

*Maladies chroniques
et traumatismes*

Étalonnage du PMSI MCO pour la surveillance des infarctus du myocarde

Année 2003

Sommaire

Abréviations	2
1. Introduction	3
2. Matériel et méthodes	4
2.1 Population	4
2.2 Sources de données	4
2.2.1 Les registres français des cardiopathies ischémiques	4
2.2.2 Le Programme de médicalisation des systèmes d'information en soins de courte durée (PMSI MCO) en 2003	4
2.3 Méthodes	5
2.3.1 Extraction des RSA avec diagnostic principal d'infarctus du myocarde dans le PMSI MCO	5
2.3.2 Levée de l'anonymat des RSA	5
2.3.3 Appariement des données des registres et des données du PMSI	5
2.3.4 Définition des épisodes de soins	6
2.3.5 Analyses statistiques	7
3. Résultats	8
3.1 Caractéristiques des IAM dans les deux sources de données	8
3.2 Sensibilité	9
3.3 Valeur prédictive positive	9
3.4 Faux positifs	10
3.5 Comparaison des nombres d'IAM comptabilisés par le PMSI et les registres en 2003	10
3.5.1 PMSI : un épisode de soin par période de 28 jours	10
3.5.2 Autres modalités d'exploitation du PMSI	11
4. Discussion	12
4.1 Sensibilité	12
4.2 Valeur prédictive positive	13
4.3 Estimation du nombre d'infarctus du myocarde	13
4.4 Difficultés	14
4.5 Limitations	14
5. Conclusion	16
Références bibliographiques	17
Annexe 1 – Formats de RSA en vigueur en 2003	19
Annexe 2 – Codes CIM-10	21
Annexe 3 – Algorithme de classement des RSA	23
Annexe 4 – Étude exploratoire des faux négatifs	24

Étalonnage du PMSI MCO pour la surveillance des infarctus du myocarde

Année 2003

Cette étude d'étalonnage a été réalisée par l'Institut de veille sanitaire (InVS), Département des maladies chroniques et des traumatismes (DMCT), en partenariat avec les registres français des cardiopathies ischémiques et leur centre coordinateur : registre de Haute-Garonne, registre de la communauté urbaine de Lille, registre du Bas-Rhin et l'Institut national de la santé et de la recherche médicale U780.

Nous remercions tous les personnels des Départements d'information médicale (DIM) qui ont contribué à cette étude en complétant les résumés de sortie anonyme par les données nominatives nécessaires pour effectuer l'appariement avec les données des registres : Docteurs Jean-Marie Balliet (clinique Saint-Joseph à Colmar), Françoise Barrère (clinique de l'Union à Saint-Jean), Nicolas Bertrand (CH Haguenau), Patrick Besse (clinique du Parc à Toulouse et clinique des Cèdres à Cornebarrieu), Olivier Boux (CHR Strasbourg), Didier Bricoteau (CH Roubaix), Fariba Coulon (CH Sainte-Catherine de Saverne), Valérie Damourette (CH Séclin), Alain Desmarets (polyclinique du Parc à Croix), Didier Fabre (CHR Toulouse), Anne-Marie Le Maux (clinique de l'Orangerie à Strasbourg), Gérard Mahieu (hôpital Saint-Philibert à Lomme), Julie Quentin (CH Armentières), Paolo de Rezende (CH Wissembourg), Norbert Roeslin (CHR Strasbourg), Sylvie Seitz (CH Saint-Gaudens), Monique Vissentin (clinique Pasteur à Toulouse), Professeurs Régis Beuscart (CHR Lille) et Franck Boutault (CHR Toulouse).

L'appariement des données a été réalisé par les Docteurs Dominique Arveiler (registre des cardiopathies ischémiques du Bas-Rhin), Michèle Montaye (registre des cardiopathies ischémiques de la communauté urbaine de Lille) et Jean-Bernard Ruidavets (registre des cardiopathies ischémiques de la Haute-Garonne).

Le contrôle des données, l'analyse et la rédaction du rapport ont été réalisés à l'InVS par Christine de Peretti et Christophe Bonaldi. Ce rapport a bénéficié des relectures attentives de Claire Fuhrman et Javier Nicolau (InVS) ainsi que du Professeur Catherine Quantin (CHRU de Dijon – DIM).

Abréviations

ARH	Agence régionale de l'hospitalisation
ATIH	Agence technique de l'information sur l'hospitalisation
CCTIRS	Comité consultatif sur le traitement de l'information en matière de recherche dans le domaine de la santé
CH	Centre hospitalier
CHR	Centre hospitalier régional
CHRU	Centre hospitalier régional universitaire
CIM	Classification statistique internationale des maladies et des problèmes de santé connexes
Cnil	Commission nationale de l'informatique et des libertés
CUDL	Communauté urbaine de Lille (dénommée Lille métropole communauté urbaine: LMCU)
DIM	Département d'information médicale
DP	Diagnostic principal
GHM	Groupe homogène de malades
HFA-DB	European Health For All DataBase (OMS-Europe)
IAM	Infarctus aigu du myocarde
MONICA	Monitoring of trends and determinants in cardiovascular disease
OCDE	Organisation de coopération et de développement économiques
OMS	Organisation mondiale de la santé
PMSI MCO	Programme de médicalisation des systèmes d'information en soins de courte durée: médecine, chirurgie et obstétrique
RSA	Résumé de sortie anonyme
RSS	Résumé de sortie standardisé
RUM	Résumé d'unité médicale
SAE	Statistique annuelle des établissements de santé
SCA	Syndrome coronaire aigu
T2A	Tarifcation à l'activité
UM	Unité médicale
Usic	Unité de soins intensifs cardiologiques
VPP	Valeur prédictive positive

1. Introduction

Les cardiopathies ischémiques, ou maladies coronaires, demeurent un problème majeur de santé publique du fait de leur grande fréquence et de leur gravité. Elles constituent en effet une cause importante de morbidité, de recours aux soins et de mortalité et représentent, à ce titre, un enjeu majeur de santé publique, tant en termes de prévention primaire (risque cardio-vasculaire), que de prise en charge thérapeutique des pathologies déclarées, et de prévention secondaire.

La surveillance épidémiologique des infarctus du myocarde et des décès coronaires est assurée depuis le milieu des années 1980 par les trois registres français des IAM et des décès coronaires : le registre de Haute-Garonne, le registre de la communauté urbaine de Lille (CUDL) et le registre du Bas-Rhin. Ces trois registres populationnels ont été créés dans le cadre du projet international MONICA qui a étudié la forte hétérogénéité géographique des taux de mortalité et d'incidence des événements coronaires aigus en Europe et en France (avec un gradient décroissant du nord vers le sud) ainsi que leurs évolutions entre 1985 et 1995. Le projet MONICA s'est terminé au milieu des années 1990, mais les trois registres français ont poursuivi la surveillance populationnelle des décès coronaires et des IAM, avec un élargissement récent à l'ensemble des syndromes coronaires aigus.

En parallèle, les bases de données médico-administratives sur les décès et les hospitalisations produisent des données nationales exhaustives qui alimentent des bases internationales (base d'Eurostat, base HFA-DB de l'Organisation mondiale de la santé (OMS) Europe ou encore base EcoSanté de l'Organisation de coopération et de développement économiques (OCDE)). Ces bases sont susceptibles d'être utilisées pour des études de tendances ou des comparaisons internationales de taux standardisés. Mais les données de ces bases médico-administratives ne sont pas systématiquement expertisées dans chaque pays. De plus, la diversité des systèmes de santé et des modalités de recueil induisent des problèmes importants de comparabilité [1].

Concernant les décès, la comparaison des données de 37 registres du projet MONICA a mis en évidence la sous-estimation notable des décès coronaires dans les statistiques françaises sur les causes de décès. Les cardiopathies ischémiques sont en effet responsables d'un nombre élevé de décès rapides et inattendus ("morts subites") pour lesquels, faute d'information médicale, la cause ne peut être formellement établie. Les données recueillies pour le projet MONICA, selon un protocole rigoureusement standardisé, ont mis en évidence la grande variabilité de la proportion de décès pour lesquels l'information est insuffisante, cette proportion s'avérant élevée en France [2]. Une étude plus récente a comparé les estimations des trois registres français à celles de la statistique des causes médicales de décès (CépiDc) ; elle a montré que la mortalité coronaire comptabilisée par le CépiDc sous-estime de 59 % la mortalité coronaire estimée par les registres après prise en compte de ces décès pour lesquels la cause coronaire ne peut être ni affirmée ni écartée, la différence étant de 3 % lorsque la comparaison se limite aux décès coronaires dûment documentés [3].

Les résumés d'hospitalisation constituent une autre source de données pérennes et quasi-exhaustives, susceptibles de contribuer à la surveillance nationale des infarctus du myocarde. En effet, toute hospitalisation en service de court séjour donne lieu à la production d'un résumé de sortie anonyme (RSA) par les établissements de soins. Les RSA sont assemblés dans la base nationale annuelle du PMSI MCO par l'Agence technique de l'information sur l'hospitalisation (ATIH). Mais ces données ont été conçues dans une logique administrative et financière, et non à des fins de surveillance épidémiologique. Des contrôles financiers existent, mais il n'y a pas de validation des RSA dans une perspective épidémiologique.

En cas d'hospitalisation pour infarctus du myocarde, il arrive qu'un même épisode donne lieu à plusieurs séjours, s'il y a un transfert entre établissements par exemple, et donc à plusieurs RSA comportant un même diagnostic d'IAM. En outre, malgré les consignes diffusées au niveau national, les habitudes de codage peuvent différer d'un établissement à l'autre ; des erreurs de codage ou la recherche d'une valorisation optimale du séjour peuvent survenir. La logique même de production du RSA, outil de valorisation des séjours à des fins de financement, implique la désignation d'un seul diagnostic principal, qui peut être complété de plusieurs diagnostics associés (voire d'un diagnostic relié). Ainsi, en cas de pathologies multiples, par exemple en cas de complication grave d'un IAM, cette dernière peut figurer comme diagnostic principal avant les modifications introduites en 2009 [4]. En cas de transferts dans plusieurs services d'un même hôpital, et donc en cas de séjours avec plusieurs résumés d'unité médicale (RUM), le logiciel de groupage attribue un, et un seul, diagnostic principal au RSA, selon un algorithme hiérarchisé automatisé (acte classant opératoire, durée de séjour dans chaque unité médicale, etc.).

Un certain nombre d'études de validation des données d'hospitalisation relatives aux IAM ont été effectuées au cours des 30 dernières années dans des pays disposant de données de référence expertisées ("gold standard"), dans le cadre de cohortes, d'études spécifiques ou de registres populationnels des infarctus du myocarde, tels des registres participant au projet MONICA [5-10]. Mais aucune étude de ce type n'avait encore été réalisée en France dans le champ cardio-vasculaire. L'exploitation de la base nationale du PMSI MCO pour la surveillance des IAM hospitalisés nécessite donc la réalisation d'études évaluant sa validité pour la surveillance épidémiologique de la population.

L'objectif de la présente étude était d'établir un système national de surveillance des IAM hospitalisés reposant sur l'expertise des trois registres MONICA et les données d'hospitalisations. Il s'agissait d'évaluer les qualités métrologiques du PMSI MCO au regard des données produites par les trois registres MONICA, pour les classes d'âge comprises entre 35 et 74 ans et les territoires correspondants, afin de déterminer les algorithmes ou facteurs correcteurs nécessaires pour produire une estimation nationale du taux d'IAM à partir du PMSI MCO.

2. Matériel et méthodes

Le projet a été soumis à l'avis du Comité consultatif sur le traitement de l'information en matière de recherche dans le domaine de la santé (CCTIRS) et a reçu une autorisation de traitement de données à caractère personnel de la Commission nationale de l'informatique et des libertés (Cnil) sous le numéro 906163.

2.1 POPULATION

La comparaison des données du PMSI aux enregistrements des registres MONICA et l'estimation de la qualité de l'identification des IAM dans le PMSI ont naturellement été restreintes à la population surveillée par ces registres. Il s'agissait de l'ensemble des personnes âgées de 35 à 74 ans résidant dans les départements du Bas-Rhin, de la Haute-Garonne et dans la CUDL, et hospitalisées pour IAM.

Pour cette étude, nous avons considéré tous les séjours liés à un diagnostic d'IAM survenu au cours de l'année 2003, qui constituait, au moment du démarrage de l'enquête, la dernière année de recueil validée par les registres des cardiopathies ischémiques.

2.2 SOURCES DE DONNÉES

2.2.1 Les registres français des cardiopathies ischémiques

Les registres des cardiopathies ischémiques du Bas-Rhin, de la Haute-Garonne et de la CUDL recueillent et analysent des données de mortalité et de morbidité coronarienne aiguë depuis 1984. Ils ont participé au projet international MONICA coordonné par l'OMS jusqu'à la fin de ce projet, en 1993 [2,11,12].

