugs that can be used chronically: when subjects report their use at a first interview and again several months later at a second interview then they may be considered as continuously exposed during the period between these two interviews. Depending on the drug class, this hypothesis may be valid to a certain extent but, in reality, exposure may be less.

The Three-City (3C) Study is a French repeated cross-sectional cohort designed to assess the risk of dementia and cognitive decline due to vascular risk factors⁹ in which participants were interviewed at two-year intervals over six years. The particularity of this cohort was that reimbursement claims data from the main national healthcare insurance system (Caisse Nationale

d'Assurance Maladie des Travailleurs Salariés, CNAM-TS) were available for the follow-up period which allowed investigation of drug exposure during inter-interview intervals. From these reimbursement data, several levels of exposure could be defined between two interviews taking into account treatment coverage with and without regularity criteria, and considering these levels of exposure as the “gold standard”, the validity of drug exposure presumed from the two cross-sectional interviews was evaluated for different drug classes.

MATERIALS AND METHODS

Methodology of the Three-City (3C) Study

Between 1999 and 2001, 9,294 subjects were included in the 3C cohort according to the following criteria: (i) to be at least 65 years old, (ii) to live in the community, (iii) to be registered on the electoral rolls of one of three French cities: Bordeaux, Dijon or Montpellier. Participants were followed for four years by means of two consecutive visits: two years after inclusion (S1: 2001-2003) and four years after inclusion (S2: 2003-2005). At inclusion and each follow-up visit, participants were interviewed by trained psychologists using a standardized questionnaire on socio-demographic characteristics, health status and drug use. The 3C Study protocol was approved by the Kremlin-Bicêtre university hospital ethics committee (Paris, France) and written informed consent was obtained from each participant. The methodology has been described in detail elsewhere.⁹

Drug use data

During 3C interviews, subjects were asked to list all medications, prescribed and OTC, used more than once a week over the preceding month by an open-ended question. A maximum of fifteen drugs could be collected. This list was validated by visual inspection of drug packages and prescription forms that subjects had to bring at interview. Drugs were registered using their brand name and coded according to the Anatomical Therapeutic Chemical (ATC) classification.¹²

In addition to interview data, reimbursement claims data were obtained from the CNAM-TS which covers approximately 85% of the French general population and 69% of those aged 65 years and over for reimbursement of prescribed drugs and other medical expenses.¹³ All CNAM-TS affiliated members were identified among the 9,294 subjects included in the 3C cohort. For each identified subject, the list of all drugs reimbursed between 1 January 2001 and 31 December 2005 at the maximum was extracted from computerized databases. The ATC code and date corresponding to each drug reimbursement claim were provided. Available drug use data from patient interviews or reimbursement claims as part of the 3C Study are illustrated in Figure 1.

Study population

According to the study objective (i.e. to evaluate the validity of drug exposure presumed from two cross-sectional interviews taking reimbursement data available between the two interviews as reference), interviews conducted at the first (S1) and the second (S2) 3C follow-up visits were considered for the present study. Thus, the 3C participants interviewed at two and four years of follow-up, and having reimbursement claims data available over the period between these two interviews were eligible for the study.

Drug exposure definitions according to patient interviews or reimbursement data

Drug classes studied were first chosen to include a majority of drugs that are reimbursed by the French healthcare insurance system and that can be used chronically whatever the regularity of use. Drug classes for which more than 90% of subjects did not report use at any of the two follow-up interviews were next excluded from the analyses. All the drug classes finally considered are listed in the Appendix in alphabetical order of ATC code.

According to interview data, subjects were classified as exposed to a given drug class between the two follow-up interviews if they reported the use of at least one drug of the class at both interviews.

In comparison, several drug exposure definitions were considered according to reimbursement claims data using two parameters: the proportion of coverage and maximal time between two consecutive claims. The proportion of coverage was calculated as the ratio between the number of observed claims and the maximal number of expected claims for the period between the two interviews. In France, most drugs are dispensed and reimbursed for a maximum of one month's treatment whatever the actual prescription. For example, a treatment prescribed for six months will be dispensed six times and a reimbursement claim will be automatically submitted by the pharmacist at each of the six dispensations. Therefore, the maximal number of expected claims for the period between the two interviews was estimated as the number of months in the period. Firstly, drug exposure according to reimbursement claims data was only defined by the proportion of coverage. For a given drug class, subjects were classified as exposed during the period between the two follow-up interviews if they had (i) at least 50% coverage or (ii) at least 80% coverage for this period. Secondly, drug exposure according to reimbursement claims data was defined by the proportion of coverage in combination to the maximal time between two consecutive claims. For a given drug class, subjects were classified as exposed during the period between the two follow-up interviews if they had (iii) at least 50% coverage with at the maximum 90 days between two consecutive claims or (iv) at least 80% coverage with at the maximum 60 days between two consecutive claims.

Statistical analyses

Prevalence of exposure within each drug class was determined from patient interviews and reimbursement claims data with all the above cited exposure definitions.

Theoretically, neither patient interviews nor reimbursement claims data could be considered as a true “gold standard” for drug exposure assessment. Nevertheless, according to the study setting, reimbursement claims data continuously available over the period of interest could reasonably represent an “alloyed gold standard” in comparison to cross-sectional patient

interviews occurring at the beginning and the end of this period.¹ For each drug class, validity of exposure presumed from patient interviews was assessed by calculating sensitivity (Se), specificity (Sp), positive predictive value (PPV) and negative predictive value (NPV) taking exposure respectively derived from each definition according to reimbursement claims data as reference (i.e. at least 50 or 80% coverage with or without a maximal time between two consecutive claims of 90 or 60 days).¹⁴ Sensitivity measured the ability of patient interviews to accurately identify exposed subjects as indicated by reimbursement claims data. Specificity measured the ability of patient interviews to accurately identify subjects not exposed as indicated by reimbursement claims data. PPV was defined as the probability for a subject to be actually exposed (according to reimbursement claims data) given exposure as indicated by patient interviews and NPV as the probability for a subject to be actually not exposed given non-exposure as indicated by patient interviews.

All analyses were performed using SAS[®] statistical software version 9.1 (SAS Institute Inc., Cary, North Carolina, USA).

RESULTS

Study population

Among the 9,294 subjects included in the 3C cohort, 4,962 (53.4%) had reimbursement claims data available from the main national healthcare insurance system (CNAM-TS). Reasons for not having reimbursement data might be that subjects were affiliated with another healthcare system or that their data were not found (name change, change of address...). There was no difference regarding age, gender, health status or drug use between subjects with or without reimbursement data. The two groups significantly differed only with regard to education level (30.4% of subjects with reimbursement data had high school level or above *vs.* 47.4% of subjects without reimbursement data).

Of these 4,962 subjects, 2,985 (60.2%) interviewed at two and four years of follow-up and having reimbursement claims data over the period between these two interviews were included in the study. At inclusion in the 3C cohort, median age of the study population was 73 years (range: 65 – 94), 65.7% were women and 71.6% reported at least one of the following cardiovascular risk factors or diseases: hypertension, diabetes, hypercholesterolemia, angina, heart rhythm disorders, myocardial infarction or stroke.

Prevalence of exposure according to patient interviews or reimbursement data (Figure 2)

From interview data, prevalence of exposure between the two interviews ranged from 3.7 to 55.2% according to the drug class. It was close to or above 20% for benzodiazepines (19.9%), hematological drugs (22.5%), lipid lowering agents (30.3%) and antihypertensives (55.2%) and below 5% for sensory system drugs (4.3%) and second line analgesics (3.7%).

From reimbursement data, prevalence of exposure measured over the period between the two interviews using a 50% proportion of coverage ranged from 6.5% (for second line analgesics) to 58.6% (for antihypertensives). For all drug classes, it was close to or higher than the prevalence of exposure presumed from interview data. The greatest relative differences between these two prevalences were observed for sensory system drugs, first and second line analgesics and non-steroidal anti-inflammatory drugs (NSAIDs). Considering a higher proportion of coverage of 80%, prevalence of exposure from reimbursement data ranged between 2.8 and 50.8% (also respectively for second line analgesics and antihypertensives). Except for sensory system drugs for which it was again higher than prevalence of exposure from interview data, prevalence of exposure measured according to reimbursement data for all other drug classes was lower than that presumed from interview data especially for mineral supplements, benzodiazepines and genito-urinary system drugs for women.

When combining a 50% proportion of coverage to a maximal time between two consecutive claims of 90 days to define exposure over the period between the two interviews

from reimbursement data, prevalence of exposure ranged between 3.0% (for second line analgesics) to 50.0% (for antihypertensives). For every drug class studied, similar prevalences of exposure were obtained according to that definition and to a 80% proportion of coverage only (see above). Conversely, considering a 80% proportion of coverage with a maximal time between two consecutive claims of 60 days, prevalence of exposure from reimbursement data ranged between 1.1 and 35.7% (again respectively for second line analgesics and antihypertensives). Except for sensory system drugs for which it was equal to prevalence of exposure from interview data, prevalence of exposure measured according to reimbursement data was lower than that presumed from interview data especially for NSAIDs, benzodiazepines, second line analgesics and mineral supplements.

Validity of exposure presumed from patient interviews

Sensitivity (Se) of exposure presumed from both interviews (i.e. ability to accurately identify exposed subjects as indicated by reimbursement data) ranged between 26.1 and 95.4% in function of the drug class studied and exposure definition according to reimbursement data (Table 1). Whatever this exposure definition, Se was always below 50% for sensory system drugs (26.1 to 35.9%) and first line analgesics (33.3 to 49.5%) and always close to or above 80% for hematological drugs (79.3 to 88.1%), benzodiazepines (80.5 to 86.0%), lipid lowering agents (82.1 to 88.0%), drugs used in cardiac therapy (i.e. cardiac glycosides, anti-arrhythmics and vasodilators used in cardiac diseases such as nitrates, 83.6 to 91.6%), systemic hormonal preparations (88.1 to 92.4%) and antihypertensives (90.9 to 95.4%). Specificity (Sp) of self-reported use at both interviews (ability to accurately identify subjects not exposed as indicated by reimbursement data) ranged between 67.1 and 99.3% according to the drug class and exposure definition from reimbursement data (Table 1). Except for antihypertensives with the most restrictive exposure definition according to reimbursement data (i.e. at least 80% coverage with a maximal time between claims of 60 days), Sp was always close to or clearly above 80%.

Positive predictive value (PPV) defined as the probability for a subject to be actually exposed (according to reimbursement data) given exposure as indicated by patient interviews was the most variable validity parameter in function of the drug class studied and exposure definition from reimbursement data (Figure 3). With the less restrictive of these definitions (i.e. at least 50% coverage), PPV was between 62.7 and 96.5% while it was between 15.8 and 61.7% with the most restrictive definition (i.e. at least 80% coverage with a maximal time between claims of 60 days). PPV was always below 50% for mineral supplements (15.9 to 43.2%), benzodiazepines (18.2 to 48.3%), analgesics (respectively 19.3 to 38.9% and 18.8 to 44.6% for first or second line analgesics) and genito-urinary system drugs for women (21.3 to 48.3%) except with the less restrictive exposure definition from reimbursement data (Table 1). Conversely, PPV was always close to or above 80% for drugs used for cardiac therapy (76.7 to 95.5%) and antihypertensives (83.0 to 96.5%) except with the most restrictive exposure definition from reimbursement data. In all cases, negative predictive value (NPV) defined as the probability for a subject to be actually not exposed given non-exposure as indicated by patient interviews was very high: between 87.6 and 99.6% (Table 1).

DISCUSSION

The results presented here show that the validity of drug exposure presumed over the period between two cross-sectional interviews by reported use at both interviews varies in function of the drug class. For certain drug classes, the hypothesis of a continuous exposure over the period may be valid but, for others, this is at best uncertain.

With respect to a 80% coverage measured from reimbursement data between the two interviews, the validity of exposure presumed from reported use at both interviews was high for certain drug classes: group 1, including antihypertensives, drugs for cardiac therapy, lipid lowering agents, drugs for acid related disorders, antidepressants, hematological drugs and systemic hormonal preparations. However, when regularity of drug reimbursement was also considered in the “gold

standard”, differences within this group were more pronounced. The two classes for which exposure presumed from interview data was the most valid were antihypertensives and drugs for cardiac therapy. Continuous intake of these drugs is required to avoid relatively immediate occurrence of symptoms. When considering 80% coverage and a maximum of 60 days between claims, the probability for a subject to be exposed in such a fashion according to reimbursement data when this was indicated by patient interviews (i.e. PPV) may seem low. Yet, it should be underlined that this represents exposure over a two-year period that comprise more than 19 months treatment and two gaps of more than 60 days at the maximum. For other classes within group 1, the PPV of exposure from interview data with reference to this most restrictive exposure definition from reimbursement data was lower. This was due to greater differences between prevalences of exposure according to both sources. Several explanations can be put forward for this differences. For instance, complete adherence to lipid lowering agents is widely reported to be low possibly owing to non-symptomatic direct consequences¹⁵ and, for antidepressants, this could be due to a certain part of short-term treatments.¹⁶ This high levels of both coverage and regularity may not reflect true patterns of use. Thus, for these, it may be of interest to modify the assumption made from interview data of continuous exposure to a 50% coverage with 90-day maximal gaps between claims that was found here to be more valid. For hematological drugs and systemic oral preparations, the relative lower PPV may be due to the presence in these large classes of drugs taken chronically and regularly (e.g. clopidogrel or thyroid therapy) and of those that can also be used for the treatment of acute conditions (e.g. heparins or corticosteroids).

Drugs in group 2 included NSAIDs, benzodiazepines, sensory system drugs, first and second line analgesics, mineral supplements and genito-urinary system drugs for women. These differentiate themselves from group 1 by the low validity of exposure presumed from reported use at both interviews already found when the “gold standard” consisted of a 80% coverage. For benzodiazepines, NSAIDs or genito-urinary system drugs for women, mostly represented by

hormone replacement therapy (HRT), PPV was the most affected validity parameter especially when regularity criteria were applied in the “gold standard”. In such situations, prevalences of exposure found from reimbursement data were particularly low explaining the low PPVs. These results are also in line with a probable chronic but irregular exposure due to non-adherence or episodic use¹⁷⁻¹⁹ that can not be well reflected using self-reported use at several cross-sectional interviews as a proxy of exposure over a period. From our findings, exposure presumed from interview data would be valid for 50% coverage without regularity criteria which is not sufficient to consider a continuous exposure over the period between the two interviews. Sensory system drugs, first and second line analgesics were also characterized by particularly low sensitivities of exposure presumed from interview data. This could be due to subjects forgetting to report drugs used topically or over short period of times (i.e. treatment of acute conditions). With regard to drugs used topically, these are often not perceived as medicines³ and this may have contributed to under-reporting of sensory system drugs. For analgesics, another explanation can be that they were actually not used at the time of each interview then not reported by subjects but taken at different times over the period between the two interviews. For such drugs, consecutive cross-sectional interviews could also be not adapted to identify exposure in the intercurrent period. Whatever the type of drugs (chronic and irregular use, topical use or repeated short-term use), reimbursement data from the CNAM-TS could be of great interest to detect subjects sufficiently exposed in terms of quantity and regularity over the period of interest while cross-sectional interviews, even repeated at regular intervals, could be limited. Nevertheless, attention must be paid to the definition of exposure from reimbursement data which must take into account known patterns of use for each drug class.

Concerning the potential presence of selection biases in this study, it should be noticed that the 3C Study was conducted in three different geographical areas in France using electoral rolls to randomly select participants which limit representativity biases. Nevertheless, the present study required to select subjects from the original cohort according to availability of reimbursement data

from the CNAM-TS. Non-salaried subjects, who are affiliated to another healthcare systems (i.e. the specific healthcare systems for farmers or independent workers), were not included in this study. Even if there was no difference regarding age, gender, health status and drug use between the 3C participants with or without reimbursement data, it remains difficult to know how the results would have changed if the whole cohort had been considered in the study. Regarding information biases, a first limit might be raised about reimbursement data which only include drugs both prescribed and approved for reimbursement. For technical reasons, non-reimbursed drugs could not be removed from interview data. A large majority of drugs available in France are listed among those approved for reimbursement and this was especially true at the time of the study (2001-2005). The recent revision of the reimbursement list with an aim to limit healthcare expenditure has thus probably not affected the study results. Moreover, this revision mostly concerns old medications whose efficacy has not been correctly demonstrated, such medications being rarely involved in pharmacoepidemiological research. At interview, subjects could also report use of OTC drugs not registered in CNAM-TS databases. Since they were not asked to specify whether each drug was prescribed or not, OTC drugs could not be identified among drugs reported as used. However, in an elderly population probably regularly followed by physicians (more than 70% of subjects in the study population presented with at least one cardiovascular disease or risk factor at inclusion in the cohort), self-medication with OTC drugs is not expected to occur frequently.²⁰ This second typical limit of reimbursement databases only affect drug classes such as NSAIDs or first line analgesics. Regarding such drug classes, the consistency of exposure measured from reimbursement data as the reference exposure can be discussed. As we also lacked information about the amount of drug reimbursed and directions of use of the corresponding prescription, it was not possible to estimate the theoretical duration of each reimbursed prescription to further identify episodes of continuous use over the period.²¹ Nevertheless, exposure defined through the proportion of coverage and maximal time between two consecutive claims could reasonably be considered as a proxy of

continuous use over the period by taking into account both a minimal number of claims and a relative regularity between these claims.

Despite such limits, our findings are consistent with earlier studies comparing drug exposure data from questionnaires to pharmacy or reimbursement claim databases over long periods of time.^{7,8,21,22} Even if these studies were mostly designed to evaluate recall accuracy of past drug use, what slightly diverges from the objective of the present study, they showed the same trends according to drug classes studied. When these studies concerned several drug classes, accuracy of drug recall was always the best for cardiovascular system drugs.^{7,21} As an example, in the study from Boudreau *et al.*, the longer the recall period (from six months to eight years), the more the sensitivity of interviews decreased but to a lesser extent for antihypertensives than for statins and than for antidepressants.²¹ Some disagreement between both sources of data was otherwise found in studies about NSAIDs, analgesics or HRT.^{8,22}

To conclude, the results suggest that for drugs used chronically and regularly, reported use at consecutive interviews reflects continuous exposure over the period between interviews. For these, etiological studies conducted as part of repeated cross-sectional cohorts may not suffer greatly from exposure misclassification. For drugs that can be used more irregularly and even intermittently, considering reported use at two interviews to assess exposure during the interval between interviews may increase error and lead to misclassification. Further investigations are needed to quantify the potential impact of exposure misclassification in function of the drug class of interest in etiological studies conducted as part of such cohorts.

