



Ministère des finances et des comptes publics  
Ministère des affaires sociales, de la santé et des droits  
des femmes  
Ministère du travail, de l'emploi, de la formation professionnelle  
et du dialogue social

santé  
protection sociale

drees

recherche, études, évaluation et statistiques

# Ouverture des données de santé: l'article 47 du projet de loi de santé

**André Loth**

*avril 2015*

# Données de santé : vue d'ensemble

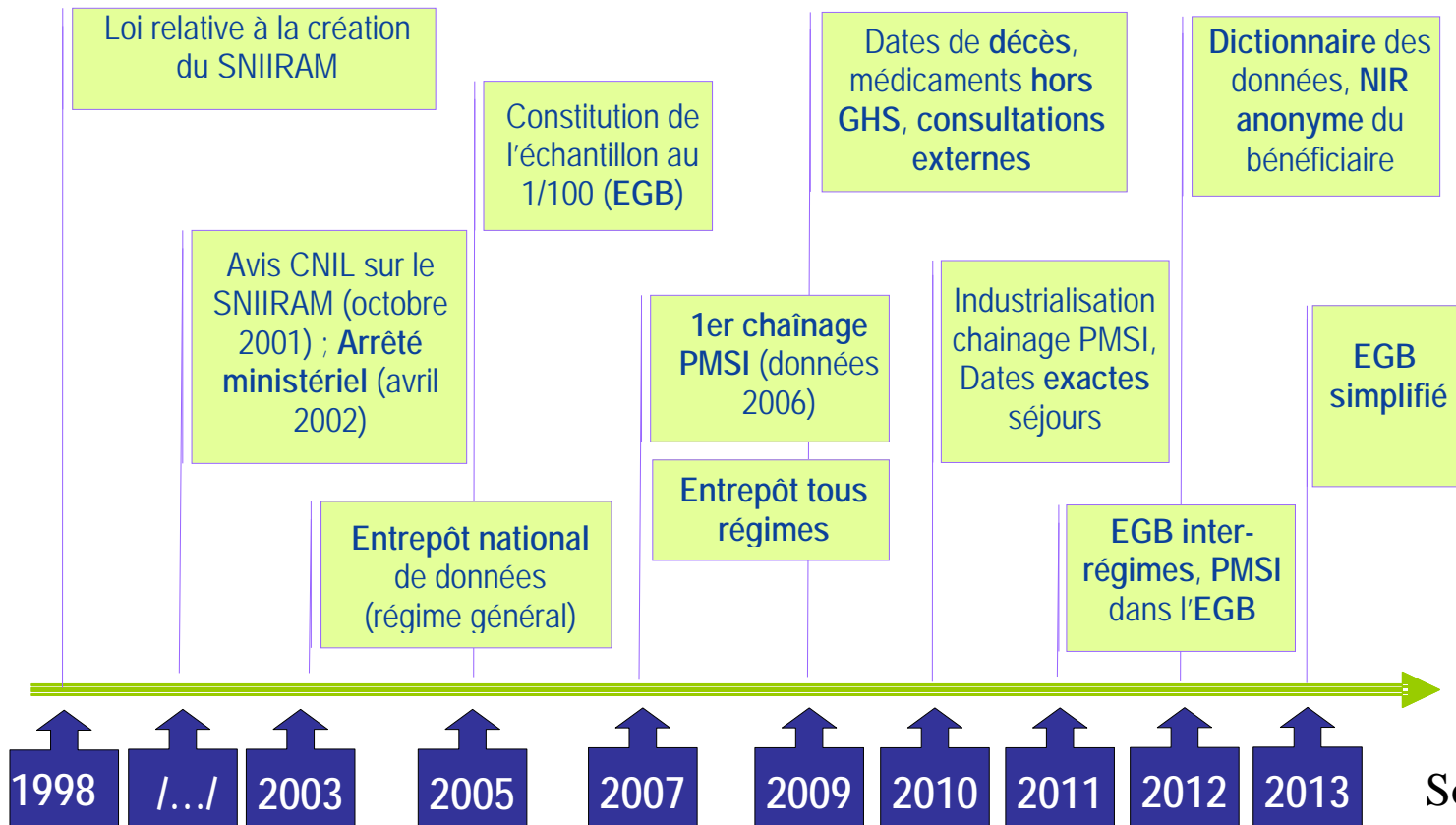
- **Données médicales à visée individuelle (pour soigner)** : dossiers médicaux
- **Données médico-administratives à visée « métier » (pour administrer, gérer, contrôler, autoriser, payer...)** : résumés de séjour hospitaliers (PMSI), « décomptes » de l'assurance maladie (SNIIR AM)...
- **Données à visée de connaissance** : collectées directement (enquêtes, panels, cohortes, recherche biomédicale...) ou **obtenues par réutilisation** :
  - de dossiers médicaux, de registres de maladies, de biobanques...
  - **de données médico-administratives « déidentifiées »** sous diverses formes (données granulaires, données agrégées, indicateurs etc.)
    - À destination des administrations elles-mêmes (système d'information « décisionnel » ou de pilotage...)
    - **A destination de tiers** (chercheurs, institutions, entreprises, journalistes...)

