



**COMMISSION NATIONALE DE L'INFORMATIQUE ET DES LIBERTÉS**

**RAPPORT RELATIF AUX MODALITÉS D'INFORMATISATION DE LA  
SURVEILLANCE ÉPIDÉMIOLOGIQUE DU SIDA**

**À PROPOS DE LA DÉCLARATION OBLIGATOIRE DE SÉROPOSITIVITÉ AU VIRUS DE  
L'IMMUNODÉFICIENCE HUMAINE**

*Rapport présenté par Monsieur Raymond FORNI  
adopté le 9 décembre 1999*

***Rédaction :***

- ***Sophie VULLIET-TAVERNIER,***  
***Chef du service "Libertés publiques, santé, protection sociale"***
- ***Jeanne BOSSI,***  
***Chargée de mission, responsable du secteur santé-recherche***

# Sommaire

## DÉLIBÉRATION N° 99-059 DU 9 DÉCEMBRE 1999

### RAPPEL DU CONTEXTE

#### I. LES FICHIERS ÉPIDÉMIOLOGIQUES

##### A. LES DIFFÉRENTS MODES DE RECUEILS D'INFORMATIONS

1. ***La déclaration obligatoire***
  - a. *Le principe de la déclaration obligatoire*
  - b. *Le circuit de transmission des déclarations obligatoires*
2. ***Les systèmes de recueil facultatif de données***
  - a. *Les registres de morbidité*
  - b. *Les différents types d'enquêtes*
  - c. *Les réseaux sentinelles*

##### B. DE LA NÉCESSITÉ DE DISPOSER DE DONNÉES DIRECTEMENT OU INDIRECTEMENT NOMINATIVES

1. ***Donnée nominative et anonymat***
2. ***La justification du recours à des données nominatives***
3. ***Etat des recueils d'informations existants : les pratiques constatées***

#### II. LA SURVEILLANCE ÉPIDÉMIOLOGIQUE DU SIDA

##### A. LE SYSTÈME ACTUEL DE SURVEILLANCE EN FRANCE

1. ***Le sida : maladie à déclaration obligatoire depuis le décret du 10 juin 1986***
2. ***Les autres systèmes d'information sur le sida***
  - a. *Les enquêtes spécifiques de cohorte*
  - b. *Les systèmes d'information dans les centres d'information et de soins pour l'immunodéficience humaine*
  - c. *L'enquête RESORS-VIH*
  - d. *Le réseau RENA VI*

##### B. LA NÉCESSAIRE ÉVOLUTION DES MODALITÉS DE SURVEILLANCE ÉPIDÉMIOLOGIQUE DU SIDA

1. ***Les systèmes de surveillance épidémiologiques à l'étranger***
2. ***La réflexion française***
  - a. *Les avis rendus en 1998*
  - b. *Le projet de l'Institut de veille sanitaire*

### III. LES RECOMMANDATIONS DE LA CNIL

#### A. QUELQUES CONSTATS PRÉLIMINAIRES

1. *Sur le principe même de la déclaration obligatoire*
2. *Sur la nécessaire dimension européenne du système de surveillance épidémiologique*
3. *Sur les fichiers des personnes séropositives*

#### B. LES RECOMMANDATIONS

1. *Une nécessaire clarification des objectifs de la surveillance épidémiologique*
2. *Le renforcement des mesures de confidentialité : vers l'adoption d'un système dit "d'anonymisation" à la source*
3. *La nature sensible des autres informations*
4. *L'anonymat, raison d'être des centres de dépistage anonymes et gratuits*
5. *La nécessité pour les pouvoirs publics d'autoriser par décret en Conseil d'Etat, pris après avis conforme de la CNIL, la collecte de données susceptibles de révéler les moeurs de la personne*
6. *La nécessité d'assurer une plus grande transparence dans la mise en place du dispositif*

#### ANNEXES

- Annexe 1 : Extrait de la loi du 1er juillet 1998 relative au renforcement de la veille sanitaire -
- Annexe 2 : Décret n° 99-362 du 6 mai 1999 fixant les modalités de transmission à l'autorité sanitaire de données individuelles -
- Annexe 3 : Décret n° 99-363 du 6 mai 1999 fixant la liste des maladies faisant l'objet d'une transmission obligatoire à l'autorité sanitaire -
- Annexe 4 : Compte rendus des auditions -
- Annexe 5 : Compte rendus des visites effectuées aux DDASS de Paris et du Val-de-Marne

**DÉLIBÉRATION N° 99-059 DU 9 DÉCEMBRE 1999 PORTANT ADOPTION DU RAPPORT ET DES RECOMMANDATIONS RELATIFS AUX MODALITÉS D'INFORMATISATION DE LA SURVEILLANCE ÉPIDÉMIOLOGIQUE DU SIDA ET EN PARTICULIER DE LA DÉCLARATION OBLIGATOIRE DE SÉROPOSITIVITÉ AU VIRUS DE L'IMMUNODÉFICIENCE HUMAINE**

La Commission Nationale de l'Informatique et des Libertés,

Vu la Convention n° 108 du 28 janvier 1981 du Conseil de l'Europe pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel ;

Vu la loi n° 78-17 du 6 janvier 1978 relative à l'Informatique, aux fichiers et aux libertés, ensemble le décret n° 78-774 du 17 juillet 1978 ;

Vu la directive 95/46/CE du Parlement européen et du Conseil du 24 octobre 1995 relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données ;

Vu la loi n° 98-535 du 1er juillet 1998 relative au renforcement de la veille sanitaire et du contrôle de la sécurité sanitaire des produits destinés à l'homme ;

Vu les décrets n° 99-362 et n° 99-363 du 6 mai 1999 fixant la liste des maladies faisant l'objet d'une transmission obligatoire de données individuelles à l'autorité sanitaire et les modalités de cette transmission ;

Après avoir entendu Monsieur Raymond FORNI, en son rapport, et Madame Charlotte-Marie PITRAT, Commissaire du gouvernement, en ses observations ;

Considérant que, par le décret n° 99-363 du 6 mai 1999, les pouvoirs publics ont souhaité inscrire sur la liste des maladies faisant l'objet d'une transmission obligatoire de données individuelles à l'autorité sanitaire, l'infection par le virus de l'immunodéficience humaine, quel que soit le stade de la maladie ;

Considérant que les conditions dans lesquelles seront collectées et traitées à des fins de surveillance épidémiologique les informations relatives aux personnes séropositives, nécessitent que soient conciliés les impératifs de santé publique et le respect de la vie privée ;

Considérant que, dans ce souci, et en application de l'article 1er du décret du 17 juillet 1978, la Commission, après avoir entrepris des auditions auprès des associations de défense des malades atteints du sida, de la Ligue des droits de l'Homme et du Conseil national de l'Ordre des médecins, effectué des visites auprès de médecins inspecteurs des Directions départementales de l'action sanitaire et sociale et avoir tenu étroitement informé l'Institut de Veille Sanitaire, émet un certain nombre de recommandations sur les modalités de l'informatisation de la déclaration obligatoire de séropositivité au virus de l'immunodéficience humaine ;

DÉCIDE :

- d'adopter le rapport et les recommandations relatifs aux modalités d'informatisation de la surveillance épidémiologique du sida et en particulier de la déclaration obligatoire de séropositivité au virus de l'immunodéficience humaine annexé à la présente délibération ;
- d'adresser ce rapport aux pouvoirs publics et organismes et associations concernés ;
- de publier ce rapport sur le site internet de la CNIL.

Le Président,

Michel GENTOT

C'est à la suite des débats intervenus sur les modalités de la mise en place de la déclaration obligatoire de séropositivité au virus de l'immunodéficience humaine (VIH), que la Commission Nationale de l'Informatique et des Libertés a décidé comme l'y incitent les dispositions de l'article 1er du décret du 17 juillet 1978<sup>1</sup>, de mener une réflexion d'ensemble sur le dispositif de surveillance épidémiologique du sida et sur la déclaration obligatoire de séropositivité au VIH en particulier.

Afin de mener ce travail de réflexion, la CNIL a entendu différentes associations de défense des malades atteints du sida, la Ligue des droits de l'Homme et le Conseil National de l'Ordre des Médecins. Elle a, en outre, recueilli sur place l'avis de médecins inspecteurs des directions départementales des affaires sanitaires et sociales qui jouent un rôle important dans le dispositif de surveillance des maladies à déclaration obligatoire et du sida en particulier.<sup>2</sup> Enfin, elle s'est rendue dans les locaux de l'Institut de Veille Sanitaire, chargé notamment de la surveillance épidémiologique du sida, afin de se rendre compte concrètement de son action et des mesures de confidentialité adoptées. La situation chez la plupart de nos voisins européens a également été étudiée.

## RAPPEL DU CONTEXTE

1 - Au cours du printemps 1998, le Ministre de la Santé annonce officiellement la décision d'inscrire la séropositivité au virus de l'immunodéficience humaine parmi la liste des maladies à déclaration obligatoire.

Un décret du 6 mai 1999, pris en application de la loi du 1er juillet 1998 sur la veille sanitaire, inscrit dans la liste des maladies à déclaration obligatoire l'infection par le virus de l'immunodéficience humaine, quelque soit le stade de l'infection ( annexe 3 ).

Un autre décret daté du même jour ( annexe 2 ), dont la CNIL n'a pas officiellement été saisie pour avis, suscite au cours de l'été 1999 de nombreuses réactions de la part d'associations de patients relayées par la presse. Celles-ci redoutent en effet que ce texte permette la mise en place d'un "fichage" nominatif des personnes séropositives et, dès lors, s'inquiètent de la possibilité de bénéficier à l'avenir d'un dépistage anonyme. En effet, alors que la loi précitée du 1er juillet 1998 dispose que "les modalités de la transmission des données à l'autorité sanitaire..., en particulier la manière dont l'anonymat est protégé, sont fixées par décret en Conseil d'Etat", le décret prévoit que la notification des déclarations est réalisée "sous la forme d'une fiche qui comporte des éléments à caractère nominatif", ce qui peut donner à penser que les données transmises seraient directement nominatives.

---

<sup>1</sup>L'article 1er du décret du 17 juillet 1978 précise dans les missions imparties à la Commission que celle-ci doit conseiller les personnes et organismes qui ont recours au traitement automatisé d'informations nominatives.

<sup>2</sup>L'ensemble des compte-rendus de ces auditions figure en annexe au présent rapport (annexes 4 et 5)

2 - De façon parallèle, l'Institut de Veille Sanitaire (IVS), organisme chargé, notamment, de la surveillance épidémiologique du sida (anciennement dénommé Réseau National de Santé Publique) saisit la CNIL, en mars 1999, d'une demande d'avis relative à l'expérimentation d'un traitement informatique des déclarations obligatoires de séropositivité dans vingt-deux départements.

Il est prévu que ces déclarations, destinées à être traitées sur informatique par l'IVS, seront, comme c'est le cas actuellement pour les déclarations obligatoires de sida avéré, indirectement nominatives<sup>3</sup>. La CNIL demande des précisions en particulier sur les modalités de transmission des informations et les dispositions prises pour assurer la confidentialité des informations.

