

## SURVEILLANCE DES MALADIES CHRONIQUES EN FRANCE : LA CONTRIBUTION DES BASES DE DONNÉES MÉDICO-ADMINISTRATIVES

// SURVEILLANCE OF CHRONIC DISEASES IN FRANCE: THE CONTRIBUTION OF HEALTH ADMINISTRATIVE DATABASES

Isabelle Grémy (i.gremy@invs.sante.fr), Anne Doussin

Institut de veille sanitaire, Saint-Maurice, France

Soumis le 26.09.2013 // Date of submission: 09.26.2013

### Résumé // Abstract

La surveillance des maladies chroniques en France s'appuie sur différentes sources de données : enquêtes en population générale, registres, causes médicales de décès et bases de données médico-administratives (BDMA). Ces dernières, exhaustives, réactives, bien qu'initialement non conçues pour l'épidémiologie, constituent désormais une source supplémentaire de données.

L'objectif est de décrire les BDMA, de montrer leurs forces, limites et complémentarités par rapport aux autres sources de données, à travers deux exemples : la surveillance du diabète et des cancers.

Les BDMA couvrent la totalité des soins de ville remboursés et des séjours hospitaliers pour la population résidant en France. Elles permettent de repérer pathologies ou processus morbides à travers leurs prises en charge, d'évaluer des actions de prévention ou de dépistage, etc. L'exhaustivité de ces bases autorise des déclinaisons à des niveaux géographiques fins et leur réactivité permet de renseigner les années les plus récentes. Les BDMA peuvent servir de contrôle des biais de mémoire lorsqu'elles sont croisées avec des enquêtes individuelles par questionnaire, ou servir de bases de redressement en comparant répondants et non-répondants (comme dans l'exemple du diabète). Mais, étant basées sur les soins pris en charge, elles évoluent avec les traitements, les politiques de remboursement, etc. Leur performance en tant qu'outil de surveillance épidémiologique doit être régulièrement réinterrogée. Plusieurs méthodes d'évaluation sont possibles en fonction des objectifs et des autres sources de données existantes.

Les BDMA peuvent contribuer à la surveillance des maladies chroniques, ce d'autant plus que leur qualité, tant de codage que de contenu, ne cesse de progresser, que les possibilités de croisement avec d'autres sources de données augmentent et que les conditions de leur accessibilité se clarifient.

*Surveillance of chronic diseases in France is supported by various data sources: surveys in general population, registers, medical causes of death, and health administrative databases (HADB). Such comprehensive and reactive databases, although they were not initially developed for epidemiological purposes, offer an additional data source.*

*The objective is to describe the HADBs, to show their strength, limitations and complementarity compared to other data sources, through two examples: surveillance of diabetes and cancers.*

*HADBs cover all reimbursed health-care in private practice and hospital admissions for population living in France. They enable to identify pathologies or disease processes through their management, and to assess prevention or screening actions, etc. The comprehensiveness of these databases allows their use at fine geographical levels, and their reactivity provides information on the most recent years. HADBs can help control recall bias when crossed with individual questionnaire surveys, or be used for adjustment when comparing respondents and non-respondents (as in the example of diabetes). However, as they are based upon reimbursed health-care, they change with treatments, reimbursement policies, etc. Their effectiveness as a tool for epidemiological surveillance needs to be regularly questioned. Several assessment methods are possible, according to the objectives and the other existing data sources.*

*HADBs can help the surveillance of chronic diseases, all the more as their quality, in terms of encoding and of content is in constant progress. The possibility of crossing these data with other data sources is increasing, and the conditions of their access are clearer.*

---

**Mots-clés :** Maladies chroniques, Surveillance épidémiologique, Bases de données médico-administratives, France  
*// Keywords:* Chronic diseases, Monitoring, Health administrative databases, France

---

## Introduction

La surveillance épidémiologique des maladies chroniques, mission confiée en France à l'Institut de veille sanitaire (InVS), est indispensable pour suivre, et si possible prévoir, les évolutions temporo-spatiales de ces maladies, de leurs déterminants et de leurs complications. Cette surveillance a notamment pour objectifs de décrire ces évolutions, de déceler d'éventuelles inflexions inattendues ainsi que de mesurer l'impact des traitements, des programmes de prévention et de dépistage et des évolutions du contexte social et économique, dans une perspective d'aide à la décision pour le pilotage des politiques de santé publique.

