

# Indicateurs de suivi de l'évolution des inégalités sociales de santé dans les systèmes d'information en santé

Collection  
*Documents*

# Indicateurs de suivi de l'évolution des inégalités sociales de santé dans les systèmes d'information en santé

Ce rapport répond à deux saisines de la DGS et la DREES, portant d'une part sur la pertinence et les possibilités de prise en compte des caractéristiques sociales dans les indicateurs transversaux et spécifiques associés aux objectifs de la loi de santé publique, et sur la hiérarchisation des indices écologiques de défavorisation sociale ; d'autre part sur la proposition d'indicateurs permettant « d'assurer le suivi des inégalités de santé dans le domaine du cancer, de ses facteurs de risque et de ses conséquences, notamment en termes de mortalité ».

Le HCSP analyse les éléments disponibles pour rendre compte de l'évolution des inégalités sociales de santé et propose des indicateurs à insérer dans les systèmes d'information en santé, ainsi que des évolutions à apporter dans ces systèmes d'information, de façon à prendre en compte ces inégalités sociales de santé dans les différents types d'usages des données de santé : systèmes de recueil de données en routine (mortalité, morbidité, institutions de soins) ; enquêtes, études et travaux de recherche ; bilans des politiques de santé.

La première partie est générale et concerne l'ensemble des états de santé et pathologies. Dans la seconde partie, qui concerne le domaine du cancer, les objectifs généraux sont plus spécifiquement appliqués aux systèmes d'information ou sources de données mobilisables sur le cancer, afin de pouvoir disposer de données suffisantes à la conduite des politiques de réduction des inégalités.



# **Indicateurs de suivi de l'évolution des inégalités sociales de santé dans les systèmes d'information en santé**

Rapport

Juin 2013

## Sommaire

Membres du groupe de travail.....	3
Contexte du rapport.....	4
Objectifs .....	5
PREMIERE PARTIE : Prise en compte des caractéristiques sociales dans les indicateurs transversaux et spécifiques dans le cadre de la réflexion en vue de la refonte des indicateurs associés aux objectifs de la loi de santé publique du 9 août 2004 .....	6
Introduction .....	6
1. DONNEES INDIVIDUELLES .....	8
Une triade essentielle .....	8
Le cas spécifique des jeunes.....	13
Les caractéristiques démographiques.....	13
Le statut vis-à-vis de la couverture du risque maladie .....	16
Les déterminants socio-économiques mobilisables pour l'étude des inégalités sociales de santé.....	17
2. INDICATEURS ECOLOGIQUES DE DEFAVORISATION SOCIALE .....	18
Avantages et limites .....	18
Quelques indices de défavorisation sociale utilisés hors de la France.....	20
Les principaux indices de défavorisation sociale utilisés en France .....	21
3. UTILISATION DES BASES DE DONNEES ADMINISTRATIVES .....	26
4. SYNTHESE : MISE EN OEUVRE DES METHODES SELON LES OBJECTIFS .....	31
Trois grandes approches pour caractériser les inégalités sociales de santé.....	31
Des approches variables selon les objectifs .....	31
5 RECOMMANDATIONS.....	34
Variables et méthodes pour caractériser les inégalités sociales de santé .....	34
Mise en œuvre des méthodes selon les objectifs.....	36
DEUXIEME PARTIE : Suivi des inégalités dans le domaine du cancer .....	38
1. Les mesures du plan cancer 2009-2013 qui nécessitent un SIS capable de mesurer les inégalités sociales de santé.....	39
2. Un système d'information sanitaire spécifique au cancer.....	39
3. Recommandations .....	41
REFERENCES.....	42
ANNEXE : .....	50

Ce rapport a été préparé et rédigé par un groupe de travail issu du groupe transversal du HCSP sur les inégalités sociales de santé et validé par le Comex le 19 juin 2013.