3 - La polémique suscitée par la parution du décret conduit la CNIL à faire paraître en juillet 1999 un communiqué de presse rappelant qu'aucun avis n'a encore été rendu et qu'en conséquence aucun traitement automatisé ne peut être mis en oeuvre par l'Institut de veille sanitaire. Un courrier est également adressé au Ministre de l'Emploi et de la Solidarité.

Le Ministre décide alors de suspendre le projet et de constituer avec les associations de patients un groupe de travail chargé de réexaminer les modalités de mise en place du dispositif des déclarations obligatoires, tant en ce qui concerne leur contenu que les procédures de déclaration et d'éventuelles solutions techniques d'anonymisation. Parallèlement, la CNIL décide, par délibération du 9 septembre 1999, qu'il y a lieu en l'état de différer l'examen de la demande d'avis présentée par l'Institut de Veille Sanitaire jusqu'à ce que cet institut arrête les modalités définitives de mise en oeuvre à titre expérimental ou définitif du dispositif de surveillance des déclarations obligatoires de séropositivité au virus de l'immunodéficience.

Au titre de sa mission de conseil, la Commission décide de procéder à une étude d'ensemble de la question afin de pouvoir faire part aux pouvoirs publics de ses conclusions et en particulier de l'opportunité ou non de disposer d'informations individuelles.

Tel est l'objet du présent rapport.

---

<sup>3</sup> Aux termes du dossier présenté, il était en effet prévu que, outre les données médicales, soient recueillis les initiales du nom et du prénom, le sexe, la date de naissance, la nationalité, le département ou pays de résidence, le code postal de domicile et le code Insee de la catégorie socio-professionnelle. Enfin devaient également être mentionnés le nom du médecin déclarant et le lieu d'exercice ainsi que la date de déclaration.

## **I - LES FICHIERS ÉPIDÉMIOLOGIQUES**

L'épidémiologie a pour objet d'étudier la fréquence et la répartition de problèmes de santé en termes de morbidité et de mortalité en fonction de leur évolution dans le temps, de leur localisation géographique, des caractéristiques démographiques, et de rechercher les facteurs de risque et les causes de maladies.

La poursuite de ces objectifs peut donc nécessiter la collecte d'informations sur les personnes et la constitution de fichiers pour certains anonymes, pour d'autres directement ou indirectement nominatifs.

Faisant essentiellement appel à des méthodes d'analyse statistique, la recherche épidémiologique recourt aujourd'hui de façon quasi systématique aux moyens informatiques pour traiter les données.

Ces fichiers épidémiologiques se caractérisent par une grande diversité en ce qui concerne tant les sources des informations et les modalités de recueil de celles-ci que la nature même des données collectées.

### **A- LES DIFFÉRENTS MODES DE RECUEIL D'INFORMATIONS**

Le problème posé est celui de savoir si la qualité, la fiabilité et la validité scientifique des analyses statistiques doit conduire à préférer un système privilégiant une collaboration volontaire ou conduire à un système qui ferait obligation aux médecins et aux patients de communiquer les informations nécessaires.

En France, les fichiers épidémiologiques sont généralement institués sur la base d'une participation volontaire des professionnels de santé, la loi n'imposant aux médecins de transmettre des informations sur leurs patients que dans un nombre très limité de situations.

Le constat est le même s'agissant du patient, le droit français privilégiant en ce domaine les libertés individuelles. Le droit de s'opposer à l'utilisation de ses données à des fins de recherche médicale a d'ailleurs été consacré par la loi du 1er juillet 1994 qui a complété sur ce point la loi du 6 janvier 1978.

L'échec relatif en France des systèmes de déclaration obligatoire tient sans doute à la conjugaison de plusieurs facteurs : une conception rigoureuse du secret médical, une approche sans doute plus individualiste que collective de ce qu'exigent les nécessités de la santé publique, le constat également qu'une information médicale de qualité s'obtient sans doute plus aisément d'un professionnel de santé que l'on convainc de participer à un système d'information sanitaire que d'un professionnel de santé contraint à accomplir une formalité qui peut alors être perçue inexactement comme purement administrative.

Ainsi, si l'on excepte les rares enquêtes de santé obligatoires menées par l'INSEE ( en particulier l'enquête décennale santé ), les statistiques individuelles de décès<sup>4</sup> et les systèmes d'information mis en place dans les hôpitaux et dans la branche assurance maladie - mais qui ne relèvent pas stricto sensu du champ de la recherche en épidémiologie - le mode de recueil obligatoire est principalement résumé aux déclarations à l'autorité sanitaire de certaines maladies.

## **1- La déclaration obligatoire**

### **a- Le principe de la déclaration obligatoire**

Le principe est inscrit dans la loi à l'article L11 du Code de la santé publique ( annexe 1).

L'obligation pour les professionnels de santé de déclarer certaines maladies date d'une loi du 30 novembre 1892 sur l'exercice de la médecine, complétée par la suite par une loi du 15 février 1902. A l'époque, le terme employé était celui de maladies "contagieuses" qui devaient être déclarées par les médecins et les sages-femmes. Etaient principalement concernées les maladies vénériennes comme la syphilis et les fièvres puerpérales au moment de l'accouchement.

Invoquant le respect du secret médical, les médecins ont longtemps manifesté des réticences à signaler les cas, élément qui a conduit les pouvoirs publics, dans un décret du 21 décembre 1936, à préciser le circuit de transmission des données collectées. C'est ce texte qui a régi, jusqu'à son abrogation par le décret du 6 mai 1999, les modalités de transmission des données à l'autorité sanitaire.

La lecture de ce décret de 1936 est, à plus d'un titre, intéressante. Il était prévu que la déclaration devait être effectuée sur des "cartes-lettres" comprenant deux feuillets, l'un rempli par le médecin, l'autre rempli par la famille. Sur chacun des feuillets devaient en particulier être inscrits le nom du malade, son prénom, la nature de la maladie désignée par un numéro d'ordre, l'adresse et la date de naissance. Au niveau départemental, l'inspecteur départemental devait tenir un registre spécial où étaient "inscrits, pour chaque arrondissement et par ordre chronologique, les cas de maladie, la date de déclaration et la désignation des localités et des immeubles où ils se sont produits". Ce registre ne comportait pas le nom du patient. A la fin de chaque mois, un registre récapitulatif établi à partir des données détenues par les inspecteurs départementaux devait être transmis au ministre de la santé publique.

---

<sup>4</sup> Les statistiques des causes médicales de décès sont réalisées par l'INSEE et par le service commun 8 de l'INSERM à partir des bulletins statistiques n°7 et des certificats médicaux de décès obligatoirement établis lors de tout décès constaté par un médecin et transmis par l'intermédiaire des médecins inspecteurs des directions départementales des affaires sanitaires et sociales. Le traitement informatique de ces documents porte sur des données indirectement nominatives : lieu de domicile, date et lieux de naissance et de décès, numéro de l'acte de décès sur le registre d'état civil.

Cependant, l'objectif d'origine, qui était d'éviter la propagation de maladies contagieuses et de mettre en place rapidement des mesures prophylactiques, a largement évolué pour s'orienter vers la surveillance épidémiologique. Une rénovation des modes de déclaration a ainsi été amorcée en 1986 tant en ce qui concerne la liste des maladies à déclaration obligatoire (le décret du 10 juin 1986 a ainsi inscrit le sida avéré sur la liste des maladies à déclaration obligatoire) que les modalités de la surveillance (circulaire du 19 novembre 1986).

De nouveaux objectifs ont ainsi été assignés au système de la déclaration obligatoire, la mise sous surveillance d'une maladie devant être justifiée soit par la nécessité d'une intervention des autorités sanitaires, en fonction de l'origine, du potentiel épidémique ou de la gravité de la maladie, soit par la nécessité d'une évaluation pour les maladies qui font l'objet d'un programme de lutte ou de prévention<sup>5</sup>. Surtout, alors que le dispositif d'origine, tombé d'ailleurs depuis en désuétude, prévoyait une déclaration nominative, les nouveaux modes de recueil qui sont mis en place dans les années 1980 se limitent pour la plupart à des données qui ne sont nominatives qu'indirectement, c'est-à-dire qui ne comportent plus le nom des personnes mais dont on ne peut exclure qu'elles puissent encore permettre l'identification des personnes concernées.

La loi sur la veille sanitaire du 1er juillet 1998 a consacré cette nouvelle orientation en prévoyant deux modes de surveillance:

- les maladies dont la surveillance est nécessaire à la conduite et à l'évaluation de la politique de santé publique et qui imposent la transmission de données individuelles à l'autorité sanitaire ;
- les maladies qui nécessitent une intervention urgente locale, nationale ou internationale et qui doivent être signalées sans délai, comme par exemple, le choléra, la listériose, la rage ou encore les toxi-infections alimentaires. Elles sont aujourd'hui au nombre de 20.

La lecture du décret 99-363 du 6 mai 1999 permet de constater que la quasi-totalité des maladies énumérées dans la première catégorie sont également visées dans la seconde catégorie, à l'exception, toutefois, du tétanos, de l'hépatite B et de l'infection à VIH.

Cette exception est importante, la nature des données susceptibles d'être communiquées à l'autorité sanitaire différant de façon notable selon le mode de signalement .

Ainsi, le décret n° 99-362 du 6 mai 1999 prévoit, s'agissant des maladies qui nécessitent une intervention urgente, la possibilité de communiquer, à la demande du médecin inspecteur de la santé publique, l'identité et l'adresse du malade, afin de mettre en place des mesures de prévention et, le cas échéant, de déclencher des investigations pour identifier l'origine de la contamination ou de l'exposition.

---

<sup>5</sup> C'est également à cette époque qu'un médecin prend la direction du "bureau des maladies transmissibles" à la Direction Générale de la Santé.

Dans une rédaction quelque peu ambiguë, et qui a suscité la polémique que l'on sait, ce même décret prévoit que s'agissant des maladies dont la surveillance est nécessaire à des fins d'évaluation de la politique de santé publique des éléments à caractère nominatif doivent être transmis

#### b- Le circuit de transmission des déclarations obligatoires

Les textes précisent que les médecins ayant diagnostiqué des cas de maladies à déclaration obligatoire doivent les transmettre aux médecins inspecteurs des directions départementales des affaires sanitaires et sociales qui constituent un niveau intermédiaire de recueil des données, de validation de celles-ci et de transmission aux autorités sanitaires nationales.

Il convient de rappeler qu'aux termes du décret du 6 décembre 1994 relatif aux missions et attributions des directions régionales et départementales des affaires sanitaires et sociales, le directeur départemental des affaires sanitaires et sociales est responsable, sous l'autorité du préfet du département de la mise en oeuvre des politiques sanitaires, médico-sociales et sociales définies par les pouvoirs publics. Ses missions comprennent notamment les actions de promotion et de prévention en matière de santé publique, ainsi que la lutte contre les épidémies et endémies.