Les bases de données administratives (BDMA) à finalité médico-économique sont désormais mobilisables pour la surveillance en santé publique. Cette évolution significative des sources de données disponibles conduit à repenser la stratégie de surveillance des maladies chroniques. Moins centré sur les méthodes employées, qui sont détaillées dans d'autres articles du même BEH, cet article a surtout pour objectif de présenter l'éventail des possibilités qu'offrent les BDMA, leurs limites, leurs enjeux et leurs perspectives pour la surveillance des maladies chroniques en France.

## Le dispositif de surveillance des maladies chroniques

La surveillance des maladies chroniques a évolué en France de façon considérable au cours des deux dernières décennies. Jusqu'à la fin des années 1990, elle reposait essentiellement sur des enquêtes déclaratives en population générale, les registres et les données de mortalité.

Depuis un peu plus de 10 ans, les BDMA complètent ce dispositif et prennent une place croissante. Ces bases, conçues pour le suivi des dépenses de santé, se révèlent précieuses pour des utilisations de surveillance épidémiologique. En effet, elles sont centralisées, permanentes et réactives, et recueillent des données de recours aux soins sur l'ensemble de la population prise en charge en France.

Le Programme de médicalisation des systèmes d'information (PMSI) hospitalier a été mis en place progressivement à partir des années 1990. Il enregistre actuellement l'ensemble des hospitalisations du secteur public et privé (« Médecine, chirurgie, obstétrique », soins de suite et de réadaptation, etc.). Le PMSI présente des limites du fait de sa finalité médico-économique mais il s'améliore de manière continue, par exemple grâce à une meilleure qualité de codage des diagnostics, la possibilité d'un chaînage des séjours d'un même patient grâce à un numéro d'anonymat irréversible unique depuis 2001, et une définition du diagnostic principal basée maintenant sur le diagnostic ayant motivé l'hospitalisation, déterminé à la sortie du patient, depuis 2009.

Le système national d'informations inter-régimes de l'assurance maladie (Sniiram) a été créé par la loi en décembre 1998. Cet entrepôt de données individuelles est alimenté en continu par les données servant au remboursement des soins de ville. Il a concerné progressivement tous les régimes d'assurance maladie et couvre actuellement la quasi-totalité de la population vivant en France (plus de 65 millions de personnes). Sont renseignées les prestations présentées au remboursement (date, actes, médicaments, dispositifs médicaux...), des informations sur les prescripteurs et les professionnels de santé ayant délivré les soins, des informations sur le patient concerné, dont la notion de bénéficiaire de la Couverture maladie universelle complémentaire (CMUc), d'affection de longue durée (ALD), de maladie professionnelle et d'invalidité, pour lesquelles les pathologies sont codées selon la Classification internationale des maladies (CIM-10). Toutefois, le motif des actes et les résultats d'examens ne sont pas renseignés. La durée de conservation des données a augmenté. Elle est aujourd'hui de trois ans plus l'année en cours et de 10 ans en « archives ».

Le PMSI est « inclus » dans le Sniiram et, grâce à l'utilisation du même numéro d'anonymat irréversible, soins de ville et hospitalisations d'un même patient peuvent être chaînés.

Depuis leur création, Sniiram et PMSI sont donc en perpétuelle évolution ; leur couverture tend vers l'exhaustivité, en termes de population mais aussi de soins, et les données disponibles s'enrichissent en permanence. De nombreuses utilisations, seules ou combinées, de ces BDMA sont devenues possibles pour la surveillance épidémiologique, notamment des maladies chroniques.