Par la suite, les trois registres français ont poursuivi leur activité selon un protocole simplifié depuis 1997. Comme auparavant, la surveillance des événements coronaires aigus se fait à partir de plusieurs sources afin d'assurer l'exhaustivité des cas, mais les IAM hospitalisés, ciblés par notre étude, sont dorénavant identifiés à partir des lettres de sortie. Les classes d'âge ont été étendues à l'ensemble des personnes âgées de 35 à 74 ans (35 à 64 ans dans le projet MONICA-OMS).

Pour chaque événement, le recueil comprend l'identité du patient (nom, nom de jeune fille et prénom, recueillis séparément à des fins de dédoublement et de suivi), la date de naissance, le sexe, un code de résidence, canton ou commune selon les cas, l'identification de l'établissement de santé, les antécédents d'infarctus ou de cardiopathie ischémique, la survie à 28 jours et la durée de survie.

Dans les registres, une base annuelle comprend tous les événements coronaires aigus pour lesquels le début des symptômes est survenu durant l'année calendaire. Par convention, les registres ne classent comme nouvel infarctus que les événements survenant plus de 28 jours après un premier épisode. Si l'intervalle de temps est plus court, le second événement n'est pas considéré comme un nouvel IAM.

2.2.2 Le Programme de médicalisation des systèmes d'information en soins de courte durée (PMSI MCO) en 2003

La base nationale du PMSI MCO est un recueil standardisé regroupant l'ensemble des RSA produits au décours de chaque hospitalisation en service de court séjour en médecine, chirurgie ou obstétrique.

Au niveau des établissements de soins, chaque séjour dans une unité médicale donne lieu à la production d'un RUM. Ces résumés sont ensuite traités par les Départements d'information médicale (DIM) qui agrègent, le cas échéant, les différents RUM produits lors d'une même venue en hospitalisation avant de procéder au groupage automatisé (RUM-RSS groupés), puis à l'anonymisation des résumés de sortie standardisés (RSS) par un module logiciel générateur de RSA. Le RSA se présente toujours comme un enregistrement unique par séjour, quel que soit le nombre de RUM du séjour.

Chaque RSA comprend un diagnostic principal codé en CIM-10. En 2003, et jusqu'en 2008, le diagnostic principal était le motif de prise en charge ayant motivé l'essentiel de l'effort médical et soignant au cours de l'hospitalisation. En cas d'hospitalisation dans une seule unité médicale, le diagnostic principal du RSA est identique au diagnostic principal du RUM à partir duquel le RSA est produit. Mais pour les séjours avec passage dans plusieurs unités médicales, le diagnostic principal du RSA est choisi parmi les diagnostics principaux des différents RUM, par un algorithme hiérarchisé automatisé tendant à sélectionner le diagnostic principal du RUM ayant mobilisé la plus forte charge de soins.

Les RSA comportent des variables administratives et des variables médicales. Les variables 2003 les plus utilisées dans une perspective de surveillance épidémiologique sont les suivantes : identification de l'établissement (numéro Finess de l'entité juridique), nombre de RUM composant le RSS d'origine, âge, sexe, mode d'entrée, provenance, mois de sortie, année de sortie, mode de sortie, destination, durée totale du séjour, durée de séjour en réanimation, code géographique de résidence (basé sur le code postal), résultats de la fonction groupage (codes CMD et GHM), diagnostic principal, le cas échéant diagnostic relié, nombre de diagnostics associés significatifs et codes CIM-10 correspondants, actes réalisés... Depuis 2002, il comprend également un numéro de chaînage anonyme permettant de relier les différentes hospitalisations d'une même personne, ainsi qu'une variable mesurant l'écart entre la date d'entrée dans l'établissement et une date virtuelle aléatoire, unique pour chaque patient, permettant d'ordonner les séjours d'une même personne (les formats des RSA 207 et 208 sont présentés en annexe 1).

Une base annuelle du PMSI MCO comprend tous les RSA pour lesquels la date de sortie est survenue durant l'année calendaire. En effet, c'est le moment de la sortie, et non la date d'entrée, qui définit l'inclusion dans la base annuelle du PMSI MCO : un IAM survenu en fin d'année 2002, dont l'hospitalisation s'est terminée au début de l'année 2003, donne lieu à un RSA contenu dans la base 2003 ; de même, le RSA d'un IAM survenu en fin d'année 2003 est inclus dans la base 2004 lorsque la sortie s'est effectuée en début d'année 2004.

2.3 MÉTHODES

2.3.1 Extraction des RSA avec diagnostic principal d'infarctus du myocarde dans le PMSI MCO

Dans un premier temps, nous avons sélectionné les RSA à partir du diagnostic principal. La liste des codes CIM-10 retenus était la suivante : codes I21 pour les IAM et I22 pour les infarctus du myocarde à répétition. Les extensions introduites au 1^{er} janvier 2002 ont été prises en compte. Le code I23 correspondant aux complications récentes d'IAM a également été retenu lorsqu'il n'y avait pas de RSA récent de prise en charge d'un IAM (délai >28 jours) (annexe 2).

Tous les RSA concernant des personnes âgées de 35 à 74 ans, résidant dans les aires géographiques couvertes par les registres (Bas-Rhin, Haute-Garonne et CUDL) et comportant un de ces codes CIM-10 en diagnostic principal ont été sélectionnés dans la base nationale du PMSI MCO 2003 et ce, quel que soit le lieu d'hospitalisation.

Deux mille un RSA répondant à ces critères de diagnostic, d'âge et de domiciliation ont été extraits de la base nationale 2003 du PMSI MCO ; 183 séjours de la base 2004 qui avaient un mois de sortie en janvier (ou février pour des durées de séjour supérieures à 31 jours) ont été ajoutés à la sélection 2003 afin de pouvoir estimer le nombre d'IAM de fin d'année avec hospitalisation prolongée en début d'année 2004.

La proportion de RSA sans numéro de chaînage anonyme était encore notable en 2003 (16,2 %), avec des différences importantes entre les trois sites : 30,3 % dans la CUDL, 16,6 % dans le Bas-Rhin et 1,0 % en Haute-Garonne. En 2004, la généralisation du chaînage était quasi-effective (2,2 % des séjours sans numéro de chaînage). Au total, 15,0 % (328) des RSA extraits n'avaient pas de numéro de chaînage.

Le chaînage des RSA a été effectué pour les 1 856 séjours ayant un numéro anonyme de chaînage. Les séjours ont été ordonnés puis classés de façon récursive en distinguant les séjours pour IAM des éventuels transferts ou réhospitalisations à moins de 28 jours et enfin, les nouveaux épisodes d'IAM survenus plus de 28 jours après le début d'une première hospitalisation (annexe 3). Cette démarche a pu être effectuée sur les données 2003, d'une part, et sur les données 2004, d'autre part, mais le numéro de chaînage anonyme ayant été modifié en 2004, il n'était pas possible de chaîner des transferts de la fin 2003, lorsque la fin du séjour correspondant était survenue début 2004.

2.3.2 Levée de l'anonymat des RSA

Il était nécessaire de procéder à l'identification nominative des séjours sélectionnés pour pouvoir réaliser l'appariement avec les données des registres. Le recours à des données nominatives ne nous a cependant semblé justifié que pour les établissements ayant accueilli plus de 5 cas dans l'année, soit 2 032 RSA (93,0 %). Les établissements ont été identifiés grâce au fichier Finess.

Ces 2 032 RSA ont été mis sous forme de fichier Excel[®] puis envoyés aux DIM des établissements concernés pour identification nominative et ajout de la date d'hospitalisation. Il était demandé :

- d'extraire les RSS correspondant aux tranches d'âges et aux zones de résidences concernées et comportant un diagnostic principal d'IAM du PMSI 2003, d'une part, et du PMSI 2004, pour les séjours ayant débuté en 2003, d'autre part, puis de les rendre nominatifs ;

- d'identifier les RSA correspondants et d'ajouter le nom de famille, le nom de jeune fille et le ou les prénom(s) ainsi que la date d'hospitalisation sur le fichier envoyé ;
- de crypter le fichier complété au moyen du logiciel fourni (GNU Privacy Guard/GPG) puis de le renvoyer par courriel à l'investigateur principal de l'étude.

Les fichiers reçus ont été contrôlés et des vérifications ou des demandes complémentaires ont été effectuées. Sur les 23 établissements concernés, 4 n'ont pas transmis de données nominatives au Département des maladies chroniques et des traumatismes de l'InVS : un a refusé et trois n'avaient plus la possibilité technique d'effectuer cette identification du fait de changement de logiciel. Le nombre de RSA provenant de ces établissements était égal à 183. Pour les 19 établissements qui ont répondu à notre requête, le taux d'identification était de 96,6 %, ce taux variant de 81,0 % à 100 % selon les établissements.

Sur l'ensemble des 2 032 RSA envoyés aux établissements, 1 787 ont été identifiés directement par les DIM. Grâce au numéro de chaînage, des RSA initialement non identifiés ont pu l'être dans un second temps en cas de séjours multiples dont un RSA avait été identifié. Au total, 1 859 RSA ont été identifiés directement ou indirectement, soit 91,5 % des RSA envoyés aux DIM et 85,1 % de l'ensemble des RSA extraits.

| TABLEAU 1 |

Identification des RSA

	N	%
RSA	2 184	
RSA non envoyés*	152	7,0
RSA envoyés aux DIM	2 032	93,0
RSA directement identifiés par les DIM	1 787	81,8
RSA identifiés (DIM ou numéro de chaînage)	1 859	85,1

* Établissements ayant eu au plus 5 séjours.

2.3.3 Appariement des données des registres et des données du PMSI

À partir des fichiers reçus des DIM, un fichier a été créé pour chaque aire de couverture des registres. Ces fichiers ont été complétés par les IAM extraits du PMSI pour lesquels nous ne disposons pas d'information nominative (établissements de soins ayant accueilli moins de 6 patients ou impossibilité technique). Toutefois, la commune de Villeneuve d'Ascq qui, par erreur, n'avait pas été extraite, a été *de facto* exclue du champ de l'analyse (cette commune a trois codes postaux, mais le code, unique, utilisé dans le PMSI diffère du code utilisé par le registre). Par ailleurs, 65 RSA ayant un code géographique non spécifique de communes de la CUDL ont été exclus du champ de l'étude (codes postaux communs à plusieurs communes n'appartenant pas toutes à la CUDL). De plus, 121 RSA avec date d'hospitalisation en 2002 ou en 2004 n'ont pas été envoyés aux registres puisqu'ils n'étaient *a priori* pas compatibles avec leur fichier 2003 d'IAM.

Finalement, 1 998 RSA, dont 1 686 étaient identifiés, ont été envoyés aux registres sous forme de fichier chiffré pour croisement avec leurs propres données sur les IAM survenus en 2003. L'appariement des bases de données a été fait par les registres à partir de la correspondance du

nom, du prénom, de la date d'hospitalisation, du sexe, de l'âge, de la commune de résidence et de l'établissement de prise en charge. Le registre devait de plus compléter chaque RSA en indiquant la catégorie diagnostique de l'événement selon la classification en vigueur au registre (IAM/décès coronaire/mort subite en moins de 24 heures sans antécédent coronaire connu/séjour non lié à un IAM). Il devait également ajouter la date du début des symptômes pour les IAM et préciser l'expertise du registre sur le RSA: IAM ou, à défaut, caractériser le faux positif: patient inconnu du registre/patient connu mais événement rejeté par le registre comme n'étant pas une coronaropathie/coronaropathie mais événement récusé comme IAM/IAM connu du registre mais RSA non retenu car ce n'est pas un nouvel événement. Il était également demandé d'indiquer si un antécédent d'infarctus avait déjà été enregistré. Les épidémiologistes des registres ont comparé les 1 686 RSA nominatifs à leurs propres données, codé leur expertise (et ajouté un item pour les IAM correspondant à de fausses adresses), puis complété le fichier par les IAM connus, mais non repérés dans le fichier de RSA.

Une fois la base complétée, chaque registre a renvoyé son fichier chiffré à l'InVS pour finalisation : les fichiers ont été contrôlés. Puis les IAM ajoutés par les registres ont été comparés aux "RSA" restés anonymes : pour 130 RSA, un appariement manuel avec les IAM ajoutés par les registres a été effectué sur la base de la correspondance complète du sexe, de l'âge, des codes de résidence, des lieux d'hospitalisation et de la compatibilité temporelle (mois de sortie et durée d'hospitalisation *versus* date de l'épisode aigu). Mais pour 182 RSA, aucun appariement n'a été possible car ces variables ne concordaient pas. Au final, le nombre d'événements ajoutés par les registres s'élevait à 374 IAM (figure 1).