Actuellement, sont ainsi transmises, sous pli confidentiel, au médecin inspecteur de la DDASS, des formulaires papier comportant les initiales du nom et du prénom de la personne ainsi que la date de naissance complète, son sexe, le code postal du lieu de résidence ou le département<sup>6</sup>, la nationalité, le pays de naissance, la localisation clinique de la maladie, la date de mise en route du traitement, les résultats biologiques, les facteurs favorisant l'apparition de la maladie et le nom et l'adresse du médecin déclarant.

Pour certaines maladies, en particulier la tuberculose et les toxi-infections alimentaires collectives, les médecins inspecteurs de santé publique des directions départementales des affaires sanitaires et sociales saisissent les données, effectuent une analyse départementale et transmettent à l'IVS de façon indirectement nominative les données nécessaires à l'épidémiologie.

Pour les autres maladies, les médecins inspecteurs de santé publique valident les fiches de déclaration et les transmettent à l'IVS.

## **2- Les systèmes de recueil facultatifs des données :**

### a - Les registres de morbidité -

Ces registres qui portent essentiellement sur les cancers (il existe aujourd'hui une quarantaine de registres en fonctionnement) permettent de recenser, dans une zone géographique déterminée, les cas de cancer à partir des données nominatives couvertes par le secret professionnel, transmises volontairement à l'organisme de recherche par les médecins et autres

---

<sup>6</sup> Certaines maladies, comme les légionelloses ou les méningites par exemple, nécessitent la collecte du code postal afin de situer géographiquement la personne.

professionnels de santé concernés. Ces registres comportent le nom des personnes. Les fichiers épidémiologiques ainsi constitués doivent permettre de connaître avec précision le nombre, l'incidence et l'étiologie des cancers dans une population donnée et d'envisager des actions de dépistage ou de prévention.

L'arrêté du 6 novembre 1995 qui définit la composition et les missions du comité national des registres, organisme chargé de se prononcer sur l'utilité de la création de nouveaux registres ou sur l'opportunité du maintien de ceux existants, définit le registre comme un instrument permettant "un recueil continu et exhaustif de données nominatives intéressant un ou plusieurs événements de santé dans une population géographiquement définie, à des fins de recherche et de santé publique, par une équipe ayant les compétences appropriées".

Les informations enregistrées dans ces fichiers, constitués sous l'égide de centres hospitaliers, d'associations loi 1901 ou de l'INSERM, portent principalement sur l'identité complète des malades, le sexe, la localisation du cancer, la date et le critère de diagnostic, le traitement des facteurs de risque et l'évolution de la maladie.

Les informations sont en général collectées par voie de "questionnaires papier" nominatifs remplis par les médecins traitants et les laboratoires d'analyses et renvoyés aux registres.

Il convient de noter que l'examen par la CNIL des caractéristiques de fonctionnement de ces registres est à l'origine de l'adoption de sa recommandation du 19 février 1985 dans laquelle la CNIL a appelé l'attention des pouvoirs publics sur l'opportunité d'une législation spécifique aux fichiers de recherche épidémiologique et qui a conduit, neuf ans plus tard, à l'adoption de la loi du 1er juillet 1994 relative aux traitements de données ayant pour fin la recherche dans le domaine de la santé.

#### b - Les différents types d'enquêtes -

Sans être exhaustif, on peut citer plusieurs types d'enquêtes caractéristiques de la recherche épidémiologique.

1- Les *enquêtes de cohorte* consistent à suivre sur une période généralement assez longue, une population déterminée, soit en demandant directement aux intéressés à intervalles réguliers des informations, soit/et en recourant aux médecins traitants.

Le principe de ces enquêtes implique que les noms et adresses des personnes concernées soient enregistrés dans la mesure où elles doivent, régulièrement ou, passé un certain délai, être interrogées à nouveau ou recontactées. Le caractère nominatif de ces enquêtes sert généralement à récapituler l'ensemble des données de suivi sur un même cas et permettre au médecin de transmettre à l'organisme de recherche les données concernant cette dernière.

Le caractère directement nominatif de telles enquêtes appellent évidemment des précautions pour garantir au niveau du traitement et de l'analyse des données, une séparation complète des données d'identification avec les renseignements médicaux.

Ainsi, de nombreuses enquêtes de cohorte sont fondées sur le principe d'une "boîte noire",

c'est-à-dire le recours à un tiers, détenteur de la seule identité des personnes, à l'exclusion donc des renseignements médicaux recueillis (par exemple, l'enquête de cohorte GAZEL "20 000 volontaires pour la recherche" réalisée par l'unité 88 de l'INSERM auprès du personnel d'EDF-GDF. La boîte noire est, en l'espèce le service de médecine du travail).

2- Les enquêtes dites *transversales* visent à connaître, à un moment précis (un jour, une semaine, un mois) et de façon instantanée, le phénomène morbide étudié : observations des cas existants au moment de l'enquête. Ces enquêtes peuvent être reconduites d'une année sur l'autre. Par exemple, il est réalisé annuellement des enquêtes un jour donné dans des services hospitaliers accueillant des malades atteints du sida. Ce type de recueil porte généralement sur des données anonymes ou très indirectement nominatives.

3- Les *enquêtes par sondage* peuvent également être réalisées auprès d'échantillons de population représentatifs, par tirage au sort, ce qui permet une observation plus fine des populations étudiées mais impose une grande vigilance dans l'analyse (risque de biais) et l'interprétation des résultats.

4- Les *enquêtes cas-témoins* permettent elles de comparer deux populations distinctes, l'une atteinte d'une pathologie déterminée et l'autre considérée comme saine. Selon les cas, les données recueillies sont directement ou indirectement nominatives.

#### c - Les réseaux "sentinelles" -

Les réseaux "sentinelles" permettent, à partir d'un réseau de médecins volontaires, de fournir à un organisme de recherche, à intervalles réguliers, des informations sur des personnes atteintes d'une pathologie déterminée afin d'en suivre les périodes de survenance, d'épidémie. (ainsi, l'INSERM et l'Institut de Veille Sanitaire suivent les périodes d'épidémie de grippe ou de bronchiolites chez les nouveaux-nés). Les données recueillies sont généralement anonymes, à l'exception des coordonnées du médecin.

## B - DE LA NÉCESSITÉ DE DISPOSER DANS LE DOMAINE DE L'ÉPIDÉMIOLOGIE DE DONNÉES DIRECTEMENT OU INDIRECTEMENT NOMINATIVES

### **1- Donnée nominative et anonymat**

Le terme "anonyme" utilisé dans le langage courant possède une double origine grecque (onoma, le nom) et latine (anonymus), et signifie dans les deux cas l'ignorance du nom. Une personne est considérée comme anonyme lorsqu'on ignore son nom.

La loi Informatique et Libertés donne une définition très large du terme "nominatif". En effet, l'article 4 de la loi du 6 janvier 1978 modifiée, précise "qu'est réputée nominative toute donnée qui permet, sous quelque forme que ce soit, directement ou non, l'identification de personnes physiques".

Ainsi une information nominative n'est pas seulement une information qui comporte le nom de la personne, mais aussi toute information qui, alors qu'elle ne serait pas associée au nom d'une personne, permettrait aussi de l'identifier.

Relèvent ainsi du champ d'application de la loi, non seulement les traitements informatiques de données directement nominatives qui comportent le nom des personnes physiques (par exemple les registres de cancer), mais également les systèmes informatiques de données indirectement nominatives, c'est-à-dire d'informations susceptibles de permettre l'identification de personnes physiques, soit par référence à des listes nominatives ou à d'autres fichiers (par exemple, numéro de téléphone, initiales du nom et du prénom ou numéro d'ordre renvoyant à une liste nominative de référence même si celle-ci est établie sur support papier et conservée par chaque médecin), soit encore par recoupement d'informations ( par exemple, date et lieu de naissance, code postal de résidence, pathologie rare...).

En conséquence, un fichier de recherche même s'il ne comporte pas le nom de la personne - s'il est au sens littéral "anonyme"- peut relever des dispositions de la loi "Informatique et Libertés". Aujourd'hui bon nombre des traitements informatiques mis en place dans le cadre de la recherche épidémiologique sont "indirectement nominatifs" au sens de l'article 4 de la loi précitée et, à ce titre, soumis au respect des dispositions de cette loi. Il en est ainsi des traitements informatiques mis en oeuvre pour exploiter les déclarations des maladies obligatoires qui revêtent un caractère "indirectement nominatif" au sens de l'article 4 de la loi du 6 janvier 1978.

Aussi quand la loi sur la veille sanitaire dispose que les modalités de transmission des données à l'autorité sanitaire doivent être fixées de telle sorte que "l'anonymat soit protégé", l'expression peut-elle passer pour ambiguë. Au sens commun, il est incontestable que cette expression signifie que les données ne doivent en aucun cas permettre d'identifier- fut-ce par recoupement- les personnes concernées.

Mais, au sens de la loi du 6 janvier 1978, cette exigence d'anonymat n'implique pas automatiquement que les données soient dépourvues de tout caractère indirectement nominatif. Aussi, lorsque le législateur précise que "les modalités de la transmission des données à l'autorité sanitaire..., en particulier la manière dont l'anonymat est protégé, sont fixées par décret en Conseil d'Etat", comprend-elle cette exigence comme devant conduire à s'assurer que toutes les dispositions de confidentialité soient prises pour garantir que le nom des personnes ne sera pas divulgué.

Dès lors et sur le fondement de la loi du 6 janvier 1978, la loi sur la veille sanitaire ne paraît pas interdire que des données individuelles et en particulier indirectement nominatives puissent être recueillies dès lors que des mesures de protection appropriées seraient adoptées.

## **2- La justification du recours à des données nominatives**

En épidémiologie, le recueil des données identifiantes sur les personnes, définies comme des sujets d'observation, et sur les professionnels de santé qui ont fourni l'information, constitue un moyen de discriminer un cas par rapport à un autre, d'éliminer les doubles enregistrements d'autant plus fréquents que les informations sont susceptibles de provenir de sources

d'information multiples ( ex : médecins hospitaliers, laboratoires...), de rassembler et de vérifier sur un patient déterminé, les renseignements obtenus, de suivre cas par cas l'évolution d'une pathologie ou d'une thérapeutique.

Les justifications du recours à des données identifiantes sont donc nombreuses dans le domaine épidémiologique.

Ainsi, l'élimination des doublons et la vérification des données soit auprès des professionnels de santé qui les ont fournies , soit auprès des patients eux-mêmes, permet de disposer de données fiables et de qualité et de statistiques scientifiquement valables même si des corrections statistiques peuvent parfois être appliquées pour pallier les risques de biais résultant d'informations insuffisantes.

De même le suivi épidémiologique de l'évolution de telle pathologie et de l'impact de telle ou telle thérapeutique peut induire le recueil en continu, sur des cas diagnostiqués, d'informations de suivi, ce qui nécessite alors de procéder à l'appariement des informations autour d'identifiants communs ( ex : nom ou initiales du nom et du prénom ).