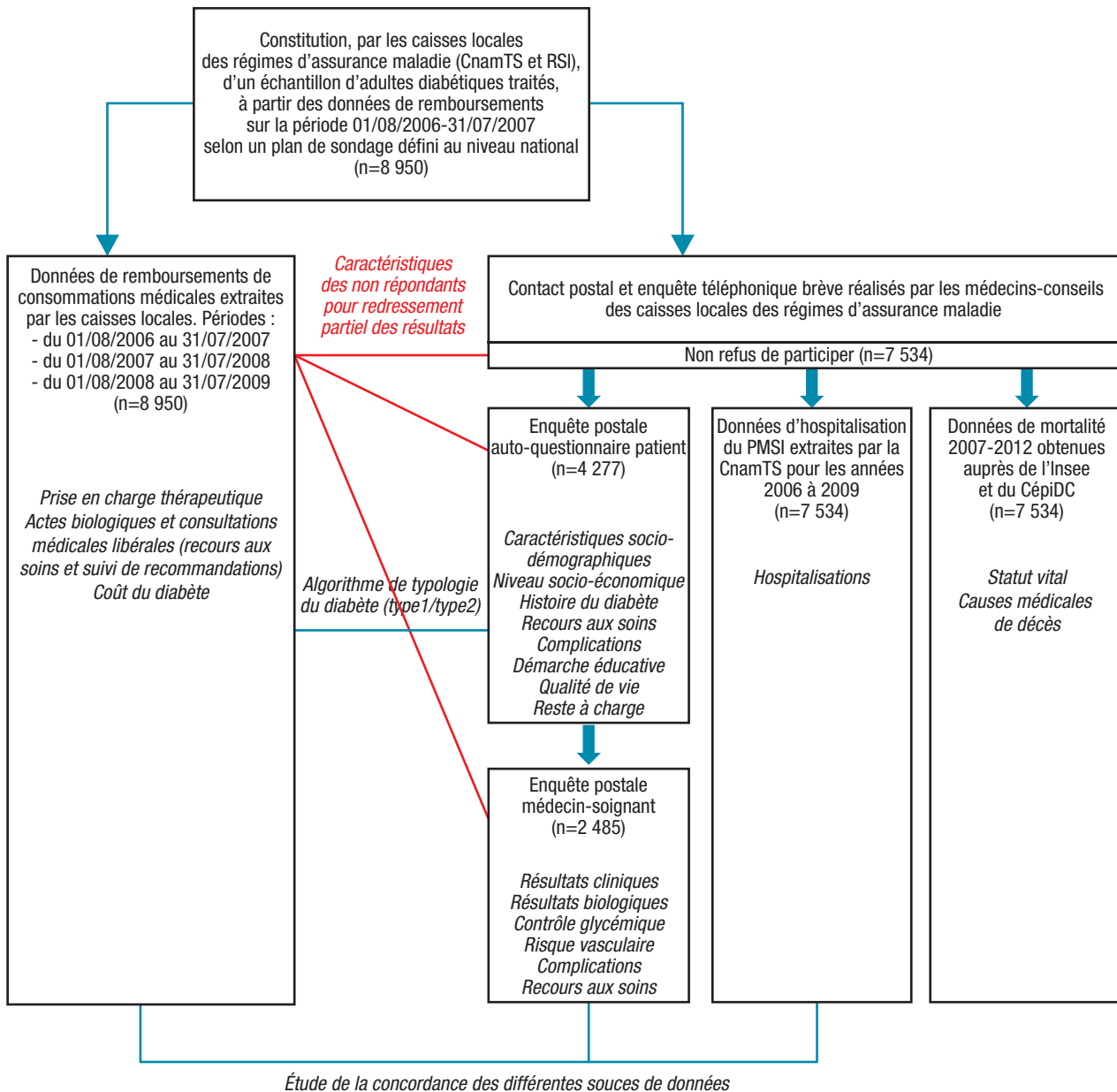
Deux exemples concrets peuvent illustrer ces différentes utilisations : la surveillance du diabète et celle des cancers.