## **MEMBRES DU GROUPE DE TRAVAIL**

Thierry Lang, HCSP, président du groupe

Chantal Cases, HCSP

Pierre Chauvin, HCSP

Eric Jouglu, HCSP

Florence Jusot, économiste, Rouen

Anne Laporte, épidémiologiste, Paris

Pierre Lombrail, médecin de santé publique, Bobigny

Gwenn Menvielle, épidémiologiste, Villejuif

### **Relecture**

Virginie Halley des Fontaines, HCSP

Isabelle Grémy, épidémiologiste, Paris

Sylvie Rey, épidémiologiste, Paris

Béatrice Tran, SGHCSP

Laurent Visier, HCSP

Denis Zmirou-Navier, HCSP

## CONTEXTE DU RAPPORT

Ce rapport est une réponse du HCSP à deux saisines

1/ Dans le cadre de la réflexion menée sous forme d'une large consultation en vue de la refonte des indicateurs associés aux objectifs de la loi de santé publique du 9 août 2004, la DGS et la DREES ont saisi le HCSP en août 2012 sur deux aspects : d'une part pour « *déterminer la pertinence et les possibilités de prise en compte des caractéristiques sociales dans les indicateurs transversaux et spécifiques [en cours de mise à jour] à partir notamment des éléments de synthèse issus de la première phase de consultation et, d'autre part, de hiérarchiser les indices de défavorisation sociale en mobilisant les travaux récents, dans la suite des conclusions de [ses] rapports sur les inégalités sociales de santé et sur les systèmes d'information en santé en 2010.* »

2/ Le plan cancer 2009-2013 prévoit dans sa mesure 8.1, relative au développement de l'épidémiologie sociale des cancers, de « *missionner le HCSP pour proposer la définition d'indicateurs pertinents, compte tenu des systèmes de recueil de données existants : pour assurer le suivi des inégalités de santé dans le domaine du cancer, de ses facteurs de risque et de ses conséquences, notamment en termes de mortalité. Le HCSP caractérisera les inégalités territoriales d'accès aux services de prévention et de prise en charge des cancers et suivra leur évolution en fonction des actions menées par les agences régionales de santé. (La mesure 8.1 s'applique aux indicateurs des grandes enquêtes nationales telles celles de [l'IRDES] (ESPS), l'enquête décennale de l'INSEE, du Baromètre cancer.* »

La DGS a saisi le HCSP en décembre 2010 pour mettre en œuvre cette action.

### **Deux saisines qui traduisent en termes opérationnels des recommandations antérieures**

L'un des axes transversaux innovants du plan cancer 2009-2013 est la prise en compte des inégalités sociales de santé. Aussi, pour achever cet objectif et permettre d'en suivre les évolutions, il est nécessaire de se doter d'outils de surveillance des déterminants sociaux et des effets différenciés des mesures du plan au sein des différentes catégories de la population.

D'autre part, le HCSP a formulé en 2010 plusieurs propositions pour prendre en compte les inégalités sociales de santé dans les politiques publiques, parmi lesquelles figuraient : le suivi des indicateurs associés aux objectifs de la loi de santé publique par catégorie socio-économique lorsque cela était pertinent (proposition 3), la production régulière de données sur les inégalités sociales de santé en France (proposition 4), nécessitant d'intégrer ou de récupérer des données sociales dans les enquêtes et systèmes d'information en santé ; enfin de développer la production de données à un échelon géographique fin, ainsi qu'un indicateur de défavorisation territorial au niveau des Iris et d'étendre les extensions régionales des grandes enquêtes (proposition 5)<sup>1</sup>. La saisine de la DGS et de la DREES remet en quelque sorte ces propositions dans un cadre opérationnel, sous-tendu par la refonte des indicateurs associés aux objectifs de la loi de 2004.

---

<sup>1</sup> HCSP. Les inégalités sociales de santé : sortir de la fatalité [Internet]. Haut conseil de la santé publique; 2009 déc. Consulté de: [http://www.hcsp.fr/docspdf/avisrapports/hcspr20091112\\_inegalites.pdf](http://www.hcsp.fr/docspdf/avisrapports/hcspr20091112_inegalites.pdf)

## OBJECTIFS

- Définir les indicateurs liés aux déterminants sociaux de la santé, les plus pertinents et utilisables en routine, dans la construction de systèmes d'information permanents aussi bien que dans la mise en place d'enquêtes ou de systèmes d'observation ponctuels, mais dont les résultats sont susceptibles d'être mis en perspective avec d'autres.
- Faire un état des lieux et hiérarchiser les indices écologiques de défavorisation sociale.
- Déterminer, en étroite collaboration avec le groupe de travail sur les systèmes d'information en santé, les évolutions nécessaires des systèmes d'informations qui permettraient de recueillir les données sociales.

