



LE RÉSEAU REDSIAM

(RÉSEAU POUR MIEUX UTILISER LES DONNÉES DU SYSTÈME NATIONAL DES DONNÉES DE SANTÉ)

Avril 2020

SOMMAIRE

1	Le contexte	3
1.1	Les bases de données de santé nationales et leur utilisation pour la recherche, la surveillance, l'évaluation et la santé publique	3
1.1.1	Le SNDS	3
1.1.2	Principales modalités d'utilisation	3
1.2	Les principales difficultés pour l'utilisation des bases de données nationales	4
1.3	L'utilisation du SNDS	4
1.4	Références	5
2	La création du réseau ReDSiam (RÉSEAU POUR MIEUX UTILISER LES DONNÉES DU SYSTÈME NATIONAL DES DONNÉES DE SANTÉ)	7
2.1	Un besoin, des compétences	7
2.2	... mais un manque de coordination et de synergie	7
2.3	Une initiative volontariste, une enquête fructueuse	7
2.4	Objectifs, composition et activités du réseau ReDSiam	8
2.4.1	Objectif général	8
2.4.2	Composition	8
2.4.3	Activités du réseau ReDSiam	8
2.5	Organisation et fonctionnement	8
2.5.1	Coordination et gouvernance	8
2.5.2	Financement	9
2.5.3	Organisation fonctionnelle	9
2.5.4	Accès aux algorithmes	10

1 LE CONTEXTE

La France dispose de bases de données médico-sociales nationales centralisées, constituées et gérées par des organismes publics, couvrant de façon exhaustive et permanente l'ensemble de la population dans divers domaines stratégiques pour la santé publique et la recherche : recours aux soins, hospitalisation, handicaps, prestations et situation professionnelle, sociale et économique. De plus, un identifiant individuel unique (le NIR : numéro d'identification au répertoire) est actuellement utilisé par pratiquement toutes les bases de données nationales. Malgré certaines limites en termes de couverture, de qualité et de validité des données, ces bases de données, concernant plus de 60 millions de personnes, constituent un patrimoine considérable, vraisemblablement sans équivalent au monde.

1.1 LES BASES DE DONNÉES DE SANTÉ NATIONALES ET LEUR UTILISATION POUR LA RECHERCHE, LA SURVEILLANCE, L'ÉVALUATION ET LA SANTÉ PUBLIQUE

1.1.1 *Le SNDS*

Les principales sources de données de santé couvrant la population française sont les données d'hospitalisation issues du PMSI (Programme de médicalisation du système d'information), les données de consommation de soins et de prise en charge de l'Assurance maladie (SNIIRAM : Système national d'information inter-régimes de l'Assurance maladie) et la mortalité et les causes de décès.

Depuis la loi de modernisation du système de santé du 26 janvier 2016, elles sont réunies au sein du Système national des données de santé (SNDS), afin de permettre des études, recherches ou évaluation présentant un caractère d'intérêt public et contribuant à l'une des finalités suivantes : l'information sur la santé ; la mise en œuvre des politiques de santé ; la connaissance des dépenses de santé ; l'information des professionnels et des établissements sur leurs activités ; l'innovation dans les domaines de la santé et de la prise en charge médico-sociale ; la surveillance, à la veille et à la sécurité sanitaire (www.snds.gouv.fr).

Depuis la création du SNIIRAM et la constitution en 2003 d'un entrepôt, limité aux remboursements du régime général, le périmètre s'est considérablement enrichi et l'accès à ses données a été facilité avec la mise en place, en 2005, d'un échantillon au 1/97^{ème} dédié aux institutions publiques, aux agences et au monde de la recherche, l'EGB (Échantillon généraliste des bénéficiaires) qui peut être utilisé à distance.

L'enrichissement du SNIIRAM, aujourd'hui SNDS, s'est traduit par la mise à disposition des utilisateurs :

- du chaînage en routine des données ambulatoires et du PMSI (en 2007),
- de l'exhaustivité de l'activité médicale (actes techniques) codée (en 2007),
- de l'activité externe des hôpitaux publics (en 2009),
- des dates de décès provenant de l'Insee (en 2009),
- des médicaments et dispositifs médicaux facturés en sus des forfaits d'hospitalisation (en 2009),
- des données provenant de l'ensemble des régimes d'assurance maladie : CNAMTS, MSA, RSI et les seize autres régimes spéciaux (y compris les sections locales mutualistes-SLM) (en 2009),
- des dates exactes d'hospitalisation (en 2010),
- les causes médicales de décès devraient être prochainement intégrées au SNDS.