Les données des trois registres ont ensuite été concaténées. Les RSA 2003 de la sélection initiale ayant une date de début d'hospitalisation en 2002 ont été réintégrés dans la base, mais les RSA de la sélection 2004 qui n'avaient pas une date de début d'hospitalisation en 2003 ont été exclus. Enfin, une variable résumant l'expertise des registres

a été créée à partir des variables d'expertise des registres, de la notion d'identification ou non des RSA et de la temporalité, puis regroupement des catégories similaires. Les catégories retenues étaient les suivantes: RSA non rendus nominatifs et, parmi les RSA identifiés, RSA correspondant à un IAM 2003/RSA correspondant à un IAM 2002/RSA avec début d'hospitalisation en 2002/RSA de patient ayant une coronaropathie ou un antécédent d'IAM connu mais ne correspondant pas à un nouvel IAM/patient connu mais séjour non lié à la coronaropathie/mort subite sans IAM ni antécédent coronaire connu/patient en réalité non domicilié dans le département et enfin, patient inconnu du registre.

Une analyse qualitative exploratoire, non exhaustive et de fait non représentative de quelques RSS, a été réalisée dans l'un des sites afin d'appréhender le profil des IAM connus des registres mais absents du PMSI. Elle est présentée en annexe 4.

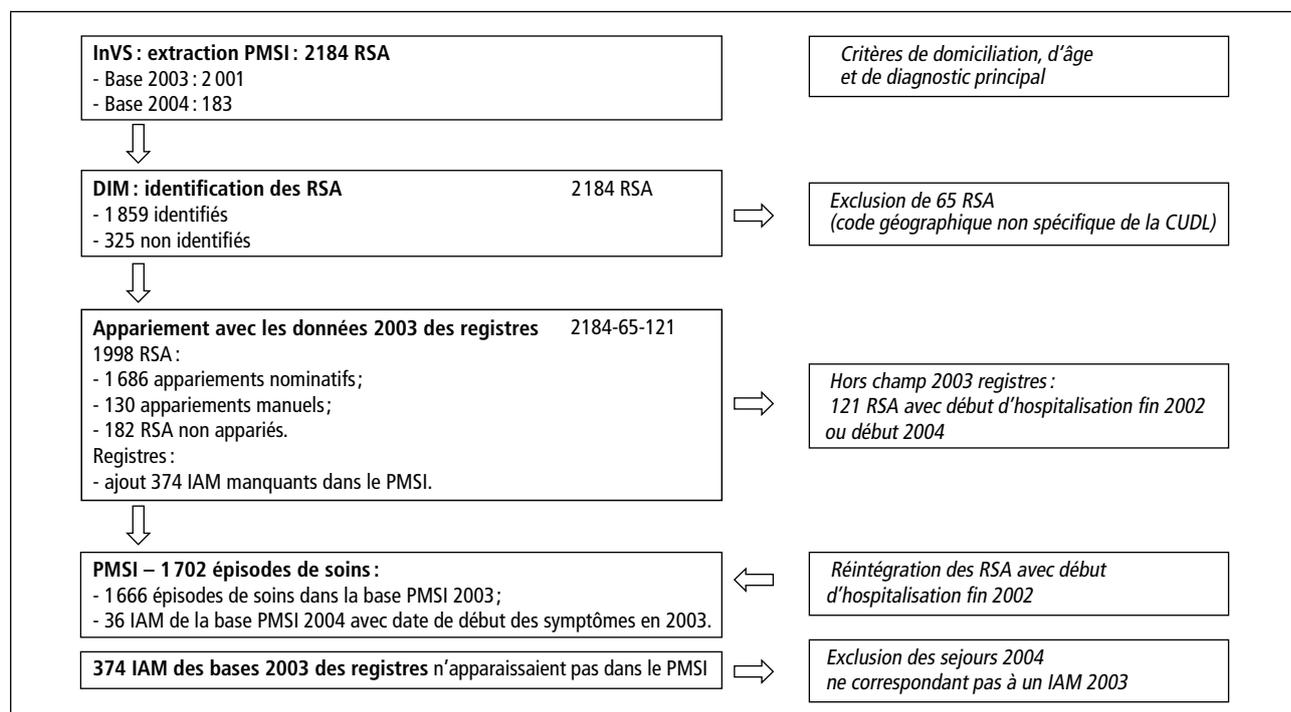
2.3.4 Définition des épisodes de soins

Par convention, les registres MONICA n'enregistrent qu'un seul événement par période de 28 jours. Pour le PMSI également, il était donc nécessaire de ne retenir qu'un épisode par période de 28 jours. L'identification des épisodes a été effectuée sur la base des dates d'hospitalisation pour les "RSA" identifiés. Pour les autres, cette procédure a été effectuée à partir du numéro de chaînage anonyme et de la variable permettant de quantifier le délai entre hospitalisations successives par la démarche décrite plus haut. Les 35 RSA 2003 qui n'avaient pas été identifiés et qui ne comportaient pas de numéro de chaînage, dont 17 avaient fait l'objet d'un appariement manuel, ont tous été considérés comme de nouveaux épisodes (2,1 %).

Cette procédure porte à 1 702 le nombre total d'épisodes relevés dans le PMSI sur la période considérée: 1 666 épisodes dans la base 2003 et 36 épisodes appariés d'IAM avec date de début des symptômes fin 2003 et sortie d'hospitalisation en 2004.

FIGURE 1

Appariement des données du PMSI à celles des registres



2.3.5 Analyses statistiques

Les données ont été analysées à l'aide du logiciel SAS version 9.1 (SAS Institute, Inc, NC, USA). La sensibilité et la valeur prédictive de la base 2003 du PMSI pour les IAM ont été estimées par comparaison avec les données 2003 des registres: elles ont été calculées à partir des épisodes de soins, avec un seul épisode d'IAM comptabilisé par période de 28 jours. La sensibilité a été calculée en rapportant le nombre d'épisodes du PMSI 2003 expertisés comme de vrais positifs 2003 par les registres au nombre total d'IAM qu'ils ont enregistré en 2003 et la VPP, en rapportant le nombre de vrais positifs au regard de la base 2003 des registres, au nombre total d'épisodes avec diagnostic principal d'IAM dans le PMSI 2003. Les intervalles de confiance à 95 % ont été calculés. La description des faux positifs a été réalisée à partir de l'expertise des registres pour le premier RSA de chaque épisode.

Dans un second temps, le nombre total d'épisodes d'IAM estimés à partir du PMSI 2003 a été comparé au nombre d'IAM enregistrés par les registres la même année. Cette comparaison a ensuite été répétée pour plusieurs modalités d'exploitation de la base PMSI :

- 1) comptabilisation de tous les RSA avec diagnostic principal d'IAM;
- 2) comptabilisation des RSA en "hospitalisation complète" ou avec mode de sortie par décès, après exclusion des séances et hospitalisations de moins d'un jour (sauf décès) et des RSA avec code GHM de transfert;
- 3) comptabilisation d'un seul épisode annuel.

Les comparaisons entre sites, sexes et groupe d'âge ont été réalisées par test du Chi-2 et complétées par des régressions logistiques pour ajuster sur l'âge et/ou le sexe. Le test de Kruskal-Wallis a été utilisé pour comparer la distribution du nombre de RSA par épisode de soins en fonction du site. Pour tous ces tests, le risque α était fixé à 0,05.

3. Résultats

3.1 CARACTÉRISTIQUES DES IAM DANS LES DEUX SOURCES DE DONNÉES

Selon les données des registres, 1723 IAM ont été enregistrés en 2003 pour la classe d'âge comprise entre 35 et 74 ans (tableau 2). Les IAM concernaient une fois sur cinq des femmes et quatre fois sur cinq des hommes. Il y avait une différence significative par site, avec une proportion d'IAM féminins un peu plus faible en Haute-Garonne ($p=0,02$). En revanche, il n'y avait pas de différence significative de la distribution par âge dans ces trois sites, pour les quatre classes d'âge décennales considérées : près de 10 % des IAM concernaient des personnes âgées de 35 à 45 ans, 24,7 % des personnes de 45 à 54 ans, 28,5 % des 55-64 ans et 37,1 % des personnes de 65 à 74 ans.

Dans le PMSI 2003, 1666 épisodes d'IAM ont été comptabilisés en 2003. La distribution par sexe était quasi-identique à celle des registres :

79,8 % des épisodes concernaient des hommes dans le PMSI, contre 79,7 % pour les registres (femmes : 20,2 et 20,3 %). La faible proportion d'IAM féminins du registre de Haute-Garonne était également observée dans le PMSI. En revanche, la distribution par âge montrait une proportion un peu plus faible d'IAM pour la classe d'âge décennale 65-74 ans dans le PMSI (35,0 *versus* 37,1 % pour les registres).

Neuf fois sur 10, le RSA initial provenait d'un établissement de soins situé dans le département de résidence et dans 9,1 % des cas, d'un établissement hors département. Il existait une association significative entre le site et le département d'hospitalisation initiale : la proportion de RSA initiaux hors département variait entre 5,5 % et 12,6 % ($p=10^{-3}$). Le nombre moyen de RSA par épisode de soins était égal à 1,17. Ce ratio était un peu plus faible en Haute-Garonne : 1,10 *versus* 1,20 dans les deux autres sites ($p=0,01$).

| TABLEAU 2 |

Caractéristiques des infarctus du myocarde selon la source

	Haute-Garonne		Lille		Bas-Rhin		Global	
	N	%	N	%	N	%	N	%
Registres 2003								
Infarctus du myocarde (IAM)	563		489		671		1723	
Sexe								
Hommes	469	83,3	385	78,7	519	77,3	1373	79,7
Femmes	94	16,7	104	21,3	152	22,7	350	20,3
Âge								
35-44 ans	60	10,7	48	9,8	59	8,8	167	9,7
45-54 ans	137	24,3	120	24,5	168	25,0	425	24,7
55-64 ans	165	29,3	141	28,8	185	27,6	491	28,5
65-74 ans	201	35,7	180	36,8	259	38,6	640	37,1
PMSI 2003								
Épisodes d'hospitalisations pour IAM	572		505		589		1666	
Sexe								
Hommes	472	82,5	399	79,0	459	77,9	1330	79,8
Femmes	100	17,5	106	21,0	130	22,1	336	20,2
Âge								
35-44 ans	59	10,3	58	11,5	53	9,0	170	10,2
45-54 ans	137	24,0	134	26,5	153	26,0	424	25,5
55-64 ans	175	30,6	143	28,3	171	29,0	489	29,4
65-74 ans	201	35,1	170	33,7	212	36,0	583	35,0
Département initial d'hospitalisation								
Dans le département de domicile	523	91,4	477	94,5	515	87,4	1515	90,9
Hors département	49	8,6	28	5,5	74	12,6	151	9,1
Nombre total de RSA	631		605		706		1942	
Nombre moyen de RSA par épisode de 28 jours	1,10		1,20		1,20		1,17	

Champ : 35-74 ans, code CIM-10 diagnostic principal en I21, I22, I23, un IAM par période de 28 jours.

3.2 SENSIBILITÉ

Afin d'estimer la sensibilité du PMSI 2003, nous n'avons retenu qu'un seul RSA par période de 28 jours pour une même personne, comme c'est le cas dans les registres MONICA. La sensibilité a été calculée en rapportant le nombre total d'épisodes de soins de la base PMSI 2003 confirmés par les registres comme IAM survenus en 2003, au nombre total d'IAM hospitalisés comptabilisés par les registres en 2003.

Sur l'ensemble des 1 723 cas d'IAM hospitalisés recensés par les trois registres en 2003, 1 313 étaient présents dans le PMSI 2003 (tableaux 3 et 4). La sensibilité du PMSI 2003 s'établissait à 76,2 % pour l'ensemble des classes d'âge comprises entre 35 et 74 ans (IC 95 % : [74,2 %-78,2 %]). Elle était égale à 80,3 % entre 35 et 64 ans, mais plus faible entre 65 et 74 ans (69,2 %, $p < 0,0001$). La sensibilité était significativement plus élevée pour les hommes (77,6 %) que pour les femmes (70,6 %, $p = 0,01$) en analyse bivariée, mais cette différence n'était plus significative après prise en compte de la structure d'âge par régression logistique.

Il n'y avait pas de différence significative entre sites : la sensibilité était égale à 73,6 % dans la CUDL, 75,1 % dans le Bas-Rhin et 79,8 % en Haute-Garonne pour l'ensemble des 35-74 ans (tableau 5).

La prise en compte des 36 IAM de la base PMSI 2004 ayant une date de survenue des symptômes en 2003 et une hospitalisation prolongée début 2004 augmente la sensibilité du PMSI à 78,3 %.