ACKNOWLEDGEMENTS

The authors wish to thank all the persons who took part in the 3C Study and the CNAM-TS for providing reimbursement data. They also thank Philip Robinson for his valuable comments.

PARTNERSHIP AND SUPPORTS

The Three-City (3C) Study was conducted under a partnership agreement between the Institut National de la Santé et de la Recherche Médicale (INSERM), the Victor Segalen – Bordeaux 2 University and the Sanofi-Synthélabo Company. The Fondation pour la Recherche Médicale funded the preparation and initiation of the study. The 3C Study was also supported by the Caisse Nationale d'Assurance Maladie des Travailleurs Salariés, Direction Générale de la Santé, Conseils Régionaux of Aquitaine, Languedoc-Roussillon and Bourgogne, Fondation de France, Ministry of Research-INSERM Programme 'Cohortes et collections de données biologiques', Mutuelle Générale de l'Education Nationale, Institut de la Longévité, Conseil Général de la Côte d'Or. The present study did not receive any specific funding.

REFERENCES

- 1- West SL, Strom BL, Poole C. Validity of pharmacoepidemiologic drug and diagnosis data. In *Pharmacoepidemiology*. Strom BL (eds). John Wiley & Sons: Chichester, 2005; 709-765
- 2- Copeland KT, Checkoway H, McMichael AJ, *et al*. Bias due to misclassification in the estimation of relative risk. *Am J Epidemiol* 1977; **105**: 488-495
- 3- Moore N, Masson H, Noblet C, *et al*. What medicines do patients really take? A comparison of free form *vs.* oriented questionnaires. *Post Marketing Surveillance* 1993; **7**: 355-362
- 4- Lau HS, de Boer A, Beuning KS, *et al*. Validation of pharmacy records in drug exposure assessment. *J Clin Epidemiol* 1997; **50**: 619-625
- 5- Noize P, Bazin F, Dufouil C, *et al*. Comparison of health insurance claims and patient interviews in assessing drug use: data from the Three-City (3C) Study. *Pharmacoepidemiol Drug Saf* 2009; **18**: 310-319
- 6- Paganini-Hill A, Ross RK. Reliability of recall of drug usage and other health-related information. *Am J Epidemiol* 1982; **116**: 114-122
- 7- Van den Brandt PA, Petri H, Dorant E, *et al*. Comparison of questionnaire information and pharmacy data on drug use. *Pharm Weekbl Sci* 1991; **13**: 91-96
- 8- West SL, Savitz DA, Koch G, *et al*. Recall accuracy for prescription medications: self-report compared with database information. *Am J Epidemiol* 1995; **142**: 1103-1112
- 9- 3C study group. Vascular factors and risk of dementia: design of the Three-City Study and baseline characteristics of the study population. *Neuroepidemiology* 2003; **22**: 316-325
- 10- Dartigues JF, Gagnon M, Barberger-Gateau P, *et al*. The Paquid epidemiological program on brain ageing. *Neuroepidemiology* 1992; **11**: 14-18
- 11- Berdot S, Bertrand M, Dartigues JF, *et al*. Inappropriate medication use and risk of falls - a prospective study in a large community-dwelling elderly cohort. *BMC Geriatr* 2009; **9**: 30
- 12- World Health Organization. Collaborating Centre for Drug Statistics Methodology. *ATC Index with DDDs*. World Health Organization, Oslo; 2000

- 13- Depont F, Fourrier A, Merlière Y, *et al.* The CADEUS study: methods and logistics. *Pharmacoepidemiol Drug Saf* 2007; **16**: 571-580
- 14- Weiss NS. Studies of diagnostic and screening tests. In *Modern Epidemiology*. Rothman KJ, Greenland S (eds). Lippincott-Raven: Philadelphia, 1998; 520-524
- 15- Jackevicius CA, Mamdani M, Tu JV. Adherence with statin therapy in elderly patients with and without acute coronary syndromes. *JAMA* 2002; **288**: 462-467.
- 16- Sihvo S, Isometsä E, Kiviruusu O, *et al.* Antidepressant utilisation patterns and determinants of short-term and non-psychiatric use in the Finnish general adult population. *J Affect Disord* 2008; **110**: 94-105
- 17- Lechevallier N, Fourrier A, Berr C. [Benzodiazepine use in the elderly: the EVA Study]. *Rev Epidemiol Sante Publique* 2003; **51**: 317-326
- 18- Moore N, Diris H, Martin K, *et al.* NSAID use profiles derived from reimbursement data in France. *Thérapie* 2004; **59**: 541-546
- 19- Serfaty D, de Reilhac P, Eschwege E, *et al.* [Compliance with hormone replacement therapy in menopausal women: results of a two-year prospective French study comparing transdermal treatment with fixed oral combination therapy]. *Gynecol Obstet Fertil* 2003; **31**: 525-533
- 20- Colomes M, Rispail Y, Berlan M, *et al.* [Drug consumption in a retired population]. *Thérapie* 1990; **45**: 321-324
- 21- Boudreau DM, Daling JR, Malone KE, *et al.* A validation study of patient interview data and pharmacy records for antihypertensive, statin, and antidepressant medication use among older women. *Am J Epidemiol* 2004; **159**: 308-317
- 22- Monster TB, Janssen WM, de Jong PE, *et al.* Pharmacy data in epidemiological studies: an easy to obtain and reliable tool. *Pharmacoepidemiol Drug Saf* 2002; **11**: 379-384

FIGURE LEGENDS

Figure 1. Representation of drug use data available from the main French national healthcare insurance system (Caisse Nationale d'Assurance Maladie des Travailleurs Salariés, CNAM-TS) and from patient interviews as part of the Three-City (3C) Study (1999-2005). Participants in the 3C cohort were interviewed two years (2001-2003) and four years (2003-2005) after inclusion (1999-2001). In parallel, for the CNAM-TS affiliated members, reimbursement claims data were extracted from computerized databases for the period between 1 January 2001 and 31 December 2003 or 2005 according to the study centre (Bordeaux, Dijon or Montpellier).

Figure 2. Prevalence of drug exposure according to interview data collected at both 3C follow-up visits and to reimbursement data from the main French national healthcare insurance system. From interview data, subjects were classified as exposed between the two interviews if they reported use at both interviews. From reimbursement data, several exposure definitions were considered according to the proportion of coverage (ratio between the number of observed claims and those expected between the two interviews) and the time between claims: at least 50% coverage with or without a maximal time between claims of 90 days or at least 80% coverage with or without a maximal time between claims of 60 days.

Figure 3. Positive predictive value of drug exposure presumed from interview data collected at 3C follow-up visits (i.e. self-reported use of the drug class at both interviews) taking exposure measured from reimbursement data for the period between the two interviews as “gold standard”. Several exposure definitions were considered according to reimbursement data (reference): at least 50% coverage (ratio between the number of observed claims and those expected between the two interviews) with or without a maximal time between claims of 90 days or at least 80% coverage with or without a maximal time between claims of 60 days.

Figure 1

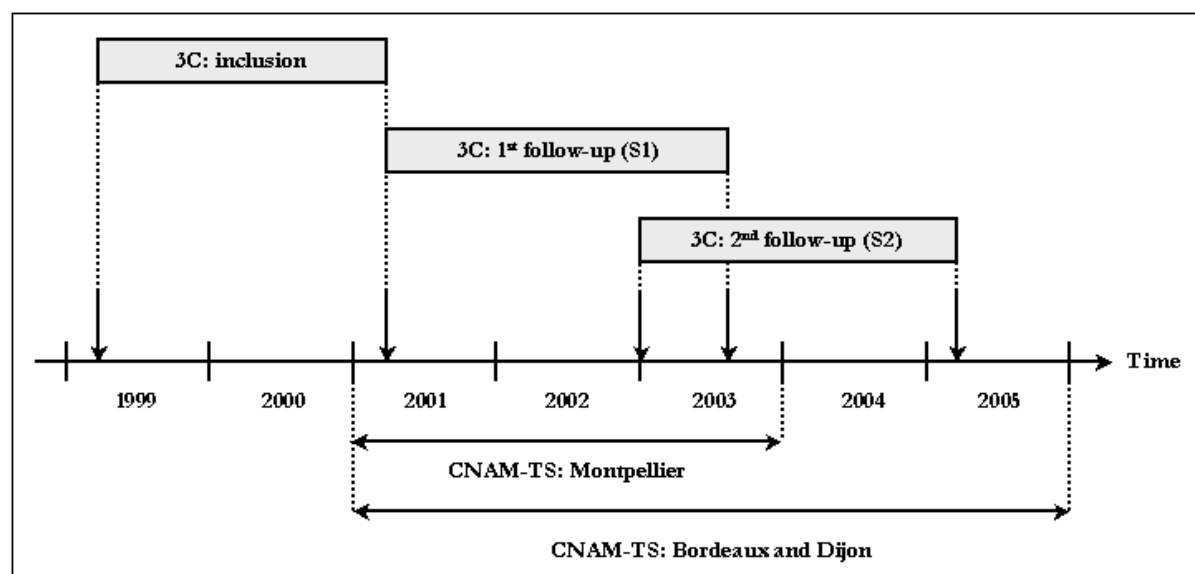


Figure 2

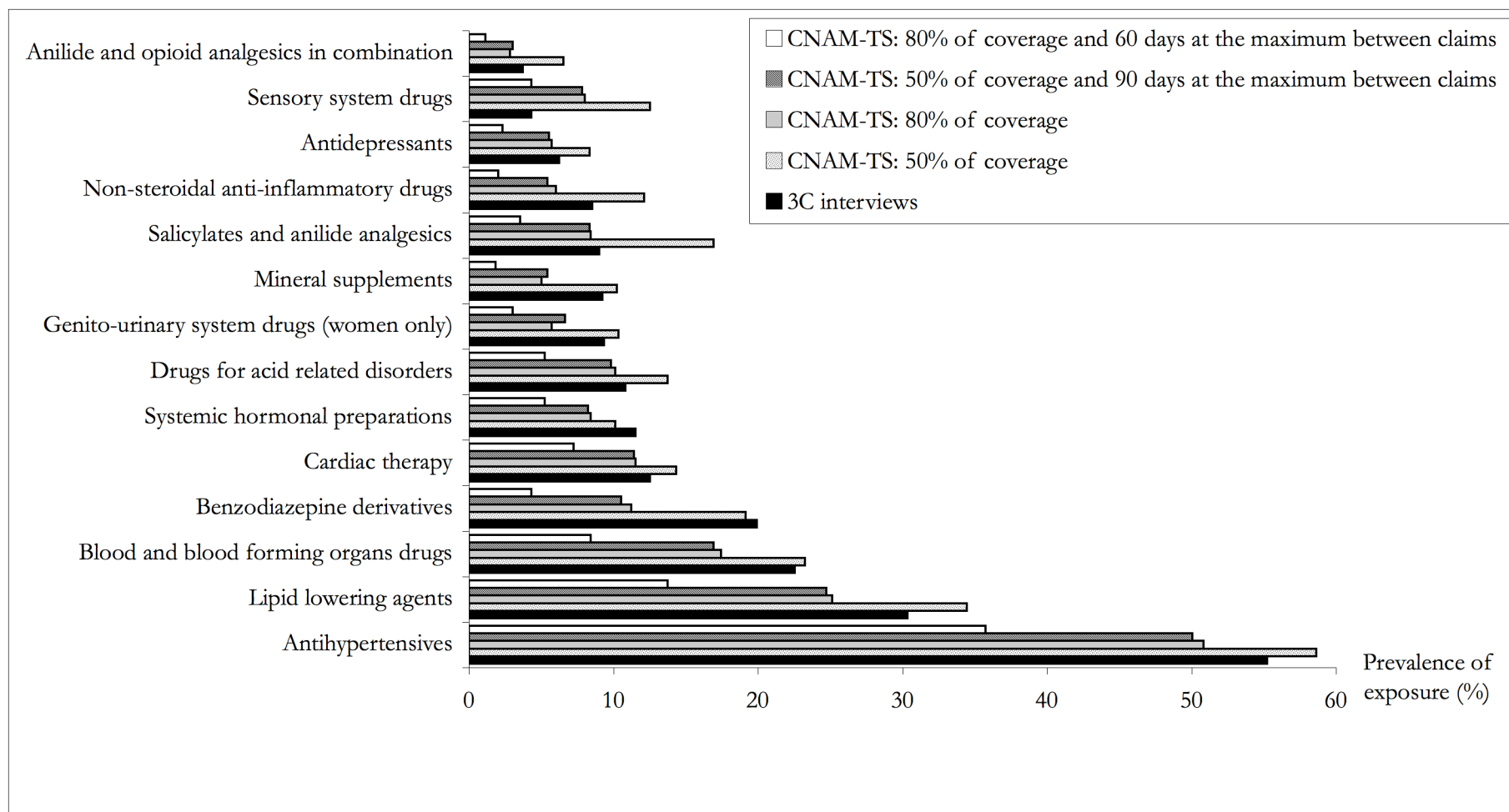


Figure 3

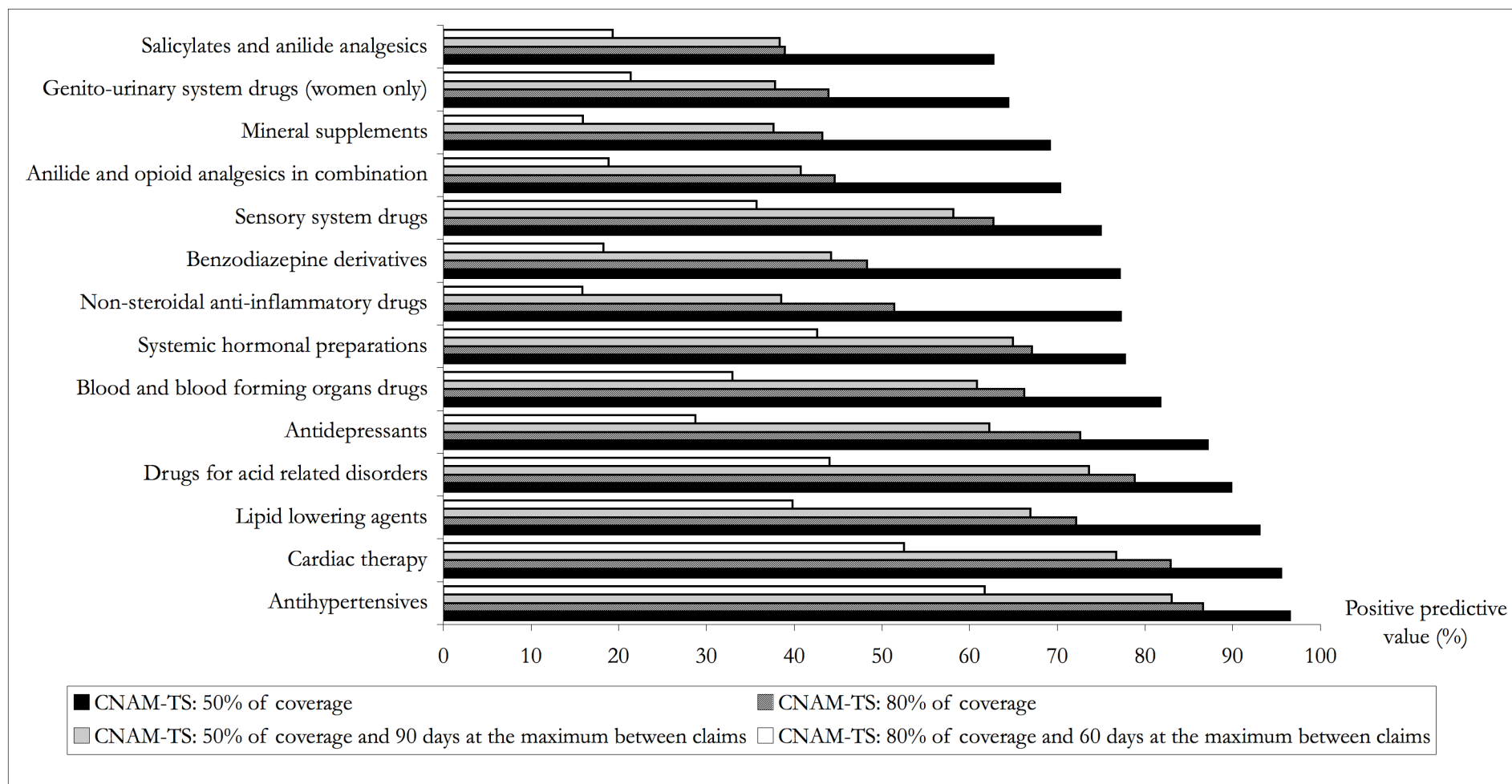


Table 1. Sensitivity (Se), specificity (Sp), positive and negative predictive values (PPVs and NPVs) of exposure from patient interviews according to different exposure definitions from reimbursement claims data used as “gold standard” over the period between the two interviews

Drug classes (n = 2 985)	Exposure definition from reimbursement data		Se (%)	Sp (%)	PPV (%)	NPV (%)
	Proportion of coverage	Maximal time between claims				
Antihypertensives	50%	-	90.9	95.3	96.5	88.1
	80%	-	94.1	85.0	86.6	93.3
	50%	90 days	91.7	81.2	83.0	90.7
	80%	60 days	95.4	67.1	61.7	96.3
Cardiac therapy	50%	-	83.6	99.3	95.5	97.3
	80%	-	90.1	97.6	82.9	98.7
	50%	90 days	84.2	96.7	76.7	97.9
	80%	60 days	91.6	93.6	52.5	99.3
Lipid lowering agents	50%	-	82.1	96.8	93.0	91.2
	80%	-	87.2	88.7	72.1	95.4
	50%	90 days	82.2	86.7	66.9	93.7
	80%	60 days	88.0	78.8	39.8	97.6
Drugs for acid related disorders	50%	-	70.7	98.7	89.8	95.5
	80%	-	83.8	97.5	78.8	98.2
	50%	90 days	80.9	96.8	73.6	97.9
	80%	60 days	91.0	93.6	44.0	99.5