1.1.2 *Principales modalités d'utilisation*

Pour l'essentiel, on rencontre deux modes d'utilisation des données du SNDS.

- utilisation de données du SNDS concernant des personnes sélectionnées uniquement sur leurs caractéristiques : Il s'agit de situations où le demandeur veut uniquement extraire des données pour des personnes sélectionnées selon des critères spécifiques correspondant à des variables

enregistrées dans le SNDS (âge, sexe, période, consommation de certains médicaments, etc.), sans croisement avec d'autres sources individuelles de données.

- utilisation de données du SNDS concernant des personnes préalablement identifiées : typiquement, il s'agit de situations où le demandeur veut enrichir des données individuelles recueillies auprès des sujets d'une étude (cohorte, cas-témoins ou transversale), de professionnels de santé ou d'autres entités, par des données de consommation de soins, d'hospitalisation et de causes de décès.

1.2 LES PRINCIPALES DIFFICULTÉS POUR L'UTILISATION DES BASES DE DONNÉES NATIONALES

L'utilisation à des fins de recherche, d'évaluation et de surveillance de ces bases de données nationales se heurte actuellement à des obstacles divers. Deux rapports du HCSP en ont fait l'analyse et ont émis diverses recommandations, concernant notamment les problèmes juridiques et organisationnels dont la résolution relève essentiellement des pouvoirs publics et des organismes producteurs de données [1]. Ultérieurement, un rapport de la Drees a complété cette analyse et proposé des améliorations d'ordre juridique, technique et institutionnel pour faciliter l'accès des administrations et des chercheurs aux bases de données administratives dans le domaine de la santé [2]. La loi de modernisation du système de santé du 26 janvier 2016 et Décret n° 2016-1871 du 26 décembre 2016 ont fixé les règles d'accès au SNDS.

Cependant, il subsiste d'importantes difficultés techniques pour une utilisation plus facile qu'actuellement des données issues du SNDS. En effet, la complexité de ses bases de données rend l'utilisation des données individuelles particulièrement difficile et nécessite des ressources scientifiques, techniques et organisationnelles complexes et de haut niveau de compétence. Disposer de l'ensemble de ces ressources dépasse largement les moyens actuellement disponibles au sein de la plupart des équipes françaises, quel que soit leur organisme d'appartenance, particulièrement quand il s'agit d'une étude ponctuelle de durée limitée.

Du point de vue de l'utilisation des données du SNDS pour la recherche, la surveillance l'évaluation et la santé publique, outre ces problèmes techniques liés à la complexité des bases de données, une importante difficulté est l'identification et la validation des maladies. En effet, dans une optique de recherche épidémiologique, de surveillance, d'économie de la santé et de santé publique en général, une de raisons majeures de l'utilisation du SNDS est l'analyse de données concernant des pathologies et des maladies spécifiques. Or, la principale limite des bases de données du PMSI de l'Assurance maladie et des causes de décès est que la validité des données de santé est de qualité variable. L'utilisation du PMSI comme source d'information sur les pathologies s'avère délicate et les diagnostics ne sont pas toujours fiables [3-8]. Les ALD ont des limites connues [8] et la qualité du codage des maladies est encore mal connue. La base de données de remboursements de l'assurance maladie est adaptée aux objectifs d'analyse des pratiques de prescription [10], d'évaluation de l'impact de campagne d'information [11] ; par contre, elle ne comporte pas d'information directe sur la nature des maladies traitées.

Au total, les bases de données du SNDS fournissent des données individuelles médicalisées, structurées et codées de manière standardisée [12], mais leur utilisation, notamment dans une optique épidémiologique, nécessite cependant un important travail méthodologique de validation [13,14].