Ce rapport est organisé en deux parties.

L'une est générale et concerne l'ensemble des états de santé et pathologies. Elle s'inscrit dans le cadre de la réflexion en vue de la refonte des indicateurs associés aux objectifs de la loi de santé publique du 9 août 2004 et vise à apporter les éléments de réponse à la saisine de la DGS et la DREES sur deux aspects : d'une part pour déterminer les possibilités de prendre en compte les caractéristiques sociales dans les indicateurs transversaux et spécifiques en cours de mise à jour ; d'autre part de hiérarchiser les indices de défavorisation sociale existants.

La seconde partie concerne la saisine portant sur le domaine du cancer. Les objectifs généraux sont plus spécifiquement appliqués aux systèmes d'information ou sources de données mobilisables sur le cancer, afin de pouvoir disposer de données suffisantes à la conduite des politiques de réduction des inégalités.

Il nous a semblé utile de présenter cette réflexion en deux étapes, car les indicateurs dans le domaine du cancer partagent un très grand nombre d'exigences méthodologiques et nécessitent des évolutions des systèmes d'information sanitaire comme l'ensemble des autres pathologies. Dans la seconde partie ne seront donc traités que les aspects spécifiques au cancer, qui doivent bénéficier par ailleurs de la réflexion d'ensemble présentée dans la première partie.

# **PREMIERE PARTIE : Prise en compte des caractéristiques sociales dans les indicateurs transversaux et spécifiques dans le cadre de la réflexion en vue de la refonte des indicateurs associés aux objectifs de la loi de santé publique du 9 août 2004**

## **INTRODUCTION**

Ce rapport fait suite à deux rapports du HCSP, l'un sur les systèmes d'information en santé, qui soulignait qu'« Il manque en France un véritable système de surveillance des inégalités sociales de santé »<sup>2</sup> et l'autre sur les inégalités sociales de santé (ISS)<sup>3</sup> qui indiquait trois pistes ou axes de travail par rapport aux ISS : 1) se fixer un objectif explicite de réduction ; 2) mesurer les ISS ; 3) agir. Ces objectifs découlaient du fait que la capacité de mesurer est un préalable à l'action et la première traduction de l'objectif. Pour mémoire, cette insuffisance de données pour surveiller, comprendre et agir sur les inégalités sociales de santé était déjà constatée et déplorée il y a 10 ans dans un chapitre du rapport de 2002 sur la santé en France du Haut Comité de la Santé Publique<sup>4</sup>.

Ce rapport soulignait que disposer d'informations fiables sur les inégalités sociales de santé, leur évolution et la surveillance des déterminants sociaux de la santé suppose de définir les variables à recueillir, mais nécessite également une réflexion sur les systèmes d'information et de recueil de données en France, devant se traduire par une évolution significative du système d'information statistique. Il est donc temps de passer, dans ce domaine, du constat à l'action et de recueillir des indicateurs pertinents.

Pour atteindre ces objectifs généraux, deux pistes complémentaires seront poursuivies :

- dans un but de surveillance, proposer des variables, indicateurs et systèmes d'information destinés à suivre le niveau et l'évolution des inégalités dans différents domaines (mortalité, morbidité, qualité de vie...) ;
- pour éclairer les choix de politiques et les évaluer, le même travail devra s'appliquer aux déterminants sociaux de la santé dont il faudra pouvoir suivre les évolutions.

Plusieurs préoccupations ont guidé le travail de ce groupe.

---

<sup>2</sup> HCSP. Groupe de travail présidé par M Goldberg. Les systèmes d'information pour la santé publique. La Doc. Française. Paris, 2009.

<sup>3</sup> HCSP. Groupe de travail présidé par T Lang. Inégalités sociales de santé : sortir de la fatalité. La Doc. Française. Paris, 2009.