En outre, le chercheur peut avoir le souci de vérifier ou d'approfondir tel ou tel résultat obtenu en procédant, à partir d'une première analyse des données recueillies, à des enquêtes complémentaires qui seront conduites tant auprès des professionnels de santé qu'auprès des patients. Tel est en particulier le cas pour certains registres de morbidité. La conservation, dans le fichier, de données identifiantes peut alors se justifier.

Enfin, une étude épidémiologique peut éventuellement s'accompagner ou être suivie d'actions de prévention ou de dépistage auprès de la population.

### **3- Etat des recueils d'informations existants : les pratiques constatées**

Selon le type d'études réalisées, la méthodologie employée, le recours ou non à l'outil informatique, les informations recueillies sont donc "plus ou moins nominatives".

Les systèmes de recueil de données épidémiologiques à caractère obligatoire font généralement appel à des données qui sont indirectement nominatives.

Le recueil et le traitement de données indirectement nominatives s'avèrent souvent indispensables, du fait même de la nature du problème étudié. Ainsi, l'étude de certaines filières de soins ou de certaines maladies contagieuses pourra nécessiter le recueil de la localisation géographique précise ( commune de résidence ) des patients et du service d'hospitalisation ou du médecin concerné. De telles données corrélées avec l'âge des patients peuvent permettre l'identification des personnes concernées, surtout si l'effectif de population est restreint ou si la source d'information est identique.

D'autres recherches épidémiologiques feront appel directement à l'identité de la personnes afin d'assurer un recensement aussi exact que possible des cas de la pathologie étudiée et permettre la réalisation d'études complémentaires : tel est le cas, notamment, des registres de cancer.

Il demeure, qu'au delà des justifications précédemment invoquées, le recueil de l'identité complète de la personne peut constituer quelquefois une "facilité" qui rend plus aisée, pour les médecins qui fournissent les informations, la recherche dans les dossiers médicaux, des cas susceptibles d'entrer dans le champ de l'étude. Or, dans bien des cas l'identité complète n'est pas nécessaire et pourrait être remplacée par les initiales du nom ou par un numéro d'ordre.

Il appartient ainsi à la CNIL et, également depuis la loi du 1er juillet 1994, au Comité consultatif sur le traitement de l'information en matière de recherche dans le domaine de la santé, d'apprécier l'utilité du recours à des données nominatives et de préconiser, dans la mesure du possible, la limitation du nombre ou de la précision des données permettant l'identification des personnes ainsi que le recours à des techniques dites "d'anonymisation".

La CNIL a toujours favorisé la mise en place et le développement de mesures de sécurité spécifiques de nature à garantir la confidentialité des données. Ces mesures peuvent être très diverses et de niveaux très différents : cryptage, séparation des données relatives à l'identité des personnes, des renseignements proprement médicaux, "appauvrissement" des données (par exemple en recueillant l'âge de la personne plutôt que sa date de naissance), élaboration de systèmes plus complexes "d'anonymisation" à la source des données d'identification.

Ainsi, l'INSERM a mis en place en 1988, avec l'accord de la CNIL un système d'information épidémiologique sur le sida faisant appel à un système d'anonymisation à la source - c'est-à-dire dans les services de maladies infectieuses hospitaliers - des données d'identification des patients atteints du sida. Ce dispositif, soumis à l'époque à l'expertise du Service central de sécurité des systèmes d'information (SCSSI)<sup>7</sup> et dit "algorithme de San Marco"<sup>8</sup>, permet, grâce à un algorithme particulier, de coder de façon irréversible les nom, prénom et date de naissance du patient, c'est-à-dire de produire un numéro non signifiant et non identifiant à partir de ces trois informations. Toute transmission ultérieure d'informations accompagnées du numéro codé, spécifique à une personne, peut permettre de savoir qu'il s'agit du même patient sans qu'il soit possible de l'identifier ni directement ni indirectement. Les informations médicales sont transmises à l'INSERM accompagnées de ce "numéro codé", ce qui permet à l'INSERM lorsqu'il reçoit les données de suivi de repérer si le cas a déjà été enregistré dans la base et dans l'affirmative de mettre à jour en conséquence les données déjà enregistrées.

---

<sup>7</sup> Le SCSSI, placé sous l'autorité du Secrétaire général de la défense nationale, a été institué par le décret du 3 mars 1986 et exerce ses compétences en particulier dans le domaine de la cryptologie et de la sécurité informatique

<sup>8</sup> Nom du responsable du laboratoire de santé publique de la faculté de médecine de Marseille qui a développé ce système.

En 1996, dans le cadre de la mise en place du programme de médicalisation des systèmes d'information (PMSI) dans les établissements de santé privés, la CNIL s'est également prononcée sur un dispositif d'anonymisation qui limite les risques d'identification tout en permettant d'assurer un lien ("chaînage") entre les différents séjours d'un patient au sein de l'établissement. Cette technique fait appel à un algorithme dit de "hachage" (SHA) qui permet de transformer de façon non réversible les données nominatives en un numéro anonyme et unique permettant, sans qu'il soit possible d'identifier le patient, d'apparier cependant sur un même individu les données relatives à ses séjours successifs dans l'établissement. Ce dispositif est développé par le Centre d'Etudes des Sécurités du Système d'Information (CESSI) qui dépend de la Caisse Nationale d'Assurance Maladie des Travailleurs Salariés et a également été expertisé par le SCSSI.

Une technique analogue est actuellement utilisée par le laboratoire d'informatique médicale du CHU de Dijon dans le cadre de systèmes d'informations épidémiologiques.

Ces dispositifs, qui sont en réalité de petits programmes informatiques, supposent que les professionnels et établissements de santé qui procèdent à la source à la collecte des données disposent de moyens informatiques<sup>9</sup>, ce qui est encore loin d'être le cas aujourd'hui pour les médecins même si la généralisation du dispositif Sésam-Vitale "incite" fortement les médecins à s'équiper.

## **II - LA SURVEILLANCE ÉPIDÉMIOLOGIQUE DU SIDA**

### **-Un contexte particulier-**

Si l'épidémie de sida a pu, dans un premier temps au moins, être un révélateur de bien des présupposés moraux ou psychologiques, elle a également, et très rapidement, contribué à faire évoluer les attitudes tant des médecins que des personnes concernées vers plus de transparence et de responsabilité.

C'est ainsi, que s'agissant des fichiers de recherche sur le sida, la Commission a pu prescrire que seul le recueil du consentement exprès des personnes concernées, c'est-à-dire un consentement écrit et précédé d'une information personnalisée, était de nature, tout à la fois, à responsabiliser le corps médical sur son devoir d'information et à garantir l'adhésion pleine et entière du patient au projet de recherche pour lequel on sollicitait sa participation.

Enfin, la recherche sur le sida a appelé une attention particulière sur les mesures de sécurité à mettre en place pour garantir aux données traitées la confidentialité nécessaire.

---

<sup>9</sup> Ou à tout le moins de petites "calcuettes" informatiques permettant de saisir les données d'identification et de calculer en conséquence un code anonyme. Un tel dispositif semble avoir été adopté en Grande-Bretagne.

## A - LE SYSTÈME ACTUEL DE SURVEILLANCE EN FRANCE

### **1- Le Sida avéré : maladie à déclaration obligatoire depuis le décret du 10 juin 1986**

Aux termes de l'article L355-22 du code de la santé publique "La définition de la politique de lutte contre l'infection par le virus de l'immunodéficience humaine appartient à l'Etat".

Le sida avéré a ainsi été inscrit sur la liste des maladies à déclaration obligatoire par un décret du 10 juin 1986.

La CNIL s'est prononcée à plusieurs reprises sur le système d'informatisation des déclarations obligatoires de sida avéré.<sup>10</sup> Les objectifs de la mise en place de cette déclaration étaient alors de connaître le nombre et les caractéristiques épidémiologiques des personnes à un stade avancé de l'infection par le VIH et de mesurer l'impact de l'accès au dépistage et au système de soins de ces personnes.

Le circuit des données recueillies et enregistrées à l'IVS est analogue à celui qui existe pour les autres maladies à déclaration obligatoire. C'est au praticien qui diagnostique un cas de sida qu'il appartient d'adresser le formulaire de déclaration obligatoire au médecin inspecteur de la santé de son département sous pli médical confidentiel. Le talon du formulaire qui comporte l'identité exacte du patient est conservé dans le dossier médical du patient.

Le médecin inspecteur de la santé du département, après vérification des renseignements fournis<sup>11</sup>, adresse le formulaire validé au médecin épidémiologiste de l'Institut de Veille Sanitaire qui saisit alors sur support informatique les données figurant sur le formulaire. Les informations saisies par l'IVS sont, s'agissant des informations d'identification : un numéro construit à partir des initiales du nom et du prénom, de la date de naissance, du département de domicile ainsi que le sexe, la date de décès, le département de naissance et le pays de domicile, la nationalité et la catégorie socio- professionnelle. Les autres informations collectées et enregistrées sur support informatique sont relatives au diagnostic du Sida et aux caractéristiques des soins.

Ces informations sont donc indirectement nominatives.

C'est le caractère indirectement nominatif de ces informations qui a permis de détecter les doublons sur l'ensemble des déclarations effectuées (depuis la mise en oeuvre du système 20% de doublons ont été ainsi repérés).

---

<sup>10</sup> Lors de la mise en place du système ( délibération n°88-91 du 6 septembre 1988), en 1995, lors de la révision de la définition de l'infection et pour permettre de détailler de façon plus précise les éléments participant au diagnostic de la maladie ( délibération n° 95-101 du 11 juillet 1995 ) et en 1997 afin de confier la responsabilité de cette surveillance au Réseau National de Santé Publique ( devenu depuis la loi du 1er juillet 1998 l'Institut de Veille Sanitaire ) et d'inclure dans la déclaration de nouveaux éléments relatifs aux critères cliniques dus à l'évolution de la pathologie du Sida ( délibération n°97-025 du 1er avril 1997 ).

<sup>11</sup>Une application informatique est mise à la disposition de ces médecins pour faciliter le travail d'analyse des formulaires qu'ils reçoivent. C'est la raison pour laquelle, la CNIL s'est également prononcée sur les applications informatiques mises en oeuvre dans les DDASS.

Après analyse des données, l'IVS retransmet aux médecins inspecteurs des DDASS, sous une forme d'une disquette cryptée des informations concernant les cas de sida répertoriés dans leur département. Les données retransmises ne comportent, s'agissant des éléments qui peuvent être considérés comme indirectement nominatifs, que le nom du médecin déclarant.

Il convient ici de relever le rôle particulier joué par les médecins inspecteurs des DDASS. Ils constituent en effet, comme pour la tuberculose et d'autres maladies à déclaration obligatoire, le relais local nécessaire pour effectuer un premier contrôle de validité des données qui leur sont transmises et prévoient ainsi les politiques de prévention.

Enfin, et c'est un point important à noter, aucune atteinte à la confidentialité des données n'a été constatée depuis la mise en place de cette déclaration, ni au sein des DDASS, ni au sein de l'IVS.