1.3 L'UTILISATION DU SNDS

Malgré les difficultés citées, une importante activité s'est développée depuis plusieurs années pour l'analyse des données constituant le SNDS, qu'il s'agisse des données de l'Assurance maladie ou de celles du PMSI (le plus souvent analysées directement à partir des données des DIM ou de la base de l'ATIH). Les données du SNDS avaient déjà donné lieu à plus de 330 publications entre 2007 et 2016 sur différentes thématiques : prévalence et incidence de pathologies, parcours de soins de patients, état de santé et

recours aux soins de populations spécifiques, usage des médicaments en vie réelle, mesure d'effet indésirables de médicament ou d'autres procédures de soins, suivi des dépenses de l'Assurance maladie... [15].

Les constats essentiels qu'on peut tirer de ces travaux sont les suivants : (1) aucune base ne peut être isolément considérée comme enregistrant exhaustivement les événements d'intérêt ; (2) il existe des problèmes pour identifier les sujets et les épisodes leur correspondant ; (3) ces bases ne contiennent pas des données médicales « validées » selon les critères épidémiologiques usuels.

Cependant, la plupart des travaux n'ont souvent concerné isolément qu'une seule base (PMSI ou Assurance maladie), ce qui présente des limites diverses. Ainsi, des études concernant les cancers menés avec la collaboration des registres du cancer montrent que, utilisés isolément, ni le PMSI ni les ALD ne permettent pas d'avoir une bonne estimation de l'incidence et de la prévalence, mais que leur utilisation conjointe peut être efficace [16] ; l'analyse des effets cardiovasculaires du benfluorex est également un exemple de l'intérêt de l'utilisation conjointe des données de l'Assurance maladie et du PMSI [17].

De plus, seul un petit nombre de maladies ont été étudiées en fonction des objectifs des investigateurs.

1.4 RÉFÉRENCES

- 1- Les systèmes d'information pour la santé publique. Haut conseil de la santé publique. La Documentation française, novembre 2009. Mieux utiliser les bases de données nationales pour la santé publique et la recherche. Rapport du Haut conseil de la santé publique, mars 2012.
- 2- Loth A, Caillol H, Chion E, Dupont J, Fontaine A, Goldberg M, Legal R, Pothier F. Mission pour faciliter l'accès des administrations et des chercheurs aux bases de données administratives dans le domaine de la santé. Direction de la recherche des études, de l'évaluation et des statistiques-Drees, Mai 2013.
- 3- Couris CM, Forêt Dodelin C, Rabilloud M, Colin C, Bobin JY, Dargent D, Raudran D, Schott AM. Sensibilité et spécificité de deux méthodes d'identification des cancers du sein incidents dans les services spécialisés à partir des données médico-administratives. *Rev Epidemiol Sante Publique* 2004, 52, 151-60.
- 4- Couris CM et al. Method of correction to assess the number of hospitalized incident breast cancer cases based on claims databases. *Journal of Clinical Epidemiology*, 2002, 55 : 386-391.
- 5- Coureau G, Baldi I, Saves M, Jaffre A, Barat C, Gruber A, Loiseau H, Kostrzewa A, Servas V, Salamon R, Gilleron V. Évaluation des performances du PMSI pour l'identification des tumeurs incidentes du système nerveux central par rapport à un registre spécialisé en Gironde, France, en 2004. *Rev Epidemiol Sante Publique*. 2012;60:295-304.
- 6- Chantry AA, Deneux-Tharoux C, Bal G, Zeitlin J, Quantin C, Bouvier-Colle MH, pour le groupe GRACE. Le programme de médicalisation du système d'information (PMSI) – Processus de production des données, validité et sources d'erreurs dans le domaine de la morbidité maternelle sévère. *Rev Epidem Santé Pub* 2012 ; 60 : 177-188.
- 7- Chantry AA, Deneux-Tharoux C, Cans C, Ego A, Quantin C, Bouvier-Colle MH. GRACE study group. Hospital discharge data can be used for monitoring procedures and intensive care related to severe maternal morbidity. *J. Clin Epidemiol* 2011, sept; 64 (9) : 1014-22.
- 8- Aboa-Eboule C, Mengue D, Benzenien E, Hommel M, Giroud M, Bejot Y, Quantin C. How accurate is the reporting of stroke in hospital discharge data? A pilot validation study using a population-based stroke registry as control. *Accepté pour publication dans Journal of Neurology*.
- 9- Païta M, Weill A. Les personnes en Affection de longue durée au 31 décembre 2008. *Points de repère* 2009, 27:1-12.
- 10- Deprez Ph-H, Chinaud F, Clech S, Germanaud J, Weill A, Cornille JL, Fender P. La population traitée par médicaments de la classe des antihistaminiques en France métropolitaine : données du régime général de l'assurance maladie, 2000. *Revue médicale de l'assurance maladie* Avril-juin 2004, 35 (1), 3-11.
- 11- Lecadet J, Vialaret K, Vidal P, Baris B, Fender P. Mesure à l'échelle d'une région des effets d'un programme national d'information sur le bon usage des antibiotiques. *Revue médicale de l'assurance maladie* Avril-Juin 2004, 35 (2) ,81-91.

