

# Les systèmes d'information pour la santé publique

## Auditions

Collection  
*Rapport*

Novembre 2009

**Réponses  
à la grille d'entretien  
avec les personnes  
auditionnées**

Institut national d'études démographiques : **Benoît Ryandey**

Institut de recherche et documentation en économie  
de la santé : **Chantal Cases**

Centre d'épidémiologie sur les causes médicales  
de décès (Cépdic) : **Eric Jougla**

**Compte-rendus  
d'audition**

Caisse nationale d'assurance vieillesse (CNAV) : **Vincent Poubelle  
Monique Beurey – Leconte**

Agence des systèmes d'information partagés de santé (Asip santé) :  
**Jean-Yves Robin**

Mission pour l'Informatisation du Système de Santé (MISS) : **André Loth**

Institut national de la statistique et des études économiques (Insee) :  
**Jean-Pierre Le Gléau - Alain Trognon**

Caisse nationale de l'assurance maladie des travailleurs salariés :  
**Yvon Merlière Alain Weill**

Institut national de prévention et d'éducation pour la santé (INPES) :  
**Pierre Arwidson**

**Glossaire**

# Benoît Riandey INED

L'Ined a une très grande ancienneté dans le domaine des enquêtes et le stock actuel est très important (plus d'une centaine). Ces enquêtes sont faites de plus en plus en coopération avec d'autres institutions et équipes scientifiques.

## Politique de diffusion

Le principe est la mise à disposition systématique des données d'enquête archivées par l'intermédiaire du réseau Quetelet. La mission de ce réseau est triple : contribuer à l'archivage des enquêtes, favoriser la diffusion des données d'enquêtes et valoriser ces mêmes données auprès de la communauté scientifique française et internationale. Les partenaires du Réseau Quetelet sont le CMH-ADISP, archives de données issues de la statistique publique du Centre Maurice Halbwachs, le Centre de données socio-politiques (CDSP), et le Service des Enquêtes de l'Institut national d'études démographiques (INED).

Calendrier de mise à disposition des nouvelles enquêtes : le principe est de prévoir un an d'apurement des fichiers, une année de priorité donnée aux équipes initiatrices de l'enquête puis la mise à disposition des équipes extérieures sur présentation de leur projet.

Cette politique de mise à disposition systématique est conforme à l'idée que des données collectées sur fond publics constituent un bien public.

## Difficultés pratiques, volume et moyens

Il peut se poser des difficultés de documentation pour des enquêtes anciennes.

Pour un nombre conséquent d'enquêtes, le volume de demandes n'est pas excessif du fait de l'intégration d'un grand nombre d'équipes dans le groupe de travail initial de l'enquête. Les moyens affectés correspondent de une à deux personnes au sein du service des enquêtes.

## Appariements de données

L'INED est particulièrement intéressé aux données du SNIIR-AM de la CNAM. En effet un identifiant stable sur le bénéficiaire permet d'approcher les migrations – données essentielles manquantes pour les démographes – par l'idée qu'une personne qui ne consomme plus est soit décédée soit sortie du territoire.

Les solutions pour l'appariement des fichiers peuvent être recherchées dans les techniques de hachage et de cryptage du NIR. Il est nécessaire de mettre en place des formations aux appariements sécurisés. Une sensibilisation est en place à l'École nationale de la statistique et de l'analyse de l'information (ENSAI) dans la filière santé et en formation permanente des attachés.

Dans certains services la mobilité excessive du personnel peut induire des problèmes de maintien de la compétence en ce domaine.

## Données localisées

Les fichiers sont anonymisés au sens de la loi informatique et libertés ce qui a pour contrepartie la suppression de la localisation fine et l'impossibilité de coupler avec des données de contexte.

Des solutions peuvent être recherchées dans les techniques d'estimation sur petits domaines.

---

## Articulation avec les systèmes internationaux

L'INED participe à un grand nombre d'enquêtes internationales : enquêtes européennes, enquêtes mondiales sur la fécondité, coopération avec les pays de sud à travers le Centre population et développement (CEPED).

## Difficultés d'accès

Le principal problème se situait au niveau de l'échantillon démographique permanent. Le CASD apporte une solution sérieuse.

## Chantal Cases - Irdes

### Le champ pour l'objet du groupe concerne l'ESPS et des enquêtes ponctuelles

#### Politique de diffusion dispositif concret

La politique d'accès et de diffusion est à distinguer de la question de la propriété des données :

Les données de l'Irdes sont en accès libre sur projet de recherche avec un délai de carence d'un an au profit des financeurs (ce qui apparaît dans les conventions avec les financeurs).

L'Irdes va entrer dans le réseau Quetelet pour les données de l'enquête mais pas pour les données appariées, ce qui nécessiterait l'accord des copropriétaires des données.

Il n'y a pas de contrôle sur le contenu des publications, mais uniquement une connaissance de ce qui est produit.

Des résultats agrégés tarifés sont diffusés sous forme payante à des organismes privés.

#### Difficultés pratiques, volume et moyens

8 conventions ont été signées concernant l'ESPS.

Il n'y a pas de difficultés pratiques car le volume est raisonnable et de plus le basculement va être réalisé sur le réseau Quetelet. En termes de moyens humains une personne est actuellement dédiée au traitement des demandes d'accès

#### Appariements

L'Irdes conduit une grosse activité d'appariement de données administratives pour l'ESPS avec les données de l'assurance maladie et pour les montants de remboursements avec la CNAV.

Les appariements sont faits par l'Insee en tant que tiers de confiance pour l'ESPS.

L'appariement ne pose pas de question pratique ou juridique, par contre il existe une question de finalité statistique non reconnue pour les données administratives qu'il faudrait faire figurer dans les COG par exemple

#### Données localisées

Des perspectives de géolocalisation existent

#### Articulation avec l'international

A noter que le minimodule d'Eurostat est intégré à ESPS et qu'Eurostat participe au comité de pilotage.

L'enquête SHARE est coordonnée pour la France par l'IRDES mais la collecte assurée par l'INSEE. Il faut noter l'énorme volume de publications suscité par cette enquête (200 répertoriées)

**Les difficultés d'accès** à des données d'autres institutions lorsqu'elles existent sont d'ordre pratique mais non pas juridique.

#### Rapports avec la CNIL

Les questions ne proviennent pas de divergences de fond mais de la faiblesse de ses moyens

---

# Eric jougla

## CépiDc Inserm

- CépiDc depuis 2001
- Inserm depuis 1968
- Auparavant, Insee

Il s'agit d'un service pérenne de l'INSERM qui en plus de la production de la statistique annuelle des causes médicales de décès en France ( le délai actuel de production est de deux ans) réalise des études et recherches sur les causes de décès et sur les méthodologies de codage des classifications de maladie. Il emploie 35 personnes.

Il n'existe pas de texte juridique en tant que tel concernant le CépiDc par contre le certificat de décès est régi au niveau du code des collectivités territoriales et a été modifié dernièrement par la loi de santé publique de 2004 (article 25) qui fixe les conditions de confidentialité et l'objet de la collecte des données (clairement à but de santé publique) et les destinataires du certificat.

Un projet majeur est actuellement le développement de la certification électronique en ligne (voir le service en ligne :

[http://www.cepidc.vesinet.inserm.fr/inserm/html/pages/certification\\_electronique\\_fr.htm](http://www.cepidc.vesinet.inserm.fr/inserm/html/pages/certification_electronique_fr.htm)

L'objectif premier est de développer la certification électronique à l'hôpital mais la certification électronique avance en cabinet libéral également. Il y a en effet nécessité de généraliser la certification électronique pour les systèmes d'alerte et pour l'amélioration de la qualité des données L'application est actuellement très bien reçue mais le déploiement encore très lent.

### Politique de diffusion difficultés pratiques, dispositif concret

1. Concernant la politique de diffusion, la base est vendue à de nombreux organismes (dont la FNORS et l'INED). Le site internet se caractérise par une grande ouverture avec des données très détaillées en accès libres.
2. Les difficultés pratiques à signaler ne concernent pas la mise à disposition des données mais des questions de production :
  - le problème du non retour des informations sur les causes de décès de certains IML ; en particulier toujours aucun retour de l'IML de Paris ce qui induit une sous-déclaration très importante des causes de décès des jeunes en Ile de France. Il y a une solution juridique à trouver.
  - Les décès non identifiés sous X (mots dans la rue)
  - La question des mort-nés.