Table 1. Continued

Drug classes (n = 2 985)	Exposure definition from reimbursement data		Se (%)	Sp (%)	PPV (%)	NPV (%)
	Proportion of coverage	Maximal time between claims				
Antidepressants	50%	-	65.6	99.1	87.1	97.0
	80%	-	79.4	98.2	72.6	98.7
	50%	90 days	70.3	97.5	62.2	98.3
	80%	60 days	77.9	95.4	28.7	99.5
Blood and blood forming organs drugs	50%	-	79.3	94.6	81.7	93.8
	80%	-	85.6	90.8	66.2	96.8
	50%	90 days	81.0	89.4	60.8	95.9
	80%	60 days	88.1	83.5	32.9	98.7
Systemic hormonal preparations	50%	-	88.1	97.2	77.7	98.6
	80%	-	91.3	95.9	67.1	99.2
	50%	90 days	90.6	95.6	64.9	99.1
	80%	60 days	92.4	93.0	42.6	99.5
Non-steroidal anti-inflammatory drugs	50%	-	54.4	97.8	77.2	94.0
	80%	-	72.8	95.6	51.4	98.2
	50%	90 days	60.5	94.5	38.5	97.7
	80%	60 days	67.2	92.7	15.8	99.3
Benzodiazepine derivatives	50%	-	80.5	94.4	77.1	95.4
	80%	-	86.0	88.4	48.3	98.0
	50%	90 days	84.0	87.6	44.2	97.9
	80%	60 days	84.5	83.0	18.2	99.2

Table 1. Continued

Drug classes (n = 2 985)	Exposure definition from reimbursement data		Se (%)	Sp (%)	PPV (%)	NPV (%)
	Proportion of coverage	Maximal time between claims				
Sensory system drugs	50%	-	26.1	98.8	74.9	90.5
	80%	-	33.8	98.3	62.7	94.5
	50%	90 days	32.2	98.0	58.1	94.5
	80%	60 days	35.9	97.1	35.7	97.1
Anilide and opioid analgesics in combination	50%	-	40.2	98.8	70.3	96.0
	80%	-	58.8	97.9	44.6	98.8
	50%	90 days	50.6	97.7	40.7	98.5
	80%	60 days	63.6	97.0	18.8	99.6
Mineral supplements	50%	-	62.4	96.8	69.1	95.8
	80%	-	79.9	94.5	43.2	98.9
	50%	90 days	64.2	93.9	37.6	97.9
	80%	60 days	81.5	92.1	15.9	99.6
Genito-urinary system drugs: women (n = 1 962)	50%	-	58.1	96.3	64.4	95.2
	80%	-	72.1	94.4	43.9	98.2
	50%	90 days	53.5	93.8	37.8	96.6
	80%	60 days	66.1	92.4	21.3	98.9
Salicylates and anilide analgesics	50%	-	33.3	96.0	62.7	87.6
	80%	-	41.6	94.0	38.9	94.6
	50%	90 days	41.4	94.0	38.3	94.7
	80%	60 days	49.5	92.5	19.3	98.1

APPENDIX – List of the drug classes considered in the study

Drug classes	Corresponding ATC codes
Drugs for acid related disorders	A02
Mineral supplements	A12
Blood and blood forming organs drugs	B
Cardiac therapy	C01A / C01B / C01D
Antihypertensives	C02 / C03 / C07 / C08 / C09
Lipid lowering agents	C10
Genito-urinary system drugs (women only)	G
Systemic hormonal preparations	H
Non-steroidal anti-inflammatory drugs	M01A
Anilide and opioid analgesics in combination	N02AA59 / N02AC54 / N02BE51 / N02AX52
Salicylates and anilide analgesics	N02BA / N02BE
Benzodiazepine derivatives	N05BA / N05CD / N05CF
Antidepressants	N06A / N06CA
Sensory system drugs	S

ATC, Anatomical Therapeutic Chemical

IV- DEUXIÈME PARTIE

**IMPACT DE LA SOURCE DE DONNÉES UTILISÉE POUR
DÉTERMINER L'EXPOSITION MÉDICAMENTEUSE SUR LA
MESURE D'UNE ASSOCIATION ENTRE UN MÉDICAMENT
ET UN ÉVÉNEMENT**

IV.1- CONTEXTE ET OBJECTIF

Les travaux présentés dans cette seconde partie ont récemment fait l'objet d'une communication orale lors du dernier congrès organisé par l'ISPE (Annexe 1) et un article scientifique est en cours de préparation. Ils ont été réalisés afin de répondre au deuxième objectif défini pour ce travail de thèse soit évaluer l'impact de la source de données utilisée pour déterminer l'exposition médicamenteuse sur la mesure d'une association entre un médicament et un événement. Compte-tenu des données disponibles, ils se rapportent plus spécifiquement au contexte des études étiologiques conduites dans des cohortes à suivis transversaux répétés comme l'étude 3C.

Dans une telle cohorte, lorsque l'événement sous étude survient au cours du suivi et que seules des données déclaratives sont utilisables pour déterminer l'exposition médicamenteuse, les sujets sont généralement classés comme exposés ou non au médicament d'intérêt selon leurs déclarations de consommation à un ou plusieurs interrogatoire(s) précédant la survenue de l'événement et ce, quelle que soit la longueur de l'intervalle de temps séparant cet(ces) interrogatoire(s) de la connaissance de l'événement.¹⁴⁴⁻¹⁴⁸ Suivant la longueur de cet intervalle, des erreurs peuvent intervenir quant à la réalité du statut d'exposition au cours de cette période. Selon leur type et leur importance, ces erreurs de classification peuvent éventuellement introduire un biais dans l'estimation de l'association entre l'exposition médicamenteuse et l'événement. Dans l'étude 3C, la disponibilité de données de remboursement au cours de cet intervalle de temps compris entre le(s) recueil(s) de données déclaratives et la survenue de l'événement permet ainsi d'explorer dans quelle mesure les estimations de risque obtenues en mesurant l'exposition médicamenteuse à partir de ces données de remboursement diffèrent de celles obtenues en considérant l'exposition médicamenteuse issue des données déclaratives.

L'objectif était donc ici :

- d'évaluer, pour différentes classes médicamenteuses, l'association entre l'exposition mesurée à partir de données déclaratives ou à partir de données de remboursement et un événement survenant dans les deux ans suivant le recueil de données déclaratives (Figure 4).

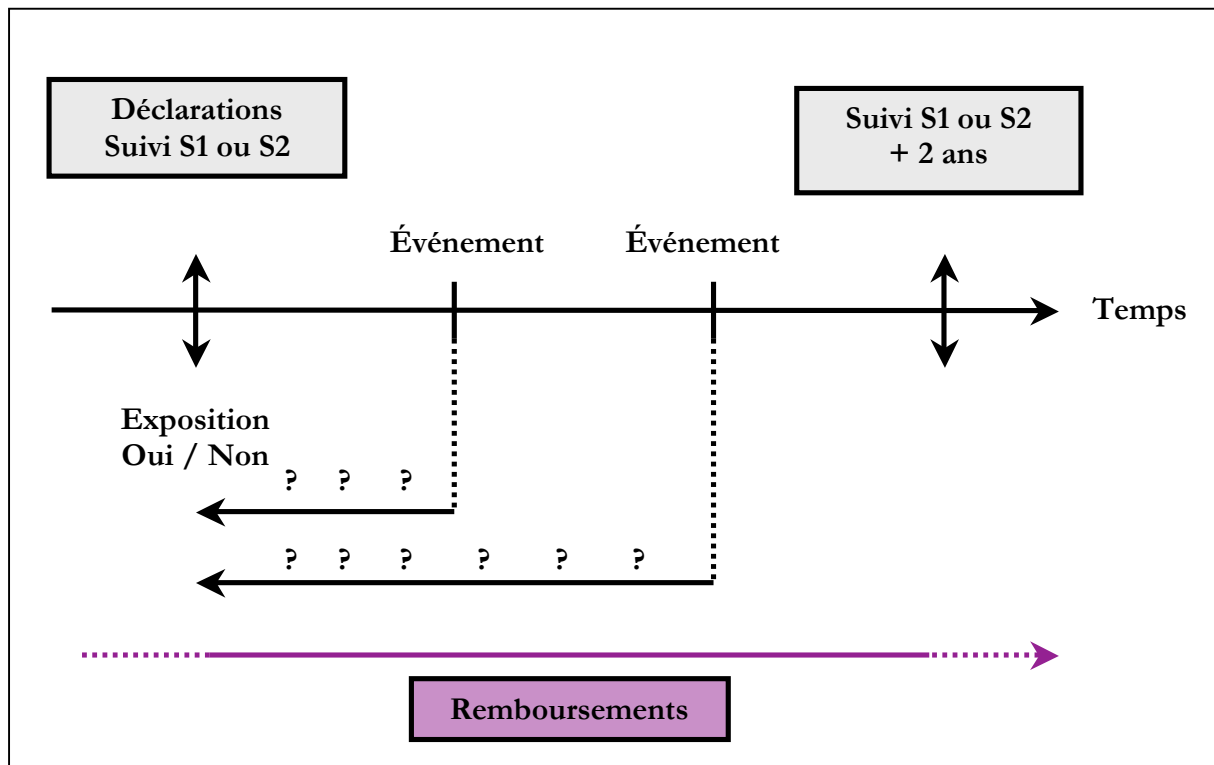


Figure 4. Illustration de la mesure de l'exposition médicamenteuse pour les études étiologiques dans une cohorte à suivis transversaux répétés : la cohorte 3C

IV.2- MÉTHODES

IV.2.1- Échantillon d'étude

Les travaux présentés dans cette partie ont porté sur un échantillon de 3 394 sujets (âge médian à l'inclusion : 73 ans ; 65,7 % de femmes) de la population source ayant des données déclaratives recueillies à un des interrogatoires de suivi de la cohorte (suivi deux ans après inclusion S1 ou quatre ans après inclusion S2) et des données de remboursement disponibles de façon continue pendant les deux années suivant cet interrogatoire.

IV.2.2- Choix des classes médicamenteuses étudiées

Les classes médicamenteuses d'intérêt dans cette seconde partie ont été choisies au vu des résultats des travaux de comparaison de l'exposition médicamenteuse mesurée à partir de données déclaratives ou de données de remboursement ayant fait l'objet de la première partie de cette thèse. En effet, que ce soit dans le contexte de la mesure d'une exposition médicamenteuse un jour donné ou au cours d'une période de temps plus longue, ces travaux ont montré que la concordance entre les expositions issues des deux sources de données ou la validité de l'exposition issue d'une des sources par rapport à celle issue de la seconde variaient considérablement en fonction de la classe médicamenteuse et, principalement, de ses modalités d'utilisation. Les résultats de concordance ou de validité étaient d'autant moins bons que les médicaments de la classe peuvent être consommés de manière irrégulière voire intermittente. Les estimations de l'association entre une exposition médicamenteuse et un événement selon que l'exposition est mesurée à partir de données déclaratives ou de remboursement risquaient, par conséquent, de davantage différer pour ce type de classes médicamenteuses. Aussi, notre choix s'est-il porté sur deux d'entre-elles : les benzodiazépines et les AINS.

Plusieurs facteurs expliquent le choix de ces deux classes médicamenteuses en particulier, à commencer par leurs profils d'utilisation sensiblement différents. Si les benzodiazépines sont connues pour être consommées de manière chronique^{142,149} mais potentiellement irrégulière¹⁵⁰ chez le sujet âgé, les AINS seraient plutôt consommés de manière intermittente en fonction des symptômes ressentis au cours de l'évolution de maladies chroniques notamment rhumatismales.^{126,151} Par ailleurs, la large utilisation de ces deux classes et l'intérêt qu'elles suscitent en pharmaco-épidémiologie les rendaient également intéressantes dans le cadre de ce travail. Pour

illustration, d'après la synthèse de la littérature présentée en introduction, près de 20 % des études étiologiques publiées, en 2008, dans la revue *Pharmacoepidemiology and Drug Safety* portaient respectivement sur des médicaments psychotropes, dont les benzodiazépines, et sur les AINS.

Cependant, il paraissait opportun de confronter ces deux classes médicamenteuses avec une de celles caractérisées, dans les travaux de la première partie, par de bons résultats de concordance ou de validité. Pour ces classes médicamenteuses utilisées au long cours et de manière régulière dans le traitement des pathologies chroniques du sujet âgé (notamment, les pathologies du système cardiovasculaire), il est probable que l'impact du choix de la source de données d'exposition médicamenteuse sur l'estimation de l'association soit plus faible. Les analyses décrites par la suite ont donc été également conduites pour la classe dite des « médicaments cardiologiques » regroupant les digitaliques, les anti-arythmiques et les dérivés nitrés.

IV.2.3- Définitions de l'exposition médicamenteuse

IV.2.3.1- Définition de l'exposition médicamenteuse selon les données déclaratives

À partir des déclarations des sujets, une seule définition de l'exposition médicamenteuse était utilisable : les sujets ayant déclaré avoir consommé au moins un médicament de la classe à l'interrogatoire de suivi (S1 ou S2) ont été considérés comme exposés selon les données déclaratives.

IV.2.3.2- Définitions de l'exposition médicamenteuse selon les données de remboursement

Comme précédemment évoqué en introduction, en pharmaco-épidémiologie, lorsqu'une association existe entre une exposition médicamenteuse et un événement, sa force peut varier selon certaines caractéristiques de l'exposition parmi lesquelles sa durée avant la survenue de l'événement, son « ancienneté » par rapport à la survenue de l'événement (par exemple, exposition « en cours », récente ou passée) ou encore sa régularité dans le temps. Lors de la conduite d'études étiologiques, plusieurs définitions de l'exposition médicamenteuse selon ces différents critères sont donc testées afin de mettre en évidence les caractéristiques de l'exposition les plus fortement associées à la survenue de l'événement.

Ainsi, étant donné le caractère continu des données extraites des bases de remboursement de la CNAM-TS, plusieurs définitions de l'exposition médicamenteuse ont été envisagées selon cette source afin d'illustrer plusieurs cas de figure pouvant être rencontrés dans les études étiologiques en pharmaco-épidémiologie. Le choix de ces cas de figure a aussi été fait en cohérence avec les travaux de la première partie puisqu'ont été considérés :

- d'une part, le cas d'une exposition « en cours » à la date de survenue de l'événement donc d'une exposition mesurée un jour donné et,
- d'autre part, le cas d'une exposition chronique et régulière au cours d'une période de temps plus longue avant la survenue de l'événement.

IV.2.3.2.1- Cas d'un événement lié à une exposition médicamenteuse « en cours »

Dans ce premier cas de figure, la survenue de l'événement pourrait se rapporter à une exposition médicamenteuse « en cours » (c'est-à-dire existante au moment de la survenue de l'événement) indépendamment de sa durée ou de sa régularité. Il s'agissait donc de mesurer cette exposition de manière transversale au moment de l'événement. Ainsi, les sujets ayant bénéficié du remboursement d'au moins un médicament de la classe dans les n jours précédant l'événement ont été classés comme exposés selon les données de remboursement (Figure 5). La fenêtre de temps de n jours utilisée correspondait, pour chaque classe médicamenteuse, à celle déterminée pour considérer les données de remboursement avant la date du premier interrogatoire dans les travaux de comparaison de l'exposition médicamenteuse mesurée à partir de données déclaratives ou de remboursement au cours de la période comprise entre les deux interrogatoires soit :

- 90 jours pour les benzodiazépines,
- 45 jours pour les AINS et,
- 60 jours pour les médicaments cardiologiques (Annexe 3).

Dans ce cas de figure, par rapport à la définition selon les données déclaratives, la définition selon les données de remboursement pouvait éviter de considérer à tort comme exposés des sujets qui ont arrêté de prendre le médicament après l'interrogatoire et ne sont donc plus exposés au moment de la survenue de l'événement ou, *a contrario*, de considérer à tort comme non exposés des sujets qui ont commencé de prendre le médicament après l'interrogatoire et qui sont en réalité exposés au moment de la survenue de l'événement.

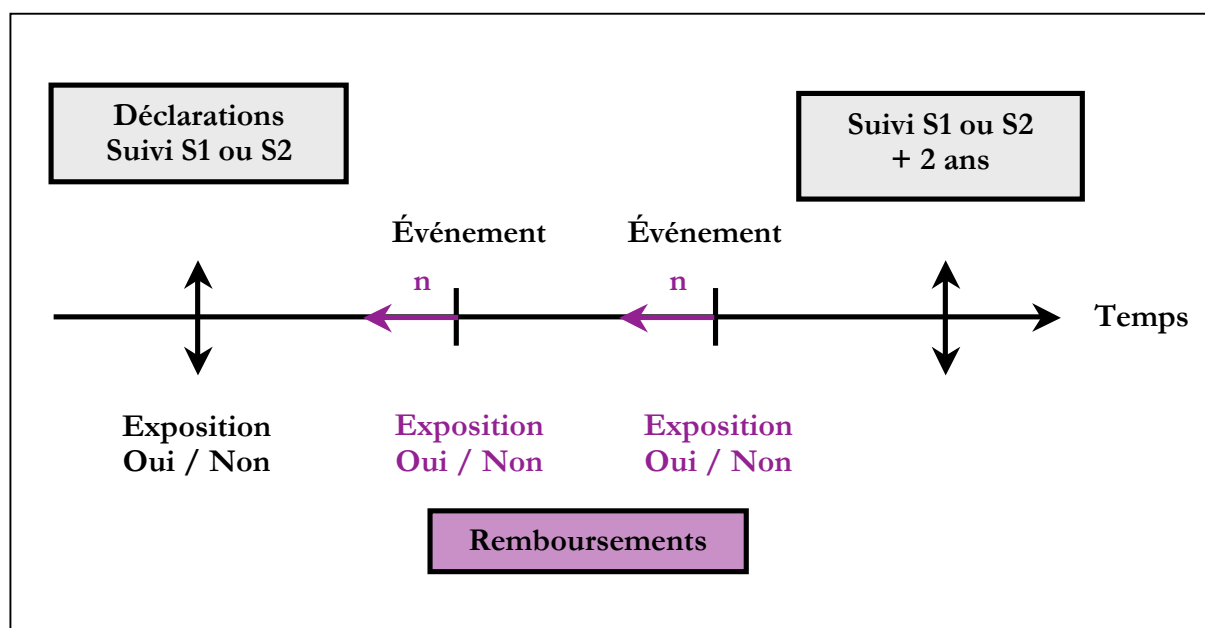


Figure 5. Illustration des définitions de l'exposition médicamenteuse selon chaque source de données dans le cas d'un événement lié à une exposition « en cours » (transversale)

IV.2.3.2.2- Cas d'un événement lié à une exposition médicamenteuse chronique régulière

Dans ce second cas de figure, la survenue de l'événement pourrait se rapporter à une exposition médicamenteuse chronique et régulière pendant une certaine période. Il s'agissait donc ici de mesurer cette exposition de manière « longitudinale » au cours d'une certaine période avant la survenue de l'événement. Par conséquent, les sujets ayant bénéficié, pour une classe médicamenteuse donnée, d'au moins 80 % des remboursements attendus pour la période comprise entre l'interrogatoire et l'événement ont été classés comme exposés selon les données de remboursement (Figure 6). La proportion de couverture a été fixée à 80 % car ce seuil est souvent choisi pour considérer qu'un sujet observe suffisamment son traitement.^{44,152}

Par rapport à la définition selon les données déclaratives, cette seconde définition pouvait éviter de considérer à tort comme exposés des sujets qui prennent le médicament par intermittence ou ne sont pas assez observants au cours de la période ou, *a contrario*, de considérer à tort comme non exposés des sujets qui ont commencé de prendre le médicament immédiatement après l'interrogatoire et ont été observants au cours de la période.