## **2- Les autres systèmes d'information sur le sida**

### **a - Les enquêtes spécifiques de cohorte -**

L'INSERM a été chargé en 1987, dans le cadre du programme national de recherche contre le sida, de constituer une cohorte de volontaires séropositifs, et de les suivre afin de déterminer les facteurs "pronostiques" de la survenue d'un sida, d'évaluer les traitements administrés lors d'essais thérapeutiques et de constituer une "sérothèque" permettant de conserver des échantillons de prélèvements sanguins à des fins de recherche. (cohorte SEROCO)<sup>12</sup>.

Le fichier ainsi constitué est directement nominatif. En outre, de nombreuses données "sensibles" sont recueillies. En effet, certaines données susceptibles d'influencer l'évolution de la séropositivité vers le sida sont recueillies qui concernent le statut familial, socio-professionnel, comportemental, clinique et biologique. Est également recueillie "l'origine raciale" afin d'étudier d'éventuelles corrélations entre cette information et les manifestations cliniques du virus. Le consentement exprès des personnes, contresigné du médecin est recueilli.

Il est prévu, par ailleurs, depuis la délibération rendue par la CNIL en 1997 sur les déclarations obligatoires de sida avéré, que soient transmises à l'unité 292 de l'INSERM, dans le cadre du suivi de la cohorte SEROCO, des données sur les personnes perdues de vue, notamment la date éventuelle du passage au stade du sida, la pathologie inaugurale, le groupe de transmission, la survenue et la date du décès, le cas échéant.

Depuis lors, d'autres enquêtes ont été réalisées, par exemple sur les risques de contamination par le VIH et l'hépatite C des professionnels de santé.

---

<sup>12</sup>Délibération n° 87-108 du 17 novembre 1987

b - Les systèmes d'information dans les centres d'information et de soins pour l'immunodéficience humaine -

En 1987, l'INSERM a été chargé de la coordination des 23 centres d'information et de soins pour l'immunodéficience humaine (CISIH) mis en place dans les établissements de soins prenant en charge des malades atteints du Sida, essentiellement dans les services de maladies infectieuses. L'INSERM a ainsi la tâche de collecter, d'exploiter et d'analyser, de façon cohérente et coordonnée l'ensemble des données épidémiologiques fournies par les centres.

La CNIL s'est prononcée à plusieurs reprises sur les fichiers informatiques mis en oeuvre tant dans ces centres qu'à l'INSERM<sup>13</sup>. Les données collectées au niveau des services sont nominatives mais font l'objet, avant leur transmission à l'INSERM d'une anonymisation à l'aide de l'algorithme de San Marco tel qu'il a été décrit précédemment.

A la suite d'une très large concertation conduite par la CNIL avec des médecins et des chercheurs, des associations de malades et des représentants du Ministère et des instances éthiques, il a été convenu que le recueil du consentement exprès des patients à l'informatisation de leurs données constituait une garantie essentielle du respect des droits des patients. Cette position a été réaffirmée par la CNIL en 1996<sup>14</sup>.

Il convient évidemment d'ajouter les recherches réalisées sous la responsabilité de l'Agence Nationale de Recherche sur le sida, et en particulier l'enquête nationale sur les comportements sexuels et le sida.

c - L'enquête RESORS-VIH -

Il s'agit d'un système régional de surveillance des découvertes de la séropositivité progressivement mis en place entre 1988 et 1996 dans 13 régions françaises à l'initiative des Observatoires régionaux de la santé (ORS). Le système abandonné en 1998 associait 71000 médecins et 2100 laboratoires et reposait sur la double participation des laboratoires et des médecins prescripteurs. Le document de collecte des données permettant de signaler la séropositivité, comportait "un volet laboratoire" et un questionnaire médical. Les deux volets étaient renvoyés à l'ORS au moyen d'enveloppes pré-adressées interdisant ainsi toute identification du laboratoire ou du médecin. Dans le souci de garantir l'anonymat des patients et le strict respect du secret professionnel, l'ORS n'avait ainsi connaissance ni du nom du patient, ni de celui du médecin prescripteur, ni de celui du laboratoire.

Les données collectées étaient relatives au sexe, mois et année de naissance et département de domicile des personnes, aux différentes sérologies effectuées, aux facteurs de risque connus du médecin et aux éventuels signes cliniques.

---

<sup>13</sup> Notamment délibération n°91-071 du 10 septembre 1991.

<sup>14</sup> Délibération n°96-055 du 18 juin 1996

#### d - Le réseau RENAVI -

Il s'agit d'un réseau national de laboratoires d'analyses de biologie médicale mis en place en 1988 pour étudier les tendances à long terme et les variations saisonnières de l'activité de dépistage du VIH en France au travers du suivi du nombre de tests réalisés par les laboratoires.

Ce réseau est composé de laboratoires privés et hospitaliers pratiquant les tests sérologiques de VIH hors don du sang. Deux recueils anonymes sont effectués : le nombre mensuel de tests sérologiques avec une répartition par sexe et le nombre mensuel de tests positifs effectués avec l'âge et le sexe des patients concernés. Toutes les régions participent à ce réseau qui compte 350 laboratoires. Il permet d'obtenir des estimations globales, nationales et régionales sur l'activité de dépistage en France mais ne permet pas de connaître les facteurs de risque des sujets testés et dépistés, ces informations n'étant pas connues des laboratoires.

### B- LA NÉCESSAIRE ÉVOLUTION DES MODALITÉS DE SURVEILLANCE ÉPIDÉMIOLOGIQUE DU SIDA

#### **1- Les systèmes de surveillance épidémiologiques à l'étranger**

Des éléments sur la prise en compte de l'infection par le virus de l'immunodéficience humaine chez la plupart de nos voisins étrangers ont été recueillis. Il ressort de ces informations que tous les pays d'Europe de l'Ouest ont mis en place depuis une dizaine d'années des systèmes de déclaration de la séropositivité au VIH, distinctes du sida déclaré, au plan national ou régional.

Ainsi, l'Allemagne dispose depuis 1988 d'un système de déclaration de la séropositivité au VIH qui comporte, s'agissant des éléments d'identification permettant l'élimination des doublons, l'année de naissance, la résidence et un code basé sur le nom. Cette déclaration est, par ailleurs reliée à la déclaration du sida avéré établie distinctement mais qui utilise le même code, ce qui permet de suivre le passage au sida déclaré.

L'Espagne a mis en place très récemment le système de déclaration du VIH après de longues discussions sur la nature des éléments d'identification. Ainsi, les initiales du nom et du prénom ont été retenues ainsi que la date de naissance complète. S'agissant du sida déclaré, le nom complet est collecté.

Au Royaume-Uni, le système national date de 1984 et comporte un code calculé à partir des consonnes du nom complété de la date de naissance. Les informations recueillies sont les mêmes pour le sida avéré.

En Italie, le système est régional et date de 1985. Sont enregistrés, les initiales, la date de naissance et la commune de naissance. Le sida avéré est lui déclaré nominativement.

On remarque cependant une évolution, dans un sens “plus identifiant”, des données collectées afin de mieux éliminer les doublons et de pouvoir affiner les politiques de prévention. Ainsi, en Suisse, les initiales du nom et du prénom ont été ajoutées en 1999 aux déclarations de VIH qui ne contenaient auparavant, comme élément d’identification, que la date de naissance et le canton de résidence. Cet ajout permettra également de rapprocher le fichier du fichier des déclarations de sida avéré.

Au Danemark, l’adoption d’un identifiant unique ( code de sécurité sociale ) pour les déclarations de VIH, actuellement en discussion, est présentée comme devant permettre de relier le fichier VIH avec d’autres fichiers de données sanitaires, pour suivre la morbidité, la mortalité et la prise en charge associées à l’infection VIH.

Par ailleurs, les différences observées dans le détail de certaines informations individuelles, recueillies s’explique par le nombre d’habitants et l’importance de l’épidémie très variable selon les pays. Par exemple, la Norvège, qui a un nombre faible de cas ( environ 100 cas par an depuis 1990 ) ne recueille que le mois et l’année de naissance sans le jour, ce qui lui suffit pour éliminer les doubles déclarations, contrairement à la plupart des autres voisins européens et à la France avec environ 4000 à 5000 cas attendus par an.

Enfin, il faut noter qu’aux Etats-Unis, 32 Etats sur 50 ont, comme pour le sida, une surveillance nominative de l’infection à VIH et qu’un tel système est envisagé dans d’autres Etats.

Au plan européen, a été créé au milieu des années 1980 un Centre Européen pour la surveillance épidémiologique du sida qui regroupe 48 pays de la région Europe de l’Organisation Mondiale de la santé utilisant la même définition du sida et dont les chiffres peuvent donc être comparés. Une seule institution par pays, responsable de la qualité des données (en France, l’Institut de Veille Sanitaire) déclare les cas de sida au Centre Européen qui est d’ailleurs hébergé en France par l’IVS ( à l’Hôpital de Saint Maurice dans le Val de Marne ). Chaque trimestre, le pays participant à ce programme européen fournit sous la forme d’une extraction de sa base de données nationale des cas de sida, des données anonymes sur tous les cas de sida déclarés depuis le début de l’épidémie. Cette base de données comporte des informations communes à l’ensemble des pays concernés notamment sur la localisation dans le pays du cas révélé et sur les modes de contamination. Une base de données européenne des cas de sida a ainsi été constituée. Au 31 décembre 1998, 218 938 cas de sida ont été déclarés dans la Région Europe de l’OMS<sup>15</sup>. 60% d’entre eux sont connus comme étant décédés. La diminution de l’incidence du sida, amorcée en 1996, se poursuit (12 853 cas diagnostiqués en 1998 comparés à 15 655 en 1997). En 1998, les incidences du sida les plus élevées étaient observées dans les ex-pays de l’Est. La transmission hétérosexuelle est responsable d’une proportion croissante des cas de sida et dans plusieurs pays dont la France, la Norvège ou la Suède, ce mode de transmission est devenu majoritaire parmi les nouveaux cas de sida.

---

<sup>15</sup> Aujourd’hui, 33,6 millions de personnes dans le monde sont touchés par le virus du sida.

## 2- La réflexion Française

En France, l'évolution de l'épidémie de sida est aujourd'hui mesurée essentiellement grâce aux résultats issus des systèmes d'information précités, et principalement de l'exploitation des déclarations obligatoires de sida avéré. Ainsi, au 31 décembre 1998, le total cumulé des cas de sida en France depuis 1983 était de 49 421. La brusque diminution du nombre de nouveaux cas de sida qui avait été observée au deuxième semestre 1996 (- 31%) s'est poursuivie en 1997 mais à un rythme plus lent (- 14%) au deuxième semestre. Une chute importante du nombre de décès a également été constatée.