### Difficultés pratiques, volume et moyens

3. Le volume est d'environ 500 demandes par an ce qui nécessite deux personnes à temps plein.

### Appariements

Les appariements se font bien avec les cohortes (une trentaine par an). Il serait très intéressant d'apparier avec l'EDP, le PMSI et l'EGB. Concernant l'EDP voir le point 8. Une étude pilote est actuellement menée avec la Drees et la Cnamts pour apparier les données PMSI et les causes de décès.

### Données localisées

Actuellement le décès est localisé à la commune mais il y aurait intérêt à disposer de données de mortalité à un niveau plus fin que la commune (Iris par exemple), ce qui suppose un géocodage des adresses de décès par l'Insee - ceci permettrait une analyse plus fine des différentiels de mortalité en milieu urbain.

### Articulation avec l'international

Le CépiDc a une activité importante de collaboration concernant la comparabilité internationale des données sur les causes de décès (OMS, Eurostat, NCHS...). Il faut noter que le codage est intégré dans un système international (modèle du certificat et règles de la classification internationale des maladies, sélection de la cause initiale de décès).

De plus, le CépiDc est le centre collaborateur de l'OMS pour la classification des maladies en langue française

### Difficultés d'accès

Il faut signaler les difficultés actuelles à avoir accès pour appariement aux échantillons Insee (EDP) qui permettrait l'analyse des différentiels sociaux de causes de décès. Par ailleurs la nécessité de réintroduire dans les fichiers CépiDc, le codage de la PCS pour tous les sujets et d'étendre la demande d'informations à la dernière profession exercée pour les retraités.

---

## Monique Beurey-Leconte et Vincent Poubelle, CNAV

Vincent Poubelle précise en préambule que la CNAV ne représente pas un institut statistique et que ses fichiers de gestion ne regroupent que des données qui servent à liquider des droits à retraite. Cette spécificité conduit le système d'information de la CNAV à suivre chaque individu de sa naissance à sa retraite, voire même après son décès, pour les pensions de réversion. Par ailleurs, le système ne réalise aucune suppression et conserve la trace de toute personne qui a cotisé au moins une fois au régime général, y compris les travailleurs immigrés. L'échantillon inter-régime des cotisants montre ainsi que 97 % des personnes qui acquièrent un droit à la retraite sont passés par le régime général.

La CNAV conserve trois fichiers nationaux :

- le Système national de gestion des identités (SNGI) regroupe toutes les personnes possédant un NIR, y compris les personnes décédées et les étrangers repartis dans leur pays d'origine, soit 97 millions d'identités. Ce fichier, alimenté par l'INSEE pour les naissances et les décès et par les CRAM pour les étrangers, sert de base de sondage ;
- le Système national de gestion des carrières (SNGC) enregistre, année par année, tous les événements susceptibles d'influencer le calcul de la retraite, tels les reports de salaires ou les périodes de maladie, d'invalidité ou de chômage. Cette chronologie est disponible depuis 1947 ;
- le Système national de gestion des dossiers (SNGD) constitue un fichier de suivi administratif de toute demande de retraite.

S'il n'existe pas de fichier central de gestion des retraités, le Système national des statistiques prestataires (SNSP) balaie trimestriellement les fichiers des CRAM pour constituer des bases exhaustives sur les retraités. Le SDDS, encore peu utilisé, constitue un fichier d'identification des employeurs. La CNAM a également délégué à la CNAV la gestion du Répertoire national inter-régimes de l'Assurance maladie (RNIAM).

Le fichier de gestion des carrières est alimenté en continu pour les salariés du régime général ou les chômeurs indemnisés. En revanche, les autres régimes, la fonction publique en particulier, ne fournissent pas tous les éléments, de sorte que ce n'est qu'à partir de 53-55 ans que la CNAV dispose de données complètes, notamment pour les 40% de poly-pensionnés.

Vincent Poubelle souligne que le niveau de détail et la qualité du système d'information s'améliorent, puisque la CNAV dispose désormais des date-à-date pour les périodes de maladie ou de chômage. Le SNSP représente actuellement le seul fichier central national. La DADS, produite par la CNAV, est envoyée aux services des impôts et à l'INSEE et injectée dans les fichiers de carrières, pour le calcul des droits à la retraite.

Quant aux statistiques, la CNAV dispose, pour les droits propres (retraites) comme dérivés (réversion), de flux exhaustifs, qui peuvent être utilisés comme base de sondage. Les données sur le secteur d'activité ne sont cependant disponibles que pour les années les plus récentes. Sur l'ensemble de ses bases de données, la CNAV tire, depuis 2002, un échantillon au vingtième, qui regroupe 4,5 millions d'individus sélectionnés selon cinq clés de NIR et permet de connaître tous les événements de la vie ainsi que le



salaires de l'individu. Il est mis à jour périodiquement pour intégrer les nouveaux retraités et renouvelé tous les deux ans. Il est d'abord utilisé pour le modèle de projection de PRISME, pour les projections à long terme, les simulations de réforme, les prévisions sur les comptes de sécurité sociale mais également pour réaliser des sous-échantillons, avec l'autorisation de l'organisme extérieur.

Marcel Goldberg s'enquiert de la disponibilité de l'information travaillée.

Vincent Poubelle explique que le modèle de projection repose sur l'échantillon issu des trois bases de données et du SNSP. Il donne, année par année, toutes les informations relatives à un individu, qui sont transformées en « reports au compte ». La CNAV ne dispose d'aucune structure dédiée à la fourniture d'échantillons ou de fichiers mais peut réaliser certaines opérations de ce type, au coup par coup. Elle a notamment répondu à une demande de l'IRDES, deux ans auparavant, pour étudier l'influence de la carrière de l'individu sur les IJ maladie et a procédé à l'appariement des données de la CNAM à celles de la CNAV. Une enquête sur les choix de départ a également été réalisée en 2005. Vincent Poubelle précise cependant que la CNAV ne s'est pas encore dotée d'une politique en la matière et qu'elle doit, du fait des informations essentielles dont elle dispose, respecter certaines règles. Il indique que la direction juridique révisait actuellement les conventions d'échanges de données avec les systèmes partenaires, pour tenir compte d'une faiblesse juridique pointée par la Cour des comptes.

Marcel Goldberg s'interroge sur d'éventuels problèmes en termes de confidentialité ou de propriété des données.

Monique Beurey-Leconte signale qu'une réflexion est en cours sur la notion de données publiques et ses influences sur les questions de propriété et de rémunération des échanges.

Vincent Poubelle indique que son principal client, en matière statistique, reste le Ministère des affaires sociales mais que la CNAV n'a pas l'habitude, hormis des obligations légales envers certains organismes, de fournir des échantillons de manière régulière. Il précise par ailleurs qu'au niveau technique, certaines pratiques peuvent être standardisées.

Marcel Goldberg observe que la CNAM offre aux statisticiens extérieurs un accès direct à son échantillon généralisé des bénéficiaires, extrait au centième du SNIRAM, et organise des formations pour faciliter cet accès.

Vincent Poubelle remarque que le contenu de l'échantillon de la CNAV s'avère complexe à comprendre et que son accessibilité se heurte à des difficultés pratiques tenant aux moyens informatiques et juridiques de l'institution. La CNAV ne dispose en effet d'aucun service de diffusion et procède au coup par coup. Quant aux services juridiques, ils doivent aujourd'hui trouver une réponse aux carences relevées par la Cour des comptes.

Eric Jouglu s'enquiert de la politique financière en matière de fourniture d'informations.

Vincent Poubelle explique que l'institution reste pour l'instant sur le principe de la gratuité mais pourrait, pour certaines extractions, demander une participation. En matière d'appariement, la CNAV préconise le critère du NIR auquel s'ajoutent le nom, le prénom et la date de naissance et s'attache à respecter les orientations de la CNIL. Il note cependant que l'existence du NIR peut poser quelques difficultés en matière de santé.