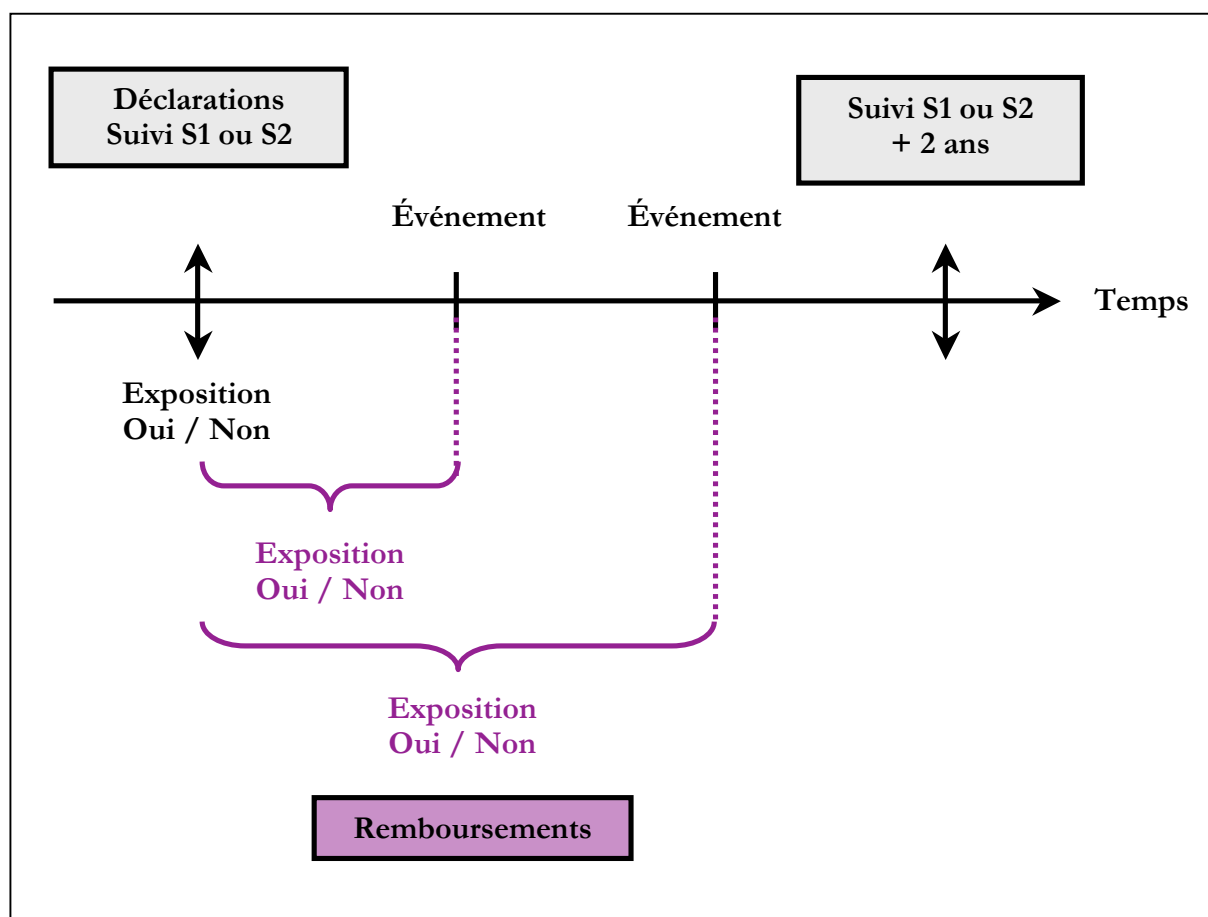


Figure 6. Illustration des définitions de l'exposition médicamenteuse selon chaque source de données dans le cas d'un événement lié à une exposition chronique régulière au cours d'une période (« longitudinale »)

IV.2.4- Simulation de la survenue d'un événement dans l'échantillon d'étude

Pour la réalisation de ces travaux, une des principales difficultés rencontrées tenait à l'absence, parmi les données disponibles pour les sujets de la cohorte 3C, d'une variable « événement » de fréquence suffisante et de date de survenue connue. Pour pallier ce manque, un événement simulé a été généré au sein de l'échantillon d'étude.

La simulation de cet événement a tenu compte des paramètres suivants :

- le nombre de sujets présentant l'événement (cas) a été déterminé en fixant arbitrairement sa fréquence dans l'échantillon d'étude,
- le nombre de sujets exposés a été fixé en utilisant la prévalence d'exposition mesurée selon les données déclaratives et,
- le nombre de cas exposés a été déterminé en fixant arbitrairement la force de l'association (rapport de cotes, RC) entre l'exposition (selon les données déclaratives) et l'événement.

Étant donné le caractère « théorique » de cette simulation d'événement, plusieurs situations pouvant être rencontrées au cours des études étiologiques en pharmaco-épidémiologie ont été à nouveau envisagées. Ainsi, pour une fréquence de l'événement arbitrairement fixée à 10 % dans l'échantillon d'étude, la force de l'association entre l'exposition (selon les données déclaratives) et l'événement pouvait correspondre à un rapport de cotes de 1, 2, 3, 4, 5 ou 6. La combinaison de ces deux paramètres a donc abouti à la simulation de six types « théoriques » d'événements plus ou moins associés à l'exposition médicamenteuse (selon les données déclaratives). La méthode appliquée pour la simulation de l'événement a été exposée en détails en Annexe 4.

Pour chacun de ces types d'événements simulés, une date de survenue a été générée de manière aléatoire en considérant une fonction de risque constante au cours de la période de deux ans suivant l'interrogatoire.

IV.2.5- Type d'étude utilisé pour évaluer l'impact de la source de données : étude cas-témoins nichée

Afin d'étudier l'impact de la source de données et de la définition utilisées pour déterminer l'exposition médicamenteuse sur les mesures d'association, une étude cas-témoins nichée dans la sous-cohorte définie comme échantillon d'étude a été conduite de façon à évaluer l'association entre l'exposition mesurée soit à partir des données déclaratives, soit à partir des données de remboursement et l'événement simulé généré comme précisé ci-dessus.

Pour chaque classe médicamenteuse étudiée et chaque type « théorique » d'événement (en fonction de la force de son association à l'exposition selon les données déclaratives), tous les cas de l'événement simulé générés au cours de la période de deux ans suivant l'interrogatoire ont été appariés à au maximum cinq témoins issus de la sous-cohorte. Conformément à la méthodologie de l'étude cas-témoins nichée, un sujet sélectionné comme témoin à un moment donné au cours de la période pouvait ultérieurement devenir un cas ou être à nouveau sélectionné comme témoin pour un autre cas. La date d'événement a été choisie comme date de référence pour chaque cas et ses témoins : l'exposition médicamenteuse selon les données de remboursement telle que schématisée en Figure 5 et Figure 6 a donc été mesurée par rapport à cette date de référence.

Puis, par régression logistique conditionnelle, un RC et son intervalle de confiance à 95 % (IC 95 %) ont été déterminés pour l'association entre chaque type d'événement simulé et :

- l'exposition selon les données déclaratives,
- l'exposition selon les données de remboursement avec la définition transversale à la date de référence,
- l'exposition selon les données de remboursement avec la définition « longitudinale » au cours de la période entre l'interrogatoire et la date de référence.

Tous les modèles de régression logistique conditionnelle ont été réalisés avec ajustement sur l'âge, le sexe, le niveau d'études et le nombre de médicaments consommés hormis la classe étudiée.

IV.2.6- Analyse complémentaire

Une analyse complémentaire a été réalisée afin d'explorer si les éventuelles différences de mesure d'association entre l'exposition selon les données déclaratives ou selon les données de remboursement et chaque type d'événement simulé pouvaient au moins en partie s'expliquer par la discordance entre les expositions issues des deux sources de données. Ainsi, pour chacun des six types « théoriques » d'événements simulés et chaque définition de l'exposition médicamenteuse selon les données de remboursement (transversale ou « longitudinale »), ont été déterminées :

- la proportion de sujets discordants entre les données déclaratives et les données de remboursement vis-à-vis de l'exposition médicamenteuse (sujets exposés selon une source et non exposés selon l'autre source) et,
- la variation relative entre les RC mesurant l'association entre l'exposition selon les données déclaratives et l'événement simulé et l'exposition selon les données de remboursement et ce même événement simulé.

Pour chaque classe médicamenteuse, 12 couples -proportion de sujets discordants ; variation relative du RC- ont été obtenus (6 types « théoriques » d'événements simulés * 2 définitions de l'exposition selon les données de remboursement). Ces 36 couples ont été représentés sur un même graphique (12 couples * 3 classes médicamenteuses) de façon à illustrer une possible relation entre ces deux paramètres. Dans cette analyse, la variation relative du rapport de cotes a été jugée importante au-delà de 30 %.

L'ensemble des analyses décrites dans cette partie « Méthodes » ont été conduites avec le logiciel SAS version 9.1 (SAS Institute Inc., Cary, North Carolina, USA).

IV.3- RÉSULTATS

Les Figures 7, 8 et 9 présentent respectivement pour les benzodiazépines, les AINS et les médicaments cardiologiques, le RC et son IC à 95 % déterminés pour l'association entre chaque type d'événement simulé et :

- l'exposition selon les données déclaratives (en noir),
- l'exposition selon les données de remboursement avec la définition transversale à la date de référence (en orange),
- l'exposition selon les données de remboursement avec la définition « longitudinale » au cours de la période entre l'interrogatoire et la date de référence (en violet).

En ce qui concerne les benzodiazépines (Figure 7), pour les forces d'association les plus faibles (soit pour des RC de 1, 2 voire 3), les résultats obtenus avec les données de remboursement étaient équivalents à ceux fixés à partir des données déclaratives et ce, quelle que soit la définition utilisée : transversale ou « longitudinale ». En revanche, pour des forces d'association plus importantes (soit pour des RC allant de 4 à 6), la mesure de l'association faite à partir des données de remboursement était plus faible que celle fixée à partir des données déclaratives.

Le même type de résultats était observé pour les AINS à la différence près que l'impact de la source de données était visible dès une force d'association correspondant à un RC de 2 (Figure 8). Par ailleurs, il existait de sensibles différences entre les deux définitions de l'exposition selon les données de remboursement : la mesure de l'association obtenue avec la définition transversale était toujours plus faible que celle obtenue avec la définition « longitudinale ».

Comparativement à ces deux classes médicamenteuses, pour les médicaments cardiologiques, les RC mesurant l'association entre chaque événement simulé et l'exposition selon les données de remboursement se rapprochaient davantage de ceux fixés avec l'exposition selon les données déclaratives (Figure 9). Les mêmes différences que pour les deux autres classes médicamenteuses étaient observées pour les forces d'associations les plus fortes mais dans de moindres proportions avec des intervalles de confiance à 95 % qui se chevauchaient systématiquement.

D'après la Figure 10, une relation linéaire semblait exister, indépendamment de la classe médicamenteuse, entre les différences de mesure d'association selon que l'exposition était issue des données déclaratives ou des données de remboursement et chaque type d'événement simulé

(symbolisées par la variation relative du RC) et la proportion de sujets discordants entre les deux sources de données quant à leur statut d'exposition (coefficient de corrélation $r = 0,59$; $p < 10^{-3}$). Ces résultats suggèrent que plus la discordance entre les deux sources de données vis-à-vis de l'exposition médicamenteuse est importante, plus la force de l'association entre l'exposition mesurée à partir de l'une ou l'autre des deux sources de données et un événement diffère. De manière attendue au vu des résultats précédemment décrits, l'ensemble des points correspondants aux médicaments cardiologiques était concentré dans une zone à la fois de faible proportion de sujets discordants et de faible variation relative du RC alors que, pour les deux autres classes médicamenteuses, les points étaient beaucoup plus dispersés dans le graphique. En outre, la variation relative du RC était supérieure à la valeur seuil de 30 % dans la moitié des cas envisagés pour les benzodiazépines et dans 10 cas sur 12 pour les AINS.

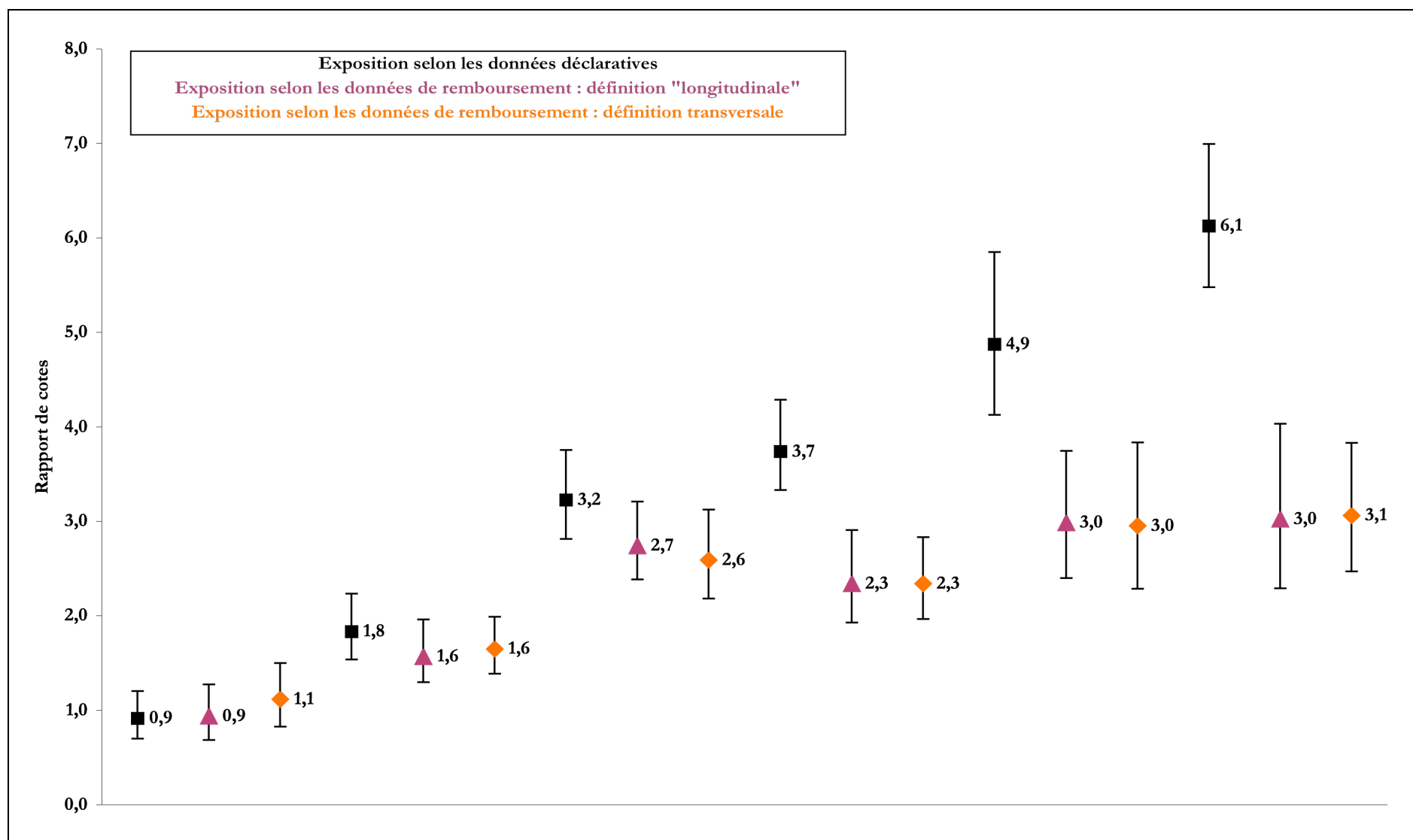


Figure 7- Influence de la source de données utilisée pour déterminer l'exposition médicamenteuse sur la mesure d'une association entre un médicament et un événement : le cas des benzodiazépines

