

# DONNÉES DE SANTÉ : CONCILIER LEUR USAGE ET LA PROTECTION DE LA VIE PRIVÉE

---

André LOTH

Ex administrateur général (ministère des  
affaires sociales et de la santé)

# SACRIFIER LA PROTECTION DE LA VIE PRIVÉE AUX USAGES DES DONNÉES MASSIVES ?

Où on verra comment l'Etat s'est doté progressivement d'un gigantesque système d'information sans en prendre d'emblée la mesure

Où on verra comment la loi a permis l'usage des données pour tous sans sacrifier la protection de la vie privée

Où on verra comment l'Administration et le Législateur se débrouillent pour concilier le domaine de la loi (la protection de la vie privée) et la complexité technique

# LES DONNÉES DONT IL EST QUESTION

- ❑ **Données de santé pour soigner** (nominatives) : dossiers médicaux, prescriptions, imagerie, prélèvements, résultats d'examens et de capteurs...
- ❑ **Données médico-administratives pour gérer, contrôler, autoriser, payer** (nominatives) :
  - **résumés hospitaliers** (Programme de Médicalisation des Systèmes d'Information - PMSI)
  - **feuilles de soins**, décomptes et fichiers de l'assurance maladie...
- ❑ **Données de santé à visée de connaissance** (« pseudonymisées ») :
  - **Recueils ad hoc** : recherche biomédicale, enquêtes, panels, cohortes, registres,... (cf. *loi Jardé sur « recherches impliquant la personne humaine »*)
  - **Réutilisation** de dossiers médicaux, de bio-banques, de données génomiques...
  - **Réutilisation de données médico-administratives individuelles (ce dont il s'agit ici)**

# LES DONNÉES « MÉDICO-ADMINISTRATIVES » : UNE AVENTURE ADMINISTRATIVE INSCRITE DANS DES LOIS

- Loi hospitalière de juillet 1991 : les hôpitaux doivent décrire leur activité (la généralisation du PMSI prépare la future tarification à l'activité de 2004)
- Loi Teulade de janvier 1993 : codage des actes et des pathologies (L. 161-29 CSS)
- Ordonnance Juppé de 1996 (L. 161-33 CSS) : codage (actes et médicaments) et dématérialisation des feuilles de soins (déploiement des cartes *Vitale* en 1998-99)
- LFSS de 1998 : crée le Système national d'information inter-régimes (Sniiram) mis en place par la Cnamts de 2003 à 2009
- Loi santé de janvier 2016 : l'article 193 crée le Système national des données de santé (SNDS) : Sniiram + PMSI + causes de décès (CépiDc)...
- Projet de loi Buzyn (article 11) : vise aujourd'hui à étendre le SNDS aux dossiers hospitaliers, registres de maladies, PMI, GIR etc.

# UNE PRISE DE CONSCIENCE TARDIVE

## **Sous estimation initiale des usages des données médico-administratives**

- Objectif initial : « ne pas être un payeur aveugle » (cf. maîtrise médicalisée des dépenses)
- Les chercheurs et la Cnamts prennent conscience du « trésor »
- Faible implication de l'Etat, textes confus sur les droits d'accès...

## **Sous-estimation initiale des risques pour la vie privée**

- Données anonymisées ? Non, pas vraiment : cf. l'exemple du PMSI (données pseudonymisées mais ré-identifiables par âge, sexe, lieu de résidence...)
- Engagements de ne pas les diffuser ? Mal respectés (pas de traçabilité des copies)

**L'affaire du palmarès des hôpitaux** à l'origine d'un nouveau chapitre de la Loi informatique et liberté (in loi CMU de 1999)

# RICHESSE DES DONNÉES ET ENJEUX (SANITAIRES, ÉCONOMIQUES, DÉMOCRATIQUES...)

Richesse sans réel équivalent dans le monde (France Etat centralisé...)

- Base exhaustive + qualité du chaînage (pseudonyme = NIR « haché ») : suivi des parcours de soins ; 17 millions de séjours hospitaliers et 1,3 milliard de feuilles de soins *par an*
- Il manque facteurs de risques (tabac, alcool, IMC...), résultats d'examens, données génomiques...
- Mais démultiplication des usages si on rend possibles des appariements avec d'autres bases (cohortes, registres de maladie, bio-banques, dossiers hospitaliers, voire *Dossier Médical Partagé...*)

Usages, Enjeux :

- Performances des prestataires de soins, évaluation des nouveaux traitements, tables de mortalité, innovations, information du public, et faire de la France un acteur majeur du traitement des données de santé : **les enjeux sanitaires et économiques**
- Qualité des soins, effets des médicaments en vie réelle : exemple du *Mediator* et **l'enjeu de sécurité sanitaire**
- Dépenses par pathologie, effectifs et revenus des professions de santé, stratégies thérapeutiques, qualité des soins, inégalités d'accès aux soins : **l'enjeu démocratique**