(Voir en annexe 4 la présentation de l'étude exploratoire et non exhaustive des faux négatifs).

3.3 VALEUR PRÉDICTIVE POSITIVE

La valeur prédictive positive (VPP) du PMSI 2003 a été déterminée en rapportant les épisodes d'hospitalisation de la base PMSI 2003 effectivement confirmés par les registres comme IAM survenus en 2003, au nombre total des épisodes d'hospitalisation pour IAM du PMSI 2003. Un seul épisode par période de 28 jours était pris en compte. Sur les 1 666 épisodes d'hospitalisation pour IAM recensés dans le PMSI 2003, 1 313 ont été confirmés comme IAM présents dans la base 2003 des registres et 353 étaient des faux positifs : la VPP était égale de 78,8 % (IC [76,8 %-80,8 %]) pour l'ensemble des classes d'âge considérées (tableau 6).

En analyse bivariée, la VPP était significativement plus élevée pour les hommes que pour les femmes (80,2 % *versus* 73,5 % $p = 0,01$). Par contre, il n'y avait pas d'association significative entre la VPP et l'âge (tableau 7). Il existait une association significative entre la VPP et les sites géographiques ($p < 10^{-4}$) : la VPP variait de 71,3 % à Lille, à 78,5 % en Haute-Garonne et 85,6 % dans le Bas-Rhin (tableau 8). L'analyse multivariée confirmait l'effet significatif du sexe ($p = 0,02$) et surtout du site ($p < 10^{-4}$).

Pour les RSA provenant des établissements dans lesquels il y a eu un appariement nominatif, la VPP s'élevait à 83,6 %, contre 58,5 % dans le cas contraire ($p < 10^{-4}$).

La VPP du PMSI 2003 était légèrement plus élevée si on prenait en compte le décalage temporel des deux sources : 79,0 % après prise

en compte des quelques IAM ayant une date de survenue en 2002 et 80,0 % après exclusion des RSA avec date d'hospitalisation en 2002 et, de ce fait, non soumis à l'expertise des registres.

| TABLEAU 3 |

Sensibilité du PMSI 2003

35-74 ans	Hommes	Femmes	Global
Registres 2003	1 373	350	1 723
Base PMSI 2003 : vrais positifs	1 066	247	1 313
Sensibilité (%)	77,6	70,6	76,2

| TABLEAU 4 |

Sensibilité de la base PMSI 2003 selon le sexe et l'âge

	Hommes	Femmes	Global
35-44 ans			
Registres 2003	144	23	167
PMSI 2003 : vrais positifs	118	20	138
Sensibilité (%)	81,9	87,0	82,6
45-54 ans			
Registres 2003	379	46	425
Base PMSI 2003 : vrais positifs	310	34	344
Sensibilité (%)	81,8	73,9	80,9
55-64 ans			
Registres 2003	412	79	491
Base PMSI 2003 : vrais positifs	328	60	388
Sensibilité (%)	79,6	75,9	79,0
65-74 ans			
Registres 2003	438	202	640
Base PMSI 2003 : vrais positifs	310	133	443
Sensibilité (%)	70,8	65,8	69,2

| TABLEAU 5 |

Sensibilité de la base PMSI 2003 selon le site

35-74 ans	Haute-Garonne	Lille	Bas-Rhin	Global
Registres	563	489	671	1 723
PMSI : vrais positifs	449	360	504	1 313
Sensibilité (%)	79,8	73,6	75,1	76,2

| TABLEAU 6 |

Valeur prédictive positive du PMSI 2003

35-74 ans	Hommes	Femmes	Global
PMSI 2003	1 330	336	1 666
IAM (bases 2003 registres)	1 066	247	1 313
VPP (%)	80,2	73,5	78,8

| TABLEAU 7 |

Valeur prédictive positive selon le sexe et l'âge

PMSI 2003	Hommes	Femmes	Global
35-44 ans			
PMSI 2003	144	26	170
IAM (bases 2003 registres)	118	20	138
VPP (%)	81,9	76,9	81,2
45-54 ans			
PMSI 2003	380	44	424
IAM (bases 2003 registres)	310	34	344
VPP (%)	81,6	77,3	81,1
55-64 ans			
PMSI 2003	413	76	489
IAM (bases 2003 registres)	328	60	388
VPP (%)	79,4	78,9	79,3
65-74 ans			
PMSI 2003	393	190	583
IAM (bases 2003 registres)	310	133	443
VPP (%)	78,9	70,0	76,0

| TABLEAU 8 |

Valeur prédictive positive par site

PMSI 2003 – 35-74 ans	Haute-Garonne	Lille	Bas-Rhin	Global
PMSI 2003	572	505	589	1 666
IAM (bases 2003 registres)	449	360	504	1 313
VPP (%)	78,5	71,3	85,6	78,8

3.4 FAUX POSITIFS

Les 353 épisodes de soins non confirmés par les registres comme IAM survenus en 2003 se répartissent de la façon suivante (tableau 9) :

- épisodes correspondant à des RSA non identifiés : 38,0% (dont 50,7% correspondaient à des hospitalisations hors départements);
- épisodes identifiés (62,0%) :
 - 41,9%, soit les deux tiers des cas identifiés, correspondaient à des hospitalisations de patients connus pour coronaropathie, mais sans notion de nouvel IAM ou, pour un cas, à une mort subite sans IAM ni antécédent coronaire connu,
 - 9,1% correspondaient à des patients ou des événements non connus des registres,
 - 6,5% à des hospitalisations ayant débuté fin 2002 (de ce fait, ces dossiers n'ont pas été comparés à la base 2003 des registres),
 - 0,6% à des IAM ayant débuté fin 2002 et présents dans les bases 2002 des registres,
 - 4,0% à d'autres éventualités (séjours non liés à une coronaropathie, patients en fait non domiciliés dans le département selon les registres).

| TABLEAU 9 |

Faux positifs

PMSI 2003 – 35-74 ans	Global
Épisodes non identifiés	134 38,0%
Épisodes identifiés :	
- patients avec début d'hospitalisation en 2002;	23 6,5%
- IAM 2002 (début d'hospitalisation en 2003, mais début des symptômes en 2002);	2 0,6%
- patient avec coronaropathie connue, mais sans nouvel infarctus;	147 41,6%
- mort subite sans IAM ni antécédent coronaire connu;	1 0,3%
- séjours non liés à une coronaropathie;	7 2,0%
- patients en réalité non domiciliés dans le département;	7 2,0%
- patients non connus des registres.	32 9,1%
Total	353

3.5 COMPARAISON DES NOMBRES D'IAM COMPTABILISÉS PAR LE PMSI ET LES REGISTRES EN 2003

Nous avons comparé différentes estimations du nombre d'IAM susceptibles d'être produites à partir de la base 2003 du PMSI au nombre effectif d'événements recensés par les registres pour l'année 2003, pour les territoires correspondants.

3.5.1 PMSI: un épisode de soin par période de 28 jours

Dans un premier temps, nous n'avons retenu dans le PMSI qu'un seul RSA par période de 28 jours pour une même personne, comme c'est le cas dans les registres MONICA. Le nombre total d'IAM hospitalisés expertisés par les registres pour les tranches d'âge surveillées s'élève à 1 723 en 2003, contre 1 666 dans le PMSI, ce qui représentait une sous-estimation globale par le PMSI de 3,3% pour cette approche (tableau 10).

Les nombres d'infarctus estimés à partir du PMSI sous-estimaient de 3,1% les valeurs comptabilisées par les registres pour les hommes et de 4% pour les femmes, l'écart devenant marginal lorsque l'on se limitait aux 35-64 ans (respectivement +0,2% et -1,4%).

| TABLEAU 10 |

Nombres d'infarctus du myocarde en 2003 selon la source

35-74 ans	Hommes	Femmes	Global
Registres	1 373	350	1 723
PMSI	1 330	336	1 666
PMSI/registres (%)	96,9	96,0	96,7

L'examen par classe d'âge décennale montrait que les estimations produites à partir du PMSI étaient très proches des nombres relevés par les registres avant 65 ans : le ratio des nombres d'événements fournis par ces deux sources (PMSI/registres) était égal à 101,8% entre 35 et 44 ans, 99,8% entre 45 et 54 ans et 99,6% entre 55 et 64 ans, soit précisément 100% pour l'ensemble des 35-64 ans. En revanche, le PMSI sous-estimait de 8,9% les événements survenus dans la classe d'âge 65-74 ans (tableau 11).

| TABLEAU 11 |

Nombres d'infarctus du myocarde en 2003 selon la source, le sexe et l'âge

	Hommes	Femmes	Global
35-44 ans			
Registres	144	23	167
PMSI	144	26	170
PMSI/registres (%)	100	113,0	101,8
45-54 ans			
Registres	379	46	425
PMSI	380	44	424
PMSI/registres (%)	100,3	95,7	99,8
55-64 ans			
Registres	412	79	491
PMSI	413	76	489
PMSI/registres (%)	100,2	96,2	99,6
65-74 ans			
Registres	438	202	640
PMSI	393	190	583
PMSI/registres (%)	89,7	94,1	91,1

Toutefois, les écarts n'étaient pas homogènes dans les trois sites. Pour l'ensemble des classes d'âge comprises entre 35 et 74 ans, on observait une légère surestimation des nombres d'infarctus du myocarde estimés à partir du PMSI en Haute-Garonne (+1,6%) et dans la CUDL (+3,3%), contrastant avec une sous-estimation de 12,2% dans le Bas-Rhin (tableau 12).

| TABLEAU 12 |

| TABLEAU 12 |

Nombres d'infarctus du myocarde en 2003 selon la source et le site

35-74 ans	Haute-Garonne	Lille	Bas-Rhin	Global
Registres	563	489	671	1 723
PMSI	572	505	589	1 666
PMSI/registres (%)	101,6	103,3	87,8	96,7

3.5.2 Autres modalités d'exploitation du PMSI

D'autres modalités d'exploitation du PMSI produisaient des estimations du nombre d'infarctus du myocarde hospitalisés un peu plus éloignées de celles fournies par les registres (tableau 13).

L'approche élémentaire consistant à comptabiliser tous les RSA avec diagnostic principal d'IAM, y compris les transferts et les réhospitalisations précoces, aboutissait à une surestimation du nombre d'IAM dans le PMSI égale à 12,7% : cette surestimation était observée dans les trois sites, variant de 5,2% dans le Bas-Rhin à 23,7% dans la CUDL.

Une seconde approche consistait à ne prendre en compte que les séjours en "hospitalisation complète", en excluant des séjours avec code GHM de transfert ainsi que les hospitalisations de moins de 24 heures dont le mode de sortie n'était pas un décès. Cette approche aboutissait à une surestimation globale de 6,2% par le PMSI.

Enfin, compte tenu du faible nombre des récidives d'infarctus survenues au cours de l'année calendaire (1,3%), nous avons également examiné l'approche basée sur le nombre de personnes hospitalisées pour infarctus. Les estimations ainsi produites par le PMSI étaient proches des valeurs recensées par les registres en Haute-Garonne (100,7%) ainsi que dans la CUDL (101,2%), mais sous-évaluaient les IAM de 14,3% dans le Bas-Rhin, soit une sous-estimation globale de 5%.

Comparaison des nombre d'infarctus du myocarde estimés par chaque source : variations selon différentes modalités d'exploitation du PMSI

2003 – 35-74 ans	Haute-Garonne	Lille	Bas-Rhin	Global
PMSI : tous RSA avec diagnostic principal d'IAM				
Registres	563	489	671	1 723
PMSI	631	605	706	1 942
PMSI/registres (%)	112,1	123,7	105,2	112,7
PMSI : seulement séjours en hospitalisation complète*				
Registres	563	489	671	1 723
PMSI	605	581	644	1 830
PMSI/registres (%)	107,5	118,8	96,0	106,2
PMSI : personnes hospitalisées pour IAM en 2003				
Registres (nombre d'événements)	563	489	671	1 723
PMSI (nombre de personnes)	567	495	575	1 637
PMSI/registres (%)	100,7	101,2	85,7	95,0

* Exclusion des GHM de transfert initial, des séances et des hospitalisations de moins d'un jour sauf si GHM de décès immédiat.

4. Discussion

De nombreuses études de validation des données hospitalières relatives aux IAM ont été réalisées depuis une vingtaine d'années [5-10;13-26]. Toutefois, seules quelques unes concernent des recueils de données postérieurs aux nouvelles définitions de l'infarctus du myocarde publiées par la société européenne et le collège américain de cardiologie pour prendre en compte l'évolution des moyens diagnostiques [27,28]. Parallèlement, peu de recueils concernent des données hospitalières avec diagnostics codés en CIM-10 [10,26]. À ce jour, aucune étude pluricentrique du PMSI pour la surveillance des infarctus du myocarde n'avait encore été effectuée en France.