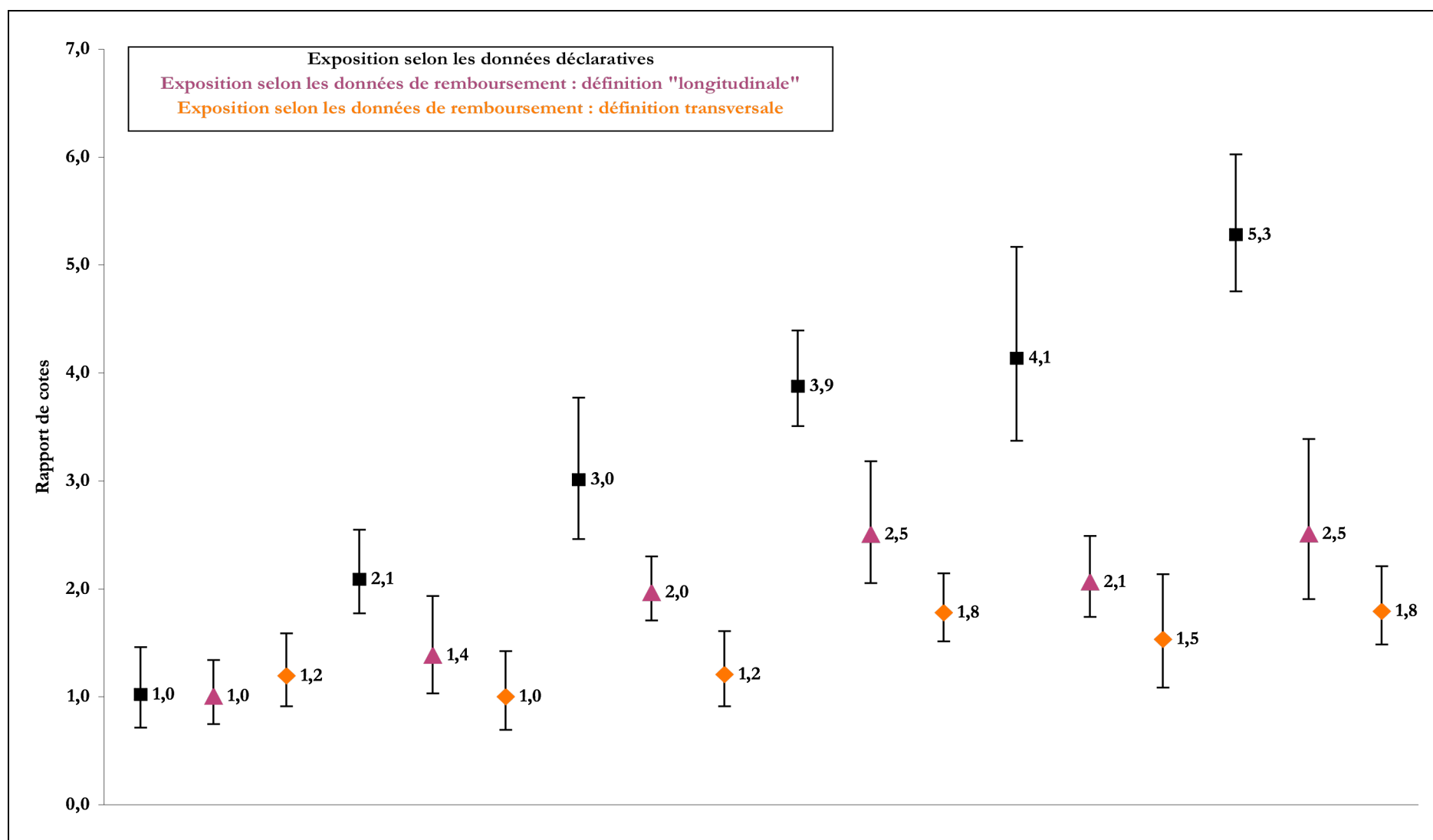


Figure 8- Influence de la source de données utilisée pour déterminer l'exposition médicamenteuse sur la mesure d'une association entre un médicament et un événement : le cas des anti-inflammatoires non stéroïdiens

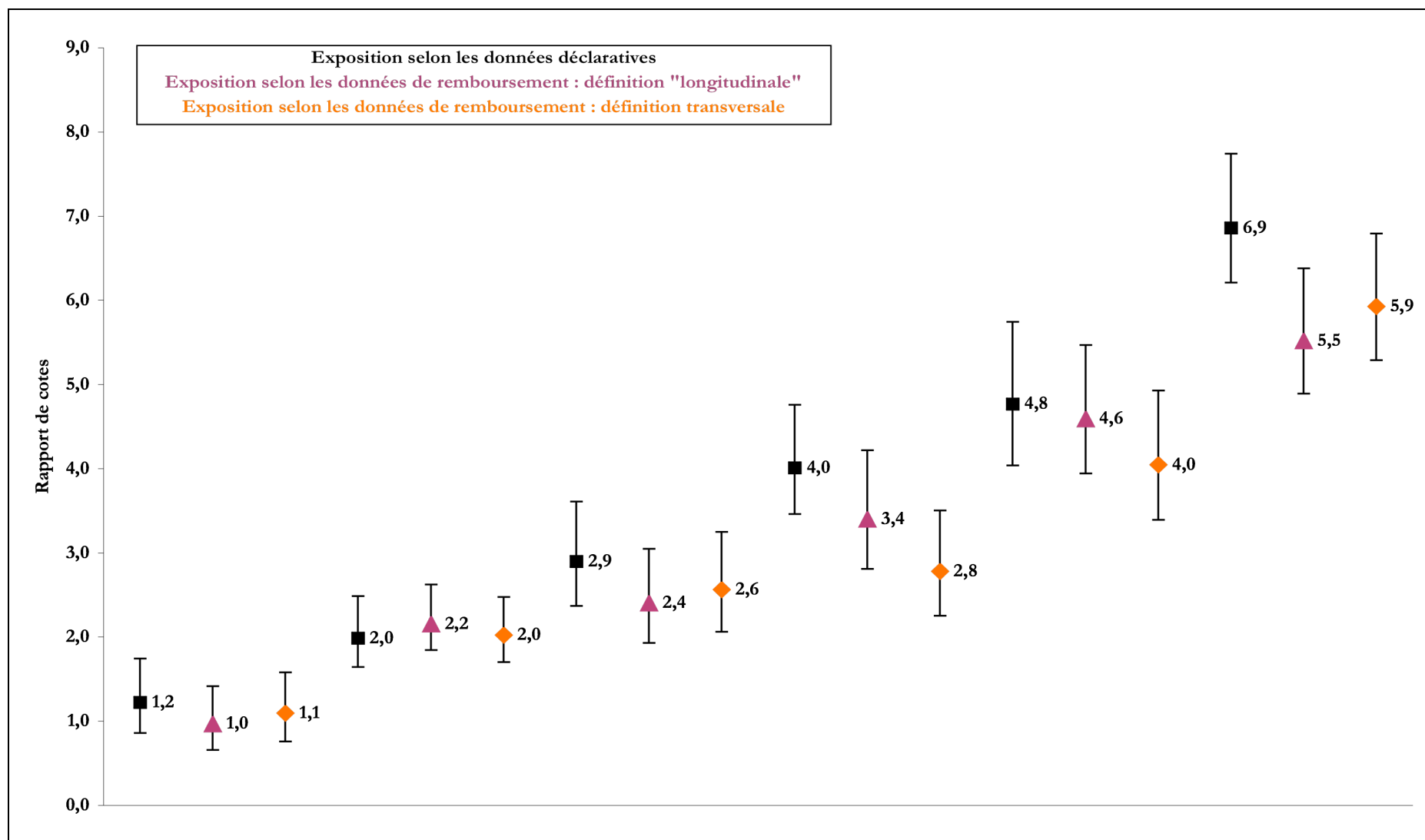


Figure 9- Influence de la source de données utilisée pour déterminer l'exposition médicamenteuse sur la mesure d'une association entre un médicament et un événement : le cas des médicaments cardiologiques

IV.4- DISCUSSION

IV.4.1- Interprétation des résultats

D'un point de vue méthodologique, il est, en premier lieu, important de noter que les travaux présentés dans cette seconde partie ne permettent pas, à proprement parler, de mettre en évidence l'existence d'un biais de classification. En épidémiologie, un biais se définit comme une erreur de raisonnement ou de procédure amenant à une représentation faussée de la réalité.³ Or, ni l'exposition appréciée à partir des interrogatoires des sujets, ni l'exposition mesurée à partir des données extraites des bases de remboursement de la CNAM-TS ne saurait exactement décrire l'exposition médicamenteuse réelle des sujets qui est finalement toujours méconnue. Néanmoins, les différences de mesure d'association entre l'exposition médicamenteuse et l'événement simulé constatées selon que l'exposition est issue des données déclaratives ou des données de remboursement apportent des pistes de réflexion quant à l'existence et l'importance d'un biais de classification pour chaque classe médicamenteuse d'intérêt.

Pour des médicaments d'utilisation chronique et régulière tels que les médicaments cardiologiques considérés (digitaliques, anti-arythmiques et dérivés nitrés), les faibles différences de mesure d'association observées suggèrent que l'exposition médicamenteuse reflétée par chacune des deux sources de données est probablement assez proche de l'exposition réelle et, par conséquent, que le biais de classification éventuellement introduit reste faible. Ainsi, pour ce type de médicaments, des données de consommation médicamenteuse recueillies, un jour donné, par interrogatoire des sujets pourraient être utilisées pour la conduite d'études étiologiques et ce, même si une longue période de temps sépare l'interrogatoire de la survenue de l'événement. L'utilisation de données de remboursement est également adaptée : elles représentent donc un potentiel intéressant pour de futures études pharmaco-épidémiologiques qui plus est, du fait de leur relative facilité d'accès par rapport à des données recueillies sur le terrain.

Des résultats très différents ont été obtenus pour les benzodiazépines et les AINS : les mesures d'association entre l'exposition médicamenteuse issue soit des données déclaratives, soit des données de remboursement et l'événement simulé variaient parfois dans de fortes proportions. Pour ce type de médicaments souvent consommés de façon irrégulière ou intermittente, les déclarations des sujets, un jour donné, peuvent être limitées pour estimer l'exposition médicamenteuse au cours de périodes de temps plus longues. Dans le cadre de cohortes à suivis

transversaux répétés comme la cohorte 3C, les hypothèses formulées vis-à-vis de l'exposition dans l'intervalle entre l'interrogatoire et un événement peuvent donc être erronées et conduire à des estimations biaisées de la force d'une association. Dans ce contexte précis, il semble légitime de considérer que l'exposition extraite des données continues de remboursement représente plus fidèlement la réalité de l'exposition médicamenteuse des sujets. Néanmoins, il convient de rester prudent quant à l'exploitation de ces données de remboursement et, en particulier, de tenir compte autant que possible des modalités d'utilisation connues du médicament dans le choix de la définition de l'exposition. Par exemple, s'agissant de la définition dite « longitudinale » dans nos travaux, le critère commun de 80 % de couverture appliqué aux trois classes médicamenteuses a été initialement choisi pour garantir un certain niveau d'exposition tout au long de la période. Or, et comme l'indiquent également les travaux de la première partie de la thèse, ce critère n'est pas le plus adapté pour décrire les modalités réelles d'utilisation de médicaments comme les benzodiazépines ou les AINS. Parmi les sujets échappant à cette définition, nombreux sont certainement ceux qui étaient effectivement exposés au cours de la période. Il serait, par conséquent, intéressant de renouveler de telles analyses en testant d'autres critères de définition de l'exposition selon les données de remboursement en termes de couverture comme de la combiner avec un délai maximal entre deux remboursements successifs.

Par ailleurs, si les benzodiazépines sont disponibles uniquement sur prescription médicale et inscrites sur les listes de médicaments remboursables, la limite des bases de données de remboursement liée à l'absence d'informations relatives aux médicaments obtenus sans ordonnance a pu affecter les résultats pour les AINS. Certains sujets ont pu être considérés à tort comme non exposés d'après cette source de données si la majorité de leur consommation d'AINS correspondait à une automédication avec des médicaments OTC. Dans une population âgée régulièrement suivie pour un certain nombre de pathologies chroniques, l'automédication et l'achat de médicaments sans ordonnance ne concernent probablement pas la majorité des sujets mais cette limite ne devrait pas être négligée dans le cadre d'études sur des populations plus jeunes.¹⁵³ Ceci étant, une étude récente a montré, par des analyses de sensibilité, que les bases de données informatisées de prescription pouvaient être utilisées pour l'estimation de l'association entre la prise d'AINS et la survenue d'un cancer colorectal malgré l'absence dans la mesure de l'exposition aux AINS de la part due aux médicaments OTC.¹⁵⁴

Pour terminer cette partie d'interprétation des résultats, un dernier point d'ordre méthodologique doit être abordé concernant non plus la mesure de l'exposition mais la génération de l'événement

simulé dans l'échantillon d'étude. Pour davantage de rigueur, l'attribution par le hasard du statut de cas ou de témoins aux sujets exposés ou non exposés de l'échantillon d'étude aurait théoriquement dû être répétée un nombre n de fois (entre 500 et 1000). Pour chaque situation déterminée par le type d'événement simulé et la source de données / définition utilisées pour estimer l'exposition médicamenteuse, n mesures de l'association entre l'exposition et l'événement auraient alors été obtenues et la mesure finalement retenue aurait été calculée comme la moyenne de ces n mesures. Ces améliorations méthodologiques sont actuellement en cours, elles devraient garantir une meilleure précision des résultats sans pour autant totalement les modifier. De plus, le caractère « théorique » de cette simulation devrait permettre dans l'avenir de tester d'autres scénaris en faisant notamment varier la fréquence de l'événement dans l'échantillon d'étude qui a été ici arbitrairement fixée à 10 %.

IV.4.2- Données de la littérature

L'importance du biais de classification et son impact sur les mesures de risque dans les études pharmaco-épidémiologiques ont été relativement peu étudiés. Ce type de biais entraînerait une sous-estimation de la force d'une association quand il est non différentiel (c'est-à-dire quand il affecte de la même manière les groupes comparés) et une sous- ou sur-estimation quand il est différentiel.⁵ Les premières études autour de cette problématique sont relativement anciennes et concernent des médicaments à visée hormonale (contraceptifs oraux ou traitement hormonal de la ménopause).^{8,9,155,156} Elles ont toutes été conduites dans le cadre d'études cas-témoins *ad hoc* où les informations concernant l'exposition médicamenteuse étaient recueillies auprès des femmes incluses et auprès de leurs médecins (interrogatoire du médecin et/ou consultation du dossier médical). La concordance entre les expositions mesurées à partir des deux sources de données était relativement bonne dans ces études et, le plus souvent, équivalente entre les cas et les témoins. Toutefois, le niveau de concordance pouvait être différent entre les deux groupes concernant certains détails tels que la durée totale d'utilisation.¹⁵⁵ Le même type de résultats a été également retrouvé dans une étude de méthodologie similaire à propos de l'exposition aux antidépresseurs.¹⁵⁷

Plus récemment, des études dont les objectifs correspondaient davantage à ceux des travaux de cette thèse méritent d'être citées.¹⁵⁸⁻¹⁶⁰ De la même manière que pour les médicaments cardiologiques, une étude menée sur l'association entre la prise de statines et un critère combiné regroupant l'infarctus du myocarde et l'accident vasculaire cérébral ou la mortalité a abouti à des

risques relatifs (RR) proches lorsque l'exposition aux statines était issue des déclarations des sujets à leur inclusion dans la cohorte (RR = 0,75 ; IC 95 % [0,41 - 1,37] pour le critère combiné et RR = 1,34 ; IC 95 % [0,66 - 2,74] pour la mortalité) ou mesurée de manière continue sur une période d'au moins deux ans à partir de données informatisées de délivrance de médicaments (RR = 0,63 ; IC 95 % [0,34 - 1,15] pour le critère combiné et RR = 0,91 ; IC 95 % [0,59 - 1,39] pour la mortalité).¹⁵⁸ Néanmoins, les auteurs mettaient en avant l'intérêt des données de délivrance avec lesquelles les mesures de risque se rapprochaient davantage de celles obtenues dans les essais cliniques en double insu avec tirage au sort ayant étudié l'efficacité des statines. Beiderbeck *et al.* ont également conclu que l'exposition issue de données déclaratives recueillies à l'inclusion dans une cohorte comme approximation de l'exposition chronique au médicament au cours du suivi pouvait introduire un biais dans la mesure d'un risque.¹⁵⁹ En effet, pour les antihypertenseurs de la classe des inhibiteurs calciques, la différence de mesure entre les données déclaratives et des données de délivrance pharmaceutique correspondait, dans leurs travaux, à un biais de classification de 0,73 [IC 95 % 0,52 - 1,02]. La différence entre les deux sources de données s'expliquait majoritairement par une forte proportion de sujets (15,5 %) n'ayant pas déclaré consommer d'inhibiteurs calciques à l'interrogatoire d'inclusion mais identifiés comme consommateurs au cours du suivi. Enfin, une étude de cohorte menée chez des sujets institutionnalisés afin d'évaluer le risque de chutes lié à la prise de benzodiazépines a retrouvé une association non significative (RR = 1,02 ; IC 95 % [0,95 - 1,10]) lorsque l'exposition était mesurée dans les 7 jours précédant l'entrée dans la cohorte et une association statistiquement significative (RR = 1,44 ; IC 95 % [1,33 - 1,56]) lorsque l'exposition était mesurée un jour donné au cours du suivi de la cohorte.¹⁶⁰ Toutefois, cette étude n'avait pas pour objectif de comparer deux sources de données pour la mesure de l'exposition aux benzodiazépines mais deux définitions à partir d'une même source (données d'administration quotidienne de médicaments).

À ce titre, en Amérique ou Europe du Nord, où des bases de données informatisées sont utilisées depuis plus longtemps pour la conduite d'études étiologiques en pharmaco-épidémiologie, un certain nombre de travaux se sont aussi intéressés à l'impact que peut avoir la définition de l'exposition médicamenteuse choisie sur les mesures de risque.^{50,161-165} Deux de ces études portaient spécifiquement sur le choix d'une fenêtre fixe d'exposition « à risque » pour les AINS.^{161,162} Les autres concernaient plutôt des médicaments d'utilisation chronique (traitement hormonal de la ménopause, antirétroviraux et hypocholestérolémiants) et exploraient notamment différentes méthodes pour considérer les espaces (« ou gaps ») entre prescriptions pour la construction d'épisodes continus d'utilisation.^{50,165}

V- CONCLUSION GÉNÉRALE ET PERSPECTIVES

Ces travaux sont, à notre connaissance, les premiers conduits en France ayant permis, d'une part, de comparer les mesures d'exposition médicamenteuse issues de données déclaratives recueillies par interrogatoire des sujets ou de données de remboursement extraites des bases de données informatisées de l'Assurance Maladie et, d'autre part, d'évaluer le retentissement du choix de l'une ou l'autre source de données sur les estimations de risque dans le cas d'études étiologiques. Un des intérêts majeurs de ces travaux est d'avoir pu explorer ces problématiques de mesure de l'exposition en pharmaco-épidémiologie à travers de nombreuses classes médicamenteuses. Il a pu être mis en évidence que les différences de mesure entre les deux sources de données et leur potentiel impact sur l'estimation d'un risque associé à l'exposition médicamenteuse varient considérablement selon la classe de médicaments et, plus particulièrement, selon ses modalités d'utilisation. Ainsi, outre un certain nombre d'autres facteurs (objectif, facilité d'accès aux sources de données, nombre de sujets à inclure, budget, délai...)⁶, la classe médicamenteuse d'intérêt dans une étude pharmaco-épidémiologique devrait rester un critère essentiel à prendre en compte pour le choix d'une source de données.