Mais, les statistiques actuellement disponibles à partir des déclarations de sida avéré ne reflètent pas l'évolution actuelle de l'épidémie du sida en France. En effet, les nouvelles modalités de prise en charge des personnes séropositives et en particulier les traitements anti-rétroviraux et les trithérapies permettent de retarder de façon significative voire même de faire régresser l'évolution de la maladie. Dès lors, une baisse relative des cas de sida déclarés ne permet pas pour autant d'en tirer des conclusions quant à l'évolution au fil du temps de l'épidémie. On peut simplement dire que l'apparition des nouvelles thérapies a eu pour effet de diminuer le nombre de cas de sida avéré sans pour autant en déduire que le nombre de séropositifs a diminué.<sup>16</sup>

La majorité des interlocuteurs rencontrés par la CNIL s'accorde à reconnaître que le système actuel de la seule déclaration du sida au stade de la maladie est donc insuffisant pour appréhender l'ensemble des aspects de l'épidémie. On estime qu'aujourd'hui, il apparaît 5000 cas nouveaux de séropositivité par an mais aucun instrument épidémiologique efficace n'existe en France pour mieux cerner les risques et adapter les politiques de prévention.

Face à ce constat, les pouvoirs publics ont souhaité avant d'arrêter leur décision quant à l'opportunité d'inscrire ou non la déclaration de séropositivité dans la liste des maladies à déclaration obligatoire, recueillir l'avis des instances scientifiques.

### a - *Les avis rendus en 1998* -

Le Conseil National du Sida, (CNS) saisi par le Directeur Général de la Santé, à la demande du Ministre, a émis le 29 janvier 1998 un avis défavorable à l'instauration d'un système de déclaration de nature obligatoire. Tout en reconnaissant les limites de la déclaration du sida avéré, le Conseil National a en effet estimé que le caractère contraignant et obligatoire du signalement de toute séropositivité au VIH comporterait en lui-même des risques d'atteinte aux libertés individuelles importants. En outre, le CNS souligne que la multiplicité des recherches menées dans le domaine du sida et l'existence d'applications informatiques spécifiques à la prise en charge des malades séropositifs permettent d'améliorer la connaissance de l'infection, dans la mesure où une mobilisation active de l'ensemble des acteurs concernés continue à exister.

---

<sup>16</sup>La prise en charge avant le passage au stade Sida donne des indications sur les pratiques antérieures de diagnostic et de suivi. Pour la période juillet 1997-juin 1998, 45% des nouveaux diagnostics de sida concernant des sujets qui ignoraient auparavant leur séropositivité. 24% des cas de Sida observés et déclarés en 1997-1998 avaient été préalablement diagnostiqués et pris en charge par les antirétroviraux dans les deux années précédant le diagnostic du sida.

L'Académie Nationale de Médecine, saisie par le Ministre de l'Emploi et de la Solidarité en février 1998, s'est, par un avis du 3 mars 1998, prononcée favorablement sur l'opportunité de mettre en place une déclaration à caractère obligatoire de l'infection à VIH "de manière strictement anonyme".

Constatant qu'aujourd'hui, "la connaissance épidémiologique de la maladie, son implantation géographique et sociale, son évolution, sa sensibilité aux mesures thérapeutiques et prophylactiques ne sont, pour le moment, que globalement appréciées par les seules données recueillies au cours de cette phase ultime", c'est-à-dire au stade du sida avéré, l'Académie de médecine a en effet estimé que la déclaration obligatoire du seul sida avéré, utile et efficace en son temps, est devenue insuffisante, négligeant le nombre beaucoup plus important des sujets infectés et qu'il convenait en conséquence d'élargir la déclaration obligatoire à l'ensemble des contaminés sérologiquement décelés qui conduirait à une meilleure connaissance de l'histoire naturelle de la maladie et de l'efficacité des traitements. L'Académie a enfin considéré que cette déclaration permettrait, en outre, aux pouvoirs publics d'adapter les actions de prévention, de dépistage et de prise en charge.

L'avis favorable rendu le 29 avril 1998 par le Conseil Supérieur d'Hygiène Publique de France, instance placée auprès du Ministre de la Santé apporte, au-delà des constats d'insuffisance du système de déclaration existant sur le sida, un éclairage sur d'autres points. En premier lieu, s'agissant du circuit de collecte des informations qui devront être recueillies, il indique, et c'est ce que le législateur retiendra dans la loi du 1er juillet 1998, que le système de recueil des données devrait associer, pour être réellement efficace et exhaustif, tant les laboratoires que les médecins. Il considère, par ailleurs, que le terme de "maladie à déclaration obligatoire" employé jusqu'à présent, notamment par le dernier texte de nature réglementaire ayant mis à jour la liste de ces maladies ( décret n° 86-770 du 10 juin 1986 ), devrait être révisé, "car il peut induire pour certains une confusion avec le terme "dépistage obligatoire"". Il insiste également sur la nécessité de procéder à une vaste campagne d'information et de sensibilisation, seule à même d'expliquer et de faire accepter le système. Enfin, il conclut sur le fait "qu'il sera nécessaire de veiller au respect des règles éthiques et de confidentialité concernant des données indirectement nominatives".

C'est à la suite de ces avis que le Ministre de la Santé a pris la décision au printemps 1998 d'inscrire la séropositivité parmi la liste des maladies à déclaration obligatoire.

Enfin, il doit être noté que, très récemment, un rapport relatif à la prise en charge thérapeutique des personnes infectées par le VIH établi à la demande du Ministère de la Santé par un groupe d'experts scientifiques<sup>17</sup> a également conclu, s'agissant de la nature de la surveillance épidémiologique du sida, à l'inefficacité d'une surveillance fondée sur les seuls cas de sida, et a recommandé en conséquence, pour permettre une nouvelle évaluation de l'épidémie de sida en France, d'étendre très rapidement le système de la déclaration obligatoire dès le stade de la séropositivité au virus.<sup>18</sup>

---

<sup>17</sup> Rapport établi sous la direction du Professeur DELFRAISSY.

<sup>18</sup> Les chiffres actuels sont en effet davantage le reflet d'une absence de dépistage et de prise en charge surtout chez les hétérosexuels. Ainsi, si les statistiques montrent que le nombre de nouveaux cas de sida diminue, cette situation est essentiellement due au fait de la connaissance de la séropositivité au VIH au moment du diagnostic et à la prise en charge précoce par des traitements antirétroviraux. Pour autant, ces

b - *Le projet de l'Institut de la Veille Sanitaire* -

Le projet de l'IVS tel qu'il a été présenté à la CNIL au cours du printemps 1999 consistait à expérimenter pour une durée d'un an l'enregistrement des cas de séropositivité au VIH déclarés afin de suivre les tendances et les caractéristiques récentes de l'épidémie, de mieux cibler les actions de prévention et d'aider à la planification des ressources au vu d'une estimation minimum du nombre de personnes infectées.

Le processus de notification retenu prévoyait un double système de notification des cas de séropositivité associant les biologistes et les médecins afin de renforcer l'exhaustivité de la collecte et l'évaluation des données ainsi recueillies. Une note d'information individuelle devait être remise à chaque personne expliquant le principe de la déclaration obligatoire et rappelant les droits ouverts au titre de la loi du 6 janvier 1978.

Comme c'est le cas actuellement pour les cas de sida déclaré, il était prévu que les déclarations seraient transmises sous pli confidentiel au médecin inspecteur de la direction départementale des affaires sanitaires et sociales du lieu d'exercice. Après validation des données par le dit médecin inspecteur, afin notamment d'éviter les doublons et de faire correspondre le cas échéant, la déclaration d'un laboratoire et la déclaration du médecin<sup>19</sup>, le médecin inspecteur aurait adressé les déclarations sous pli confidentiel avec la mention secret médical au médecin épidémiologiste chargé de la surveillance du VIH à l'Institut de veille sanitaire ( IVS ).

Les informations qui devaient alors figurer sur le questionnaire et qui étaient destinées à faire l'objet d'un traitement automatisé par l'IVS étaient les suivantes : les initiales du nom et les initiales du prénom (le prénom complet figure sur le questionnaire mais seule l'initiale devait être enregistrée), le sexe, la date de naissance, la nationalité, le département ou pays de résidence, le code postal de domicile et le code Insee de la catégorie socio-professionnelle (la profession figure en clair sur le questionnaire). Devaient être également collectées des données propres à caractériser l'infection, la date de prélèvement de la sérologie positive, la date des sérologies antérieures, la date de contamination estimée, le motif de dépistage, le mode de transmission présumé et le pays d'origine du partenaire en cas de contamination hétérosexuelle probable. Enfin le nom du médecin déclarant et le lieu d'exercice ainsi que la date de déclaration devaient également figurer sur le questionnaire.

Il convient de noter que dans les centres de dépistage anonymes et gratuits, le prénom complet n'aurait pas été collecté, ni le code postal du domicile, remplacé par le département de domicile.

---

chiffres ne signifient pas que l'infection par le VIH est en régression en France.

<sup>19</sup> A cet effet, l'IVS mettra à la disposition des médecins inspecteurs des DDASS un logiciel pouvant faciliter le couplage des questionnaires issus des laboratoires et des prescripteurs.

A la suite des inquiétudes manifestées par les associations de défense des droits des malades atteints du sida notamment sur les risques de fichage des séropositifs, les pouvoirs publics ont donc décidé de revoir entièrement les modalités de mise en place de ce dispositif et en particulier les mesures de confidentialité.

Dans un communiqué de presse du 5 novembre 1999, établi à l'issue des consultations engagées avec l'Institut de la Veille Sanitaire et les associations de lutte contre le sida et de défense des droits de l'homme et dont le contenu a été confirmé par un courrier du 24 novembre 1999 adressé au Président de la CNIL, le Ministre de l'Emploi et de la Solidarité et le Secrétaire d'Etat à la Santé ont fait savoir que le futur système mis en oeuvre par l'IVS devra présenter toutes les garanties pour préserver l'anonymat des personnes séropositives. Il est ainsi envisagé une anonymisation à la source par codage informatique des données recueillies par les biologistes. Un comité de pilotage associant l'ensemble des acteurs du système devra définir les conditions de fonctionnement du système.

En outre, le décret n°99-632 du 6 mai 1999 fixant les modalités de transmission à l'autorité sanitaire des données individuelles sera modifié pour tenir compte de la réflexion intervenue depuis lors.

L'ensemble du dispositif devrait être opérationnel à l'été 2000.

### **III - LES RECOMMANDATIONS DE LA CNIL**

#### **A- QUELQUES CONSTATS ET RAPPELS S'IMPOSENT**

##### **1- Sur le principe même de la déclaration obligatoire**

Les pouvoirs publics ont inscrit la séropositivité au virus de l'immunodéficience humaine sur la liste des maladies à déclaration obligatoire. Il convient de relever que les différentes associations consultées par la CNIL sont dans leur grande majorité favorables au caractère obligatoire de la déclaration, quelques unes d'entre elles exprimant cependant le souhait que l'établissement de la déclaration soit soumis au consentement de la personne.

La décision prise par les pouvoirs publics d'inscrire la séropositivité sur la liste des maladies à déclaration obligatoire implique cependant non seulement l'obligation pour les médecins et les laboratoires de signaler les cas diagnostiqués mais également l'impossibilité pour le patient de s'opposer à cette collecte d'informations. Dès lors, le dispositif est, dans son principe même, incompatible avec le recueil du consentement préalable de la personne concernée avant toute collecte d'informations sur sa séropositivité.