<sup>4</sup> HCSP. Groupe de travail présidé par J Lebas. Les inégalités et disparités de santé. Paris, La Documentation française, 2002, pp. 163-243.

Les indicateurs retenus et leur recueil doivent prévoir la possibilité de comparer le niveau des inégalités sociales de santé dans l'espace (y compris en matière de comparaisons internationales) et dans le temps (évolutions).

En bonne pratique épidémiologique, les mesures choisies doivent correspondre à un objectif. Les recommandations varient donc selon les objectifs poursuivis par les systèmes d'information. Il s'agit de distinguer clairement les informations et indicateurs qu'il est indispensable de recueillir dans les systèmes d'information permanents existants et celles qui peuvent être recueillies dans des enquêtes, des travaux de recherche ou des études, dans lesquels le recueil des données peut-être plus approfondi.

Plusieurs échelons d'indicateurs sont nécessaires au pilotage d'une politique de lutte contre les inégalités sociales de santé. Le groupe a donc inclus dans sa réflexion la nécessité de disposer d'informations au niveau national, mais également au niveau local et territorial. Naturellement, la réflexion est différente selon qu'il s'agit de disposer d'informations sur les inégalités sociales de santé au niveau de la population ou au niveau des institutions sanitaires (hôpitaux, système de soins ambulatoires...).

Le groupe s'est attaché à distinguer ce qui dans ses recommandations paraissait faisable à court terme de ce qui était réalisable dans une situation à plus longue échéance. Il a tenu à dégager un certain nombre de solutions pragmatiques, tenant compte de la faisabilité et de la nécessité de disposer d'informations utilisables. Certaines propositions concernent des mesures pouvant être disponibles relativement rapidement, d'autres nécessitent des refontes techniques du système d'information sanitaire et/ou une réflexion juridique.

Le texte consacré aux bases de données administratives s'est inspiré très largement des travaux produits par le groupe système d'information sanitaire du Haut Conseil de la Santé Publique, présidé par Marcel Goldberg.

Inégalités sociales et territoriales de santé, sont à l'évidence liées, mais dans les réflexions sur les inégalités sociales de santé, le lien avec le territoire n'est pas toujours très clair, alors que nombre d'interventions sur les inégalités sociales de santé s'inscrivent sur une base territoriale. En revanche, les approches par territoire occultent parfois la question des inégalités induites par les différences sociales. A titre d'exemple, la mesure 14 du plan cancer 2009-2013 vise à lutter contre les inégalités de recours au dépistage ; mais l'indicateur chiffré invite à cibler les « départements rencontrant le plus de difficultés ». A l'inverse, une approche des inégalités sociales de santé non liée au territoire peut les minorer, si par exemple l'analyse de l'accès aux soins et des obstacles financiers n'intègre pas aussi la dimension géographique d'accès au secteur 1 dans un territoire.

Dans une première partie, nous aborderons successivement trois types d'information permettant de documenter les inégalités sociales de santé :

-Le premier est de proposer des variables individuelles qui pourraient être recueillies en routine dans les systèmes d'information sanitaires (SIS) ou recueillies plus systématiquement que c'est le cas aujourd'hui dans les grandes études, les enquêtes et les travaux de recherche d'épidémiologie et de santé publique.

-Le deuxième présente la possibilité d'approximer les caractéristiques individuelles à partir des indices écologiques de défavorisation sociale de la zone géographique de résidence des personnes.

-Le troisième traite de l'utilisation des bases de données administratives qui permettent de caractériser la situation socio-économique individuelle ou qui pourraient être enrichies dans cet objectif. Cette orientation a déjà été développée dans un précédent rapport<sup>5</sup>, mais nous le compléterons de propositions complémentaires.

Dans un deuxième temps, une synthèse permettra de préciser quelles variables et sources de données peuvent être mobilisées selon les objectifs visés.