Cette étude apporte un certain nombre d'informations sur les conditions d'exploitation du PMSI MCO pour la surveillance des taux de survenue des IAM hospitalisés en France, ainsi que sur les limites de cette source. Toutefois, il ne s'agit pas à proprement parler d'une étude de validation qui impliquerait un "gold standard" avec retour au dossier médical pour expertiser les symptômes, dosages biologiques et électrocardiogrammes selon des critères précis [29]. Il s'agit plutôt d'un étalonnage du PMSI MCO sur les données des trois registres français des cardiopathies ischémiques dont le protocole a été simplifié depuis 1997 : la recherche de l'exhaustivité est restée très systématisée, mais le diagnostic est basé sur celui des cliniciens (lettre de sortie) ; toutefois, un recueil détaillé d'informations sur les symptômes, marqueurs biologiques et électrocardiogrammes est effectué chaque année pendant quelques semaines pour permettre l'étude des évolutions à critères constants [30,31].

Dans la base nationale 2003 du PMSI MCO, les infarctus du myocarde ont été appréhendés à partir du diagnostic principal des RSA. Les classes d'âge concernées étaient limitées aux 35-74 ans, âges auxquels la quasi-totalité des IAM non immédiatement létaux sont hospitalisés. Un seul séjour par période de 28 jours a été retenu. Les critères géographiques étaient la domiciliation dans les départements du Bas-Rhin ou de la Haute-Garonne ou la CUDL, quel que soit le lieu d'hospitalisation. À noter toutefois que les résidents de la CUDL domiciliés à Villeneuve d'Ascq ou dans une commune n'ayant pas un code postal spécifique de l'appartenance à la CUDL ont été exclus du champ de l'étude dans les deux sources de données (PMSI et registres).

4.1 SENSIBILITÉ

La sensibilité du PMSI MCO par rapport aux registres MONICA a été estimée à 76,2 % entre 35 et 74 ans. La sensibilité était relativement homogène dans les trois sites (73,6 %, 75,1 %, 79,8 %). Elle était toutefois plus faible pour les personnes âgées de 65 à 74 ans (69,2 %) que pour celles de 35 à 64 ans (80,3 %).

Cette valeur de la sensibilité est assez faible. Dans les études publiées au cours des 20 dernières années, il existait des variations de la sensibilité entre 68 % [23] et 97 % [9] pour les IAM "certains", les valeurs étant plus faibles lorsque la sensibilité était rapportée à l'ensemble des infarctus "probables" ou "possibles". De fait, il existe des différences méthodologiques notables portant sur

différents aspects. En premier lieu, la méthode définissant le "gold standard" est en général basée sur des critères diagnostics précis avec catégorisation selon le niveau de probabilité de l'IAM (certain, probable ou possible) et rarement, sur le diagnostic porté par les cliniciens à la sortie [20], comme c'est le cas dans cette étude, ce qui ne permet pas d'établir le niveau de probabilité de l'IAM sur des critères explicites.

En second lieu, la plupart des études ont évalué la sensibilité du PMSI, non seulement à partir du diagnostic principal, mais aussi à partir d'un nombre variable de diagnostics associés. Madsen [9] a étudié l'impact de la prise en compte des diagnostics associés sur la sensibilité qui variait entre 92,2 % (diagnostic principal uniquement) et 96,6 % (diagnostics principal et associés). Les diagnostics associés n'ont pas été étudiés dans la présente étude : leur prise en compte aurait impliqué l'examen de nombreux RSA supplémentaires (soit 2007) de VPP probablement médiocre [26].

Les faux négatifs ont été appréhendés de façon exploratoire, qualitative, non systématique et non représentative (annexe 4). Cette étude exploratoire a permis de faire les constatations suivantes : la proportion de RSA issus de plusieurs RUM était élevée. Trois types de situations prédominaient : d'une part, des séjours avec complications cardiaques, d'autre part, des séjours avec diagnostic principal d'angor instable ou de cardiopathie ischémique chronique et, enfin, des séjours longs avec pathologies non cardiaques graves, l'IAM n'ayant alors été cause de soins que pour une faible part de l'hospitalisation. Les faux négatifs ont été étudiés dans plusieurs études [5,7,8,21] : les diagnostics portés sur les résumés d'hospitalisation de ces faux négatifs se répartissaient de façon variable selon les études entre des cardiopathies ischémiques autres que les IAM, des pathologies cardiaques non coronaires, notamment l'insuffisance cardiaque, des douleurs thoraciques codées comme simples symptômes et enfin, des causes non cardiaques. Compte tenu des consignes de codage du diagnostic principal en vigueur en France jusqu'en 2008 et des critères de sélection du diagnostic principal en cas d'hospitalisation multi-unités, la sensibilité du diagnostic principal d'IAM ne pouvait atteindre l'exhaustivité. Le codage des complications cardiaques graves était inévitable lorsque celles-ci exigeaient des soins longs. De même, les IAM survenant au cours d'hospitalisations pour pathologie non coronaire grave (avec acte chirurgical classant par exemple) ne peuvent pas figurer en diagnostic principal.

Il a été observé une diminution de la sensibilité avec l'âge. Cette association a été observée par Boyle [7] qui observe parallèlement une diminution de la sensibilité en cas d'antécédent coronaire connu. De fait, le risque de complications ou de pathologies intercurrentes augmente avec l'âge. Les co-morbidités observées lors des hospitalisations pour IAM ont été étudiées récemment par Lichtman [32] : des pathologies non cardiaques aiguës ne sont pas exceptionnelles : elles seraient présentes dans 7 % des hospitalisations pour IAM, les plus fréquentes étant des pneumonies sévères, des hémorragies digestives sévères, des accidents vasculaires cérébraux et des sepsis.

Ces difficultés devraient diminuer avec les évolutions récentes du PMSI. D'une part, les diagnostics principaux de chacun des RUM du RSA sont identifiables depuis 2008 dans le PMSI, permettant dorénavant l'identification des IAM survenus lors de séjours pour des motifs non coronaires, sous réserve de les rechercher. Surtout, la modification de la définition du diagnostic principal introduite en 2009, dorénavant "le problème de santé qui a motivé l'admission du patient dans l'Unité médicale (UM) déterminé à la sortie de l'UM" devrait diminuer les faux négatifs dus au codage des complications. Mais ces mesures ne résoudre pas la question des erreurs de codage des IAM, qui pourraient toutefois être minimisées par le rappel des consignes de codages, les vérifications inhérentes à la T2A...

4.2 VALEUR PRÉDICTIONNELLE POSITIVE

La VPP du PMSI 2003 par rapport aux données 2003 des registres était globalement égale à 78,8 % (un seul séjour retenu dans le PMSI par période de 28 jours). Dans les études publiées, la VPP varie entre 56 % [23] et 96 % [20]. Les différences méthodologiques évoquées ci-dessus expliquent une part de cette variabilité avec, en premier lieu, l'utilisation ou non d'un algorithme catégorisant les séjours selon le niveau de probabilité de l'infarctus en fonction des critères cliniques, biologiques et ECG. La VPP varie entre 56 % et 87 % lorsqu'elle est référencée aux infarctus du myocarde "certains" et entre 75 % et 96 % lorsque la définition du standard est élargie aux infarctus "certains, probables ou possibles". Le nombre de diagnostics pris en compte a également un impact sur la VPP qui est plus élevée quand seul le diagnostic principal est pris en considération (88 % à 96 %).

D'autres facteurs de variation ont été évoqués telles les pratiques de diagnostic et de codage [21]. Des différences dans la formulation du diagnostic (par exemple IAM ou syndrome coronaire aigu) peuvent constituer un élément d'explication. De fait, l'année 2003 se situe durant la période de transition caractérisée par la généralisation progressive du dosage des troponines et des nouveaux critères diagnostiques de l'IAM. De fait, il existait des différences significatives entre sites : la valeur la plus faible était observée dans la CUDL (71,3 %) et la plus élevée dans le Bas-Rhin (85,6 %). Sur Lille, quelques cas de syndromes coronaires aigus à troponines positives n'étaient pas retenus comme IAM. Des différences géographiques ont également été observées en Finlande [10] avec, toutefois, diminution de cette variabilité lorsque tous les syndromes coronaires aigus, y compris l'angor instable, étaient pris en compte dans l'analyse.

Il n'est pas habituel de prendre en compte les codes I23 qui correspondent à des complications récentes d'infarctus du myocarde. En pratique, ce codage est peu utilisé dans le PMSI puisque seulement 8 RSA présentaient ce code. Parmi ceux-ci, 4 ont été confirmés par les registres comme IAM.

21,2 % des épisodes d'hospitalisation ont été considérés comme des faux positifs. Parmi eux, la moitié correspondait à des RSA non identifiés (134 soit 38 %) ou, moins souvent, à des patients identifiés mais non connus des registres (32 épisodes de soins, soit 9,1 %). Parmi les faux positifs correspondant à des RSA non identifiés, 18 n'avaient pas de numéro de chaînage anonyme et ont donc été considérés comme des épisodes de soins. Dans l'hypothèse, extrême, où ces 18 RSA correspondraient tous à des rehospitalisations précoces, et non à de nouveaux épisodes de soins, la VPP serait augmentée de 1 point (79,7 % *versus* 78,8 %). Par ailleurs, la moitié (50,75 %) de

ces épisodes non identifiés a été hospitalisée hors des départements de résidence. Un constat similaire a été fait par Mahonen [8] qui a observé que les hospitalisations survenues en dehors des zones couvertes par les registres représentaient le tiers ou la moitié des faux positifs selon les périodes. On ne peut exclure que certains de ces cas correspondent à des IAM non repérés par les registres, ce qui aurait alors pour effet d'augmenter la VPP. Rappelons toutefois que l'objectif de l'étude consiste à étalonner le PMSI au regard des données recueillies par les registres.

4.3 ESTIMATION DU NOMBRE D'INFARCTUS DU MYOCARDE

À un même infarctus du myocarde peuvent correspondre plusieurs RSA. En effet, il n'est pas rare que des personnes hospitalisées pour IAM soient transférées dans un autre établissement et qu'un même épisode de soins donne lieu à plusieurs RSA. Des rehospitalisations brèves peuvent également survenir après la sortie, pour examens de contrôle par exemple ; en 2003, avant l'introduction de la T2A (et des contrôles inhérents), il était probablement moins rare que les RSA correspondants comportent un code d'IAM.

La meilleure estimation globale du nombre d'infarctus du myocarde survenus dans l'année a été produite sur la base du nombre d'épisodes aigus, un seul épisode étant comptabilisé par période de 28 jours, conformément aux conventions des registres MONICA : le ratio du nombre d'épisodes comptabilisés dans le PMSI en 2003 au nombre d'infarctus validés par les registres la même année est égal à 96,7 % pour les personnes d'âge compris entre 35 et 74 ans et à 100 % pour les 35-64 ans. Sur cette base, le nombre d'IAM survenus en métropole en 2003 a été estimé à partir de la base du PMSI, soit 23 127 épisodes d'IAM pour la classe d'âge comprise entre 35 et 64 ans, 37 525 pour celle des 35-74 ans et 64 527 tous âges confondus.

Par contre, le nombre d'IAM était surestimé de 12,7 % par la prise en compte de tous les séjours comportant un diagnostic principal d'IAM (35-74 ans). La surestimation était de 6,2 % lorsque l'on considérait les "hospitalisations complètes", en excluant les séjours avec code GHM de transfert ainsi que les hospitalisations de moins d'un jour (sauf lorsque le mode de sortie était un décès). Enfin, l'estimation du nombre d'IDM basée sur le nombre de personnes ayant eu au moins une hospitalisation avec diagnostic principal d'IAM aboutissait à une estimation proche de celle produite à partir du nombre d'épisodes de soins, soit une sous estimation de 5 %. De fait les récurrences survenues au cours de la même année calendaire sont peu fréquentes : 1,3 % des IDM expertisés par les registres en 2003.

Dans l'ensemble, les études de validation des IAM dans les bases de données d'hospitalisation sont centrées sur l'étude de la sensibilité et de la VPP de cette source. Mais peu de publications ont étudié la validité des nombres ou taux d'événements produits par ces sources [23] ou leurs variations selon différentes modalités d'exploitation des données d'hospitalisation. Le traitement des séjours multiples est assez peu décrit. Quelques études effectuées par des équipes impliquées dans le projet MONICA ont considéré un seul séjour par période de 28 jours [8,10,17], d'autres, uniquement le premier séjour dans l'année [25]. De fait, si les séjours multiples ne sont pas rares dans le PMSI, la survenue de nouveaux IAM au cours de la même année calendaire l'est beaucoup plus (1,3 % en 2003 pour les trois registres des cardiopathies ischémiques).