- 12- Fender P, Weill A. Epidémiologie, santé publique et bases de données médico-tarifaire. (Éditorial) Rev Epidemiol Santé Publique, 2004, 52,113-117.
- 13- Bernier MO, Mezzarobba M, Maupu E, Caer-Lorho S, Brisse HJ, Laurier D, Brunelle F, Chatellier G. Utilisation des données du Programme de Médicalisation des Systems d'Information (PMSI) dans les études épidémiologiques : application à la Cohorte Enfant Scanner. RESP, oct 2012, n°5, 60, 363-71.
- 14- Quantin C, Velten M, Huet F, Benzenine E, Dupoux A, Hamadi A, Farrington CP, Tubert-Bitter P. Paediatric vaccine pharmacoepidemiology: a validation study of febrile convulsions identification from French automated hospital databases. International Society for Clinical Biostatistics. Conference, Mini symposium on Statistics in vaccine research, 19-23 August 2012.
- 15- Tuppin P, Rudant J, Constantinou P, et al. Value of a national administrative database to guide public decisions: From the système national d'information interrégimes de l'Assurance Maladie (SNIIRAM) to the système national des données de santé (SNDS) in France. Rev Epidemiol Sante Publique. 2017;65 Suppl 4:S149–S167.
- 16- Grosclaude et al et Lauzeille et al, BEH-Numéro thématique - Surveillance épidémiologique des cancers en France : utilisation des bases de données médico-administratives. 31 janvier 2012/n° 5-6.
- 17- Weill A, Paita M, Tuppin P, Fagot JP, Neumann A, Simon D, Ricordeau P, Montastruc JL, Allemand H. Benfluorex and valvular heart disease: a cohort study of a million people with diabetes mellitus. Pharmacoepidemiology and Drug Safety 2010; 19: 1256–1262.

2 LA CRÉATION DU RÉSEAU REDSIAM (RÉSEAU POUR MIEUX UTILISER LES DONNÉES DU SYSTÈME NATIONAL DES DONNÉES DE SANTÉ)

2.1 UN BESOIN, DES COMPÉTENCES...

La politique d'ouverture de la CNAM, gestionnaire du SNDS, et l'amélioration régulière des possibilités d'accès aux bases de données génèrent des besoins grandissants dans les milieux de la recherche épidémiologique, de l'économie de la santé, de l'évaluation en santé, de la surveillance et de la santé publique, et il existe maintenant de nombreuses compétences appuyées sur des expériences variées couvrant des domaines diversifiés.

2.2 ... MAIS UN MANQUE DE COORDINATION ET DE SYNERGIE

Les spécialistes compétents sont des personnes formées et disposant d'une importante expérience dans un domaine qui demande une longue appropriation et une continuité du travail sur des bases de données particulièrement complexes et évolutives.

Cependant, les équipes qui travaillent dans ce domaine, appartenant à des organismes publics, parapublics ou privés (CNAM, InVS, Inserm, Irdes, DIM, CHU, CRO...) sont le plus souvent isolées et ne travaillent habituellement pas ensemble. De ce fait, les compétences disponibles sont largement éparpillées dans des équipes au statut et à l'appartenance institutionnelle divers, sans coordination organisée. Cette situation est l'origine de travaux qui ne génèrent pas de cumul des connaissances, des expériences et des savoir-faire, et qui sont parfois redondants.