En un mot, on ne peut tout à la fois se déclarer favorable à la déclaration obligatoire de séropositivité et favorable au principe du consentement de la personne.

Mais, l'impossibilité pour les personnes de pouvoir exercer un droit d'opposition rend d'autant plus impérieuse la nécessité d'une information claire et complète des personnes concernées sur le système mis en place, la mise en oeuvre de garanties de confidentialité, l'appréciation rigoureuse, au regard des nécessités de santé publique, de la pertinence des

informations collectées et de manière plus spécifique le sort à réserver aux résultats de tests de séropositivité pratiqués par les centres de dépistage anonymes et gratuits.

## **2 - Sur la nécessaire dimension européenne du système de surveillance épidémiologique**

Ainsi qu'il a été indiqué, un système européen de surveillance épidémiologique du sida a été mis en place, ce qui nécessite à l'évidence que le centre européen chargé du système qui - il convient de le rappeler, est géré par la France - dispose, de la part de l'ensemble des pays participant au système, d'informations répondant à des définitions et critères communs<sup>20</sup>.

Or actuellement, si la France est en mesure de communiquer au centre européen des données sur les cas de sida avéré, il n'en est pas de même pour la séropositivité au VIH. Il doit être relevé que les données actuellement traitées par le centre européen sont les suivantes : numéro séquentiel propre à chaque pays, pays, sexe, âge, date du diagnostic, département, statut vital, modes de contamination et caractéristiques de la pathologie.

## **3- Sur les fichiers de personnes séropositives**

Des fichiers directement nominatifs comportant des informations relatives à des personnes séropositives existent, qu'il s'agisse des dossiers médicaux tenus dans les services hospitaliers, les cabinets médicaux ou les laboratoires d'analyses, des fichiers de gestion de remboursement détenus par les organismes de sécurité sociale (aujourd'hui, toute personne découverte séropositive au VIH et qui suit un traitement médical est prise en charge à 100%) ou encore des fichiers constitués par les associations d'aide aux personnes atteintes par le VIH. Ces fichiers, dont la légitimité n'a jamais été mise en cause nécessitent, bien entendu, des mesures de protection particulières. Leur légitimité n'a cependant jamais été remise en cause. Mais il est vrai qu'il s'agit dans la plupart des cas, de fichiers qui ne sont pas spécifiques aux personnes séropositives au VIH.

Le système de surveillance épidémiologique des cas de séropositivité tel qu'il est envisagé suppose, lui, la constitution d'un fichier spécifique aux personnes séropositives. C'est cette spécificité, cette "spécialité" qui fait naturellement de ce fichier un fichier très "sensible".

Cependant, comme pour le dispositif des déclarations obligatoires de sida avéré, les traitements informatiques ne sont pas appelés à être directement "nominatifs" mais indirectement nominatifs dans la mesure - et dans la mesure seule - où ils comporteront les initiales du nom et du prénom, la date de naissance ainsi que, le cas échéant, le code postal de domicile et la catégorie socio-professionnelle. En outre, à l'égal des systèmes de déclaration déjà existants, les traitements d'informations doivent être gérés sous la responsabilité d'un établissement public spécifiquement

---

<sup>20</sup> Or, si la France participe effectivement au dispositif en signalant les cas de sida avéré, il n'en est pas de même aujourd'hui pour les cas de séropositivité.

chargé de cette surveillance, par des médecins et chercheurs dûment habilités et astreints au secret professionnel. Il est à cet égard important de noter qu'aucune dérive ayant entraîné une rupture de la confidentialité n'a été constatée depuis la mise en place en 1988 de la déclaration obligatoire du sida avéré. Les associations de malades le confirment : les risques de rupture de la confidentialité et de discrimination sociale ou professionnelle en raison de l'état de santé ont pu être constatés dans les relations employeur/employé ou dans le domaine des assurances ou de la banque, aucune ne l'a été à partir d'un "mésusage" des données traitées par l'IVS.

Mais si aucune dérive n'a été constatée jusqu'à présent, les techniques informatiques et la libéralisation de l'utilisation des moyens de cryptologie permettent aujourd'hui de concevoir et de diffuser des dispositifs de sécurité pouvant assurer par des techniques dites "d'anonymisation" une confidentialité renforcée des fichiers de recherche médicale.

## B - LES RECOMMANDATIONS

### **1- Une nécessaire clarification des objectifs de la surveillance épidémiologique**

Si le principe de la déclaration obligatoire ne paraît pas contesté<sup>21</sup>, il apparaît toutefois nécessaire que les pouvoirs publics clarifient les objectifs de ce recueil d'informations.

En effet, s'agit-il seulement d'obtenir, année par année, un état exhaustif du nombre de personnes séropositives en France permettant de connaître les tendances et l'évolution de l'épidémie de sida et d'évaluer, de façon globale, l'impact des actions de prévention ? Ou bien l'objectif est-il également d'instituer une véritable surveillance épidémiologique de l'évolution des cas d'infection par le VIH, du stade de la découverte de la séropositivité à l'apparition éventuelle du sida avéré ? Cette dernière hypothèse dans laquelle il deviendrait possible de mesurer, de façon fine, l'impact des actions thérapeutiques et de prévention nécessiterait alors un suivi des cas et en particulier un rapprochement avec le système des déclarations obligatoires des cas de sida.

Ce choix d'objectifs qu'il appartient aux pouvoirs publics de définir clairement au regard des impératifs de santé publique emporte des conséquences importantes tant sur la nature des données susceptibles d'être collectées que sur leur durée de conservation et les liens éventuels avec d'autres systèmes de surveillance. Il implique en conséquence des choix en termes de protection des données.

Si, en effet, il s'agit d'assurer un dénombrement annuel des cas de séropositivité pour disposer d'une "photographie" de l'infection en France, il n'est nul besoin de traiter d'informations qui puissent peu ou prou présenter un caractère nominatif. Il ne paraît pas non plus nécessaire de conserver sur une longue période ces données.

Si en revanche, la finalité du système des déclarations obligatoires, telle qu'elle semble se dessiner actuellement, est d'assurer un suivi épidémiologique des cas de séropositivité au virus du

---

<sup>21</sup> Encore que, comme on l'a vu, les points de vue exprimés par certaines associations soient sur ce point contradictoires.

VIH et de mesurer l'évolution de l'épidémie de sida en France, elle ne peut être réalisée que par la collecte de données qui doivent être, à la source, individualisées, afin de permettre de détecter les doublons, de vérifier les données auprès des professionnels de santé et, le cas échéant, d'apparier les données avec celles collectées dans le cadre du sida avéré.

Un tel choix nécessiterait de disposer de données suffisamment discriminantes, de telle sorte qu'on puisse distinguer un cas d'un autre, de conserver les informations pendant une durée plus longue.

A cet égard, et compte tenu notamment de l'allongement du délai aujourd'hui constaté entre la découverte d'une séropositivité et la déclaration d'un sida, les données devraient nécessairement être conservées longtemps, même si les positions exprimées par les associations sont divergentes. Il convient de noter que la durée de conservation par l'IVS des données indirectement nominatives collectées dans le cadre du système de la déclaration obligatoire de sida avéré est, aujourd'hui, de 10 ans.

Par ailleurs, le lien qui serait effectué, le cas échéant, avec les données contenues dans le fichier du sida déclaré, afin notamment d'évaluer le temps qui s'écoule entre la découverte d'une séropositivité et un éventuel passage au stade du sida ou encore de connaître le statut vital de la personne, nécessiterait d'adopter pour la séropositivité le même système d'identification que pour le sida.

Sur ce point, l'Institut de Veille Sanitaire, à partir des déclarations obligatoires de sida avéré dont il dispose, a tenté d'analyser l'impact d'un "appauvrissement"<sup>22</sup> des informations indirectement nominatives figurant sur les déclarations de sida, à savoir actuellement : les initiales du nom et du prénom, la date de naissance complète et le département de domicile.

Actuellement, le traitement de ces informations indirectement nominatives permet de repérer dans une proportion de plus de 99% les cas de doublons supposés. En revanche, la suppression des initiales conduirait à enregistrer dans la base 20% des doublons correspondant en fait à une seule et même personne. Ce pourcentage atteindrait plus de 75% si les initiales et le jour de naissance étaient supprimés.

Un "appauvrissement" trop important des données peut donc fausser de façon conséquente les statistiques et remettre en cause la fiabilité scientifique de la surveillance épidémiologique.

Scientifiquement, le recueil à la source des initiales du nom et du prénom ainsi que de la date de naissance, qui pourraient certes être transmis sous une forme codée, paraît dès lors indispensable. En outre, la mention sur la déclaration des coordonnées des médecins déclarants<sup>23</sup> qui confère, certes, un certain caractère nominatif à la déclaration, se justifie par la nécessité de

---

<sup>22</sup> Appauvrir des données, sur le plan statistique, ne veut pas dire, leur enlever toute valeur significative, mais recourir à des données moins discriminantes. Ainsi, par exemple, enregistrer l'âge plutôt que la date de naissance ou encore l'année de naissance sans collecter le jour et le mois de naissance.

<sup>23</sup> Selon l'IVS un tiers des déclarations de sida qui lui parviennent nécessite un retour auprès des médecins déclarants pour compléter les données manquantes ou vérifier telle ou telle information.

pouvoir, le cas échéant, vérifier les informations fournies et assurer ainsi des statistiques de qualité.

## **2 - Le renforcement des mesures de confidentialité : vers l'adoption d'un système dit "d'anonymisation" à la source :**

S'il est dans l'intérêt de la santé publique de disposer de statistiques fiables sur l'évolution de l'infection par le VIH en France et donc de faire appel à un système fondé sur le recueil à la source de données indirectement nominatives, il demeure que, eu égard à la sensibilité des informations traitées, toutes garanties doivent être prises pour assurer la confidentialité des informations et protéger l'anonymat des personnes, ainsi que le prescrit la loi sur la veille sanitaire.

Le recours à des techniques dites "d'anonymisation" à la source est de nature à répondre aux besoins particuliers de confidentialité que nécessite la mise en place d'un tel système en évitant que ne soient conservées tant au plan local que national des données qui, par recoupement, pourraient permettre l'identification des personnes.

Ainsi qu'il a été indiqué précédemment, certains algorithmes dits de "hachage"<sup>24</sup>, déjà utilisés et expertisés permettent de coder les nom, prénom et date de naissance du patient, c'est-à-dire de produire un numéro non signifiant et non identifiant à partir de ces trois informations. Toute transmission ultérieure d'informations accompagnées du même numéro codé, peut permettre de savoir qu'il s'agit du même patient sans qu'il soit possible de l'identifier ni directement ni indirectement. Ainsi, les données concernant une même personne peuvent être appariées sans que son identité n'ait été connue.

Mais si les pouvoirs publics s'orientent effectivement vers une telle solution, comme semble l'annoncer le communiqué de presse paru le 5 novembre dernier, confirmé par le courrier que le Ministère a adressé au Président de la CNIL le 24 novembre dernier, encore convient-il de déterminer celles des informations qui seront utilisées par l'algorithme - pour produire le numéro - et d'évaluer les conséquences de la mise en place d'un tel dispositif.