Marcel Goldberg remarque que dans le secteur de la santé, l'appariement peut concerner soit des bases de données existantes, soit des sujets pour lesquels le NIR existe à un certain endroit et peut être retrouvé.

Chantal Cases ajoute qu'une troisième utilisation consiste dans la réalisation d'enquêtes sur un échantillon de la CNAV, qui ne nécessite pas de NIR. Elle observe que le problème des systèmes d'informations de santé réside dans l'absence de données sociales, en particulier sur les carrières professionnelles et que l'analyse des inégalités de santé représente l'une des carences des indicateurs de la loi de santé publique.

Marcel Goldberg rappelle que la future loi de santé publique devrait mettre en exergue la question des inégalités de santé et qu'il conviendra de trouver une manière d'enrichir les bases de données de santé.

---

Annette Leclerc demande si le bénéficiaire de l'allocation adulte handicapé qui n'aurait jamais travaillé peut tout de même entrer dans le fichier des carrières.

Vincent Poubelle répond qu'il figure dans le fichier des identités mais non dans celui des carrières. Il est cependant possible de déterminer qu'il n'a jamais cotisé. Un appariement est possible avec la CNAF, qui alimente les données concernant l'assurance retraite des parents au foyer.

Chantal Cases note que malgré certaines limites, les fichiers de la CNAV s'avèrent extrêmement riches.

Marcel Goldberg constate que les sources de données de santé ne disposent d'aucune information en matière de carrière et que la CNAV constitue donc la source essentielle en la matière. Il remarque que la CNAV n'a pas mis en place de politique sur les appariements.

Vincent Poubelle admet que l'institution étudie toutes les demandes.

Marcel Goldberg suggère que le groupe de travail émette une recommandation forte sur l'accessibilité des données de la CNAV. Il s'interroge sur l'éventualité d'organiser un service dédié pour faire face à une croissance de la demande.

Chantal Cases ajoute qu'avait été évoquée la possibilité de recommander la mention, dans les conventions d'objectifs, de la nécessité d'une diffusion des données, de manière à justifier l'octroi de moyens spécifiques.

Vincent Poubelle répond que la convention 2009-2012 a déjà été signée.

Eric Jouglà souligne qu'une réflexion pourrait être menée sur la mutualisation, une opération pouvant être utilisée par plusieurs institutions.

Marcel Goldberg signale également que l'assurance maladie s'oriente actuellement vers la création d'une structure d'interface qui permette de désengorger les services internes et facilite l'accès aux données complexes.

Alain Weill insiste sur la complexité du contenu. Les utilisateurs reçoivent l'intégralité du SNIRAM et sollicitent fortement la CNAM pour des conseils.

Vincent Poubelle rappelle que la diffusion externe des données n'entre pas dans les missions de la CNAV. Quant aux données localisées, la CNAV dispose d'un fichier contenant l'adresse des retraités. Le recensement des non retraités s'avère plus difficile mais le système du droit à l'information apporte des améliorations.

Marcel Goldberg s'enquiert des relations internationales en termes de classification et nomenclature.

Vincent Poubelle indique que la CNAV correspond avec les systèmes de retraite des autres pays mais uniquement à des fins de gestion.

Gérard Badéyan demande si ces échanges représentent des flux importants.

Vincent Poubelle explique qu'il n'existe aucun flux automatisé. C'est le demandeur qui fournit lui-même les éléments.

Alain Trugeon demande si les personnes décédées figurent dans l'échantillon au vingtième et si une limite a été fixée au volume du fichier des identités.

Vincent Poubelle explique qu'aucune suppression n'est prévue pour l'instant sur le fichier des identités mais que l'échantillon ne comprend que les personnes dont la CNAV n'a pas trace du décès, auxquelles sont accrochés les éléments de carrière disponibles.

## Jean-Yves ROBIN, ASIP Santé, et André LOTH, MISS

André Loth indique qu'il est aujourd'hui responsable de la Mission pour l'informatisation du système de santé (MISS), rattachée au Secrétaire général, qui devrait devenir une délégation aux systèmes d'information chargée d'animer le futur Conseil national des systèmes d'information dont les contours ont été dessinés par le rapport Gagneux de mai 2009. Il a également été responsable du PMSI au sein de la Direction des hôpitaux et a accompagné la création d'un pôle d'expertise sur les nomenclatures de santé et la première base de données du PMSI puis a dirigé le projet SESAM Vitale et contribué à la mise en place du codage des médicaments.

Jean-Yves Robin explique que l'ASIP s'inscrit elle aussi dans le domaine des données de santé, dont le pilotage opérationnel a été réformé par les pouvoirs publics. Les missions de maîtrise d'ouvrage existantes du CPS et du DMP seront en effet réunies au sein d'une agence dont le périmètre se trouvera élargi aux secteurs sanitaire et médico-social et qui pourra également assister l'INVS dans son travail de veille, tout en construisant des référentiels et en contribuant à la progression de l'interopérabilité des systèmes d'informations. Le Ministère, l'Assurance maladie et la CNSA constituent les trois administrateurs de cette agence.

Marcel Goldberg suggère que les données du PMSI sont facilement accessibles.

André Loth reconnaît que les difficultés, liées au problème d'identification par référence, ont été résolues. Lors de la mise en place du format du Résumé de sortie anonyme (RSA) dans les années 1990, la CNIL avait admis sa diffusion libre mais le PMSI devenant exhaustif, l'identification des personnes s'est avérée possible. C'est pourquoi des restrictions à l'accès aux bases de données exhaustives avaient été posées.

Alain Weill remarque qu'au sein même du SNIRAM, il est possible d'obtenir des données qui ne devraient pas être accessibles.

André Loth admet que ce problème existe.

Marcel Goldberg note que l'accessibilité au PMSI aux fins d'études est admise.

Chantal Cases rappelle que le coût d'accès s'avérait non négligeable mais qu'il existe aujourd'hui une tarification spécifique.

Lucile Olier ajoute que le coût a effectivement fortement baissé.

André Loth regrette le faible nombre de chercheurs qui étudient les questions de tarification des soins et de systèmes de gestion en France.

Chantal Cases observe que le PMSI est accessible mais que l'étude de coût l'est beaucoup moins, les établissements s'étant prononcé contre la disponibilité des fichiers de tarification.

André Loth rappelle que l'accessibilité tient souvent à une insuffisance des moyens des établissements et que le rapport Gagneux se prononce en faveur d'une amélioration sur ce point.

Marcel Goldberg reconnaît que toutes les institutions auditionnées font le même constat. Leur mission ne consistant pas à fournir des informations à l'extérieur, les moyens internes dédiés à cette activité sont extrêmement faibles. Il rappelle que le groupe de travail pourrait recommander que l'Etat veille à ce que les COG des institutions comprennent une obligation de diffusion et les moyens que celle-ci implique.

André Loth remarque que cette mise à disposition peut s'avérer très consommatrice de ressources.

Marcel Goldberg admet qu'une réflexion doit être menée sur l'organisation de cette activité. Il sollicite l'avis de Monsieur ROBIN, les textes régissant le DMP interdisant formellement l'accès aux données.

Jean-Yves Robin considère que l'étape de production et de stockage des données représente un passage obligé et, s'il ignore les finalités qui seront attribuées au DMP par le législateur dans le futur, qu'une évolution est possible. Les pouvoirs publics et les utilisateurs du système doivent éprouver de la confiance envers ce nouveau dispositif, c'est pourquoi ses finalités doivent être clairement définies. Lorsqu'une autorisation est recueillie pour une certaine finalité, il s'avère toujours délicat d'utiliser la donnée dans un autre cadre. L'intention du législateur, dans le cadre du DMP, visait d'abord à créer de la confiance. Le DMP ne doit pas cependant être conçu comme un entrepôt de données mais comme un projet doté d'un effet structurant sur la production des données en général. Il fait en sorte que les acteurs produisent plus de données en introduisant une exigence d'interopérabilité, créant, de ce fait, l'opportunité de disposer d'informations plus aisément. Jean-Yves Robin insiste sur l'importance, dans la production de données, de l'usage des nomenclatures internationales, afin de structurer ces données. Il estime que la T2A, au contraire, introduit une nomenclature à vocation de gestion, qui se révèle complexe à corrélérer avec des nomenclatures standards de description de l'activité médicale.