## Surveillance du diabète

La surveillance du diabète repose, entre autres, sur le dispositif Entred (Échantillon national témoin représentatif des personnes diabétiques) qui associe, pour un échantillon de diabétiques traités par antidiabétiques, enquête par auto-questionnaire et données des BDMA (soins ambulatoires ou de ville et hospitalisations *via* le PMSI) (figure 1). Les objectifs sont de décrire, notamment, l'état de santé des diabétiques traités, leur prise en charge ainsi que les déterminants et les complications associés au diabète. Cette étude, réalisée à deux reprises selon une méthode similaire, a permis d'établir un état des lieux de l'état de santé de la population diabétique et de sa prise en charge médicale en début et fin de la précédente décennie. Elle est complétée par des études de mortalité à 5 et 10 ans. Ce dispositif montre par exemple une amélioration sensible, qui reste cependant insuffisante au regard des recommandations, de la prise en charge des personnes diabétiques<sup>1</sup> (figure 2), et met en évidence les inégalités sociales en termes d'état de santé et de qualité des soins. Limite inévitable d'un tirage au sort à partir des BDMA, et donc lié au recours au système

Figure 1

Sources de données et principales analyses dans l'enquête Entred-Métropole, adultes, France 2007-2010



CnamTS : Caisse nationale d'assurance maladie des travailleurs salariés ; RSI : régime social des indépendants ; PMSI: Programme de médicalisation des systèmes d'information ; CépiDC : Centre d'épidémiologie sur les causes médicales de décès (Inserm).

de soins, ce dispositif ne fournit aucun renseignement sur les personnes diabétiques non diagnostiquées et/ou non traitées par antidiabétiques.

**Surveillance des cancers**

La surveillance des cancers est assurée par le dispositif des registres. *Nec plus ultra* de la qualité des données (exhaustivité, recherche active des cas, codage standardisé, programme de formation continue pour les codeurs, panel d'experts pour le codage des cas difficiles, retour possible aux dossiers médicaux, etc.), ce dispositif présente cependant deux écueils majeurs, inhérents à la qualité recherchée :

- une couverture géographique relativement faible (moins de 20% de la population, variable selon le cancer considéré) ;

- un délai de trois ans minimum entre la collecte des données et leur publication.