## **1. DONNEES INDIVIDUELLES**

La situation socio-économique (SSE) d'un individu est une notion multifactorielle dont la caractérisation nécessite le recueil de plusieurs dimensions et ne peut pas être résumée par un seul indicateur. C'est également une notion dynamique, même si elle ne peut être caractérisée qu'à des moments donnés. Si certaines de ses dimensions peuvent être considérées comme acquises une fois pour toutes (le genre, le niveau d'éducation), la plupart renvoient de fait à des situations largement changeantes au cours de la vie d'un individu. Enfin, les différentes caractéristiques décrivant la situation socio-économique d'une personne sont souvent corrélées entre elles (ce qui ne veut pas dire qu'elles le sont complètement et que les unes peuvent se substituer aux autres). Par exemple, le niveau d'éducation détermine dans une certaine mesure, non seulement la profession d'un individu, mais également sa probabilité d'occuper un emploi et par là (au moins partiellement) ses revenus ou sa couverture maladie. De la même façon, son niveau d'éducation, son milieu professionnel et son entourage ont une incidence sur l'adoption de certains comportements de santé. Si aucun de ces facteurs (ni individuellement, ni pris tous ensemble) ne détermine à 100% un comportement de santé, ni a fortiori une maladie ou un événement de santé, chacun d'entre eux peut avoir des effets propres mais peut aussi entrer en interaction avec d'autres. Le plus souvent, c'est bien le constat ou l'étude de ces cumuls de facteurs sociaux qui sont pertinents et signifiants d'un point de vue de santé publique.

Dans le texte qui suit, sachant que tout recueil de variables complémentaires a un coût, certaines variables seront présentées comme essentielles, indispensables à tout suivi des inégalités sociales de santé. D'autres variables seront présentées comme complémentaires, permettant de mieux comprendre les phénomènes.

### **UNE TRIADE ESSENTIELLE**

Classiquement, le statut socioéconomique d'un individu est interrogé – au minimum – par une triade « éducation, revenus, situation professionnelle ». Le lien entre santé et position socio-économique, en utilisant ces trois mesures, reste un des résultats les plus cohérents et constants de la littérature épidémiologique. L'analyse par sexe est aujourd'hui essentielle, tant les inégalités sociales de santé affectent différemment, dans nombre de domaines, les

---

<sup>5</sup> HCSP. Groupe de travail présidé par M Goldberg. Les systèmes d'information pour la santé publique. Op. cit.

hommes et les femmes. Les classes d'âge doivent être adaptées au cycle de vie des personnes, qui peut être différent d'un groupe social à l'autre et qui peut changer, globalement, avec les transformations et les évolutions sociétales (entrée dans la vie active, âge de la retraite, proportion et qualité de vie des personnes âgées...). Être capable d'identifier les populations immigrées et issues de l'immigration sera également un enjeu important, pour faire la part, dans les inégalités sociales observées, des origines socioéconomiques (liées aux conditions de vie), culturelles, en lien avec des phénomènes de discrimination (liés au patronyme, au lieu de résidence, à la couleur de peau, ...) ou encore les conditions de migration elles-mêmes.

### *Niveau d'instruction*

Le niveau d'instruction est utilisé comme une approximation et un reflet des ressources cognitives et intellectuelles d'un individu. En particulier, on considère que le niveau de connaissances et de compétences acquises pendant les années de formation permet d'être plus réceptif aux messages et recommandations de santé, d'intégrer plus facilement les normes médicales et d'avoir un accès plus facile aux services de santé. On considère également que l'éducation apporte les ressources nécessaires pour être en capacité (et avoir le sentiment) de pouvoir, au moins partiellement, contrôler sa vie, de la maîtriser dans un environnement de contraintes données, ce qui constitue la base de comportements ultérieurs, notamment dans le domaine de la santé.

Le niveau d'instruction - en tant que source de capital humain- détermine aussi largement les revenus d'activité de la personne et traduit donc également l'effet direct des revenus sur la santé, en termes notamment d'accès financier aux soins. A l'inverse, les effets du niveau d'instruction sur la situation économique et la santé dépendent de la classe d'âge, du sexe et de la position sociale. Le niveau d'instruction est en partie conditionné par la situation sociale des parents (ou de la famille d'origine) et plus spécialement des mères et détermine dans une large mesure la trajectoire professionnelle (et financière) ultérieure. Au niveau populationnel, le niveau d'instruction est un bon prédicteur de morbidité et mortalité.