André Loth observe que la CIM et la CCAM ne représentent pas des nomenclatures à visée économique de gestion mais s'attachent à décrire l'activité médicale. Il rappelle ainsi que la CCAM a été introduite sur la base des recommandations du comité européen de normalisation sur la manière de décrire les actes médicaux et qu'elle n'avait aucune visée tarifaire.

Marcel Goldberg pense que s'il s'impose, le DMP aura forcément un effet d'entraînement sur une production de données plus structurée mais qu'à court terme, il ne peut servir d'appui aux recommandations pour la future loi. Par ailleurs, pour certaines spécialités comme l'épidémiologie, le DMP n'aura vraisemblablement pas d'autre utilité que permettre la validation d'un diagnostic.

Jean-Yves Robin reconnaît qu'à moyen-long terme, le DMP n'a pas d'utilité directe pour certaines disciplines mais peut s'avérer utile, par exemple, pour l'alimentation des registres des cancers.

Marcel Goldberg approuve. Il observe cependant que l'émergence du DMP s'est couplée à celle de l'identifiant national de santé, qui fait craindre l'interdiction de tout appariement des données.

André Loth rappelle que la plupart des acteurs du dispositif sont très sensibles à l'utilité des données dans le domaine de la santé publique. Par ailleurs, si la loi Kouchner de mars 2002 a conditionné l'utilisation des données nominatives destinées aux soins à l'accord des personnes intéressées, il n'en est pas de même pour le PMSI ou le SNIRAM, qui contiennent des données anonymes ou à finalité économique. Des dispositifs techniques peuvent prévoir par la suite l'usage de données, issues du DMP et des dossiers hospitaliers, rendues anonymes, pour la recherche, mais il convient d'attendre que la confiance s'installe.

Jean-Yves Robin précise que dans les bases de l'assurance maladie, la personne est identifiée par le NIR. L'INS est irréversible et produit à partir du NIR de naissance. Selon les réflexions en cours, la CNAV pourrait être chargée de la production de l'INS à partir de sa base NIR et détiendrait donc la base d'équivalence. Elle jouerait ainsi le rôle de tiers de confiance en permettant la réalisation d'appariements de données.

Marcel Goldberg admet que le groupe craignait que le passage à l'INS soit irréversible et qu'un appariement entre les données de la CNAM et de la CNAV devienne impossible.

Jean-Yves Robin indique que la situation ne sera pas irréversible.

André Loth reconnaît que l'idée d'une substitution de l'INS au NIR avait été émise par la commission *ad hoc*, compte tenu du fait que le caractère signifiant du NIR était considéré comme un défaut, mais qu'elle est aujourd'hui étudiée avec prudence.

Jean-Yves Robin ajoute que le NIR est entaché de défauts et la perspective de le remplacer, à plus ou moins longue échéance, par un numéro unique non signifiant doit être étudiée. Actuellement, l'introduction de l'INS pour la sphère santé des données de la personne contraint le système d'appariement de données sans toutefois le rendre impossible.

Chantal Cases souligne qu'autour de la question de l'information utilisée en aval du DMP se pose la problématique d'un important manque d'informations dans le système, quant aux diagnostics posés ou aux pathologies en médecine de ville. Il n'existe que des solutions ponctuelles, à travers notamment le codage des médicaments ou la disponibilité de fichiers privés. La question de fond reste la possibilité d'identifier de la morbidité traitée en ville.

Jean-Yves Robin explique que le DMP, en imposant certains référentiels et l'ouverture des bases de données de production à l'ensemble des acteurs de soins, offre la possibilité d'extraire des informations de santé publique. Se pose cependant la question des missions et de la rémunération des médecins dans ce cadre, comme l'a montré l'exemple du compte-rendu d'anatomo-pathologie.

Marcel Goldberg observe que la structuration de l'information non seulement constitue une mission complémentaire des médecins mais contribue surtout à l'amélioration de la qualité des soins.

Paul Avillach s'enquiert de l'infrastructure mise en place dans le cadre de la surveillance épidémiologique de la grippe.

Jean-Yves Robin explique que les données saisies dans l'ensemble des centres 15 de France sont structurées afin d'être consolidées autour d'un référentiel national. L'infrastructure reste légère. Les dispositifs de saisie de données dans les centres 15 ont été sécurisés. Leur applicatif a été modifié mais une réflexion devrait être engagée par la suite sur leur standardisation, comme pour les centres de dépistage du cancer.

Marcel Goldberg note que le rapport Gagneux émet plusieurs propositions en termes de gouvernance générale des systèmes d'information de santé, dans l'optique de faciliter, pour les pouvoirs publics et le monde de la recherche, un accès aux données, qui se négocie encore aujourd'hui au cas par cas.

André Loth précise que le chapitre du rapport relatif à la gouvernance des données de santé est issu du constat ancien de carence en matière d'expertise et de culture universitaire sur les sujets de gestion des données, à des fins de pilotage comme à des fins épidémiologiques. La tarification et le suivi de l'activité à travers les procédures de financement montrent un fractionnement des compétences entre l'Etat, l'Assurance maladie ainsi que la DRESS, la CNSA ou l'IRDES. De ce constat a émergé l'idée que le regroupement des expertises faciliterait l'accès aux données pour les chercheurs, accès qualifié d'essentiel par le rapport. Cette mise en synergie des forces, qui se heurte à la glaciation des relations entre le Ministère et l'Assurance maladie quant à la régulation du système, s'opèrera à l'issue d'un processus assez long et qui ne se limitera pas à la création d'un gendarme qui décidera de l'octroi d'un accès. Le rapport indique également qu'il conviendrait de regrouper les organismes de concertation, dont l'articulation n'est pas aisément identifiable pour un œil extérieur.

Chantal Cases regrette que le rapport s'intéresse davantage à la gouvernance qu'à l'analyse des besoins et des finalités du système d'information.

Marcel Goldberg approuve le constat d'absence actuellement de lisibilité et de gouvernance et suggère que le groupe de travail s'attache à étudier ces besoins.

Jean-Yves Robin propose trois stratégies quant à l'accès aux données : la centralisation, la mutualisation et la standardisation à la source.

Marcel Goldberg estime que les recommandations du groupe de travail se révéleront sans doute moins techniques mais préconiserons la participation du HCSP aux hautes instances qui pourraient se mettre en place.

Chantal Cases craint que l'absence de prise en compte des besoins n'engendre la constitution de structures incomplètes.

Marcel Goldberg précise que les missions du HCSP en matière d'évaluation devront être prises en compte au sein de ces instances, où le monde de la recherche devra être également représenté.

Jean-Yves Robin signale que les avis du HCSP seront pris en compte dans la comitologie de l'ASIP et qu'une interface entre les deux organismes devra être trouvée.

Marcel Goldberg ne remet pas en cause l'esprit général du rapport mais insiste sur l'importance de veiller à tous ces aspects dans le cadre de sa mise en œuvre.

---

## Jean-Pierre Le Gleau et Alain Trognon, INSEE

Jean-Pierre Le Gleau indique que l'INSEE produit des statistiques, diffuse des études et assure la coordination du service statistique public, qui est composé de l'INSEE et de 14 services statistiques ministériels.

La politique générale de diffusion des données agrégées a été définie il y a cinq ou six ans. Ces informations sont gratuites et accessibles en ligne. Si des personnes souhaitent disposer de données qui ne sont pas disponibles sur Internet, elles peuvent en faire la demande. Une prestation « sur mesure », tarifée, est alors envisageable.

Les données agrégées sont construites à partir de données individuelles. Ces dernières sont protégées par le secret statistique. Dans ce domaine, le texte de référence est la loi 51-711 du 7 juin 1951. Son article 6 précise que les renseignements individuels ne peuvent faire l'objet d'aucune communication de la part du service dépositaire, en l'occurrence l'INSEE. Il existe néanmoins quelques exceptions. Ainsi, les données concernant les entreprises ne sont plus couvertes par le secret après 25 ans. Le délai est désormais de 75 ans pour les personnes physiques. Jusqu'à la loi sur les archives du 15 juillet 2008, il était de 100 ans. Le secret concernant des données individuelles peut également être levé après avis du comité du secret statistique, accord du service dont émanent les données et décision favorable de la Direction des Archives de France.