La procédure de codage des données d'identification pourrait être effectuée à la source par les biologistes qui transmettraient aux médecins prescripteurs, avec le résultat du test pratiqué, le formulaire épidémiologique comportant pour toute référence le code produit par l'algorithme, à charge pour les médecins de compléter les variables épidémiologiques de la notification. Un tel dispositif pourrait être de nature à lever, au moins en partie, les réserves exprimées par certaines associations auditionnées par la CNIL sur la participation des laboratoires au dispositif.

Mais au-delà de l'anonymisation à la source, des mesures de protection devraient être adoptées pour, à la fois, sécuriser les fichiers informatiques (notamment par un cryptage) et trancher le sort des déclarations papier en particulier au niveau des DDASS . La transmission aux médecins inspecteurs des DDASS des déclarations devrait ainsi s'effectuer, dans tous les cas, dans des conditions garantissant la confidentialité, telles que par exemple l'envoi sous pli confidentiel

---

<sup>24</sup> Ainsi, en est t'-il, du dispositif développé par le Centre d'Etudes des Sécurités du Système d'Information (CESSI) qui dépend de la Caisse Nationale d'Assurance Maladie des Travailleurs Salariés.

des déclarations au médecin chargé de la surveillance épidémiologique du sida. En outre, une fois le travail de validation effectué par les médecins au niveau départemental et les données transmises à l'IVS, aucune déclaration ne devrait être conservée.

Une attention particulière doit, enfin, être portée à la diffusion des statistiques établies à partir des données collectées et traitées afin d'éviter tout risque d'identification des personnes. En effet, toutes précautions doivent être prises pour que la diffusion de statistiques sur l'état de la séropositivité en France n'aboutisse pas à publier des données qui corrélées avec d'autres pourraient permettre d'identifier un individu.

### **3 - La nature sensible des autres informations :**

La mise en place d'une "anonymisation" à la source et de mesures de sécurité adéquates ne dispense pas de s'interroger sur la pertinence des autres informations appelées à figurer sur la déclaration de séropositivité.

Il s'agit, en particulier du code postal de résidence, de la profession et de l'origine géographique.

#### **- Le code postal de domicile -**

Cette variable qu'il était initialement prévu de mentionner sur la déclaration de séropositivité était présentée comme ayant pour objet de permettre aux médecins inspecteurs des DDASS de mieux cibler les actions de prévention locale en fonction, notamment, de l'importance relative du nombre de personnes séropositives. Par exemple, la connaissance, dans une zone géographique donnée, d'un nombre relativement important de toxicomanes séropositifs pourrait conduire les autorités sanitaires à décider de l'implantation dans cette zone de distributeurs de seringues stériles.

Mais force est de constater que la pertinence du recueil de cette donnée, notamment au plan national, n'est pas à ce jour réellement démontrée. En outre, sa collecte et son exploitation pourraient être de nature à permettre une localisation géographique précise des cas de séropositivité, surtout dans les communes de moyenne importance, avec tous les risques de rupture de la confidentialité et de stigmatisation qui pourraient en résulter.

Dès lors le recueil sous une forme aussi détaillée que le code postal de la commune de résidence des personnes séropositives peut paraître excessif au regard des objectifs recherchés.

#### **- La profession -**

Cette variable peut présenter un intérêt dans la mesure où elle constitue une indication de la situation sociale des personnes séropositives. S'il peut effectivement être utile sur le plan épidémiologique de mieux connaître les catégories de situation socio-professionnelles des personnes et en particulier la proportion de chômeurs, d'étudiants, ou de cadres ... afin de mieux cibler les actions de prévention, il ne paraît pas nécessaire de disposer du détail de la profession précise des personnes concernées. Dès lors, seule la mention des catégories socio-professionnelles selon la nomenclature de l'INSEE ou d'une nomenclature inspirée de celle-ci paraît pertinente.

L'adoption d'une liste simplifiée des catégories socio-professionnelles - la nomenclature de l'INSEE comportant plus de trente catégories - aurait le mérite de simplifier la tâche du médecin chargé de compléter le formulaire, l'assimilation d'une profession à une catégorie socio-professionnelle donnée n'allant pas toujours de soi.

- L'origine géographique -

Le formulaire de déclaration initial prévoyait la collecte d'une information sur l'origine géographique de la personne séropositive et du partenaire sexuel en distinguant l'Afrique Sub-Saharienne et les Caraïbes des autres pays. Cette formulation aurait pu donner à penser que c'est l'origine "raciale" des personnes ou de leurs partenaires qui est recherchée.

Tel n'est pas en réalité l'objet du recueil de cette donnée. Il s'agit, de déterminer sur le plan épidémiologique si les personnes ont vécu ou séjourné dans des pays qui connaissent une forte épidémie de sida et où la transmission hétérosexuelle du virus est actuellement prédominante.

Une rubrique de même nature figure d'ailleurs parmi les données transmises au centre européen pour la surveillance épidémiologique du sida. Cependant, dans ce dernier cas, il n'est pas fait référence à telle ou telle zone géographique du monde ; il est demandé si la personne est originaire d'un pays où la transmission hétérosexuelle est prédominante ou a eu des relations sexuelles avec une personne originaire ou ayant vécu dans un pays où la contamination hétérosexuelle est prédominante. La même rubrique figure également sur les déclarations obligatoires de sida avéré.

Les différentes associations auditionnées par la CNIL ont, dans leur majorité, exprimé leur opposition à la collecte de l'information telle qu'elle figure actuellement sur les déclarations.

La Commission ne conteste pas l'utilité épidémiologique d'une telle information. Il lui apparaît toutefois nécessaire dans le souci de lever toute ambiguïté sur sa signification, de faire exclusivement référence à la notion de séjour dans un pays ou zone géographique connaissant une prédominance de la transmission hétérosexuelle du virus du sida, une liste indicative de ces pays ou zones d'origine pouvant être jointe afin de faciliter l'établissement de la déclaration.

#### **4 - L'anonymat, raison d'être des centres de dépistage anonymes et gratuits**

L'activité des centres de dépistage anonymes et gratuits a connu depuis leur mise en place en 1987 une progression constante. Aujourd'hui, une personne sur cinq se fait dépister dans un CDAG. Dans ce cas, les personnes n'ont pas à révéler ni leur nom, ni leur adresse. Toute personne qui souhaite bénéficier d'un dépistage anonyme dans ces centres se voit attribuer, lors de sa venue un numéro séquentiel reportée sur une fiche, qu'elle devra présenter à nouveau lors de la remise des résultats de sa sérologie.

L'instauration d'une déclaration obligatoire fondée sur une l'identification même indirecte de la personne, et fut-elle codée, est, par définition, contraire au principe de l'anonymat des centres de dépistage anonymes et gratuits.

Inclure les CDAG dans le dispositif de la déclaration obligatoire de séropositivité serait, d'une part, prendre le risque de briser la relation de confiance instaurée entre le professionnel de santé et le patient, d'autre part remettre en cause le droit pour la personne de choisir le lieu et les conditions de son dépistage, peut-être et surtout, la dissuader de se faire dépister.

Or, l'exclusion des CDAG du dispositif ne paraît pas de nature à nuire à la recherche épidémiologique dans la mesure où, aux termes des informations recueillies par la Commission, les personnes qui se révèlent séropositives à l'issue d'un test pratiqué dans un CDAG, sont systématiquement orientées vers un service de soins. C'est à ce stade, et à ce stade seulement, dans une deuxième étape, que la séropositivité de la personne concernée pourra être déclarée. Elle ne doit pas l'être dans les CDAG.

Telle est d'ailleurs la pratique dans les pays qui disposent de tels centres, le dispositif des déclarations obligatoires ne s'y applique jamais.

### **5 - La nécessité pour les pouvoirs publics d'autoriser par décret en Conseil d'Etat pris après avis conforme de la CNIL, la collecte de données susceptibles de révéler les moeurs de la personne.**

Le formulaire de déclaration de la séropositivité au VIH prévoit le recueil de données sur les modes de contamination et en particulier sur les relations sexuelles de la personne. Ainsi, seraient collectées des données sur la nature de ces relations : homosexuelles, hétérosexuelles ou bisexuelles.

Or, le recueil à des fins de surveillance épidémiologique des modes de contamination sexuels revêt un intérêt public incontestable pour la recherche dans la mesure où il permet de caractériser l'évolution de l'épidémie. Il convient donc de permettre l'enregistrement et la conservation de ces informations dans le fichier dès lors qu'il serait mis en oeuvre dans le cadre de l'exercice des missions confiées à l'IVS et dans le respect des dispositions de la loi du 6 janvier 1978.

Il pourrait être soutenu que certaines données ayant fait l'objet d'un codage par algorithme, le fichier épidémiologique serait dépourvu de tout caractère directement ou indirectement nominatif.

Une telle argumentation ne saurait être retenue pour deux raisons : en premier lieu, un certain nombre d'informations en clair dans le fichier confère au traitement un caractère très indirectement - certes - mais indirectement nominatif. Tel est le cas des coordonnées du médecin déclarant, de la catégorie socio-professionnelle et, le cas échéant, du code postal de domicile, s'il était maintenu. En deuxième lieu et surtout malgré son appellation, l'algorithme dit "d'anonymisation" n'anonymise pas, comme on pourrait l'entendre au sens commun, il est plus exact de dire qu'il code en créant un numéro à partir des initiales des nom et prénom et de la date de naissance.

Toute éventuelle application de l'algorithme aux initiales et à la date de naissance d'une personne déterminée produira le même numéro et ce numéro permettrait, le cas échéant, de rechercher dans le fichier si la personne y figure ou non et à quel stade d'évolution du virus elle se situe.

A ces deux égards, il n'est pas contestable que malgré les mesures de sécurité prises, le fichier conserve un caractère indirectement nominatif.

A ce titre, et enregistrant des données sur les moeurs des personnes, le fichier relève des dispositions de l'article 31 de la loi du 6 janvier 1978.

## **6 - La nécessité d'assurer une plus grande transparence dans la mise en place du dispositif**

La réflexion entreprise par la CNIL démontre l'utilité d'une concertation avec tous les acteurs du dispositif. En effet, la très grande et légitime sensibilité manifestée par l'opinion publique à l'égard de la mise en place de fichiers dans le domaine du sida rend indispensable la plus grande transparence tant à l'égard des patients que dans le cadre des relations qu'ils entretiennent avec le corps médical.

Le caractère obligatoire de la déclaration de l'infection par le virus de l'immunodéficience humaine, quel que soit le stade de la maladie, doit être accompagné d'une information très claire des personnes sur les objectifs et les modalités du dispositif de sorte que chacun puisse être pleinement convaincu à la fois de l'importance majeure que revêt, en termes de santé publique, la lutte contre l'épidémie de sida et que le fichier épidémiologique mis en place par l'IVS l'est dans des conditions garantissant, au-delà de la confidentialité des données recueillies, le respect de la vie privée de chacun.

Telles sont les recommandations de la Commission.