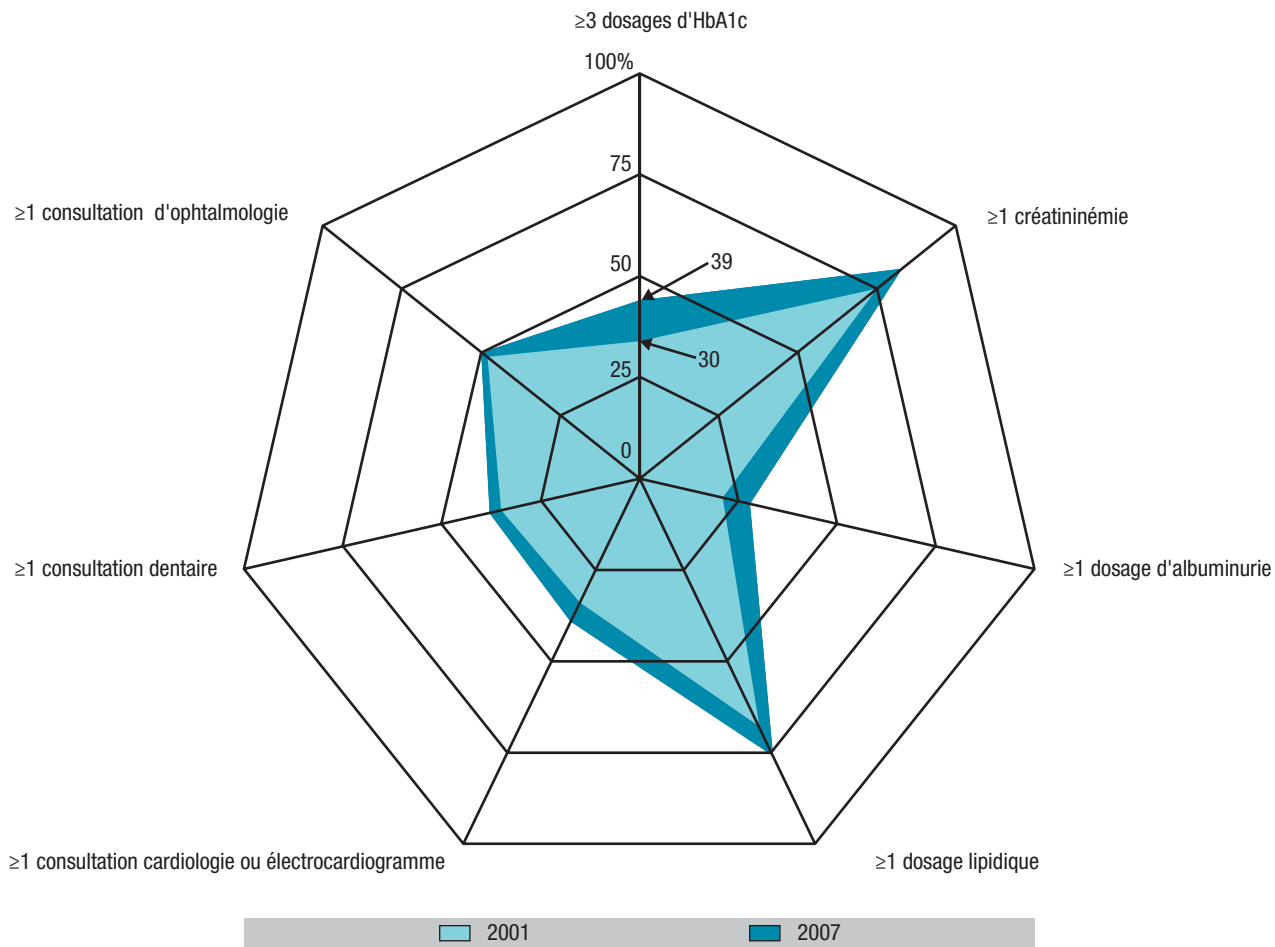
Dans ces conditions, les BDMA peuvent contribuer à la surveillance des cancers pour fournir :

- des informations sur les tendances récentes de l'incidence nationale, c'est-à-dire sur les trois dernières années ;
- des estimations de l'incidence de cancers au niveau régional et départemental dans des zones non couvertes par les registres.

Concernant les tendances récentes de l'incidence nationale, les courbes d'évolution des nouvelles prises en charge pour ALD cancer montrent, pour la plupart des localisations de cancers, des évolutions stables et parallèles aux courbes d'incidence

Figure 2

Évolution entre 2001 et 2007 du suivi des actes recommandés annuellement dans le diabète de type 2



Lecture : 39% des personnes diabétiques enquêtées dans Entred 2007 avaient bénéficié de 3 dosages d'HbA1c dans l'année, contre 30% dans Entred 2001, témoignant d'une amélioration du suivi de l'hémoglobine glyquée chez les personnes diabétiques.

issues des registres. Sur les années les plus récentes, pour lesquelles les données de registres ne sont pas encore disponibles, on suppose que cette relation entre les deux bases de données se maintient, et ce d'autant plus que les courbes obtenues corroborent des connaissances et hypothèses épidémiologiques fiables : par exemple, l'inflexion des ALD notée pour le cancer de la prostate était déjà observée avec la stabilisation de l'incidence des données issues des registres, et elle est cohérente avec les observations européennes de diminution de l'incidence.

Pour les estimations d'incidence infra-régionale ou départementale, utiliser les BDMA (PMSI ou ALD)<sup>2</sup> nécessite de vérifier au préalable qu'elles fournissent de bons proxys de l'incidence et, si oui, de déterminer laquelle des deux, PMSI ou ALD, produit les estimations les plus fiables. La performance de chacune doit être impérativement évaluée par comparaison, pour les départements couverts par les registres, entre les incidences produites à partir des BDMA et celles des registres, qui servent de *gold standard* : une méthode statistique *ad hoc* compare les rapports I/PMSI ou I/ALD pour chaque

département de la zone registre (I est l'incidence départementale issue du registre et PMSI ou ALD est « l'incidence administrative » issue du PMSI ou des ALD). De plus, un algorithme de critères, non présenté ici, a également été mis en œuvre pour permettre d'identifier la plus performante des deux bases.