Entre les données individuelles et les données agrégées, il existe une autre catégorie. Il s'agit de fichiers individuels non nominatifs. L'INSEE en a rendu accessible un certain nombre sur son site Internet. Ils ne contiennent aucune variable permettant l'identification des personnes. S'agissant de la localisation géographique par exemple, ne figure que la région.

En réponse à Marcel Goldberg, Jean-Pierre Le Gleau précise que le principal fichier individuel non nominatif disponible en ligne est l'Enquête Emploi. Il était initialement prévu de développer ce type d'informations mais cette démarche a été freinée, des craintes subsistant quant aux possibilités d'identification des individus. Une mission de l'Inspection générale de l'INSEE doit être réalisée à ce sujet. Le principal risque concerne les couples interrogés de manière séparée. L'un reconnaissant ses réponses pourrait, par déduction, connaître celles de l'autre.

Un certain nombre de chercheurs ont estimé que ces fichiers individuels non nominatifs étaient insuffisants. Des fichiers de production et de recherche peuvent donc être réalisés. Ils sont plus détaillés, sans pour autant permettre l'identification des individus. L'INSEE estimant ne pas avoir la compétence pour reconnaître leur qualité de chercheur, les informations sont mises à la disposition des personnes effectuant une demande par l'intermédiaire du Réseau Quetelet. Evidemment, celles-ci ne sont pas autorisées à diffuser les données auxquelles elles ont pu avoir accès.

Marcel Goldberg demande si les chercheurs qui déposent une demande d'accès à des fichiers de production et de recherche doivent le faire dans un cadre formalisé. Pour en avoir fait l'expérience récemment, Eric Jouglu confirme qu'il est nécessaire de remplir un dossier.

Pour les enquêtes auprès des ménages, Jean-Pierre Le Gleau indique que les fichiers de production et de recherche contenaient, jusqu'en 2008, les informations les plus détaillées qu'il était possible d'obtenir. Des évolutions ont été introduites par la loi sur les archives, qui a modifié certaines dispositions de la loi de 1951. Le délai couvert par le secret a été ramené de 100 à 75 ans. Par ailleurs, une procédure dérogatoire, nécessitant de recueillir l'avis d'un comité du secret statistique a été créée.

Le décret d'application ayant tardé, le comité du secret statistique n'a été mis en place qu'il y a un mois. Présidé par un conseiller d'Etat, désigné par le vice-président du Conseil d'Etat, il devrait se réunir pour la première fois au mois d'octobre. La fréquence des séances devrait être d'une ou deux fois par trimestre.

Marcel Goldberg demande s'il est possible, en utilisant cette nouvelle procédure, d'avoir accès à des données personnelles. Jean-Pierre Le Gleau le confirme.

Dans la mesure où il s'agit d'une ouverture importante, Jean-Pierre Le Gleau pense que le comité du secret statistique connaîtra probablement un afflux massif de demandes au cours des prochains mois.

Les chercheurs qui estimeront que les fichiers de production et de recherche – dont l'utilisation doit néanmoins continuer à être privilégiée – ne sont pas suffisants devront déposer une demande auprès du comité du secret statistique. Ils devront remplir un dossier substantiel, précisant notamment l'objet de la recherche, ainsi que les raisons pour lesquelles les informations souhaitées sont nécessaires. Ils devront en outre apporter des garanties personnelles. Si l'avis rendu par le comité du secret statistique est favorable et que l'INSEE et la Direction des Archives de France donnent leur accord, ils auront accès aux données mais pour une durée limitée et dans le cadre d'une procédure sécurisée.

Les fichiers mis à disposition ne seront pas directement nominatifs. La loi l'autoriserait mais l'INSEE a choisi de ne pas le faire. Il sera néanmoins possible de retrouver, sans trop de difficulté, les individus.

Alain Trognon explique qu'il existe plusieurs systèmes permettant d'accéder à des données confidentielles de manière sécurisée. Certains pays, comme les Etats-Unis ou le Canada, ont choisi d'installer des sortes de « bunkers » dans les universités. D'autres, comme le Danemark ou la Hollande, ont privilégié l'accès informatique à distance. La France a retenu ce second modèle, moins coûteux et offrant une meilleure égalité entre les chercheurs.

Une expérimentation, portant sur 11 recherches, est en cours depuis un an. Des terminaux, permettant de se connecter à un serveur de l'INSEE sur lequel sont stockées les données souhaitées, ont été installés dans les laboratoires de recherche concernés. Les informations ne peuvent pas être téléchargées et seules les personnes autorisées – qui seront identifiées grâce à leurs empreintes digitales – peuvent les consulter.

Les chercheurs étrangers pourront s'inscrire dans ce dispositif, à condition que leur pays d'origine entretienne des relations avec la France. L'objectif est en effet qu'ils puissent faire l'objet de poursuites en cas de rupture du secret.

Les terminaux permettent d'effectuer des calculs directement sur le serveur, les résultats étant ensuite logés dans un espace dédié. Ceux-ci sont vérifiés par un statisticien de l'INSEE, puis mis à la disposition des chercheurs. Mais aucune opération n'est réalisée sur les systèmes informatiques des centres de recherche. Il s'agit d'une spécificité du système français.

Alain Trognon reconnaît que la communauté des chercheurs a des difficultés à comprendre la rigueur qui entoure le dispositif. Avant de donner accès aux terminaux, une séance d'explication est donc organisée. A cette occasion, une charte est d'ailleurs signée.

Il est possible de réaliser des appariements. Alain Trognon précise, en réponse à Marcel Goldberg, que ces opérations sont effectuées par l'INSEE à la demande des chercheurs.

Chantal Cases demande s'il est possible que l'INSEE n'ait pas le droit d'accéder aux données du second fichier nécessaire pour la réalisation de l'appariement. Eric Jouglu signale que ce cas de figure existe, notamment pour l'EDP et les causes de décès. Alain Trognon indique qu'il est possible, pour cette situation particulière, de faire appel à un tiers de confiance.

Alain Trognon rappelle que le dispositif est, pour le moment, en phase d'expérimentation. Dans un premier temps, Jean-Pierre Le Gleau estime qu'il pourra traiter une vingtaine de demandes en parallèle.

Marcel Goldberg note que le système mis en place augmente sensiblement la charge de travail des équipes de l'INSEE. Alain Trognon le reconnaît. Un renforcement des moyens sera nécessaire, l'objectif étant, à terme, de pouvoir traiter une centaine de projets de recherche de manière concomitante.

---

Marcel Goldberg souhaite avoir des précisions sur le périmètre d'intervention du comité du secret statistique. Jean-Pierre Le Gleau indique que celui-ci couvre l'ensemble de la statistique publique.

Marcel Goldberg s'interroge sur l'articulation de la loi Informatique et Libertés avec les dispositions sur la protection sur le secret statistique. Jean-Pierre Le Gleau explique que les contraintes issues de ces différents textes, dont la philosophie est assez proche mais pas tout à fait identique, s'ajoutent. Ainsi, la CNIL a posé des limites à la communication des résultats du recensement car certaines informations auraient permis de cibler des populations. Pourtant, le secret statistique n'était pas remis en cause. De la même façon, le comité du secret statistique peut s'opposer à la diffusion de certaines informations, alors que la CNIL les auraient autorisées car ne portant pas atteinte à la vie privée. Marcel Goldberg en conclut qu'il existe, de chaque côté, une sorte de droit de veto.

Jean-Pierre Le Gleau rappelle que tous les traitements intégrant des données à caractère personnel doivent être déclarés à la CNIL. L'INSEE le fait systématiquement, ce qui pourrait en dispenser les chercheurs utilisant ensuite les fichiers. Des discussions sont en cours à ce sujet. Si le principe d'une dispense n'était pas retenu, la procédure serait au moins simplifiée. Marcel Goldberg estime qu'il s'agit d'un point important compte tenu des délais de réponse de la CNIL. Jean-Pierre Le Gleau souligne néanmoins que les appariements devront continuer à faire l'objet d'une demande d'autorisation.