Les conclusions des études de validation sur la possibilité d'utiliser les bases de données sur les hospitalisations pour le suivi des tendances sont variables : parfois positives [9,10,24], mais souvent mitigées [8,23], voire réservées [5,7,13]. De fait, le nombre de diagnostics pris en compte dans les études de validation est source de disparités. Ainsi, Boyle et Rosamond ont observé une surestimation des séjours sur la base d'études dans lesquelles plusieurs diagnostics étaient pris en compte [7,23]. Austin, sur la base d'un seul diagnostic (pathologie ayant mobilisé le plus de soins), observe un nombre d'infarctus très similaire dans les deux sources comparées [24].

4.4 DIFFICULTÉS

La comparaison des cas d'IAM dans les deux sources de données nécessitait une identification préalable des données du PMSI. Puisque l'objectif de l'étude consistait à étalonner la base nationale des RSA, et compte tenu de la diversité des logiciels utilisés dans les DIM, d'une part, et des différences structurelles des fichiers de RSS et de RSA, d'autre part, il a été choisi de demander aux DIM d'identifier directement les RSA sur la base de la correspondance avec les RSS extraits selon les critères précisés. Plusieurs informations contenues, directement ou indirectement, dans ces deux fichiers devaient permettre le rapprochement des enregistrements : mois de sortie, durée de séjour, sexe, âge, code postal/géographique. Il est toutefois possible que quelques erreurs soient survenues.

L'identification des séjours n'a pas pu être exhaustive : 15 % des RSA n'ont pas été identifiés, pour des raisons techniques dans un certain nombre de cas (changement de logiciel) ou lorsque les hospitalisations se sont déroulées hors des départements de résidence. Parmi eux, 130 RSA ont fait l'objet d'un appariement sur la base de la compatibilité du sexe, de l'âge, des dates d'événement et du lieu de résidence. Au final, 134 épisodes de soins de la base PMSI 2003 n'ont pas pu être identifiés directement ou indirectement, dans la mesure où ils ne correspondaient à aucun des faux négatifs signalés par les registres. Ces épisodes ont tous été considérés comme des faux positifs, même lorsqu'ils provenaient d'établissements de soins situés hors du département de résidence. Ces RSA non identifiés ont pu entraîner une sous-estimation de la VPP, dans l'éventualité où certains d'entre eux correspondraient à des infarctus du myocarde hospitalisés lors de déplacements en province n'ayant pas été portés à la connaissance des registres (si les suites ont été simples et les visites ou examens de contrôles également effectués en province). Par contre, leur impact sur la sensibilité est probablement faible, compte tenu des vérifications effectuées par comparaison avec les faux négatifs et ils n'ont eu aucun impact sur la comparaison des nombre d'IAM produits par ces deux sources.

De fait, les registres MONICA ne comptabilisent qu'un IAM par période de 28 jours. Il était donc nécessaire de définir également des épisodes de soins d'une durée de 28 jours dans le PMSI. Cette procédure a été effectuée à partir des dates d'hospitalisation pour les RSA identifiés et, à défaut, du numéro de chaînage anonyme quand il existait. Le chaînage du PMSI n'a été introduit en France qu'en 2002 et n'était pas encore totalement généralisé début 2003. Aux débuts du chaînage, les erreurs étaient nombreuses et le taux d'anomalie élevé [33]. Dans la base nationale des RSA, ce retard à l'introduction du numéro de chaînage limite les possibilités de chaînage des séjours d'une même personne. La proportion de RSA avec numéro de chaînage absent ou erroné était égale à 12,5 % dans la base nationale 2003 pour les IAM. Les années suivantes

ont été marquées par une amélioration sensible de la qualité du chaînage : 5,5 % en 2004 et 1,1 % en 2007. Dans l'étude, l'impact de ces RSA sans numéro de chaînage anonyme est marginal dans la mesure où la proportion de RSA sans numéro de chaînage et non identifiés est faible : 2,1 % en 2003, dont la moitié ont fait l'objet d'un appariement manuel avec des cas des registres (correspondance multicritères). Sous l'hypothèse d'un ratio de séjours multiples par épisode identique pour les RSA identifiés ou avec numéro de chaînage et pour les RSA non identifiés et sans numéro de chaînage, la sensibilité serait inchangée, le volume d'épisodes serait diminué de 0,3 % (2,1 %-2,1 %/1,17, avec 1,17=nombre moyen de RSA par épisodes de 28 jours) et la VPP augmenterait en conséquence. Sous l'hypothèse extrême où tous les RSA non appariés, non identifiés et sans numéro de chaînage seraient des doublons d'épisodes déjà comptabilisés, la VPP serait alors augmentée, ce qui constitue un maximum de 1 %.

Une difficulté supplémentaire tient au décalage de la temporalité dans ces deux bases. Ce décalage entraîne une légère sous-estimation de la sensibilité et de la VPP réelle. Pour la sensibilité, la comparaison de la base PMSI 2003 à la base registre 2003 exclut *de facto* les IAM survenus en fin d'année 2003 dont l'hospitalisation s'est prolongée en début d'année 2004. L'ajout de ces épisodes augmente la sensibilité réelle du PMSI de 76,2 % à 78,3 %. À l'inverse, certains RSA du début de l'année ne peuvent pas être dans la base 2003 des registres si la dates de début de symptômes/début d'hospitalisation était en 2002. Ainsi, la prise en compte des IAM avec date de survenue en fin d'année 2002 augmente la VPP de 78,8 % à 79,0 % et l'exclusion des RSA avec date d'hospitalisation en 2002 et, de ce fait, non soumis à l'expertise les registres, porte la VPP à 80,0 %. De plus, il n'était pas possible de chaîner les séjours d'une même personne entre les années 2003 et 2004 en raison d'une modification de l'identifiant anonyme en 2004. Cette difficulté n'a pas d'impact sur les analyses principales qui comparent les années 2003 dans les deux sources. Mais elle contribue probablement à sous-estimer la sensibilité dans l'analyse secondaire prenant en compte les IAM survenus en fin d'année pour lesquels la sortie d'hospitalisation a eu lieu au début de l'année 2004 dans les cas où les RSA n'ont pas été identifiés, mais de façon marginale, compte tenu des contrôles et des appariements manuels effectués.

Enfin, l'exhaustivité du PMSI n'était pas totale en 2003. L'exhaustivité globale du PMSI a été évaluée à 97 % par la Direction de la recherche, des études, de l'évaluation et des statistiques en 2003, pour atteindre 99 % à partir de 2004. Cet aspect n'a pas pu être évalué précisément dans cette étude. L'examen des données mensuelles montre qu'un défaut d'exhaustivité ne peut pas être exclu sur un établissement, pour un semestre. Mais il s'agit d'un établissement à faible recrutement (<50 en 2004) avec, en conséquence, un impact potentiel limité sur la sensibilité ainsi que sur le nombre total d'IAM.

4.5 LIMITATIONS

Les classes d'âge investiguées ont été limitées aux 35-74 ans, ce qui ne permet pas de généraliser ces résultats aux IAM survenant chez des personnes plus âgées. De fait, nous avons observé une diminution de la sensibilité et une sous-estimation du nombre d'infarctus chez les 65-74 ans. Il est probable que ce phénomène concerne également les IAM survenant chez les personnes de 75 ans et plus. Quelques études ont étudié la VPP des données d'hospitalisations chez les plus

âgés, qui s'avère élevée [10,20,25] ; une seule a également examiné la sensibilité, mais n'a pas observé de diminution de la sensibilité ni de la VPP pour les personnes âgées de 75 ans ou plus, comparativement aux 35-74 ans [10].

De plus, il existait des disparités entre sites qui limitent les possibilités de déclinaisons départementales voire régionales des estimations basées sur le PMSI. Il est probable que ces disparités soient en partie liées à des pratiques de diagnostic et de codage différentes. L'année 2003 correspond en effet à une période de mutation

marquée par la généralisation progressive du dosage des troponines et l'évolution des critères diagnostics de l'infarctus du myocarde (distinction entre IAM avec ou sans sus-décalage ST, élargissement de la problématique coronaire aiguë au concept de syndrome coronaire aigu).

Enfin, les deux sources de données ne sont pas indépendantes. Elles sont toutes deux basées sur le diagnostic des cliniciens à la fin de l'hospitalisation et non sur une expertise effective des dossiers selon des critères explicites.

5. Conclusion

Cette étude avait pour objectif d'étalonner les estimations des taux de survenue des IAM produites à partir du PMSI MCO au regard des données recueillies par les registres des trois registres des cardiopathies ischémiques ce, pour les classes d'âge comprises entre 35 et 74 ans. Les données ont été extraites de la base nationale 2003 du PMSI MCO, à partir du diagnostic principal.

La comparaison de la base 2003 du PMSI aux données de la même année des registres a permis d'estimer sa sensibilité à 76,2 % et sa VPP à 78,8 %. Pour les sites et les tranches d'âge concernées par l'étude, le nombre d'infarctus du myocarde était sous-estimé de 3,3 % par la prise en compte d'un séjour par période de 28 jours, alors que la prise en compte de toutes les hospitalisations d'au moins 24 heures (sauf décès immédiat) et non classées dans un GHM de transfert conduisait à surestimer les IAM de 6,2 %. *A fortiori*, la comptabilisation de tous les séjours avec diagnostic principal d'IAM, y compris les transferts ou les hospitalisations de moins de 24 heures, surestimait le nombre total d'IAM de 12,7 %.

De telles études sont nécessaires pour contribuer à la réflexion sur la validité des estimations produites à partir du PMSI [34]. Elles soulignent les limites de cette source de données médico-administratives, contrainte par ses propres finalités et règles, qui peuvent toutefois différer en fonction de la problématique étudiée. Elles permettent d'éclairer les conditions d'utilisation de cet outil pour produire des estimations nationales, en l'occurrence, sur les taux de survenue des IAM. Les variations géographiques observées incitent toutefois à la prudence pour des niveaux géographiques

fins (départements, petites régions) qui sont plus dépendants des pratiques de diagnostic et de codage d'un petit nombre d'établissements de soins.

Cette étude a été réalisée sur les données de l'année 2003, dans un contexte fortement évolutif. De nombreuses évolutions du PMSI sont survenues depuis cette date. La mise en place de la tarification à l'activité a probablement entraîné une amélioration du recueil et du codage du diagnostic principal du fait de son impact sur le financement des établissements de soins et des contrôles réalisés ; ces modifications sont susceptibles d'améliorer l'exhaustivité, d'une part, et la VPP, d'autre part (diminution de l'utilisation du code IAM pour des séjours pour surveillance et/ou contrôles). De plus, plusieurs modifications récentes sont susceptibles d'avoir amélioré la sensibilité du PMSI : il est en effet devenu possible depuis 2008 d'identifier les RSA comportant au moins un RUM avec diagnostic principal d'IAM, ce qui pourrait diminuer les faux négatifs dus à des IAM survenus au cours de séjours liés à des pathologies graves non coronaires. Enfin, la définition du diagnostic principal a été modifiée en 2009 : c'est dorénavant "le problème de santé qui a motivé l'admission du patient dans l'UM déterminé à la sortie de l'UM". Cette modification devrait diminuer les faux négatifs dus au codage des complications. De ces évolutions importantes découle la nécessité de renouveler périodiquement l'étude des qualités métrologiques du PMSI. Il semble par ailleurs souhaitable d'élargir cette surveillance à l'ensemble des syndromes coronaires aigus (que les registres des cardiopathies ischémiques surveillent depuis 2006) et non seulement aux IAM.