# LE CONTENU DU SYSTÈME NATIONAL DES DONNÉES DE SANTÉ (v.1 : LOI DE 2016)

- **Données individuelles, pseudonymisées, par patient :**
  - Age, sexe, commune de résidence, date de décès
  - Causes de décès (ajout par loi de 2016)
  - Date et nature des soins de ville (y c. médicaments) présentés au remboursement
  - Dates et nature des soins hospitaliers (y c. diagnostics, chirurgies...)
  - Tarifs appliqués, prise en charge par assurance maladie obligatoire...
  - Affections de longue durée, CMUc...
  - Arrêts de travail, rentes AT-MP, invalidité...
  - *[Médico-social ? Echantillon assurance complémentaire ? Prévus par loi de 2016]*
- **Chaînage (suivi longitudinal des individus y c. enfants) et exhaustivité**
- **Données sur prestataires de soins :** diplômes et spécialités, mode d'exercice âge, sexe, localisation, activités, recettes, prescriptions...

# LA POLÉMIQUE, LE RAPPORT DE 2013, LA LOI DE 2016

- 2005 à 2009 les chercheurs découvrent la base mais aussi les lenteurs de la procédure d'accès aux données (notamment l'exigence d'un décret en CE pour l'usage du NIR seul moyen d'apparier des données dé-identifiées)
- En 2012-2013 critiques virulentes, pétition, tribunes dans la presse sur :
  - refus de la Cnamts d'ouvrir ses données même anonymes à des acteurs privés (assureurs, industriels du médicament) par crainte du mésusage
  - souci de protéger la vie privée vu comme un prétexte pour freiner l'open data, la recherche et l'innovation (prise de position en ce sens du président de l'Institut des données de santé et du Pr. Didier Sicard ancien pdt du CCNE)
  - Inversement, constat d'une diffusion imprudente des données hospitalières avec l'accord de la Cnil et du ministère et d'un défaut de doctrine et de clarté
  - Au-delà des obstacles juridiques, délais liés aussi à l'énormité et à la complexité des données.
- Commande d'un rapport à Pierre-Louis Bras qui m'associe à sa rédaction



# LE RAPPORT BRAS-LOTH DE 2013 DÉFINIT UNE DOCTRINE QUI SERA REPRISE DANS LA LOI

**1- Ouvrir sans restriction les données anonymes** (quel qu'en soit l'usage) **et les données nominatives sur les tarifs** moyens des professionnels de santé (concertation pour les autres données d'activité)

**2- Accès restreint aux données de santé personnelles**

- À des administrations dont la mission l'exige (typiquement ANSM, HAS, DREES...),
- Et à des organismes publics ou privés pour des études ayant une finalité d'intérêt public quels que soient les commanditaires, dans des conditions de sécurité convenables (traçabilité...)

**3- Traiter le NIR comme le nom** et remplacer le « décret en CE » par une autorisation de la CNIL pour permettre les appariements). La fétichisation du NIR ne protège pas les citoyens ; elle freine la recherche

# LA LOI DE 2016 : CLARIFICATION INCOMPLÈTE

- Article 193 : une alouette d'open data (pour les données anonymes), et un cheval d'accès contrôlé (pour les données à caractère personnel)
- Accès permanent pour certaines administrations
- Autorisations d'accès (Cnil après instruction par Institut national des données de santé)
  - **pour des bonnes raisons** (intérêt public, et seulement aux données techniquement nécessaires...) mais qu'est-ce que l'intérêt public ? Et si toutes les finalités doivent être décrites à l'avance, pas d'appariements pérennes possibles ?
  - **pour des bonnes personnes** (privé ou public mais pas d'accès direct des industriels et assureurs)
  - **dans de bonnes conditions techniques** (traçabilité, audits...)
- Tentative de simplification de la loi informatique et libertés pour la recherche en santé (fusion des chapitres, NIR, appariements...).
- Demi-échec, d'où le nouveau projet de loi : faciliter les appariements, étendre le SNDS aux données cliniques et étendre les missions de la Plateforme des Données de Santé (alias *Health Data Hub*)

# A NOTER

- ❑ Compatibilité plutôt que compromis entre utilisation des données et protection de la vie privée. A cet égard, le Législateur n'a pas suivi le CCNE (réduire la protection des uns pour accroître le bénéfice des autres ? )
- ❑ Compétence du Législateur (vie privée = domaine de la loi) vs complexité des sujets et incompétence technique du Parlement (exemple de sujets à la fois politiques et techniques : faisabilité du droit d'opposition vs inanité des droits d'accès ou de rectification),
- ❑ RGPD et particularités françaises (maintien de l'autorisation préalable et maintien d'un régime particulier du NIR : cf. article 30 LIL)
- ❑ Différences des régimes d'accès pour les chercheurs entre :
  - données de santé (Cnil...)
  - données INSEE ou autres données sociales ou fiscales (Cnis, Comité du secret...)

**MERCI DE VOTRE ATTENTION**