Eric Jouglu note qu'un certain nombre de demandes d'appariements seront probablement identiques. Une fois que le fichier aura été réalisé, il demande s'il sera possible d'en ouvrir l'accès. Alain Trognon confirme qu'il faudra aller dans ce sens, tout en s'assurant à chaque fois que les chercheurs concernés remplissent les conditions pour consulter ces informations.

Les ressources de l'INSEE étant limitées, Chantal Cases estime qu'il serait souhaitable d'identifier quelques thèmes prioritaires. Elle souligne en particulier l'urgence de disposer d'indicateurs sur les inégalités sociales de santé. Marcel Goldberg confirme qu'il existe une très grande attente dans ce domaine. Chantal Cases signale d'ailleurs qu'il a été envisagé de préconiser, dans la future loi de santé publique, la réalisation d'un fichier de mortalité apparié.

Jean-Pierre Le Gleau reconnaît que le nombre de demandes sera probablement supérieur aux capacités de traitement de l'INSEE. Il sera donc indispensable de définir des priorités. Chantal Cases estime que la loi de santé publique pourrait peut-être en déterminer quelques-unes. Pour sa part, Jean-Pierre Le Gleau préconise que leur fixation soit confiée au comité du secret statistique. Marcel Goldberg se demande si ce dernier est réellement à même de jouer ce rôle. Au regard de sa composition, Jean-Pierre Le Gleau le pense.

Marcel Goldberg craint que le dispositif mis en place ne comporte des limites dans la réalisation des appariements, notamment pour des suivis de très longue durée, avec des données recueillies à de nombreuses sources. Il paraît difficile d'imaginer que ces informations puissent être traitées par l'INSEE.

Jean-Pierre Le Gleau reconnaît que le dispositif ne permettra probablement pas de traiter tous les cas de figure. Son existence constitue néanmoins une avancée majeure par rapport à la situation passée.

Marcel Goldberg confirme que les évolutions en cours sont positives et très significatives.

Chantal Cases souligne qu'il serait intéressant de disposer de données localisées, au moins dans le cadre des fichiers de production et de recherche. La mention du département n'est pas pertinente. Chantal Cases demande si des réflexions sont en cours sur une codification en fonction des types de zonage. Ces caractéristiques ne seraient pas identifiantes. Jean-Pierre Le Gleau indique que les représentants du Réseau Quetelet ont déjà fait remonter cette attente. Il est donc envisagé de proposer un maillage plus fin.

Jean-Pierre Le Gleau signale qu'un seul fichier de production et de recherche existe par enquête. S'il en était autrement, des rapprochements entre les données pourraient être effectués sans aucun contrôle.



## Yvon Merlière et Alain Weill, CNAM

Yvon Merlière indique que la CNAM dispose de bases de données individuelles concernant les remboursements et les pathologies, en particulier dans le cadre des ALD (Affections de Longue Durée). Il existe en outre un chaînage avec les données du PMSI.

Les accès aux données en clair du patient sont très limités. Ces informations ne sont pas chaînées avec le PMSI. Seules les données anonymisées le sont. La CNIL a néanmoins fixé des règles quant à leur utilisation. A l'exception des médecins conseils de la CNAM, personne n'est autorisé à croiser le code d'une commune de résidence, un mois de naissance et une date de soins.

Des bases de données agrégées ont été constituées. Depuis peu, il existe en outre un échantillon général de bénéficiaires, qui apporte une exhaustivité des informations sur 500 000 personnes couvertes par le régime général. Il devrait être étendu dès 2010 au secteur agricole et aux travailleurs indépendants, ce qui porterait l'échantillon à 600 000 personnes.

L'accès aux données est règlementé par la loi du 2 décembre 1998. Il s'agit de l'article L. 161-28-1 du Code de la Sécurité sociale. Un arrêté ministériel a défini les organismes utilisateurs de la sphère publique. Ils sont nommément cités, y compris les organismes de recherche. Pour chacun d'entre eux est désigné une autorité compétente, responsable pénalement en cas de litige. Les habilitations sont accordées après une formation préalable.

D'un point de vue technique, l'accès aux données se fait via Internet depuis n'importe quel poste de travail. Il est néanmoins nécessaire de disposer d'une ligne sécurisée. Pour l'avenir, la mise en place d'une connexion « https » est à l'étude. Elle devrait être mise en œuvre dès la fin de l'année. Cette solution sera à la fois plus simple et moins coûteuse.

Les travaux menés à partir des données de la CNAM peuvent avoir les trois finalités suivantes :

- la connaissance des dépenses de l'Assurance Maladie ;
- la connaissance des prescriptions faites par les professionnels de santé ;
- la construction et le suivi des indicateurs de santé publique.

Alain Weill note que le système a également vocation à transmettre un certain nombre de données aux professionnels de santé et au fisc.

Yvon Merlière indique que l'EGB existe depuis 2003. L'historique des données exhaustives est théoriquement récupérable pendant un maximum de trois ans (deux ans et l'année en cours). Or des demandes de chercheurs nécessiteraient d'avoir des délais plus longs. Une demande va être faite en ce sens. Depuis deux ans, la CNAM signale en effet à la CNIL qu'elle n'a pas détruit les données car des enquêtes doivent être réalisées.

La CNAM reçoit un certain nombre de demandes d'enquête *ad hoc*. L'EGB ne répond pas à tous les besoins, notamment en ce qui concerne les événements rares.

Toutes ces demandes spécifiques doivent recueillir l'avis de l'IDS. Ce dernier a donc mis en place un comité d'experts et élaboré une procédure d'évaluation et de hiérarchisation des dossiers. Les extractions doivent en effet être réalisées par la CNAM, qui ne dispose que d'une équipe de trois personnes pour travailler sur ce sujet. Cette situation constitue un fort goulot d'étranglement. Il n'est pas possible de

---

traiter plus de quatre à cinq fichiers par an. Il est possible d'effectuer des appariements, dès lors que les données ayant servi à l'anonymisation ont été saisies initialement.

Marcel Goldberg précise que l'IDS doit donner son autorisation pour l'accès aux données du SNIIRAM. Celui-ci est néanmoins automatique si le CNIS ou la CNIL se sont déjà prononcés par ailleurs. Il existe encore beaucoup d'incertitudes sur le volume de demandes qui pourraient être soumises au comité d'experts. Ce dernier a d'ailleurs manifesté des réticences quant au rôle qu'il devra jouer. Un certain nombre d'ajustements restent nécessaires, notamment en ce qui concerne la définition des priorités.

Yvon Merlière signale que l'IDS a donné un premier accord à un chercheur de l'université de Lyon sur l'utilisation des statines. Marcel Goldberg note que ce dossier a été problématique et n'a fait que renforcer les réticences du comité d'experts. En effet, il n'a pas été possible de se prononcer en fonction de sa qualité scientifique. Selon les règles établies par l'IDS, le refus ne pouvait en effet s'appuyer que sur des raisons déontologiques ou juridiques.

Yvon Merlière souligne que le SNIIRAM est fait pour la sphère publique. Industrie pharmaceutique et les assurances complémentaires seraient très intéressées. La CNAM n'est pas favorable à son ouverture aux premiers, qui auraient ainsi une connaissance complète du marché du médicament. Or beaucoup de recherches sont aujourd'hui financées par des fonds privés. Chantal Cases estime que ces liens n'empêchent pas forcément l'indépendance.

Yvon Merlière signale qu'une charte d'utilisation interdit la diffusion des données à des tiers. Une clause concernant l'obligation de les détruire à l'issue d'un certain délai va être ajoutée. Le système est cependant beaucoup plus souple que celui mis en place par l'INSEE. Marcel Goldberg note que cette solution reste néanmoins conforme à la loi.

Yvon Merlière indique que les données en clair sont utilisées pour des appariements. Il est alors fait appel à un tiers de confiance, en l'occurrence l'INSEE. La procédure a été développée avec la CNIL.