Références bibliographiques

- [1] Madsen M, Gudnason V, Pajak A, Palmieri L, Rocha E, Salomaa V, Sans S, Steinbach K, Vanuzzo D, on behalf of the EUROCISS research group. Population-based register of acute myocardial infarction: manual of operations. *Eur J Cardiovasc Prev Rehabil* 2007;14 (Suppl 3):S3-S22.
- [2] Tunstall-Pedoe H, Kuulasmaa K, Mähönen M, Tolonen H, Ruokokoski E, Amouyel P, for the WHO MONICA (Monitoring trends and determinants in cardiovascular disease) project. Contribution of trends in survival and coronary-events rates to changes in coronary heart disease mortality: 10-year results from 37 WHO MONICA project populations. *Lancet* 1999;353:1547-57.
- [3] Ducimetière P, Jouglu E, Haas B, Montaye M, Ruidavets JB, Amouyel P, Arveiler D, Ferrières J, Bingham A. Mortalité coronaire en France selon les sources d'information. *Rev Epidemiol Santé Publique* 2006;54:453-61.
- [4] Guide méthodologique de production des informations relatives à l'activité médicale et à sa facturation en médecine, chirurgie, obstétrique et odontologie applicable au 1^{er} mars 2009 – fascicule spécial n° 2009/5 bis du Bulletin officiel (BO), www.atih.sante.fr/?id=000240009BFF.
- [5] Beaglehole R, Stewart AW, Walker P. Validation of coronary heart disease hospital discharge data. *Aust NZ J Med* 1987;17:43-6.
- [6] Palomäki P, Miettinen H, Mustaniemi H, Lehto S, Pyörälä K, Mähönen M, Tuomilehto J. Diagnosis of acute myocardial infarction by MONICA and FINMONICA diagnostic criteria in comparison with hospital discharge diagnosis. *J Clin Epidemiol* 1994;47:659-66.
- [7] Boyle CA, Dobson AJ. The accuracy of hospital records and death certificates for acute myocardial infarction. *Aust NZ J Med* 1995;25:316-23.
- [8] Mähönen M, Salomaa V, Brommels M, Molarius A, Miettinen H, Pyörälä K, Tuomilehto J, Arstila M, Kaarsalo E, Ketonen M, Kuulasmaa K, Lehto S, Mustaniemi H, Niemelä M, Palomäki P, Torppa J, Vuorenmaa T. The validity of hospital discharge register data on coronary heart disease in Finland. *Eur J Epidemiol* 1997;13:403-15.
- [9] Madsen M, Davidsen M, Rasmussen S, Abildstrom SZ, Osler M. The validity of the diagnosis of acute myocardial infarction in routine statistics: a comparison of mortality and hospital discharge data with the Danish MONICA registry. *J Clin Epidemiol* 2003;56:124-30.
- [10] Pajunen P, Koukkunen H, Ketonen M, Jerkkola T, Immonen-Räihä P, Kärjä-Koskenkari P, Mähönen M, Niemelä M, Kuulasmaa K, Palomäki P, Mustonen J, Lehtonen A, Arstila M, Vuorenmaa T, Lehto S, Miettinen H, Torppa J, Tuomilehto J, Kesäniemi YA, Pyörälä K, Salomaa V. The validity of the Finnish hospital discharge register and causes of deaths register data on coronary heart disease. *Eur J Cardiovasc Prev Rehabil* 2005;12:132-7.
- [11] Tunstall-Pedoe H, Kuulasmaa K, Tolonen H, Davidson M, Mendis S. MONICA – monograph and multimedia source book – world's largest study of heart disease, stroke, risk factors and population trends 1979-2002. Geneva: WHO, 2003.
- [12] Tunstall-Pedoe H, Kuulasmaa K, Amouyel P, Arveiler D, Rajakangas AM, Pajak A. Myocardial infarction and coronary deaths in the World Health Organization MONICA project. Registration procedures, event rates, and case-fatality rates in 38 populations from 21 countries in four continents. *Circulation* 1994;90:583-612.
- [13] Heliövaara M, Reunanen A, Aromaa A, Knekt P, Aho K, Suhonen O. Validity of hospital discharge data in a prospective epidemiological study on stroke and myocardial infarction. *Acta Med Scand* 1984;216:309-15.
- [14] Lindblad U, Rastam L, Ranstam J, Peterson M. Validity of register data on acute myocardial infarction and acute stroke: the Skaraborg hypertension project. *Scand J Soc Med* 1993;21:3-9.
- [15] Mascioli SR, Jacobs DR, Kottke TE. Diagnostic criteria for hospitalized acute myocardial infarction: the Minnesota experience. *Int J Epidemiol* 1989;18:76-83.
- [16] Pietilä K, Tenkanen L, Mänttari M, Manninen V. How to define coronary heart disease in register-based follow-up studies: experience from the Helsinki heart study. *Ann Med* 1997;29:253-9.
- [17] Hammar N, Alfredsson L, Rosen M, Spetz CL, Kahan T, Ysberg AS. A national record linkage to study acute myocardial infarction incidence and case fatality in Sweden. *Int J Epidemiol* 2001;30:S30-4.

- [18] Pladevall M, Goff DC, Nichaman MZ, Chan F, Ramsey D, Ortiz C, Labarthe DR. An assessment of the validity of ICD code 410 to identify hospital admissions for myocardial infarction: the Corpus Christi heart project. *Int J Epidemiol* 1996;25:948-52.
- [19] McAlpine R, Pringle S, Pringle T, Lorimer R, MacDonald TM. A study to determine the sensitivity and specificity of hospital discharge diagnosis data used in the MICA study. *Pharmacoepidemiol Drug Saf* 1998;7:311-8.
- [20] Levy AR, Tamblyn RM, Fitchett D, McLeod PJ, Hanley JA. Coding accuracy of hospital discharge data for elderly survivors of myocardial infarction. *Can J Cardiol* 1999;15:1277-82.
- [21] Cox JL, Melady MP, Chen E, Naylor CD. Towards improved coding of acute myocardial infarction in hospital discharge abstracts: a pilot project. *Can J Cardiol* 1997;13:351-8.
- [22] Heckbert S, Kooperberg C, Safford M, Psaty B, Hsia J, McTiernan A, Gaziano JM, Frishman WH, Curb JD. Comparison of self-report, hospital discharge codes, and adjudication of cardiovascular events in the women's health initiative. *Am J epidemiol* 2004;160:1152-8.
- [23] Rosamond WD, Chambless LE, Sorlie PD, Bell EM, Weitzman S, Smith JC, Folsom AR. Trends in the sensitivity, positive predictive value, false-positive rate, and comparability ratio of hospital discharge diagnosis codes for acute myocardial infarction in four US communities, 1987-2000. *Am J Epidemiol* 2004;160:1137-46.
- [24] Austin PC, Daly PA, Tu JV. A multicenter study of the coding accuracy of hospital discharge administrative data for patients admitted to cardiac care units in Ontario. *Am Heart J* 2002;144:290-6.
- [25] Kiyota Y, Schneeweiss S, Glynn RJ, Cannuscio C, Avorn J, Salomon DH. Accuracy of medicare claims-based diagnosis of acute myocardial infarction: estimating positive predictive value on the basis of review of hospital records. *Am Heart J* 2004;148:99-104.
- [26] Joensen AM, Jensen MK, Overvad K, Dethlefsen C, Schmidt E, Rasmussen L, Tjonneland A, Johnsen S. Predictive values of acute coronary syndrome discharge diagnoses differed in the Danish National Patient Registry. *J clin epidemiol* 2009;62:188-94.
- [27] Myocardial infarction redefined – a consensus document of the Joint European Society of Cardiology/American College of Cardiology Committee for the redefinition of myocardial infarction. *Eur Heart J* 2000;21:1502-13.
- [28] Thygesen K, Alpert JS, White HD, on behalf of the joint ESC/ACCF/AHA/WHF task force for the redefinition of myocardial infarction. Universal definition of myocardial infarction. *Circulation* 2007;116:2634-53.
- [29] Luepker RV, Apple FS, Christenson RH *et al.* Case definitions for acute coronary heart disease in epidemiology and clinical research studies: a statement from the AHA Council on epidemiology and prevention; AHA statistics Committee; World Heart Federation Council on epidemiology and prevention; the European Society of Cardiology Working Group on epidemiology and prevention; Centers for Disease Control and Prevention; and the National Heart, Lung, and Blood Institute. *Circulation* 2003;108:2543-9.
- [30] Montaye M, Ducimetière P, Ruidavets JB, Arveiler D, Dallongeville J, Bingham A, Ferrières J, Wagner A, Amouyel P. Le gradient Nord-Sud de la morbidité et de la mortalité coronaires en France : données récentes des registres français des cardiopathies ischémiques, 1997-2002. *BEH* 2006;8-9:62-4.
- [31] Arveiler D, Wagner A, Ducimetière P, Montaye M, Ruidavets JB, Bingham A, Ferrières J, Amouyel P, Haas B. Trends in coronary heart disease in France during the second half of the 1990s. *Eur J Cardiovasc Prev Rehabil* 2005;12:209-15.
- [32] Lichtman JH, Spertus JA, Reid KJ, Radford MJ, Rumsfeld JS, Allen NB, Masoudi FA, Weintraub WS, Krumholz HM. Acute non cardiac conditions and in-hospital mortality in patients with acute myocardial infarction. *Circulation* 2007;116:1925-30.
- [33] Trombert-Paviot B, Couris CM, Couray-Targe S, Rodrigues JM, Colin C, Schott AM. Qualité et utilité d'un identifiant patient anonyme et unique pour le chaînage des séjours hospitaliers dans les bases de données médico-économiques françaises. *Rev Epidemiol Santé Publique*. 2007;55:203-11.
- [34] Goldberg M, Quantin C, Guéguen A, Zins M. Bases de données médico-administratives et épidémiologie : intérêts et limites. *Courrier des statistiques* 2008;124:59-70.

Annexe 1 – Formats de RSA en vigueur en 2003

RSA 207

	Taille	Début	Fin
Numéro Finess de l'entité juridique	9	1	9
Numéro de version du format du RSA (207)	3	10	12
N° d'index du RSA	10	13	22
Numéro de version du format du "RSS-groupé" (ou du RSS) lu	3	23	25
Numéro de version de GENRSA (067)	3	26	28
Groupage lu sur le "RSS-groupé" : version, CMD, GHM, code retour	9	29	37
Groupage obtenu par GENRSA : version, CMD, GHM, code retour	9	38	46
Nombre de RUM composant le RSS d'origine	2	47	48
Âge en années	3	49	51
Âge en jours	3	52	54
Sexe	1	55	55
Mode d'entrée dans le champ du PMSI-MCO	1	56	56
Provenance	1	57	57
Mois de sortie	2	58	59
Année de sortie	4	60	63
Mode de sortie du champ PMSI-MCO	1	64	64
Destination	1	65	65
Type de séjour	1	66	66
Durée totale du séjour dans le champ du PMSI (vide si séances)	3	67	69
Durée de la période sur laquelle s'étalent les séances (si séances)	3	70	72
Code géographique de résidence	5	73	77
Poids à la naissance (en grammes)	4	78	81
Nombre de séances	2	82	83
IGS 2	3	84	86
Code de prise en charge pour les cliniques sinon ' '	1	87	87
Facture associée au RSA à 0 euros pour les cliniques sinon ' '	1	88	88
Durée de séjour en réa	3	89	91
Filler 2	3	92	94
Diagnostic principal (DP)	6	95	100
Diagnostic relié (DR)	6	101	106
Nombre de diagnostics associés significatifs (nDAS) dans ce RSA	2	107	108
Nombre de zones d'actes (nA) dans ce RSA	2	109	110
DAS n° 1	6		
.....			
DAS n° nDAS	6		
Zone d'acte n°1	Acte n° 1 (CDAM)	4	
	Nombre de réalisation de l'acte n° 1	2	
...	...		
	...		
Zone d'acte n° nA	Acte n° nA (CDAM)	4	
	Nombre de réalisation de l'acte n° nA	2	

RSA 208

	Taille	Début	Fin
Numéro Finess de l'entité juridique	9	1	9
Numéro de version du format du RSA (208)	3	10	12
N° d'index du RSA	10	13	22
Numéro de version du format du "RSS-groupé" (ou du RSS) lu	3	23	25
Numéro de version de GENRSA	3	26	28
Groupage lu sur le "RSS-groupé" : version, CMD, GHM, code retour	9	29	37
Groupage obtenu par GENRSA : version, CMD, GHM, code retour	9	38	46
Nombre de RUM composant le RSS d'origine	2	47	48
Âge en années	3	49	51
Âge en jours	3	52	54
Sexe	1	55	55
Mode d'entrée dans le champ du PMSI-MCO	1	56	56
Provenance	1	57	57
Mois de sortie	2	58	59
Année de sortie	4	60	63
Mode de sortie du champ PMSI-MCO	1	64	64
Destination	1	65	65
Type de séjour	1	66	66
Durée totale du séjour dans le champ du PMSI (vide si séances)	3	67	69
Durée de la période sur laquelle s'étalent les séances (si séances)	3	70	72
Code géographique de résidence	5	73	77
Poids à la naissance (en grammes)	4	78	81
Nombre de séances	2	82	83
IGS 2	3	84	86
Code de prise en charge pour les cliniques sinon ' '	1	87	87
Facture associée au RSA à 0 euros pour les cliniques sinon ' '	1	88	88
Durée de séjour en réa	3	89	91
Filler 2	3	92	94
Diagnostic principal (DP)	6	95	100
Diagnostic relié (DR)	6	101	106
Nombre de diagnostics associés significatifs (nDAS) dans ce RSA	2	107	108
Nombre de zones d'actes (nA) dans ce RSA	2	109	110
DAS n° 1	6		
.....			
DAS n° nDAS	6		
Acte n° 1 (CCAM)	7		
Zone d'acte n° 1			
Phase de l'acte n° 1	1		
Nb de réalisation de l'Acte n° 1	2		
...			
...			
...			
Acte n° nA (CCAM)	7		
Zone d'acte n° nA			
Phase de l'acte n° nA	1		
Nombre de réalisation de l'acte n° nA	2		