Pour des médicaments d'utilisation chronique et régulière chez le sujet âgé (par exemple, antihypertenseurs, digitaliques, anti-arythmiques, dérivés nitrés, hypolipémiants, médicaments hématologiques, hormones thyroïdiennes), les mesures d'exposition étaient très peu discordantes entre les deux sources de données quel que soit le contexte considéré (mesure un jour donné ou au cours d'une période de temps plus longue). D'après l'analyse menée sur les digitaliques, anti-arythmiques et dérivés nitrés, le biais de classification pouvant être lié au choix de l'une ou l'autre source serait faible. Pour ce premier type de médicaments, les résultats de ces travaux confirment l'utilité du patient comme source de données d'exposition médicamenteuse (y compris lorsqu'il est uniquement interrogé sur sa consommation médicamenteuse actuelle). Ils mettent aussi en évidence le potentiel représenté par les données de remboursement de l'Assurance Maladie à condition d'utiliser des critères de définition adaptés. À titre d'exemple, il a été constaté que, pour la mesure d'une exposition un jour donné, le choix de la fenêtre de temps à utiliser pour capturer les sujets exposés d'après les bases de données de remboursement conditionnait beaucoup les résultats.

Pour des médicaments pouvant être consommés chroniquement mais de façon irrégulière ou par intermittence (par exemple, AINS, benzodiazépines, antalgiques de palier 1 ou 2, suppléments minéraux...), la concordance entre les deux sources de données était moyenne voire mauvaise et les mesures d'association entre l'exposition issue soit des données déclaratives, soit des données

de remboursement et un événement variaient parfois dans de fortes proportions. Pour ce second type de médicaments, l'interrogatoire des sujets sur leur consommation médicamenteuse actuelle peut s'avérer limité pour refléter leur exposition médicamenteuse au cours de périodes de temps plus longues. Dans ce contexte précis, l'utilisation de données de remboursement semble d'un réel intérêt à condition de veiller, encore une fois, à choisir une définition de l'exposition pertinente et notamment adaptée aux modalités réelles d'utilisation du médicament.

En perspective à ces travaux, il serait intéressant de les renouveler dans d'autres contextes que celui d'une cohorte à suivis transversaux répétés, comme la cohorte 3C, et surtout sur d'autres populations de sujets. Même si le sujet âgé est particulièrement étudié en pharmaco-épidémiologie du fait de sa polymédication, pour des médicaments concernant également des sujets jeunes (AINS, antalgiques, psychotropes notamment), il est probable que des profils de consommation du médicament et d'utilisation du système d'Assurance Maladie différents modifient les résultats obtenus ici.

Des travaux comparables devraient pouvoir être réalisés prochainement sur une population différente, celle de l'étude CESIR-U (Combinaison d'Études sur la Santé et l'Insécurité Routière - Urgences) actuellement en cours. Cette étude conduite dans le cadre du Programme Hospitalier de Recherche Clinique (PHRC) inter-régional dans les Centres Hospitaliers Universitaires (CHU) de Bordeaux, Limoges et Toulouse, vise à décrire la consommation de médicaments à risque présumé pour la conduite automobile (médicaments à pictogrammes) par les conducteurs impliqués dans des accidents graves de la circulation, hospitalisés au moins 24 heures. Afin de déterminer l'exposition médicamenteuse antérieure à l'accident, il a été prévu de recueillir des données déclaratives auprès des sujets inclus au cours d'un interrogatoire en « face à face » et auprès de leur médecin traitant par questionnaire postal (taux de réponse actuel : 66,4 %) et également de demander une extraction de leurs données de remboursement pour les 6 mois précédant la date de l'accident. En outre, les sujets étant interrogés sur les dates de prise de chaque médicament déclaré être consommé, il devrait être possible d'étudier plus précisément les problématiques de fenêtres de temps à considérer dans les données de remboursement pour prédire l'exposition un jour donné telle que déclarée par les sujets.

De la même manière, afin d'améliorer la validité des définitions utilisées à partir de données de remboursement pour mesurer l'exposition médicamenteuse au cours de longues périodes de temps, contexte où ces données trouvent tout leur intérêt, il serait utile de poursuivre la

réalisation de travaux visant à décrire précisément les modalités d'utilisation des médicaments en fonction également du contexte clinique. Comme ceci a déjà été évoqué dans d'autres pays¹⁶⁶, les études dites « multi-sources », dont l'étude CADEUS¹⁶⁷ (COX-2 et AINS classiques : DEscription des UtiliSateurs) ou une étude à propos de la consommation de psychotropes chez les enfants¹⁶⁸ sont des exemples récents, représentent très certainement un outil méthodologique d'avenir en pharmaco-épidémiologie.

VI- RÉFÉRENCES BIBLIOGRAPHIQUES

- (1) Guillemot D, Maison P, Carbon C, *et al.* Enjeux de la pharmaco-épidémiologie en France. *Presse Med* 1999; 28: 493-9
- (2) Montastruc JL. Les cinq « trop » des essais cliniques : plaidoyer pour une pharmacoépidémiologie pour tous. *La Revue Prescrire* 2006; 26: 471
- (3) Bégaud B. Dictionnaire de pharmaco-épidémiologie. ARME-Pharmacovigilance Editions: Bordeaux, 1998
- (4) West SL, Strom BL, Poole C. Validity of pharmacoepidemiologic drug and diagnosis data. In *Pharmacoepidemiology*. Strom BL (eds). John Wiley & Sons: Chichester, 2005; 709-65
- (5) Copeland KT, Checkoway H, McMichael AJ, *et al.* Bias due to misclassification in the estimation of relative risk. *Am J Epidemiol* 1977; 105: 488-95
- (6) Depont-Hébert F. Le patient, le médecin et les bases de données de l'Assurance Maladie comme sources de données en pharmaco-épidémiologie. Thèse pour le Doctorat de l'Université Bordeaux 2, Sciences Biologiques et Médicales, Épidémiologie et Intervention en Santé Publique. Bordeaux: Université Victor Segalen Bordeaux 2, 2007
- (7) Moore N, Masson H, Noblet C, *et al.* What medicines do patients really take? A comparison of free form *vs.* oriented questionnaires. *Post Marketing Surveillance* 1993; 7: 355-62
- (8) Horwitz RI, Feinstein AR, Stremlau JR. Alternative data sources and discrepant results in case-control studies of estrogens and endometrial cancer. *Am J Epidemiol* 1980; 111: 389-94
- (9) Goodman MT, Nomura AM, Wilkens LR, *et al.* Agreement between interview information and physician records on history of menopausal estrogen use. *Am J Epidemiol* 1990; 131: 815-25
- (10) Gama H, Correia S, Lunet N. Questionnaire design and the recall of pharmacological treatments: a systematic review. *Pharmacoepidemiol Drug Saf* 2009; 18: 175-87
- (11) Paganini-Hill A, Ross RK. Reliability of recall of drug usage and other health-related information. *Am J Epidemiol* 1982; 116: 114-22
- (12) Harlow SD, Linet MS. Agreement between questionnaire data and medical records. The evidence for accuracy of recall. *Am J Epidemiol* 1989; 129: 233-48
- (13) Klungel OH, de Boer A, Paes AH, *et al.* Influence of question structure on the recall of self-reported drug use. *J Clin Epidemiol* 2000; 53: 273-7
- (14) Ademi Z, Turunen JH, Kauhanen J, *et al.* A comparison of three questionnaire-based measures of analgesic use over 11 years in adult males: a retrospective analysis from a prospective population-based cohort study. *Clin Ther* 2007; 29: 529-34

- (15) Mitchell AA, Cottler LB, Shapiro S. Effect of questionnaire design on recall of drug exposure in pregnancy. *Am J Epidemiol* 1986; 123: 670-6
- (16) West SL, Savitz DA, Koch G, *et al.* Recall accuracy for prescription medications: self-report compared with database information. *Am J Epidemiol* 1995; 142: 1103-12
- (17) Feldman Y, Koren G, Mattice K, *et al.* Determinants of recall and recall bias in studying drug and chemical exposure in pregnancy. *Teratology* 1989; 40: 37-45
- (18) Landry JA, Smyer MA, Tubman JG, *et al.* Validation of two methods of data collection of self-reported medicine use among the elderly. *Gerontologist* 1988; 28: 672-6
- (19) West SL, Savitz DA, Koch G, *et al.* Demographics, health behaviors, and past drug use as predictors of recall accuracy for previous prescription medication use. *J Clin Epidemiol* 1997; 50: 975-80
- (20) Caskie GI, Willis SL. Congruence of self-reported medications with pharmacy prescription records in low-income older adults. *Gerontologist* 2004; 44: 176-85
- (21) Kimmel SE, Lewis JD, Jaskowiak J, *et al.* Enhancement of medication recall using medication pictures and lists in telephone interview. *Pharmacoepidemiol Drug Saf* 2003; 12: 1-8
- (22) Beresford SA, Coker AL. Pictorially assisted recall of past hormone use in case-control studies. *Am J Epidemiol* 1989; 130: 202-5
- (23) Norell SE, Boethius G, Persson I. Oral contraceptive use: interview data versus pharmacy records. *Int J Epidemiol* 1998; 27: 1033-7
- (24) Van Ganse E. Utilisation de données informatisées en pharmaco-épidémiologie. *Thérapie* 2000; 55: 123-6
- (25) Hennessy S. Use of health care databases in pharmacoepidemiology. *Basic Clin Pharmacol Toxicol* 2006; 98: 311-3
- (26) Saunders KW, Stergachis A, Von Korff M. Group Health Cooperative of Puget Sound. In *Pharmacoepidemiology*. Strom BL (eds). John Wiley & Sons: Chichester, 1994; 171-85
- (27) Christensen DB, Williams B, Goldberg HI, *et al.* Comparison of prescription and medical records in reflecting patient antihypertensive drug therapy. *Ann Pharmacother* 1994; 28: 99-104
- (28) Carson JL, Strom BL. Medicaid databases. In *Pharmacoepidemiology*. Strom BL (eds). John Wiley & Sons: Chichester, 1994; 199-216
- (29) McKenzie DA, Semradek J, McFarland BH, *et al.* The validity of medicaid pharmacy claims for estimating drug use among elderly nursing home residents: The Oregon experience. *J Clin Epidemiol* 2000; 53: 1248-57

- (30) Friedman GD. Kaiser Permanente medical care program: Northern California and other regions. In *Pharmacoepidemiology*. Strom BL (eds). John Wiley & Sons: Chichester, 1994; 187-97
- (31) Strand LM, Downey W. Health databases in Saskatchewan. In *Pharmacoepidemiology*. Strom BL (eds). John Wiley & Sons: Chichester, 1994; 217-29
- (32) Tamblyn R, Lavoie G, Petrella L, *et al.* The use of prescription claims databases in pharmacoepidemiological research: the accuracy and comprehensiveness of the prescription claims database in Quebec. *J Clin Epidemiol* 1995; 48: 999-1009
- (33) Leufkens HG, Urquhart J. Automated record linkage and other approaches to pharmacoepidemiology in The Netherlands. In *Pharmacoepidemiology*. Strom BL (eds). John Wiley & Sons: Chichester, 1994; 231-44
- (34) Boethius G, Wiman F. Recording of drug prescriptions in the county of Jämtland, Sweden. I. Methodological aspects. *Eur J Clin Pharmacol* 1977; 12: 31-5
- (35) García-Rodríguez LA, Perez Gutthann S. Use of the UK General Practice Research Database for pharmacoepidemiology. *Br J Clin Pharmacol* 1998; 45: 419-25
- (36) Lewis JD, Schinnar R, Bilker WB, *et al.* Validation studies of the health improvement network (THIN) database for pharmacoepidemiology research. *Pharmacoepidemiol Drug Saf* 2007; 16: 393-401
- (37) Monster TB, Janssen WM, de Jong PE, *et al.* Pharmacy data in epidemiological studies: an easy to obtain and reliable tool. *Pharmacoepidemiol Drug Saf* 2002; 11: 379-84
- (38) Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol* 2005; 58: 323-37
- (39) Friedman GD, Flick ED, Udaltsova N, *et al.* Screening statins for possible carcinogenic risk: up to 9 years of follow-up of 361,859 recipients. *Pharmacoepidemiol Drug Saf* 2008; 17: 27-36
- (40) Sukel MP, Breekveldt-Postma NS, Erkens JA, *et al.* Incidence of cardiovascular events in breast cancer patients receiving chemotherapy in clinical practice. *Pharmacoepidemiol Drug Saf* 2008; 17: 125-34
- (41) Tamim HM, Hanley JA, Hajeer AH, *et al.* Risk of breast cancer in relation to antibiotic use. *Pharmacoepidemiol Drug Saf* 2008; 17: 144-50
- (42) Holst L, Nordeng H, Haavik S. Use of herbal drugs during early pregnancy in relation to maternal characteristics and pregnancy outcome. *Pharmacoepidemiol Drug Saf* 2008; 17: 151-9

- (43) Ibáñez L, Sabaté M, Ballarín E, *et al.* Use of granulocyte colony-stimulating factor (G-CSF) and outcome in patients with non-chemotherapy agranulocytosis. *Pharmacoepidemiol Drug Saf* 2008; 17: 224-8
- (44) Perreault S, Dragomir A, Blais L, *et al.* Population-based study of the effectiveness of bone-specific drugs in reducing the risk of osteoporotic fracture. *Pharmacoepidemiol Drug Saf* 2008; 17: 248-59
- (45) Hammad TA, Graham DJ, Staffa JA, *et al.* Onset of acute myocardial infarction after use of non-steroidal anti-inflammatory drugs. *Pharmacoepidemiol Drug Saf* 2008; 17: 315-21
- (46) Kelly JP, Coogan P, Strom BL, *et al.* Lung cancer and regular use of aspirin and nonaspirin nonsteroidal anti-inflammatory drugs. *Pharmacoepidemiol Drug Saf* 2008; 17: 322-7
- (47) Lewis JD, Strom BL, Localio AR, *et al.* Moderate and high affinity serotonin reuptake inhibitors increase the risk of upper gastrointestinal toxicity. *Pharmacoepidemiol Drug Saf* 2008; 17: 328-35
- (48) Derijks HJ, Heerdink ER, De Koning FH, *et al.* The association between antidepressant use and hypoglycaemia in diabetic patients: a nested case-control study. *Pharmacoepidemiol Drug Saf* 2008; 17: 336-44
- (49) Said Q, Gutterman EM, Kim MS, *et al.* Somnolence effects of antipsychotic medications and the risk of unintentional injury. *Pharmacoepidemiol Drug Saf* 2008; 17: 354-64
- (50) Nielsen LH, Løkkegaard E, Andreassen AH, *et al.* Using prescription registries to define continuous drug use: how to fill gaps between prescriptions. *Pharmacoepidemiol Drug Saf* 2008; 17: 384-8
- (51) Guiguet M, Mary-Krause M, Ename B, *et al.* Influence of control selection in nested case-control studies: the example of exposure to antiretroviral treatment (ART) and the risk of myocardial infarction (MI) in the French Hospital Database on HIV (FHDH-ANRS CO4). *Pharmacoepidemiol Drug Saf* 2008; 17: 468-74
- (52) Russo A, Autelitano M, Bisanti L. Spironolactone and gastrointestinal bleeding: a population based study. *Pharmacoepidemiol Drug Saf* 2008; 17: 495-500
- (53) Meropol SB, Chan KA, Chen Z, *et al.* Adverse events associated with prolonged antibiotic use. *Pharmacoepidemiol Drug Saf* 2008; 17: 523-32
- (54) Delaney JA, Moodie EE, Suissa S. Validating the effects of drug treatment on blood pressure in the General Practice Research Database. *Pharmacoepidemiol Drug Saf* 2008; 17: 535-45
- (55) Knol MJ, Derijks HJ, Geerlings MI, *et al.* Influence of antidepressants on glycaemic control in patients with diabetes mellitus. *Pharmacoepidemiol Drug Saf* 2008; 17: 577-86

- (56) Cunnington M, Webb D, Qizilbash N, *et al.* Risk of ischaemic cardiovascular events from selective cyclooxygenase-2 inhibitors in osteoarthritis. *Pharmacoepidemiol Drug Saf* 2008; 17: 601-8
- (57) Weiner MG, Xie D, Tannen RL. Replication of the Scandinavian Simvastatin Survival Study using a primary care medical record database prompted exploration of a new method to address unmeasured confounding. *Pharmacoepidemiol Drug Saf* 2008; 17: 661-70
- (58) Tannen RL, Weiner MG, Xie D. Replicated studies of two randomized trials of angiotensin-converting enzyme inhibitors: further empiric validation of the 'prior event rate ratio' to adjust for unmeasured confounding by indication. *Pharmacoepidemiol Drug Saf* 2008; 17: 671-85
- (59) Haukka J, Tiihonen J, Härkänen T, *et al.* Association between medication and risk of suicide, attempted suicide and death in nationwide cohort of suicidal patients with schizophrenia. *Pharmacoepidemiol Drug Saf* 2008; 17: 686-96
- (60) Margolis DJ, Hoffstad O, Strom BL. Association between serious ischemic cardiac outcomes and medications used to treat diabetes. *Pharmacoepidemiol Drug Saf* 2008; 17: 753-9
- (61) Walker AM, Koro CE, Landon J. Coronary heart disease outcomes in patients receiving antidiabetic agents in the PharMetrics database 2000-2007. *Pharmacoepidemiol Drug Saf* 2008; 17: 760-8
- (62) Källén B, Olausson PO. Maternal use of selective serotonin re-uptake inhibitors and persistent pulmonary hypertension of the newborn. *Pharmacoepidemiol Drug Saf* 2008; 17: 801-6
- (63) Yang YX, Hennessy S, Probert K, *et al.* Chronic statin therapy and the risk of colorectal cancer. *Pharmacoepidemiol Drug Saf* 2008; 17: 869-76
- (64) Skurtveit S, Furu K, Bramness JG, *et al.* Benzodiazepine use in all alcohol consumers predicts use of opioids in patients 20 years later--a follow-up study of 13,390 men and women aged 40-42 years. *Pharmacoepidemiol Drug Saf* 2008; 17: 926-33
- (65) García-Rodríguez LA, Massó-González EL, Wallander MA, *et al.* The safety of rosuvastatin in comparison with other statins in over 100,000 statin users in UK primary care. *Pharmacoepidemiol Drug Saf* 2008; 17: 943-52
- (66) García-Rodríguez LA, González-Pérez A, Stang MR, *et al.* The safety of rosuvastatin in comparison with other statins in over 25,000 statin users in the Saskatchewan Health Databases. *Pharmacoepidemiol Drug Saf* 2008; 17: 953-61