Cette étude de validation montre que PMSI ou ALD peuvent fournir séparément des estimations d'incidence départementale ou régionale pour 14 des localisations de cancers étudiées (11 pour les ALD). En revanche, pour plusieurs localisations (thyroïde, mélanome, leucémies, système nerveux central) ces bases, en l'état actuel, ne permettent pas des estimations valides. Croiser au niveau individuel PMSI, ALD et/ou d'autres données du Sniiram pourrait permettre d'identifier, par des algorithmes plus fins, les types de cancers pour lesquels les bases ALD ou PMSI séparées se sont révélées peu performantes. L'article de Y. Kudjawa et coll. dans ce même numéro illustre cette possibilité.

## Forces et limites des bases médico-administratives

Si les BDMA ne répondent pas à tous les besoins de surveillance, elles offrent des possibilités nouvelles, comme le montrent les deux exemples ci-dessus et l'ensemble des articles de ce numéro. En effet, Sniiram et PMSI sont, avec les certificats de décès, les principales sources de données médicalisées nationales françaises exhaustives, permanentes et disponibles à un niveau géographique fin (la commune de résidence du patient). Réunies au sein d'un même système d'information, appelé ici « Sniiram-PMSI », elles couvrent quasiment l'ensemble des soins pris en charge par l'assurance maladie de la population résidant en France. Le processus d'amélioration continue, en termes de contenu (nouvelles variables, nouvelles définitions, qualité du codage, etc.) et de couverture populationnelle (quasi-totalité des régimes), les rend utilisables à des fins de surveillance, notamment des maladies chroniques. Les évolutions récentes du cadre réglementaire (arrêté du 19 juillet 2013), qui donnent un accès élargi au Sniiram aux agences sanitaires françaises, et en particulier à l'InVS, vont permettre de développer cette utilisation.

Les BDMA permettent de repérer certaines pathologies, événements de santé ou processus morbides. Elles permettent de distinguer les cas incidents des cas prévalents : grâce à la date des soins et à l'historique disponibles, il est possible de repérer la première occurrence d'un recours aux soins qui définit un cas incident, toute occurrence subséquente pour une même pathologie étant considérée comme prévalente. Le Sniiram-PMSI permet aussi d'apprécier la prise en charge et d'évaluer certaines actions de prévention ou de dépistage comme la couverture vaccinale, les mammographies ou le dosage des PSA.

Un autre intérêt majeur, bien illustré par Entred, mais aussi dans des études de cohortes telles que Constances ([www.constances.fr](http://www.constances.fr)) ou Coset ([www.coset.fr](http://www.coset.fr)), dans les enquêtes transversales santé et protection sociale (ESPS : [www.irdes.fr/EspaceRecherche/Enquetes/ESPS/index.html](http://www.irdes.fr/EspaceRecherche/Enquetes/ESPS/index.html)) ou dans le projet Hygie ([www.irdes.fr/EspaceRecherche/Partenariats/Hygie/index.htm](http://www.irdes.fr/EspaceRecherche/Partenariats/Hygie/index.htm)), réside dans la possibilité de croiser ces bases, au niveau individuel, avec d'autres sources de données, et de créer ainsi de nouveaux systèmes d'information « multisources ». Le croisement des BDMA avec des données d'enquête permet de pallier l'absence de données sociodémographiques, professionnelles et sur les habitudes de vie dans les BDMA ou, à l'inverse, de contrôler certaines données déclaratives (biais de mémoire), d'assurer un suivi passif, y compris des non-répondants et des perdus de vue, et éventuellement d'opérer des redressements. L'appariement avec d'autres sources de données administratives, comme celles de la Caisse nationale d'assurance vieillesse (Cnav), est une autre façon d'accéder à des informations sociales, économiques et professionnelles.