Annexe 2 – Codes CIM-10

Codes en I21 de la CIM-10 correspondant aux diagnostics d'infarctus du myocarde. La sélection des séjours a porté sur les nouveaux codes CIM-10 et sur les codes antérieurs au 1 ^{er} janvier 2002		
Antérieurs au 1 ^{er} janvier 2002	Postérieurs au 1 ^{er} janvier 2002	Diagnostic
I21.0		Infarctus (transmurale aigu) du myocarde (de la paroi) antérieure
	I21.00	Prise en charge initiale
	I21.000	Prise en charge initiale, infarctus de 24 heures ou moins
	I21.08	Autres prises en charge
I21.1		Infarctus (transmurale aigu) du myocarde (de la paroi) inférieure
	I21.10	Prise en charge initiale
	I21.100	Prise en charge initiale, infarctus de 24 heures ou moins
	I21.18	Autres prises en charge
I21.2		Infarctus (transmurale aigu) du myocarde d'autres localisations
	I21.20	Prise en charge initiale
	I21.200	Prise en charge initiale, infarctus de 24 heures ou moins
	I21.28	Autres prises en charge
I21.3		Infarctus (transmurale aigu) du myocarde de localisation non précisée
	I21.30	Prise en charge initiale
	I21.300	Prise en charge initiale, infarctus de 24 heures ou moins
	I21.38	Autres prises en charge
I21.4		Infarctus sous-endocardique (aigu du myocarde)
	I21.40	Prise en charge initiale
	I21.400	Prise en charge initiale, infarctus de 24 heures ou moins
	I21.48	Autres prises en charge
I21.9		Infarctus (aigu) du myocarde sans précision
	I21.90	Prise en charge initiale
	I21.900	Prise en charge initiale, infarctus de 24 heures ou moins
	I21.98	Autres prises en charge

Codes en I22 de la CIM-10 correspondant aux diagnostics d'infarctus du myocarde à répétition. La sélection des séjours a porté sur les nouveaux codes CIM-10 et sur les codes antérieurs au 1 ^{er} janvier 2002		
Antérieurs au 1 ^{er} janvier 2002	Postérieurs au 1 ^{er} janvier 2002	Diagnostic
I22.0		Infarctus du myocarde à répétition (de la paroi) antérieure
	I22.00	Prise en charge initiale
	I22.000	Prise en charge initiale, infarctus de 24 heures ou moins
	I22.08	Autres prises en charge
I22.1		Infarctus du myocarde à répétition (de la paroi) inférieure
	I22.10	Prise en charge initiale
	I22.100	Prise en charge initiale, infarctus de 24 heures ou moins
	I22.18	Autres prises en charge
I22.8		Infarctus du myocarde à répétition d'autres localisations
	I22.80	Prise en charge initiale
	I22.800	Prise en charge initiale, infarctus de 24 heures ou moins
	I22.88	Autres prises en charge
I22.9		Infarctus du myocarde à répétition de localisation non précisée
	I22.90	Prise en charge initiale
	I22.900	Prise en charge initiale, infarctus de 24 heures ou moins
	I22.98	Autres prises en charge

Codes en I23 de la CIM-10 correspondant aux complications récentes d'un infarctus du myocarde

Codes CIM-10	Diagnostic
I23.0	Hémopéricarde comme complication récente d'un infarctus aigu du myocarde
I23.1	Communication interauriculaire comme complication récente d'un infarctus aigu du myocarde
I23.2	Communication interventriculaire comme complication récente d'un infarctus aigu du myocarde
I23.3	Rupture de la paroi cardiaque sans hémopéricarde comme complication récente d'un infarctus aigu du myocarde
I23.4	Rupture des cordages tendineux comme complication récente d'un infarctus aigu du myocarde
I23.5	Rupture du muscle papillaire comme complication récente d'un infarctus aigu du myocarde
I23.6	Thrombose de l'oreillette, de l'auricule et du ventricule comme complication récente d'un infarctus aigu du myocarde
I23.8	Autres complications récentes d'un infarctus aigu du myocarde

Annexe 3 – Algorithme de classement des RSA

Algorithme de classement des séjours avec DP d'IAM d'un même patient en "IAM" (premier épisode ou récurrence à plus de 28 jours) "transfert" et "ré-hospitalisation à moins de 28 jours" :

- ces séjours étant préalablement ordonnés chronologiquement par le "numéro de séjour", posons:
 - n ($n \geq 1$) le nombre total de séjours avec un DP d'IAM identifiés pour un même patient,
 - v_i le "numéro de séjour" du i^{e} séjour ($i = 1, \dots, n$),
 - $\delta(i, j) = v_j - v_i, i < j$, la différence des numéros de séjours entre le j^{e} séjour et le i^{e} séjour (= nombre de jours entre les 2 séjours),
 - d_i la durée du séjour i .

Étape 1 : le premier séjour de la série est considéré comme étant relié à un IAM et est à cette étape le séjour de référence $s^*=1$.

Puis pour $j = 1, \dots, n$:

- étape j :
 - si $\delta(j-1, j) = d_{j-1} \Rightarrow$ séjour $j =$ transfert ;
 - si $\delta(j-1, j) > d_{j-1}$:
 - si $\delta(s^*, j) \leq 28$ jours \Rightarrow séjour $j =$ ré-hospitalisation à moins de 28 jours,
 - si $\delta(j-1, j) > 28$ jours \Rightarrow séjour $j =$ IAM : récurrence survenue plus de 28 jours après le début d'une première hospitalisation et le j^{e} séjour devient le nouveau séjour de référence : $s^*=j$.

Annexe 4 – Étude exploratoire des faux négatifs

Pour tenter d'appréhender le profil des faux négatifs, une étude qualitative exploratoire a été réalisée dans un des trois registres : les premiers dossiers d'infarctus du myocarde connus des registres, mais non repérés dans la base nationale du PMSI de deux établissements ont été examinés, soit 31 RSS. Ces RSS étaient constitués de tous les résumés d'Unités médicales correspondant à l'hospitalisation au cours de laquelle un infarctus du myocarde avait été identifié par le registre.

La proportion de séjours multi-unités était élevée dans cet échantillon exploratoire ; elle atteignait 74 %, contre 43,8 % pour l'ensemble des RSA extraits du PMSI.

Ces séjours n'apparaissant pas avec un diagnostic principal d'infarctus du myocarde dans la base nationale des RSA correspondaient pour l'essentiel aux trois situations suivantes :

- des séjours durant lesquels des complications cardiaques étaient survenues (insuffisance cardiaque, choc, arrêt cardiaque...) (n=13) ;

- des séjours pour lesquels les résumés d'unité médicale comportaient un diagnostic principal d'angor instable, voire de cardiopathie ischémique chronique, mais pas d'infarctus du myocarde. Cette situation était presque aussi fréquente que la précédente (n=11) ;

- moins souvent (n=4), il s'agissait d'infarctus du myocarde survenant au cours d'hospitalisations motivées par des pathologies non coronaires ayant nécessité des actes chirurgicaux "classants" ou des séjours prolongés multi-unités : selon l'algorithme de détermination des diagnostics principaux des séjours multi-unités, ce sont ces pathologies non coronaires consommatrices de soins qui devaient apparaître en diagnostic principal lors de la phase de production du RSA (sélection automatique par algorithme d'un seul diagnostic principal en cas de séjours multi-unités).

Étalonnage du PMSI MCO pour la surveillance des infarctus du myocarde Année 2003

L'objectif de cette étude était d'étalonner la base nationale des hospitalisations de court séjour (PMSI MCO) pour la surveillance des infarctus du myocarde hospitalisés. Pour ce faire, le PMSI MCO a été comparé aux données recueillies par les trois registres MONICA, pour les tranches d'âge et les territoires correspondants.

Ont été inclus tous les cas d'infarctus du myocarde hospitalisés survenus en 2003, chez des personnes âgées de 35 à 74 ans et domiciliées dans un territoire couvert par un registre MONICA (communauté urbaine de Lille, départements de la Haute-Garonne et du Bas-Rhin).

Dans un premier temps, les résumés de sortie anonyme (RSA) du PMSI MCO 2003 comportant un diagnostic principal d'infarctus du myocarde et respectant ces critères d'âge et de domiciliation, ainsi que les séjours de la base 2004 susceptibles d'avoir débuté fin 2003, ont été sélectionnés. Dans un deuxième temps, ces RSA ont été envoyés aux départements d'information médicale des hôpitaux ou cliniques concernés pour complément d'information sur l'identité et la date d'hospitalisation des patients, à l'exception des RSA provenant d'établissements ayant eu moins de 6 séjours pour infarctus du myocarde. Ces fichiers complétés ont ensuite été envoyés aux registres pour comparaison avec leurs propres recueils, avant retour à l'InVS pour vérifications et analyses. Lorsqu'il y avait plusieurs RSA par événement, un seul épisode a été comptabilisé par période de 28 jours, comme c'est le cas pour les registres MONICA. Les échanges de données nominatives ont été cryptés.

Parmi les infarctus du myocarde hospitalisés expertisés par les registres en 2003, 76,2% ont été repérés dans le PMSI 2003 : la sensibilité était un peu plus élevée entre 35 et 64 ans (80,3%), mais diminuait entre 65 et 74 ans (69,2%, $p < 10^{-4}$).

Parmi les infarctus du myocarde repérés dans le PMSI, 78,8% étaient enregistrés dans la base 2003 des registres (valeur prédictive positive) et 21,2% ont été expertisés comme des faux positifs.

Le nombre d'infarctus du myocarde estimés par le PMSI sur la base d'un seul épisode par période de 28 jours était légèrement inférieur au nombre d'IAM collectés par les registres MONICA (-3,3%). Ce résultat variait avec l'âge, la sous-estimation étant plus importante entre 65 et 74 ans (-8,9%). À l'inverse, la prise en compte de tous les RSA avec diagnostic principal d'infarctus du myocarde, y compris les transferts ou les hospitalisations de moins de 24 heures, entraînait une surévaluation du nombre d'infarctus du myocarde survenus entre 35-74 ans (+12,7%).

La comparaison du PMSI MCO 2003 aux données des trois registres MONICA a permis d'estimer la sensibilité du PMSI à 76,2% et sa valeur prédictive positive à 78,8% pour les classes d'âge comprises entre 35 et 74 ans. À ces âges, le nombre d'infarctus du myocarde calculé sur la base d'un épisode par période de 28 jours sous-estimait légèrement le nombre total d'infarctus du myocarde (-3,3%), alors que la prise en compte de tous les RSA comportant un diagnostic principal d'infarctus du myocarde surestimait ce nombre de 12,7%.

Mots clés : infarctus du myocarde, étalonnage, PMSI MCO, registres, court-séjour, surveillance épidémiologique, France

The accuracy of hospital records for acute myocardial infarction in 2003

The objective of this study was to compare estimates of myocardial infarction (MI) rates obtained from the French national hospital discharge records database (known as "PMSI") to data collected by the three French Monica registries.

Records with a main (most responsible) diagnosis of MI and relative to patients aged 35 to 74 years living in one of the three registers areas were selected from the 2003 national anonymous database. Records were sent to each corresponding hospital medical information department for completing names and hospitalization dates, except for hospitals having had 5 or less MI records. When there were several hospitalizations, only one event was counted per 28 days as in Monica registers. Data were encrypted and sent to registries for comparison with their own data, and for adding false negatives; then completed data were returned to InVS for analysis.

The sensitivity of PMSI main diagnosis was no greater than 76.2%, but was slightly higher in the 35-64 year-olds (80.3%) than in the 65-74 year-olds (69.2%). Specificity only attained 78.8%. Overall, the estimate of MI number calculated from PMSI was 3.3% lower than the estimate from registries (even 8.9% lower between 65 and 74 years).

In 2003, PMSI sensitivity and specificity were only 76.2% and 78.8%, respectively, in the 35-74 years age group. Besides, PMSI underestimated the total number of MI.

Citation suggérée :

Étalonnage du PMSI MCO pour la surveillance des infarctus du myocarde – Année 2003. Saint-Maurice (Fra) : Institut de veille sanitaire, juillet 2010, 24 p. Disponible sur : www.invs.sante.fr

INSTITUT DE VEILLE SANITAIRE

12 rue du Val d'Osne

94 415 Saint-Maurice Cedex France

Tél. : 33 (0)1 41 79 67 00

Fax : 33 (0)1 41 79 67 67

www.invs.sante.fr

ISSN : 1 956-6964

ISBN : 9 78-2-11-099194-2

ISBN-NET : 9 78-2-11-099266-6

Tirage : 120 exemplaires

Réalisé par Diadeis-Paris

Dépôt légal : juillet 2010