## Données médico-administratives et SNIIR-AM

- Données médico-administratives à vocation exhaustive :
  - Décomptes assurance maladie (SNIIR AM)
  - Résumés de sortie hospitaliers (PMSI)
  - Certificats de décès (CépiDc)
  - Certificats de santé des enfants...
  - Accidents corporels
  - Autorisations d'activité (référentiels : FINESS, RPPS...)
  - etc.

- SNIIRAM-PMSI :
  - Toute la population
  - Age, sexe, commune de résidence, date de décès
  - Soins de ville (y c. médicaments)
  - Soins hospitaliers
  - Arrêts de travail...
  - ALD, CMU...
  - **Chaînage (suivi longitudinal par individus)**
  - prestataires de soins

## SNIIRAM-PMSI : une base récente



Source  
CNAMTS

## Une richesse sous-utilisée et un débat public

- Une base encore jeune et déjà riche donc mais difficile d'accès :
  - obstacles juridiques (obscurité, complexité et problème du NIR...)
  - obstacles techniques
- Critiques des chercheurs (rapports du HCSP)
- Critiques de l'Institut des données de santé et polémique début 2013
- Rapport Bras sur la gouvernance et l'utilisation des données de santé (octobre 2013) : vers une doctrine claire...
- « Commission open data » : vers le consensus
- Projet de loi de santé : enjeux sanitaires, démocratiques, économiques

## Utilisations des données pour études et recherches (exemples)

- Qualité de la prise en charge dans notre système de soins
- Effets des médicaments en vie réelle (pharmaco-épidémiologie) : exemple du Mediator
- Parcours de soins
- Cartographie des dépenses par pathologie et analyse médicalisée des dépenses (diabètes, cancers, dépression...)
- Etudes de patientèle ou de performance des établissements de santé, suivi de la diffusion des nouveaux traitements, production de tables de mortalité (segmentation du risque)...
- On n'a pas besoin d'une base exhaustive sur le plan statistique. C'est souvent vrai, mais :
  - études d'événements rares => La France peut être le principal fournisseur de données de qualité au niveau mondial
  - études régionales (disparités de prise en charge, inégalités territoriales, études de parcours...)
- A quoi s'ajoutent les études faites par appariements avec d'autres jeux de données :
  - cohortes et recueils complémentaires auprès des personnes (exemple 200 000 personnes dans la cohorte CONSTANCES à terme...)
  - enquêtes (ex : enquête de la DREES auprès des personnes âgées dépendantes en 2015-2016, enquêtes santé de l'Irdes et de la DREES...)
  - ou appariements avec autres bases administratives
    - le système national de gestion des carrières de la CNAV : impact des professions sur la santé...
    - les bases de la CNAF ou de la DGFIP : impact sur la santé de la composition et des revenus du ménage...