Toutefois, ces bases, en particulier celles de l'assurance maladie, sont complexes, comportent de très nombreuses variables et sont lourdes à utiliser. Par construction, le Sniiram-PMSI ne concerne que les patients pris en charge et les soins remboursés et présentés au remboursement. On n'y trouve ni motif de recours, ni résultat d'examen. Limite ou avantage, il est impossible de remonter au dossier du patient en raison de l'anonymisation irréversible. Ces BDMA évoluent en permanence, notamment en fonction des évolutions de la prise en charge et des politiques de remboursement. La « qualité » des informations est liée à leur finalité *princeps* : par exemple, l'inscription en ALD est « professionnel et patient dépendante » ; elle peut dépendre, entre autres, de l'existence et de la qualité d'une couverture complémentaire de patient ou de la sévérité et du stade de la maladie.

Enfin, dans une perspective d'analyse et de surveillance épidémiologiques, le repérage des patients atteints d'une pathologie peut s'avérer difficile, voire impossible, en particulier pour les pathologies peu ou pas hospitalisées ou qui ne font pas l'objet d'une prise en charge en ALD. L'élaboration d'algorithmes, plus ou moins complexes, avec les cliniciens et les producteurs de données, combinant la consommation de médicaments, la réalisation d'actes médicaux spécifiques et les hospitalisations, peut permettre d'améliorer ce repérage.

Pour ces différentes raisons, les indicateurs d'incidence et de prévalence « administratives » issus de ces bases et leurs évolutions ne peuvent être directement assimilés à des indicateurs épidémiologiques. D'un point de vue épidémiologique, les qualités métrologiques de ces BDMA doivent être évaluées. En fonction des objectifs, plusieurs types d'études sont possibles :

- par comparaison au cas par cas, pour estimer la sensibilité et la valeur prédictive positive, soit par retour au dossier médical au niveau local et sur des effectifs restreints (pour le PMSI, le retour au dossier est possible au niveau d'un établissement et sous réserve de l'obtention des autorisations requises), soit par croisement avec des données d'enquêtes ou de cohortes ;
- par comparaison avec des données agrégées mais reposant sur des données dont la qualité est éprouvée (comme c'est le cas dans l'exemple ci-dessus sur les cancers).

Une veille prospective des évolutions des prises en charge et des bases elles-mêmes est indispensable et peut conduire, en cas de modification importante, à réinterroger la validité des algorithmes utilisés.

Malgré les avantages des BDMA, la surveillance des maladies chroniques ne peut reposer uniquement sur elles. Elle nécessite le recours à d'autres sources de données : enquêtes, examens de santé, etc., notamment pour l'étude des déterminants et habitudes de vie, des stades précoces des pathologies, des recours différentiels au système de soins ou encore pour des populations spécifiques peu ou pas

repérables dans les BDMA (personnes sans domicile par exemple).

## Enjeux et perspectives

Du fait du vieillissement de la population, au-delà de la surveillance épidémiologique thématique des maladies chroniques et afin de guider les politiques de prévention et de prise en charge, une surveillance appropriée de la co-occurrence des maladies, de leurs déterminants et de leurs conséquences en termes de handicap et de dépendance reste à définir et à développer. Dans cette perspective, il paraît indispensable qu'un lien puisse être fait entre les BDMA et les sources de données médico-sociales dans le domaine du handicap et de la dépendance (maisons départementales des personnes handicapées, allocation personnalisée d'autonomie...).

Le contexte économique actuel, particulièrement défavorable, requiert plus que jamais de mettre en œuvre une surveillance efficace, peu coûteuse mais réactive, dans laquelle l'évaluation de l'impact du contexte économique sur la santé, ainsi que la mesure et le suivi des inégalités de santé, en particulier socio-économiques, représentent des priorités. Les BDMA sont pour cela des sources de données particulièrement intéressantes, à condition de pouvoir disposer de variables socio-économiques et professionnelles, soit par le recours à des variables proxys telles que les indices de désavantage social, qui approchent le statut socio-économique individuel sur la base d'une combinaison validée d'indicateurs économiques et sociaux de la zone de résidence, soit par l'appariement avec des données administratives comme celles de la Caisse nationale d'assurance vieillesse (Cnav).