Alain Weill note que l'autorisation du recueil du NIR constitue un vrai enjeu pour le développement des appariements. Marcel Goldberg confirme que la situation actuelle est extrêmement bloquante. Les seules exceptions supposent un décret en Conseil d'Etat. Il serait nécessaire de modifier les textes législatifs pour ouvrir la procédure de manière très encadrée. Plusieurs réunions ont été organisées à ce sujet avec la CNIL. Cette dernière n'y serait a priori pas défavorable. Chantal Cases craint toutefois qu'il n'existe une forte opposition de la part des usagers.

Marcel Goldberg souhaite avoir des précisions sur le volume de demandes – potentiellement recevables – transmises à la CNAM et sur leur nature.

Alain Weill rappelle que la diffusion de données ne fait pas partie des missions de la CNAM, à la différence de l'INSEE. Il s'agit d'un problème de fond dans l'affectation des ressources.

Alain Weill note par ailleurs qu'un certain nombre de demandes pourraient être traitées directement par une analyse plus approfondie de l'EGB. Il existe un problème d'appropriation de ces informations par un certain nombre d'organismes.

## Pierre Arwidson de l'INPES

**Pierre Arwidson présente l'INPES, créé à partir d'une association, le CFES, créée en 1952. Pour accomplir ses missions, cette association avait l'habitude de réaliser des enquêtes. Après la première enquête CFES 1, il a été décidé de faire une seule enquête multithématique. La première enquête Baromètre santé a été réalisée en 1992 ; elle a notamment été financée par l'assurance-maladie.**

Certaines enquêtes sont multithématiques, d'autres s'adressent aux professionnels de santé. Depuis 2000, une alliance avec l'Observatoire français de toxicomanie permet de réaliser des enquêtes de plus grande échelle. Les enquêtes multithématiques sont désormais réalisées tous les cinq ans et la prochaine sera administrée en 2010. L'INPES réalise d'autres enquêtes : une campagne nutritionnelle, un Baromètre Cancer, une enquête sur la prévalence de la dépression ou encore un Baromètre Santé Environnement.

Marcel Goldberg souhaite savoir qui décide de réaliser telle ou telle enquête.

Pierre Arwidson indique que le Baromètre santé est de l'initiative de l'INPES alors que les autres enquêtes ont été commandées par les autorités sanitaires.

Gérard Badéyan remarque qu'il existe également une enquête sur le tabac.

Pierre Arwidson précise qu'une cohorte de fumeurs a été recrutée au niveau international, dans le cadre d'un partenariat avec une université canadienne.

Marcel Goldberg souhaite savoir si ces enquêtes sont discutées dans le cadre d'un comité scientifique et s'enquiert de la gouvernance de l'INPES.

Pierre Arwidson indique que tous ces sujets sont présentés en conseil d'administration et en conseil scientifique. En 2002, Jean-Paul Moatti, président du conseil scientifique, a cherché à instaurer des enquêtes continues, avec une recommandation méthodologique. En revanche, le choix des thèmes est plutôt lié aux demandes, sans être validé par le conseil scientifique. L'INPES ne vient pas de la statistique publique même s'il est de plus en plus sollicité par la DREES pour participer au suivi d'indicateurs.

Alain Weill s'enquiert du choix des thèmes d'intervention.

Pierre Arwidson répond que l'INPES travaille sur onze programmes validés par la tutelle. Le comité des programmes réunit l'ensemble des parties prenantes de l'action de l'INPES.

Alain Trugeon ajoute que, pour chaque enquête, un comité de pilotage décide de l'ordonnement du questionnement.

Eric Jouglà rappelle que le CFES avait des organes décentralisés.

Pierre Arwidson confirme le CFES avait un réseau de comités départementaux et régionaux qui bénéficiait de financements locaux. Comme toutes les associations locales, ce réseau se transforme actuellement.

Alain Trugeon explique que les relais locaux de l'INPES pour les enquêtes sont plutôt les ORS que les Instances Régionales d'Education et de Promotion de la Santé.

A la première question relative à la politique de diffusion, Pierre ARWIDSON répond que le premier axe, à la demande du conseil scientifique de l'INPES, est d'apporter un financement aux scientifiques qui voudraient réaliser des recherches secondaires. En outre, les fichiers sont donnés gratuitement à toute équipe de recherche du public, sous réserve de la signature d'une convention sur l'usage des données.

---

Marcel Goldberg demande si les données sont rendues anonymes.

Pierre Arwidson répond que les fichiers de sondage fournis par les instituts de sondage sont totalement anonymes. Pour fonctionner, les instituts de sondage ont effectivement un accord avec la CNIL et déclarent annuellement les procédures d'anonymisation.

Eric Jouglà a déjà sollicité l'INPES car il avait besoin de données précises sur l'alcool et il les a obtenues très rapidement, sous 15 jours, le temps de signer la convention.

Pierre Arwidson indique qu'un site Internet a été créé pour permettre, après signature de la convention, de télécharger les fichiers sous quatre modalités. Il est possible pour un internaute de faire des tris à plat simples ou des tris croisés. Une base de données recense toutes les questions posées ce qui permet de générer une nouvelle enquête.

Gérard Badéyan précise qu'il existe une base de données pour une trentaine de pays.

Pierre Arwidson connaît cette source même s'il n'est pas toujours à la recherche de questions nouvelles puisque l'enquête est alors plus difficile à administrer. Le volume n'est pas très important. L'enquête Baromètre Santé s'adresse à 30 000 personnes. L'INPES a contracté des conventions d'usage de ses données avec 50 équipes depuis 2000.

Eric Jouglà demande si une personne est dédiée à la diffusion des données.

Pierre Arwidson répond par l'affirmative. L'INPES ne fait pas de publicité.

Paul Avillach indique que cette question a été posée dans toutes les auditions ; elle vise à connaître les portes d'entrée prévues.

Marcel Goldberg rappelle que le but de ce groupe de travail est de faciliter l'accès aux ressources de santé publique. Quand un chercheur souhaite accéder à des données, la procédure peut s'avérer très difficile. L'idée est de promouvoir l'usage de ces données de manière plus large.

Pierre Arwidson estime qu'il est toutefois simple de créer un onglet. Il rappelle que l'enquête relève toutefois d'un lourd processus pour les personnes qui le réalisent.

Marcel Goldberg suggère de mettre en place d'une politique d'exclusivité qui protégerait les droits des producteurs durant un certain temps. Les recommandations du groupe de travail devront intégrer les préoccupations de ceux qui produisent les données et leur souci légitime de les protéger.

Pierre Arwidson indique que l'équipe dédiée aux Baromètres comprend 11 personnes alors que l'équipe enquêtes compte 30 à 40 ETP.

Marcel Goldberg s'enquiert du budget.

Pierre Arwidson répond que le budget études s'élève à 4,5 millions d'euros pour 30 à 40 enquêtes liées notamment à l'évaluation des programmes de prévention. Le prochain Baromètre Cancer, qui portera sur un échantillon de 5 000 personnes, coûtera 300 000 euros. La partie Baromètres doit, au final, coûter moins d'un million d'euros par an, lissé sur cinq ans.

Gérard Badéyan constate qu'un administrateur de l'INSEE est détaché à l'INPES.

Marcel Goldberg constate que ces données, anonymes, sont inappareillables.

Alain Trugeon indique que les régions souhaitent souvent obtenir une extension régionale des enquêtes et des données appareillables, suivies sur la région. Un certain nombre d'indicateurs de la loi relative à la santé publique figure dans le Baromètre Santé.

Pierre Arwidson explique qu'une partie de l'enquête est administrée aux 30 000 personnes mais qu'une autre partie est uniquement adressée à un tiers des personnes interrogées.

Alain Trugeon précise que des questionnements spécifiques à une région peuvent être rajoutés en fin de questionnaire, selon les demandes des régions exprimées par le GRPS ou l'ORS. L'INPES a pris en charge la moitié de l'échantillonnage régional.

Pierre Arwidson explique que la tutelle a demandé à l'INPES d'avoir une couverture nationale et exprimé son souhait que les données soient davantage régionales.

Marcel Goldberg comprend que cette demande qui vise à connaître les disparités entre les régions.

Pierre Arwidson indique que, sur l'enquête relative à l'alcool, il est possible d'avoir des données régionales lors d'événements comme la journée d'appel.