2.3 UNE INITIATIVE VOLONTARISTE, UNE ENQUÊTE FRUCTUEUSE

Le constat esquissé ci-dessus est aujourd'hui largement partagé. C'est dans ce contexte que la décision de mettre en place un réseau consacré à une meilleure coordination des travaux sur l'utilisation de ce qui était alors le SNIIRAM (données de l'Assurance maladie et du PMSI) a été prise lors d'une réunion qui s'est tenue le 29 juin 2012 à laquelle participaient des personnes de la CNAM, de Santé publique France et d'enseignants-chercheurs.

Composition du groupe de travail initial

Jean Claude Desenclos, Anne Doussin, Anne Fagot-Campagna, Marcel Goldberg, Isabelle Grémy, Javier Nicolau, Catherine Quantin, Philippe Ricordeau, Alain Weill

La première étape a consisté en une enquête auprès de la communauté des épidémiologistes et des spécialistes de l'information médicale. Un document de présentation du réseau et une fiche d'information à compléter pour chaque utilisation des bases médico-administratives ont été envoyés en février 2013 aux personnes figurant sur les listes de diffusions de l'Adelf, d'Epiter, et des participants aux congrès EMOIS. Plus de 50 réponses comprenant plus de 100 fiches ont été retournées ; elles provenaient d'équipes hospitalières, d'organismes de recherche, d'agences de santé, et de consultants privés. Les bases de données utilisées couvraient l'ensemble du dispositif existant à l'époque : SNIIRAM (DCIR, EGB), PMSI (hospitalier, régional, national), Erasme régional, bases de la MSA et du RSI. Les affections étudiées concernaient pratiquement tout le spectre de la pathologie : tumeurs, endocrinologie, nutrition et métabolisme, maternité, périnatalité, infectieux, appareil respiratoire, maladies du système nerveux, traumatologie, troubles mentaux et du comportement, ostéo-articulaire, muscles et tissus conjonctif, appareil digestif, appareil circulatoire ; des travaux concernaient aussi la pharmacovigilance et la pharmaco-épidémiologie, l'analyse des parcours de soins et du système de soins. Les résultats de cette enquête ont été présentés le 21 mars 2013 lors du congrès EMOIS de Nancy, et la création du réseau ReDSiam a été annoncée à cette occasion.

2.4 OBJECTIFS, COMPOSITION ET ACTIVITÉS DU RÉSEAU REDSIAM

2.4.1 Objectif général

L'objectif général du réseau ReDSiam est de favoriser par le partage d'expérience et l'interaction entre équipes le développement, la validation et la promotion des méthodes d'analyse des données médico-administratives à des fins de recherche, de surveillance, d'évaluation et de santé publique.

2.4.2 Composition

Le réseau est constitué sur la base du volontariat par des personnes appartenant essentiellement à des organismes publics et parapublics (et plus récemment d'organismes privés à but lucratif), de disciplines diverses (épidémiologie, économie de la santé, évaluation, professionnels de l'information médicale, statisticiens, informaticiens), travaillant dans le domaine de l'utilisation des données médico-administratives à des fins de recherche, de surveillance et de santé publique.

2.4.3 Activités du réseau ReDSiam

Les activités du réseau concernent le partage d'expérience, la mutualisation et la capitalisation des connaissances dans le champ de l'identification des maladies :

- Discussion critique *via* une grille d'analyse et mise à disposition d'algorithmes spécifiques de maladies basés sur des combinaisons de diagnostics du PMSI et/ou d'ALD, de données sur les actes médicotechniques, les consommations de médicaments ou d'autres produits de santé plus ou moins spécifiques de la maladie concernée, les consultations de professionnels, causes de décès, etc.
- Le cas échéant, développement de nouveaux algorithmes, mise à jour régulière d'algorithmes en fonction de l'évolution des techniques médicales, de l'introduction de nouveaux médicaments et des modifications des nomenclatures d'actes ;
- Validation des algorithmes par des méthodes diverses : retour au médecin traitant, confrontation avec des dossiers médicaux hospitaliers, des questionnaires remplis par les sujets, croisement avec d'autres sources, telles que registres de maladie, cohortes...

Différentes activités spécifiques accompagnent cette activité de fond :

- Mise en place et maintenance d'une bibliothèque centralisée d'algorithmes
- Manifestations scientifiques régulières du réseau ReDSiam (séminaires techniques, symposium associé à EMOIS, ...)
- Mise en place et maintenance d'un site Internet (incluant un forum d'échanges)

D'autres activités seront à développer mais nécessiteront des ressources dédiées non disponibles pour le moment.