Le pilotage régional de la politique de santé nécessite de disposer d'informations à un niveau géographique infra-régional pertinent. Les informations sont disponibles au niveau communal dans le Sniiram-PMSI, mais une géolocalisation plus fine pourrait permettre des études contextuelles et environnementales au niveau géographique le plus adapté.

Le projet Amphi (Analyse de la mortalité post-hospitalisation), à la recherche d'indicateurs de mortalité représentatifs de la qualité des soins, a montré que l'appariement de la base de données PMSI-MCO à celle des causes médicales de décès *via* le Sniiram était possible. Cet appariement peut et doit maintenant sortir du stade de l'expérimentation. En effet, outre les appariements déjà évoqués, réunir au sein d'un même système d'information Sniiram-PMSI et causes médicales de décès permettrait d'étudier la mortalité par cause pour certaines pathologies identifiables selon des critères de prise en charge et ferait de ce système un outil d'une très grande puissance.

À plus long terme, la mise en place du dossier médical personnel (DMP) et du cadre d'interopérabilité des

systèmes d'informations en santé pourrait conduire à une structuration et une standardisation des informations de santé enregistrées par les professionnels pour la prise en charge. Elle pourrait ainsi permettre une utilisation de ces informations pour la santé publique, en particulier la surveillance. Compte tenu de l'importance des BDMA, mais aussi de leurs limites, il est impératif que ces données produites pour le soin puissent, à terme, être appariées aux données du Sniiram et du PMSI pour tout ou partie de la population. Cela permettrait par exemple de pouvoir renseigner motifs de consultation et résultats d'examen biologiques, mais aussi habitudes de vie et autres déterminants.

La France dispose de bases de données administratives médicalisées exhaustives nationales enrichies en continu. La possibilité de les appairer à d'autres bases de données d'envergure nationale représente un atout majeur face aux enjeux de santé publique actuels, notamment concernant les maladies chroniques.

Des évolutions du cadre législatif et réglementaire sont indispensables pour rendre ces appariements possibles dans des délais raisonnables et en garantissant aux citoyens la sécurité et la confidentialité des données personnelles de santé ainsi que leurs utilisations dans l'intérêt général.

Ces enjeux et perspectives font l'objet de réflexions stratégiques au niveau national, notamment au Haut Conseil de la santé publique<sup>3</sup> et dans un très récent rapport remis à la Ministre de la santé<sup>4</sup>. ■

## Références

[1] Pornet C, Bourdel-Marchasson I, Lecomte P, Eschwège E, Romon I, Fosse S, *et al*, for the Entred Scientific Committee. Trends in the quality of care for elderly people with type 2 diabetes: the need for improvements in safety and quality (the 2001 and 2007 Entred studies). *Diabetes Metab.* 2011;37(2):152-61.

[2] Colonna M, Mitton N, Remontet L, Belot A, Bossard N, Grosclaude P, *et al*. Méthode d'estimation de l'incidence régionale des cancers à partir des données d'incidence des registres, des données de mortalité par cancer et des bases de données médico-administratives. *Bull Epidémiol Hebd.* 2013;(43-44-45): 566-74.

[3] Haut Conseil de la santé publique. Pour une meilleure utilisation des bases de données administratives et médico-administratives nationales pour la santé publique et la recherche. Paris:HCSP; 2012. 56 p. <http://www.hcsp.fr/Explore.cgi/avisrapportsdomaine?clefr=258>

[4] Bras PL, Loth A. Rapport sur la gouvernance et l'utilisation des données de santé. Paris:Ministère de la Santé et des Affaires sociales;2013. 128 p. <http://www.drees.sante.gouv.fr/rapport-sur-la-gouvernance-et-l-utilisation-des-donnees-de,11202.html>

## Citer cet article

Grémy I, Doussin A. Surveillance des maladies chroniques en France : la contribution des bases de données médico-administratives. *Bull Epidémiol Hebd.* 2013;(Hors-série):9-14.