Marcel Goldberg imagine que la prochaine loi de santé publique pourrait comprendre des indicateurs sur les inégalités, y compris géographiques.

Alain Trugeon indique qu'un travail sur la convention pluriannuelle d'objectifs cherche à mettre en évidence les inégalités sociales de santé. Cet outil peut être très performant s'il est diffusé sur des régions.

Marcel Goldberg estime que l'enquête devra permettre de mettre en évidence les différences géographiques.

Alain Trugeon relève que le Baromètre Santé Environnement vise justement à déterminer ces différences régionales. Les régions PACA et Ile-de-France parviennent à disposer de données régionales sans payer de suppléments.

Marcel Goldberg imagine que les ARS permettront de disposer de davantage de données régionales, soit par la multiplication d'enquêtes régionales, soit par des enquêtes nationales qui prennent davantage en compte les données régionales.

Eric Jouglà souhaite savoir si certaines demandes sont très régionales.

Pierre Arwidson reçoit très peu de demandes en ce sens.

Gérard Badéyan estime que les collectivités territoriales s'adressent plutôt aux ORS.

Alain Trugeon ajoute que les départements demandent constamment à disposer des données issues des enquêtes régionales. Il faut toutefois noter que les données sont alors difficiles à interpréter.

Annette Leclerc constate que le niveau local ne permet plus l'anonymisation des données.

Alain Trugeon indique que le regroupement de différentes études sur plusieurs années permet toutefois de produire certaines données à un niveau infra-territorial.

Au niveau international, Pierre Arwidson indique que les enquêtes ont été versées au niveau européen mais sans avoir de contraintes en termes de formats. La DREES est correspondante pour la France sur les modules européens.

Marcel Goldberg demande si l'INPES a des contraintes de production imposées au niveau européen.

Pierre Arwidson répond par la négative.

Gérard Badéyan indique qu'un Baromètre Santé sera bientôt diffusé au niveau européen et qu'un règlement sur les enquêtes santé imposera en 2012 des modules. Il souhaite savoir si l'INPES participe à des enquêtes européennes.

Pierre Arwidson répond par la négative. L'INPES répond toutefois aux questions qui lui sont posées, notamment par l'OMS.

Marcel Goldberg déplore que l'Europe n'ait pas de politique structurée qui impose des contraintes aux producteurs de données.

Pierre Arwidson indique que l'INPES participe financièrement à l'enquête européenne HBSC sur les jeunes âgés de 11 à 15 ans, sous l'égide de l'OMS ainsi qu'à l'enquête ESPAD sur la consommation de substances psychoactives pour les lycéens, en lien avec l'OFDT.

Sur la recherche de données extérieures, Pierre Arwidson répond que l'INPES n'a pas vraiment le temps de rechercher de telles données.

Marcel Goldberg reconnaît que cette question s'adresse plutôt aux utilisateurs. Il revient sur la procédure de la CNIL.

---

Pierre Arwidson répond que les instituts de sondage se chargent d'effectuer les formalités à l'égard de la CNIL, par des déclarations annuelles. Il est en revanche de la responsabilité de l'INPES de concevoir des enquêtes qui comprennent des questions anonymes.

Eric Jouglà s'enquiert des éventuelles contraintes rencontrées par l'INPES.

Pierre Arwidson évoque le manque de ressources humaines, surtout si les enquêtes régionales sont appelées à se développer.

Marcel Goldberg pourrait imaginer que le HCSP puisse formuler des demandes spécifiques pour intégrer des questionnements qui permettent de répondre aux indicateurs de la loi.

Pierre Arwidson explique que tel est déjà le procédé habituel : l'INPES ajoute dans le Baromètre Santé les indicateurs utiles.

Alain Trugeon indique que, dans l'optique de la nouvelle loi de santé publique, les producteurs ont été interrogés pour savoir si tel ou tel indicateur pourrait être renseigné.

Marcel Goldberg constate que ce processus est informel. Il conviendrait sans doute qu'un représentant du HCSP siège dans au sein des instances de pilotage de l'INPES pour veiller à formaliser les demandes.

Marcel Goldberg imagine de formuler des recommandations générales et ponctuelles. Le groupe de travail a été créé à la suite d'un constat de manque de données. Une des recommandations devra souligner le manque de structuration des besoins de la loi de santé publique vis-à-vis des producteurs de données. Actuellement, les demandes sont ponctuelles et précipitées. Il faudrait donc sans doute que le Haut Conseil se dote d'une structure permanente sur ce point.

Gérard Badéyan demande si l'INPES veille au contenu du questionnaire seul.

Pierre Arwidson répond par l'affirmative.

Marcel Goldberg rappelle que la loi Informatique et Libertés interdit certaines questions, notamment sur les opinions philosophiques et politiques ou les croyances religieuses.

Pierre Arwidson précise que cette interdiction ne s'applique qu'aux données à caractère personnel (ou nominatif), selon l'article 8-1 de la loi, et non aux données anonymes. Les questions peuvent être posées si elles s'avèrent pertinentes dans le questionnaire.

Annette Leclerc indique qu'il est possible d'obtenir plus de précisions sur cette question auprès de la CNIL ou du Cnis.

Gérard Badéyan constate que les questions sur la religion figuraient dans le baromètre santé de 2005.

## Glossaire

|            |  |
|------------|--|
| ASIP Santé | Agence des systèmes d'information partagés de santé                        |
| CCAM       | Classification Commune des Actes Médicaux                                  |
| CDSP       | Centre de données socio-politiques   |
| CépiDc     | Le Centre d'épidémiologie sur les causes médicales de décès                |
| CFES       | Comité français d'éducation pour la santé                                  |
| CIM        | Classification internationale des maladies                                 |
| CNAV       | Caisse nationale d'allocations vieillesse                                  |
| CNAM       | Caisse nationale d'assurance maladie                                       |
| CNIL       | Commission nationale de l'informatique et des libertés                     |
| Cnis       | Conseil national de l'information statistique                              |
| CNSA       | Caisse nationale de solidarité pour l'autonomie                            |
| COG        | Conventions d'objectifs et de gestion                                      |
| CPS        | Caisse de prévoyance sociale   |
| CRAM       | Caisse régionale d'assurance maladie                                       |
| DADS       | Déclaration automatisée des données sociales                               |
| DMP        | Dossier médical personnel  |
| DREES      | Direction de la recherche, des études, de l'évaluation et des statistiques |
| ENSAI      | Ecole nationale de la statistique et de l'analyse de l'information         |
| ESPAD      | European school survey on Alcohol and other drugs                          |
| ESPS       | Enquête Santé et Protection sociale  |
| ETP        | Equivalent temps plein   |
| FNORS      | Fédération nationale des observatoires de santé                            |
| GRPS       | Groupement Régional de Promotion de la Santé                               |
| HBSC       | Health behaviour in school-aged children                                   |
| IML        | Institut médico-légal  |
| INED       | Institut national études démographiques                                    |
| Inpes      | Institut national de prévention et d'éducation pour la santé (Inpes)       |
| IRDES      | l'Institut de recherche et documentation en économie de la santé           |
| INS        | Identifiant national de santé  |
| INSERM     | Institut national de la santé et de la recherche médicale                  |
| InVS       | Institut de veille sanitaire   |
| MISS       | Mission pour l'informatisation du système de santé                         |
| NCHS       | National center for health statistics                                      |
| NIR        | Numéro d'inscription au répertoire   |
| OFDT       | Observatoire français des drogues et des toxicomanies                      |
| OMS        | Organisation mondiale de la santé  |
| ORS        | Observatoire régional de santé   |
| PCS        | Profession et catégorie socioprofessionnelle                               |
| PMSI       | Programme de médicalisation des systèmes d'information                     |
| RNIAM      | Répertoire national inter-régimes de l'Assurance maladie                   |
| RSA        | Résumé de sortie anonyme   |
| SDDS       | Suivi des déclarants de données sociales                                   |
| SNGC       | Système national de gestion des carrières                                  |
| SNGD       | Système national de gestion des dossiers                                   |
| SNGI       | Système national de gestion des identités                                  |
| SNIIR-AM   | Système national d'information inter-régime de l'Assurance maladie         |
| SNSP       | Système national des statistiques prestataires                             |