- (67) Patterson MK, Castellsague J, Walker AM. Hospitalization for peptic ulcer and bleeding in users of selective COX-2 inhibitors and nonselective NSAIDs with special reference to celecoxib. *Pharmacoepidemiol Drug Saf* 2008; 17: 982-8
- (68) Koro CE, Fu Q, Stender M. An assessment of the effect of thiazolidinedione exposure on the risk of myocardial infarction in type 2 diabetic patients. *Pharmacoepidemiol Drug Saf* 2008; 17: 989-96
- (69) Assimes TL, Elstein E, Langleben A, *et al.* Long-term use of antihypertensive drugs and risk of cancer. *Pharmacoepidemiol Drug Saf* 2008; 17: 1039-49
- (70) Martin K, Latry P, Moride Y, *et al.* Où trouver des données pour la pharmaco-épidémiologie en France ? Quelques idées. *Rev Épidémiol Santé Publique* 2007; 55: 58-63
- (71) Goldberg M, Leclerc A, Chastang JF, *et al.* Mise en place d'une cohorte épidémiologique à Electricité de France - Gaz de France. Recrutement des volontaires. Principales caractéristiques de l'échantillon. *Rev Épidémiol Santé Publique* 1990; 38: 265-8 et 378-80
- (72) Dartigues JF, Gagnon M, Michel P, *et al.* Le programme de recherche Paquid sur l'épidémiologie de la démence. Méthodes et résultats initiaux. *Rev Neurol* 1991; 147: 225-30
- (73) Clavel-Chapelon F, van Liere MJ, Giubout C, *et al.* E3N, a French cohort study on cancer risk factors. E3N Group. Etude Epidémiologique auprès de femmes de l'Education Nationale. *Eur J Cancer Prev* 1997; 6: 473-8
- (74) 3C study group. Vascular factors and risk of dementia: design of the Three-City Study and baseline characteristics of the study population. *Neuroepidemiology* 2003; 22: 316-25
- (75) Weill A, Ricordeau P, Vallier N, *et al.* Analyse des prestations remboursées par l'Assurance Maladie : matériel et méthodes. *Diabetes Metab* 2000; 26: 49-54
- (76) Micallef J, Pradel V, Thirion X, *et al.* Utilisation des bases de données de l'Assurance Maladie par les Centres d'Évaluation et d'Information sur la Pharmacodépendance : exemples, intérêts et perspectives. *Thérapie* 2004; 59: 581-8
- (77) De Roquefeuil L, Fivaz C, Le Guen C, *et al.* L'utilisation des systèmes d'information de l'Assurance Maladie pour l'analyse des pratiques de prescription : quelques exemples. [en ligne]. Disponible à l'adresse : <http://www.ameli.fr/l-assurance-maladie/statistiques-et-publications/points-de-repere/n-4-systemes-d-information.php>. Consulté le 5 novembre 2009
- (78) Fender P, Weill A. Épidémiologie, santé publique et bases de données médico-tarifaires. *Rev Épidémiol Santé Publique* 2004; 52: 113-7

- (79) Verger P, Clavaud H, Bidaud S, *et al.* Les prescriptions d'antalgiques, d'anti-inflammatoires ou le nombre d'actes radiologiques sont des facteurs prédictifs d'un traitement par antidépresseur. *Presse Med* 2007; 36: 50-6
- (80) Lugardon S, Roussel H, Sciortino V, *et al.* Triptan use and risk of cardiovascular events: a nested-case-control study from the French health system database. *Eur J Clin Pharmacol* 2007; 63: 801-7
- (81) Goldberg M, Zins M, Lert F. Rapport pour l'Institut de recherche en santé publique. Proposition pour la création d'une PLATE-FORME scientifique et technique pluri-organismes pour l'aide à la gestion de cohortes et de grandes enquêtes épidémiologiques. Le projet *Plastico*. [en ligne]. Disponible à l'adresse : <http://www.constances.fr/PLASTICO-IRESP.pdf>. Consulté le 5 novembre 2009
- (82) Tuppin P, Kusnik-Joinville O, Weill A, *et al.* Primary health care use and reasons for hospital admissions in dementia patients in France: database study for 2007. *Dement Geriatr Cogn Disord* 2009; 28: 225-32
- (83) Tuppin P, Neumann A, Danchin N, *et al.* Combined secondary prevention after hospitalization for myocardial infarction in France: analysis from a large administrative database. *Arch Cardiovasc Dis* 2009; 102: 279-92
- (84) Lacroix I, Hurault C, Sarramon MF, *et al.* Prescription of drugs during pregnancy: a study using EFEMERIS, the new French database. *Eur J Clin Pharmacol* 2009; 65: 839-46
- (85) Furher R, Rouillon F. La version française de l'échelle CES-D (Center for Epidemiologic Studies - Depression Scale). Description et traduction de l'échelle d'auto-évaluation. *Psychiatrie et psychobiologie* 1989; 4: 163-6
- (86) Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist* 1969; 9: 179-86
- (87) Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975; 12: 189-98
- (88) World Health Organization. Collaborating Centre for Drug Statistics Methodology. ATC Index with DDDs. World Health Organization, Oslo; 2000
- (89) Enlund H, Tuomilehto J, Turakka H. Patient report validated against prescription records for measuring use of and compliance with antihypertensive drugs. *Acta Med Scand* 1981; 209: 271-5
- (90) Van den Brandt PA, Petri H, Dorant E, *et al.* Comparison of questionnaire information and pharmacy data on drug use. *Pharm Weekbl Sci* 1991; 13: 91-6

- (91) Johnson RE, Vollmer WM. Comparing sources of drug data about the elderly. *J Am Geriatr Soc* 1991; 39: 1079-84
- (92) de Jong-van den Berg LT, Waardenburg CM, Haaijer-Ruskamp FM, *et al.* Drug use in pregnancy: a comparative appraisal of data collecting methods. *Eur J Clin Pharmacol* 1993; 45: 9-14
- (93) Heerdink ER, Leufkens HG, Koppedraaijer C, *et al.* Information on drug use in the elderly: a comparison of pharmacy, general-practitioner and patient data. *Pharm World Sci* 1995; 17: 20-4
- (94) Lau HS, de Boer A, Beuning KS, *et al.* Validation of pharmacy records in drug exposure assessment. *J Clin Epidemiol* 1997; 50: 619-25
- (95) Saunders K, Simon G, Bush T, *et al.* Assessing the feasibility of using computerized pharmacy refill data to monitor antidepressant treatment on a population basis: a comparison of automated and self-report data. *J Clin Epidemiol* 1998; 51: 883-90
- (96) Sjahid SI, van der Linden PD, Stricker BH. Agreement between the pharmacy medication history and patient interview for cardiovascular drugs: the Rotterdam elderly study. *Br J Clin Pharmacol* 1998; 45: 591-5
- (97) Klungel OH, de Boer A, Paes AH, *et al.* Agreement between self-reported antihypertensive drug use and pharmacy records in a population-based study in The Netherlands. *Pharm World Sci* 1999; 21: 217-20
- (98) Olesen C, Søndergaard C, Thrane N, *et al.* Do pregnant women report use of dispensed medications? *Epidemiology* 2001; 12: 497-501
- (99) Metlay JP, Hardy C, Strom BL. Agreement between patient self-report and a Veterans Affairs national pharmacy database for identifying recent exposures to antibiotics. *Pharmacoepidemiol Drug Saf* 2003; 12: 9-15
- (100) Al-Windi A. The validity of a questionnaire on medicines used in health care practice: comparison of a questionnaire and computerized medical record survey. *Eur J Clin Pharmacol* 2003; 59: 321-9
- (101) Strom BL, Schinnar R. An interview strategy was critical for obtaining valid information on the use of hormone replacement therapy. *J Clin Epidemiol* 2004; 57: 1210-3
- (102) Boudreau DM, Daling JR, Malone KE, *et al.* A validation study of patient interview data and pharmacy records for antihypertensive, statin, and antidepressant medication use among older women. *Am J Epidemiol* 2004; 159: 308-17

- (103) Løkkegaard EL, Johnsen SP, Heitmann BL, *et al.* The validity of self-reported use of hormone replacement therapy among Danish nurses. *Acta Obstet Gynecol Scand* 2004; 83: 476-81
- (104) Caskie GI, Willis SL, Warner Schaie K, *et al.* Congruence of medication information from a brown bag data collection and pharmacy records: findings from the Seattle longitudinal study. *Exp Aging Res* 2006; 32: 79-103
- (105) Glintborg B, Poulsen HE, Dalhoff KP. The use of nationwide on-line prescription records improves the drug history in hospitalized patients. *Br J Clin Pharmacol* 2007; 65: 265-9
- (106) Haukka J, Suvisaari J, Tuulio-Henriksson A, *et al.* High concordance between self-reported medication and official prescription database information. *Eur J Clin Pharmacol* 2007; 63: 1069-74
- (107) Nielsen MW, Søndergaard B, Kjølner M, *et al.* Agreement between self-reported data on medicine use and prescription records vary according to method of analysis and therapeutic group. *J Clin Epidemiol* 2008; 61: 919-24
- (108) Herings RM, Stricker BH, Leufkens HG, *et al.* Public health problems and the rapid estimation of the size of the population at risk. Torsades de pointes and the use of terfenadine and astemizole in The Netherlands. *Pharm World Sci* 1993; 15: 212-8
- (109) Chastagner M, Gault M, Coudène JP, *et al.* Respect des contre-indications à la prescription de quinolones en médecine ambulatoire. *Rev Med Ass Maladie* 2000; 31: 27-32
- (110) Hérique A, Proy M-O, Truchi L, *et al.* Prescription des médicaments chez la femme enceinte : analyse descriptive dans le département de l'Aude. *Rev Med Ass Maladie* 2001; 32: 25-30
- (111) Parjoie R, Cherrier C, Hérique A, *et al.* Les antibiotiques. Analyse descriptive des prescriptions dans la région Nord-Est en 1998. *Rev Med Ass Maladie* 2001; 32: 227-34
- (112) Berchery D, Roussel H, Bourrel R, *et al.* Utilisation des héparines en médecine de ville : données de l'Assurance Maladie de Midi-Pyrénées. *Thérapie* 2003; 58: 431-7
- (113) Perearnau P, Hessemann H, Weill G. Le recours aux vasoconstricteurs dans les affections ORL bénignes de l'enfant : une pratique banale aux risques sous-estimés. *Rev Med Ass Maladie* 2003; 34: 33-40
- (114) Lecadet J, Vidal P, Baris B, *et al.* Médicaments psychotropes : consommation et pratiques de prescription en France métropolitaine. I. Données nationales, 2000. *Rev Med Ass Maladie* 2003; 34: 75-84

- (115) Deprez Ph-H, Chinaud F, Clech S, *et al.* La population traitée par des médicaments de la classe des antiasthmatiques en France métropolitaine : données du régime général de l'Assurance Maladie, 2000. *Rev Med Ass Maladie* 2004; 35: 3-11
- (116) Schick J, Willem P, Weill G. Prescriptions de médicaments psychotropes aux enfants et adolescents, région Alsace, 2002. *Rev Med Ass Maladie* 2004; 35: 13-8
- (117) Allenbach D, Montagnier B, Souche A, *et al.* La population traitée par médicaments antituberculeux en 2003 : les données du régime général de l'Assurance Maladie. *Rev Med Ass Maladie* 2004; 35: 223-32
- (118) Damagnez M, Magnouat A, Niezborala J, *et al.* Médicaments contre-indiqués chez l'enfant : analyse des prescriptions en médecine de ville dans le département de l'Aude. *Thérapie* 2005; 60: 507-13
- (119) Karsenty D, Bourdel H, Cherrier C, *et al.* La population traitée par médicaments antiépileptiques en France métropolitaine en 2000 : les données du régime général de l'Assurance Maladie. *Rev Med Ass Maladie* 2005; 36: 123-32
- (120) Lévy P, Fanello S, Pivette J, *et al.* Anti-inflammatoires non stéroïdiens et risques iatrogènes : analyse des données de l'Assurance Maladie. *Rev Med Ass Maladie* 2005; 36: 153-61
- (121) Levy-Neumand O, Gillet A, Christophe B, *et al.* Migraine : traitement de fond et abus médicamenteux. *Prat Organ Soins* 2006; 37: 325-32
- (122) Simiand-Erdociain E, Lapeyre-Mestre M, Bagheri-Charabiani H, *et al.* Drug consumption in a very elderly community-dwelling population. *Eur J Clin Pharmacol* 2001; 57: 691-2
- (123) Pradel V, Lapierre V, Thirion X, *et al.* Étude de la « méthadone-relais » dans le département des Bouches-du-Rhône. *Bulletin Épidémiologique Hebdomadaire* 2001; 32: 151-3
- (124) Thirion X, Lapierre V, Micallef J, *et al.* Buprenorphine prescription by general practitioners in a French region. *Drug Alcohol Depend* 2002; 65: 197-204
- (125) Frauger E, Thirion X, Chanut C, *et al.* Détournement d'usage du trihexyphénidyle (Artane[®], Parkinane[®]) : tendances récentes. *Thérapie* 2003; 58: 541-7
- (126) Moore N, Diris H, Martin K, *et al.* Profils d'achats des AINS à partir des données d'Assurance Maladie en France. *Thérapie* 2004; 59: 541-6
- (127) Martin K, Bégaud B, Verdoux H, *et al.* Patterns of risperidone prescription: a utilization study in South-West France. *Acta Psychiatr Scand* 2004; 109: 202-6
- (128) Fourier-Réglat A, de Baillencourt S, Molimard M, *et al.* Prescription du célécoxib (Célébrex[®]) lors de sa commercialisation : une étude conduite à partir de la base de données du régime général de l'Assurance Maladie de la région Aquitaine. *Thérapie* 2005; 60: 25-9

- (129) Mancini J, Thirion X, Masut A, *et al.* Anxiolytics, hypnotics, and antidepressants dispensed to adolescents in a French region in 2002. *Pharmacoepidemiol Drug Saf* 2006; 15: 494-503
- (130) Lugardon S, Roussel H, Bourrel R, *et al.* Profil des utilisateurs d'anti-inflammatoires non stéroïdiens. *Thérapie* 2006; 61: 109-14
- (131) Huot L, Couris CM, Tainturier V, *et al.* Trends in HRT and anti-osteoporosis medication prescribing in a European population after the WHI study. *Osteoporos Int* 2008; 19: 1047-54
- (132) Hurault C, Lacroix I, Bourrel R, *et al.* Prescriptions et délivrances d'anti-inflammatoires non stéroïdiens après le 6^e mois de grossesse : impact d'une lettre d'information envoyée par les autorités de santé. *Presse Med* 2008; 37: 767-74
- (133) Brefel-Courbon C, Grolleau S, Thalamas C, *et al.* Comparison of chronic analgesic drugs prevalence in Parkinson's disease, other chronic diseases and the general population. *Pain* 2009; 141: 14-8
- (134) Schwarz A, Faber U, Borner K, *et al.* Reliability of drug history in analgesic users. *Lancet* 1984; 2: 1163-4
- (135) Moore N, Pierfitte C, Pehourcq F, *et al.* Comparison of patient questionnaires, medical records, and plasma assays in assessing exposure to benzodiazepines in elderly subjects. *Clin Pharmacol Ther* 2001; 69: 445-50
- (136) Pierfitte C, Macouillard G, Thicoïpe M, *et al.* Benzodiazepines and hip fractures in elderly people: case-control study. *BMJ* 2001; 322: 704-8
- (137) Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159-74
- (138) Wacholder S, Armstrong B, Hartge P. Validation studies using an alloyed gold standard. *Am J Epidemiol* 1993; 137: 1251-8
- (139) Weiss NS. Studies of diagnostic and screening tests. In *Modern Epidemiology*. Rothman KJ, Greenland S (eds). Lippincott-Raven: Philadelphia, 1998; 520-4
- (140) Salles-Montaudon N, Fourier A, Dartigues JF, *et al.* Évolution des traitements médicamenteux des personnes âgées vivant à domicile. *Rev Med Interne* 2000; 21: 664-71
- (141) Bourdel-Marchasson I, Dubroca B, Letenneur L, *et al.* Incidence and predictors of drug-treated diabetes in elderly French subjects. The PAQUID Epidemiological Survey. *Diabet Med* 2000; 17: 675-81
- (142) Fourier A, Letenneur L, Dartigues JF, *et al.* Benzodiazepine use in an elderly community-dwelling population. Characteristics of users and factors associated with subsequent use. *Eur J Clin Pharmacol* 2001; 57: 419-25