2.5 ORGANISATION ET FONCTIONNEMENT

Le réseau a été mis en place et fonctionne de façon « informelle » selon une Charte. La participation est à titre individuel et implique une Déclaration d'engagement et d'intérêts. ReDSiam n'a pas d'existence juridique, ni de rattachement à un organisme.

2.5.1 Coordination et gouvernance

- Comité directeur (CD) du réseau
 - Rôle : le CD a comme rôle de définir la stratégie générale de ReDSiam, le programme de travail, de prendre les décisions organisationnelles, de rechercher des financements pour le fonctionnement du réseau, de valider la Charte du réseau qui précise les modalités de participation, de mutualisation, les règles de publication, les modalités de fonctionnement des groupes de travail thématiques, etc.

ReDSiam – Mise en place

- Composition : le groupe de travail initial a été élargi et le CD est actuellement composé de représentants des organismes suivants :
 - Organismes producteurs des données du SNDS : Caisse nationale d'assurance maladie (CNAM), Mutualité sociale agricole (MSA), Agence technique de l'information sur l'hospitalisation (ATIH).
 - Organismes utilisateurs des données du SNDS : Agence nationale de sécurité du médicament et des produits de santé (ANSM), Caisse nationale militaire de l'assurance maladie (CNMSS), Direction de la recherche, des études, de l'évaluation et des statistiques (Drees), Fédération nationale des observatoires de santé (FNORS), Haute autorité de santé (HAS), Institut national du cancer (INCA), Plateforme des données de santé (INDS), Institut national de la santé et de la recherche médicale (Inserm), Institut de recherche et documentation en économie de la santé (Irdes), Santé Publique France, Epi-Phare.
- Un Bureau a été constitué et gère le fonctionnement du réseau.
- Portage opérationnel : le secrétariat, l'hébergement du site web, la gestion sont assurés par l'Unité Cohortes épidémiologiques en population (UMS 011 UVSQ-Inserm – dirigée par le Pr. Marie Zins).

2.5.2 *Financement*

Les organismes participant au réseau assurent une subvention annuelle permettant de couvrir les dépenses liées au fonctionnement (secrétariat, site Internet, documentation, réunions...).

2.5.3 *Organisation fonctionnelle*

- Groupes de travail : les activités de ReDSiam reposent essentiellement sur le fonctionnement de groupes de travail animés par un ou deux coordonnateurs. En 2020, le réseau compte plus de 250 membres, participant aux groupes de travail suivants, animés par 20 coordonnateurs :
 - Appareil cardiovasculaire
 - Appareil locomoteur
 - Appareil respiratoire
 - Endocrinologie, nutrition et métabolisme
 - Handicap
 - Infectieux
 - *Maladie rénale (en cours)*
 - Maladies neurodégénératives
 - Maternité, périnatalité
 - Mesure de la qualité des soins
 - Méthodologie
 - Polypathologie-Multimorbidité
 - Troubles mentaux et du comportement
 - Tumeurs

Les groupes de travail évoluent avec le temps ; la liste des groupes de travail tenue à jour est consultable sur le site de ReDSiam : [Page de présentation ReDSiam](#)

- [Site Internet](#) : information sur les activités de ReDSiam, espace coopératif de travail, forum d'échanges ; les règles de fonctionnement du site seront fixées dans la Charte.
- Un Comité technique associant les responsables des Groupes de travail et le Bureau organise le travail et assure le suivi régulier des activités de ReDSiam.

L'organisation du travail de ReDSiam est précisée dans la Charte qui fixe les règles de fonctionnement du réseau.

2.5.4 *Accès aux algorithmes*

L'accès aux algorithmes de la bibliothèque ReDSiam se fait sur demande motivée auprès du Bureau en complétant un formulaire ci-joint

L'accès est gratuit. Lorsque le demandeur appartient à un organisme public pour un usage sans lien avec un organisme privé à but lucratif, l'accès est accordé automatiquement. Lorsque le demandeur appartient à un organisme privé à but lucratif, la demande est examinée par le Bureau qui peut autoriser l'accès. En cas d'avis négatif du Bureau, la demande est transmise pour décision au Comité directeur.