## Une doctrine claire : deux types d'accès aux données

- Pour les données anonymes : *open data* (le développer !)
  - Cas particulier des données nominatives sur l'activité des professionnels (autoriser la réutilisation des données déjà publiées par l'assurance maladie)
  
- Pour les données indirectement nominatives : accès restreint, projets autorisés par la CNIL, pour de bonnes raisons (besoin technique vérifié + intérêt public) et dans de bonnes conditions (traçabilité) pour des demandeurs publics ou privés
  - Procédures simplifiées à
  - Cas particulier des organismes dont la mission requiert un accès permanent



## L'ouverture des données de santé dans le projet de loi de santé (article 47)

- Création d'un système national des données de santé (SNDS = SNIIRAM + PMSI + causes de décès + données médico-sociales + données d'assurance maladie complémentaire)
- Règles d'accès (open data / accès restreint...) et Gouvernance (gestion confiée à la CNAMTS, Institut national des données de santé, pilotage par l'Etat)
- En finir avec la contrainte du décret en CE pour les appariements nécessitant le NIR (en vue d'une recherche ou étude en santé)
- Fusion des chapitres IX et X de la loi informatique et libertés : unification et simplification



## Données « déidentifiées » ou « pseudonymisées », données anonymes

- Feuilles de soins et dossiers médicaux sont évidemment nominatifs à l'origine
- Les données du SNIIRAM-PMSI sont « dé-identifiées » (pas de nom + NIR chiffré de manière irréversible) mais pas vraiment anonymes :
  - Si les données sont précises et exhaustives, on y reconnaît des personnes sur lesquelles on dispose d'informations par ailleurs (proches, célébrités, employés, clients...)
  - Si par ex. on connaît hôpital, mois de sortie, durée du séjour, date (m/a) de naissance, sexe, code géographique, alors 89% des hospitalisés sont identifiables dans la base PMSI et 100 % des personnes hospitalisées plus d'une fois dans la période
- Les données du PMSI et les extractions du SNIIRAM sont aujourd'hui en accès restreint mais sans traçabilité :
  - Où sont les données ?
  - Plus il y a de personnes dans le secret, moins il y a de secret...

## Risques de réidentification : comment s'en prémunir ?

- Open data :
  - Oui pour les données anonymisées: Comment ? Lesquelles ?
  - Non pour les données *personnelles* de santé
- Accès restreint aux données personnelles : le droit à la vie privée n'est pas une idée dépassée
- Pas de sanction ni dissuasion possible sans traçabilité...
- L'objectif c'est l'ouverture la plus large possible (maximum de données, maximum d'utilisations) compatible avec la protection de la vie privée
- Mais quel équilibre ? Craintes des deux côtés...

## Gouvernance : quatre aspects

- Où se discutent les orientations, les résultats : un GIP Institut national des données de santé
- Qui donne l'autorisation d'accès à des données personnelles ?
  - **Autorisation** CNIL après avis :
    - **Avis** du comité d'expertise (évalue techniquement le besoin d'accéder aux données),
    - **Avis** du futur INDS (avis sur l'utilité publique, *si la question se pose*)
    - *Ou avis d'un comité de protection des personnes (pour les recherches biomédicales etc.)*
  - Un portail unique pour aider les porteurs de projet à s'y retrouver
  - Des procédures simplifiées : simple engagement de conformité
- Qui assemble et gère concrètement les données ? CNAMTS, ATIH...
- Pilotage par un Etat plus présent et plus attentif aux demandes

## A lire :

- *Rapport sur la gouvernance et l'utilisation des données de santé (« rapport Bras »)*
- *Rapport de la Commission open data en santé (y c., son annexe 9: Rapport sur les risques de réidentification)*
- *Rapport 2014 du Conseil d'Etat, sur le numérique et les droits fondamentaux*
- *article 47 du projet de loi sur la santé (+ dossier législatif, sur le site Legifrance)*

**Merci de votre attention**

**Des questions ?**