- (143) Soudry A, Dufouil C, Ritchie K, *et al.* Factors associated with changes in antidepressant use in a community-dwelling elderly cohort: the Three-City Study. *Eur J Clin Pharmacol* 2008; 64: 51-9
- (144) Fourrier A, Letenneur L, Bégaud B, *et al.* Nonsteroidal antiinflammatory drug use and cognitive function in the elderly: inconclusive results from a population-based cohort study. *J Clin Epidemiol* 1996; 49: 1201
- (145) Lagnaoui R, Bégaud B, Moore N, *et al.* Benzodiazepine use and risk of dementia: a nested case-control study. *J Clin Epidemiol* 2002; 55: 314-8
- (146) Dartigues JF, Carcaillon L, Helmer C, *et al.* Vasodilators and nootropics as predictors of dementia and mortality in the PAQUID cohort. *J Am Geriatr Soc* 2007; 55: 395-9
- (147) Pariente A, Dartigues JF, Bénichou J, *et al.* Benzodiazepines and injurious falls in community dwelling elders. *Drugs Aging* 2008; 25: 61-70
- (148) Berdot S, Bertrand M, Dartigues JF, *et al.* Inappropriate medication use and risk of falls - a prospective study in a large community-dwelling elderly cohort. *BMC Geriatr* 2009; 9: 30
- (149) Lagnaoui R, Depont F, Fourrier A, *et al.* Patterns and correlates of benzodiazepine use in the French general population. *Eur J Clin Pharmacol* 2004; 60: 523-9
- (150) Lechevallier N, Fourrier A, Berr C. Utilisation de benzodiazépines chez le sujet âgé : données de la cohorte EVA. *Rev Épidémiol Santé Publique* 2003; 51: 317-26
- (151) Fosbøl EL, Gislason GH, Jacobsen S, *et al.* The pattern of use of non-steroidal anti-inflammatory drugs (NSAIDs) from 1997 to 2005: a nationwide study on 4.6 million people. *Pharmacoepidemiol Drug Saf* 2008; 17: 822-33
- (152) Simpson SH, Eurich DT, Majumdar SR, *et al.* A meta-analysis of the association between adherence to drug therapy and mortality. *BMJ* 2006; 333: 15
- (153) Colomes M, Rispail Y, Berlan M, *et al.* Consommation médicamenteuse d'une population de retraités. *Thérapie* 1990; 45: 321-4
- (154) Yood MU, Campbell UB, Rothman KJ, *et al.* Using prescription claims data for drugs available over-the-counter (OTC). *Pharmacoepidemiol Drug Saf* 2007; 16: 961-8
- (155) Stolley PD, Tonascia JA, Sartwell PE, *et al.* Agreement rates between oral contraceptive users and prescribers in relation to drug use histories. *Am J Epidemiol* 1978; 107: 226-35
- (156) Nischan P, Ebeling K, Thomas DB, *et al.* Comparison of recalled and validated oral contraceptive histories. *Am J Epidemiol* 1993; 138: 697-703
- (157) Cotterchio M, Kreiger N, Darlington G, *et al.* Comparison of self-reported and physician-reported antidepressant medication use. *Ann Epidemiol* 1999; 9: 283-9

- (158) Maitland-van der Zee AH, Klungel OH, Stricker BH, *et al.* Comparison of two methodologies to analyze exposure to statins in an observational study on effectiveness. *J Clin Epidemiol* 2004; 57: 237-42
- (159) Beiderbeck AB, Sturkenboom MC, Coebergh JW, *et al.* Misclassification of exposure is high when interview data on drug use are used as a proxy measure of chronic drug use during follow-up. *J Clin Epidemiol* 2004; 57: 973-7
- (160) Ray WA, Thapa PB, Gideon P. Misclassification of current benzodiazepine exposure by use of a single baseline measurement and its effects upon studies of injuries. *Pharmacoepidemiol Drug Saf* 2002; 11: 663-9
- (161) Van Staa TP, Abenham L, Leufkens H. A study of the effects of exposure misclassification due to the time-window design in pharmacoepidemiologic studies. *J Clin Epidemiol* 1994; 47: 183-9
- (162) McMahon AD, Evans JM, McGilchrist MM, *et al.* Drug exposure risk windows and unexposed comparator groups for cohort studies in pharmacoepidemiology. *Pharmacoepidemiol Drug Saf* 1998; 7: 275-80
- (163) Csizmadi I, Collet JP, Benedetti A, *et al.* Defining hormone replacement therapy in longitudinal studies: impact on measures of effect. *Pharmacoepidemiol Drug Saf* 2004; 13: 215-25
- (164) Gross R, Bilker WB, Strom BL, *et al.* Validity and comparison of two measures of days supply in Medicaid claims data. *Pharmacoepidemiol Drug Saf* 2008; 17: 1029-32
- (165) Mantel-Teeuwisse AK, Klungel OH, Verschuren WM, *et al.* Comparison of different methods to estimate prevalence of drug use by using pharmacy records. *J Clin Epidemiol* 2001; 54: 1181-6
- (166) De Jong van den Berg LT, Feenstra N, Sorensen HT, *et al.* Improvement of drug exposure data in a registration of congenital anomalies. Pilot-study: pharmacist and mother as sources for drug exposure data during pregnancy. EuroMAP Group. European Medicine and Pregnancy Group. *Teratology* 1999; 60: 33-6
- (167) Depont F, Fourrier A, Merlière Y, *et al.* The CADEUS study: methods and logistics. *Pharmacoepidemiol Drug Saf* 2007; 16: 571-80
- (168) Lévy L, Martin-Guehl C, Lechevallier-Michel N, *et al.* Use of psychotropic drugs in 0 to 5 years old children in Aquitaine (France): prevalence and associated factors. *Pharmacoepidemiol Drug Saf* 2006; 15: 504-9

- (169) Centre National Hospitalier d'Information sur le Médicament. Banque de données sur le médicament Thériaque. [en ligne]. Disponible à l'adresse : <http://www.theriaque.org>. Consulté le 5 novembre 2009

VII- ANNEXES

Annexe 1- Liste des communications en rapport avec le sujet de thèse

INTRODUCTION

Noize P, Bazin F, Pariente A, Moore N, Fourrier-Réglat A. Assessment of drug exposure in etiological studies: a review from *Pharmacoepidemiology and Drug Safety*.

Communication affichée, 25th International Conference on Pharmacoepidemiology and Therapeutic Risk Management, Providence, 16-19 August 2009. *Pharmacoepidemiol Drug Saf* 2009; 18: S63 [abstract 145].

PREMIÈRE PARTIE

Noize P, Moore N, Lechevallier-Michel N, Fourrier-Réglat A. Comparison of patient interviews and computerized database for determination of drug exposure in pharmacoepidemiology: data from the “Three-City” (3C) Study.

Communication orale, 22nd International Conference on Pharmacoepidemiology and Therapeutic Risk Management, Lisbon, 24-27 August 2006. *Pharmacoepidemiol Drug Saf* 2006; 15: S58 [abstract 124].

Noize P, Pariente A, Sokolowski J, Bazin F, Moore N, Fourrier-Réglat A and the 3C Study Group. Validity of benzodiazepine exposure measured from patient interview: data from the 3C cohort.

Communication affichée, 11^{ème} Congrès de la Société Française de Pharmacologie et de Thérapeutique, Toulouse, 11-13 avril 2007. *Fundam Clin Pharmacol* 2007; 21: 49 [abstract 243].

Pariente A, **Noize P**, Sokolowski J, Bazin F, Moore N, Fourrier-Réglat A. Validité de la mesure de l'exposition aux benzodiazépines par interrogatoire des patients: données de la cohorte 3C.

Communication orale, 3^{ème} Séminaire Scientifique de Pharmaco-épidémiologie de l'ADELF, Paris, 24-25 mai 2007. *Rev Epidemiol Sante Publique* 2007; 55: S3.

Noize P, Pariente A, Sokolowski J, Moore N, Fourier-Réglat A. Validity of benzodiazepine exposure measured from patient interview: data from the 3C cohort.

Communication affichée, 23rd International Conference on Pharmacoepidemiology and Therapeutic Risk Management, Quebec, 19-22 August 2007. *Pharmacoepidemiol Drug Saf* 2007; 16: S233 [abstract 492].

Noize P, Bazin F, Pariente A, Moore N, Fourier-Réglat A. Description of drug use according to interview and reimbursement data in repeated cross-sectional cohort studies: data from the Three-City (3C) Study.

Communication affichée, 25th International Conference on Pharmacoepidemiology and Therapeutic Risk Management, Providence, 16-19 August 2009. *Pharmacoepidemiol Drug Saf* 2009; 18: S258 [abstract 602].

DEUXIÈME PARTIE

Noize P, Bazin F, Pariente A, Bégaud B, Moore N, Fourier-Réglat A on behalf of the 3C Study group. Agreement for drug exposure assessment between reimbursement data (French national health insurance database, CNAM-TS) and patient interviews in the framework of etiological studies: data from the Three-City (3C) cohort.

Communication orale, 13^{ème} Congrès de la Société Française de Pharmacologie et de Thérapeutique, Marseille, 15-17 avril 2009. *Fundam Clin Pharmacol* 2009; 23: 89 [abstract 448].

Bazin F, **Noize P**, Pariente A, Moore N, Fourier-Réglat A. Impact de la source de données pour la mesure de l'exposition médicamenteuse sur les estimations de risque dans les études étiologiques : données de la cohorte des Trois-Cités (3C).

Communication orale, 4^{ème} Séminaire Scientifique ADELFAFCROs de Pharmaco-épidémiologie, Paris, 14-15 mai 2009.

Bazin F, **Noize P**, Pariente A, Moore N, Fourier-Réglat A. Impact of exposure data sources on risk estimates in etiological studies: application to benzodiazepines in the Three-City (3C) Study.

Communication orale, 25th International Conference on Pharmacoepidemiology and Therapeutic Risk Management, Providence, 16-19 August 2009. *Pharmacoepidemiol Drug Saf* 2009; 18: S103 [abstract 241].

Bazin F, **Noize P**, Pariente A, Moore N, Fourrier-Réglat A. Impact of exposure data sources on risk estimates in etiological studies: application to non-steroidal anti-inflammatory drugs (NSAIDs) in the Three-City (3C) Study.

Communication affichée, 25th International Conference on Pharmacoepidemiology and Therapeutic Risk Management, Providence, 16-19 August 2009. *Pharmacoepidemiol Drug Saf* 2009; 18: S64 [abstract 148].

Annexe 2- Description des étapes de préparation des données de remboursement

Les données initiales se présentaient sous la forme d'une vingtaine de fichiers informatiques comportant une ligne par sujet et par médicament remboursé. Chaque fichier contenait les quatre variables suivantes :

- numéro d'identification 3C du sujet,
- date de remboursement du médicament,
- code CIP du médicament remboursé,
- code ATC du médicament remboursé.

Après avoir harmonisé les formats des variables dans chaque fichier et, plus particulièrement, les formats de dates, parfois différents dans un même fichier, un important travail a été effectué sur le codage ATC des médicaments selon les étapes suivantes :

➤ Vérification des codes ATC « incomplets »

Il s'agit de codes ne comportant que quatre ou cinq caractères sur les sept habituels. D'après la classification ATC, ces codes désignaient des catégories bien définies de médicaments et ne représentaient donc pas des données aberrantes (par exemple, A02X pour « autres anti-acides, médicaments de l'ulcère peptique et des flatulences »).

➤ Recherche des codes ATC manquants

Si le codage ATC présentait un certain nombre de données manquantes, le codage CIP était lui exhaustif mais prenait diverses formes (nombres à 5, 6 ou 7 chiffres). Après contact auprès de la CNAM-TS, seuls les codes CIP à 7 chiffres et débutant par 3 devaient être retenus. La présence d'autres formes de codes CIP dans les données de remboursement pouvait, soit provenir d'erreurs de saisie, soit correspondre à des « anciens » médicaments (dont des médicaments homéopathiques) n'ayant pas véritablement de code. L'ensemble des lignes pour lesquelles, à la fois le code ATC manquait et le code CIP n'était pas conforme ont été identifiées : elles indiquaient l'existence d'un remboursement sans que la nature du médicament remboursé puisse être retrouvée.

Lorsque le code CIP était conforme, le code ATC correspondant était déterminé de la manière suivante :

1- recherche du nom de spécialité du médicament à partir de son code CIP au moyen d'un document fourni par l'Assurance Maladie et répertoriant, entre autres, le nom de spécialité, le code CIP et le code ATC d'un grand nombre de médicaments. Pour les quelques codes CIP non retrouvés dans ce document, la banque de données Thériaque disponible via internet a été consultée¹⁶⁹,

2- recherche, grâce au nom de spécialité du médicament, du code ATC lui ayant été attribué dans les données déclaratives de l'étude 3C.

➤ Vérification de tous les autres codes ATC

De la même manière, toutes les lignes pour lesquelles, un code CIP conforme et un code ATC étaient d'emblée disponibles, ont été vérifiées. Certains médicaments ont fait l'objet de modifications du code ATC du fait de différences entre le code présent dans les données de remboursement et le code attribué dans les données déclaratives de l'étude 3C.

Annexe 3- Mesure de l'exposition médicamenteuse selon les données de remboursement entre les deux visites de suivi de la cohorte 3C : détermination de la fenêtre de temps à considérer avant la date de la première visite

Afin de déterminer la fenêtre de temps la plus appropriée pour considérer les données de remboursement avant la date de la première visite de suivi S1, les étapes suivantes ont été réalisées pour chacune des classes médicamenteuses :

1- Construction d'un tableau de contingence 2×2 incluant le nombre de sujets :

- ayant déclaré consommer au moins un médicament de la classe à l'interrogatoire et ayant bénéficié du remboursement d'au moins un médicament de la classe dans les n jours précédents,
- n'ayant déclaré consommer aucun médicament de la classe à l'interrogatoire et n'ayant bénéficié du remboursement d'aucun médicament de la classe dans les n jours précédents,
- ayant déclaré consommer au moins un médicament de la classe à l'interrogatoire et n'ayant bénéficié du remboursement d'aucun médicament de la classe dans les n jours précédents,
- n'ayant déclaré consommer aucun médicament de la classe à l'interrogatoire et ayant bénéficié du remboursement d'au moins un médicament de la classe dans les n jours précédents avec n successivement égal à 15, 30, 45, 60, 75 ou 90 jours.

2- Détermination, pour chacune des six fenêtres de temps n considérées, de la concordance entre les données déclaratives issues de l'interrogatoire et les données de remboursement dans les n jours précédant l'interrogatoire estimée par le coefficient kappa (κ).¹³⁷

Pour chaque classe médicamenteuse, la fenêtre de temps n pour laquelle le coefficient κ présentait sa valeur maximale a été jugée la plus appropriée pour considérer les données de remboursement avant la date de la première visite de suivi S1 (Tableau 4).

Tableau 4. Concordance entre les données déclaratives recueillies à l'interrogatoire de suivi S1 et les données de remboursement dans les n (15 à 90) jours précédant ce même interrogatoire déterminée à l'aide du coefficient κ (n = 2 985)

Classe médicamenteuse	Coefficient κ selon la fenêtre de temps considérée avant S1					
	15 jours	30 jours	45 jours	60 jours	75 jours	90 jours
Anti-acides, médicaments de l'ulcère peptique (...)	0,49	0,68	0,71	0,70	0,68	0,66
Suppléments minéraux	0,40	0,61	0,70	0,73	0,74	0,76
Médicaments du sang et des organes hématopoïétiques	0,46	0,67	0,73	0,75	0,76	0,76
Médicaments en cardiologie ^a	0,56	0,83	0,86	0,89	0,89	0,88
Antihypertenseurs	0,43	0,77	0,85	0,87	0,89	0,90
Hypolipidémiants	0,47	0,75	0,82	0,86	0,88	0,90
Médicaments du système génito-urinaire : femmes ^b	0,39	0,61	0,67	0,70	0,69	0,69
Hormones systémiques	0,43	0,63	0,64	0,64	0,63	0,61
Anti-inflammatoires non stéroïdiens	0,49	0,61	0,63	0,61	0,60	0,57
Paracétamol et aspirine	0,31	0,42	0,41	0,40	0,38	0,36
Paracétamol associé à un antalgique opiacé	0,46	0,55	0,57	0,53	0,51	0,49
Benzodiazépines et dérivés	0,39	0,59	0,67	0,71	0,73	0,74
Antidépresseurs	0,55	0,76	0,80	0,82	0,82	0,82
Médicaments des organes sensoriels	0,34	0,43	0,43	0,41	0,38	0,37

^aDigitaliques, anti-arythmiques et dérivés nitrés

^bn = 1 962

Annexe 4- Méthode utilisée pour la simulation de la survenue d'un événement

Pour une classe médicamenteuse donnée, la simulation de l'événement dépend de sa fréquence dans la population et de la force de l'association souhaitée avec l'exposition selon les données déclaratives.

	Cas	Témoins	Total
Exposés	CE	TE	l1
Non exposés	CNE	TNE	l2
Total	c1	c2	N

Sont connus N, le nombre de sujets exposés l1 et le nombre de sujets non exposés l2 :

- N est l'effectif de la population d'étude,
- le nombre de sujets exposés est donné par la prévalence d'exposition selon les données déclaratives,
- le nombre de sujets non exposés est déduit des deux paramètres précédents ($N - l1$).

Le nombre de cas c1 et de témoins c2 dépend de la fréquence p fixée pour l'événement :

- Nombre de cas : $c1 = N * p$
- Nombre de témoins : $c2 = N * (1-p)$.

Pour que la force de l'association entre l'exposition (selon les données déclaratives) et l'événement simulé corresponde à un rapport de cotes (RC) fixé, on doit avoir :

$$RC = \frac{CE * TNE}{TE * CNE}$$

$$TE = l1 - CE$$

$$CNE = c1 - CE$$

$$TNE = c2 - l1 + CE$$

$$RC = \frac{CE \times (c2 - l1 + CE)}{(l1 - CE) \times (c1 - CE)}$$

$$\Leftrightarrow (RC - 1) \times CE^2 + (-RC \times l1 - RC \times c1 - c2 + l1) \times CE + RC \times l1 \times c1 = 0$$

Si on pose:

$$a = (RC - 1)$$

$$b = (-RC \times l1 - RC \times c1 - c2 + l1)$$

$$c = RC \times l1 \times c1$$

on obtient:

$$\Delta = b^2 - 4ac \text{ et si } \Delta > 0, CE = \frac{-b \pm \sqrt{\Delta}}{2a}$$

On conserve la solution CE plausible (positive ou inférieure à la taille de la population). On obtient ainsi un nombre théorique de cas exposés pour une fréquence de l'événement p dans la population d'étude et une force de l'association entre l'exposition et l'événement fixée RC. Les nombres de témoins exposés, de cas non exposés et de témoins non exposés peuvent alors en être déduit.

Connaissant ces effectifs théoriques, un événement simulé peut être généré au sein de la population d'étude en tirant au sort, selon une loi binomiale, chez les sujets exposés et non exposés (selon les données déclaratives), le nombre théorique de cas calculé précédemment. Les sujets exposés et non exposés n'ayant pas été tirés au sort constituent donc les témoins exposés et non exposés.