L'éducation est probablement l'indicateur le plus utilisé dans les enquêtes épidémiologiques et données de surveillance parce qu'il présente différents avantages : il est facile à obtenir, il est stable tout au long de la vie (une fois les études terminées), il ne dépend pas de la situation actuelle sur le marché du travail et permet donc d'apprécier le statut socioéconomique des personnes inactives, il n'est plus affecté par l'état de santé ultérieur et il est facilement comparable dans les études internationales (voir par exemple [Mackenbach et al, 2007]). En pratique, de nombreuses études interrogent ou analysent le niveau d'éducation en 3 classes : primaire, secondaire, supérieur (ce qui a, de plus, l'avantage de correspondre à des expressions courantes en français : « Avez-vous fait des études supérieures ? » ...). Le niveau secondaire peut être lui-même éventuellement subdivisé en 2 classes (secondaire avant et après le brevet). Ces subdivisions en 3 ou 4 groupes sont des regroupements des classes plus détaillées de la Classification internationale type de l'éducation (CITE, [Unesco, 2011]).

Une des limites de cet indicateur est que, d'une génération à l'autre, l'accroissement continu du nombre d'années d'étude, induit une homogénéisation du niveau de formation dans la population. Ainsi, son interprétation pour suivre l'évolution des inégalités sociales de santé sur de longues périodes doit donc être particulièrement précautionneuse.

### *Situation professionnelle : distinguer la situation vis-à-vis de l'emploi et la profession*

La situation professionnelle peut avoir plusieurs conséquences sur la santé : directement, par les expositions professionnelles (physiques, chimiques, mais également psychosociales [Niedhammer, 2006, Leclerc et al, 2011]), et indirectement par les ressources qu'elle procure : financières, matérielles mais également en termes de sociabilité ou de statut social et identitaire.

La situation professionnelle, à un instant donné, comprend deux premières dimensions distinctes, qu'il est intéressant d'interroger successivement : la situation vis-à-vis de l'emploi et la profession.

La plupart des études utilisent la situation professionnelle en cours, plus rarement celle qui a été la plus longue. Pour les personnes qui ne sont pas en emploi, on renseigne la dernière profession occupée (ou, plus rarement, la plus longue). Dans certaines situations, il peut être intéressant de renseigner également la profession du conjoint ou de la « personne de référence du ménage » si la personne interrogée n'est ni cette personne, ni son conjoint (par exemple, dans le cas d'un enfant adulte vivant chez ses parents ou d'un retraité vivant chez ses enfants). On prendra en compte ensuite, au choix, la situation professionnelle de la personne interrogée, celle de la personne de référence du ménage, la plus haute des deux et/ou conjointement la combinaison de celle de la personne interrogée et celle de son conjoint.

En pratique, une première question interroge la situation vis-à-vis de l'emploi et distingue les personnes :

- en emploi (quel que soit le type et le statut d'emploi : salarié, à son compte, déclaré ou non)
- en apprentissage sous contrat ou en stage rémunéré
- étudiant(e) ou en stage non rémunéré
- au chômage (inscrit ou non à Pole emploi, percevant ou non des allocations chômage)
- retraité(e) ou préretraité(e)
- au foyer
- dans une autre situation

Une seconde question interroge la profession - ou la dernière profession - occupée, soit en clair (ce qui nécessitera un recodage ultérieur), soit en utilisant une classification existante. En France, la classification de référence est, depuis 1982, celle des « professions et catégories sociales » (PCS) de l'INSEE qui comporte trois niveaux d'agrégation emboîtés : les groupes socioprofessionnels (en 8 catégories, y compris les retraités et les personnes sans activité professionnelle), les catégories socioprofessionnelles (en 2 versions qui comprennent respectivement 24 ou 42 catégories) et les professions (en 486 catégories). Pour des raisons de commodité, le groupe socioprofessionnel est le plus utilisé.

On conseille fortement de renseigner la dernière profession exercée, pour les retraités (contrairement à l'usage introduit par l'INSEE) comme pour les chômeurs ou tous les inactifs au moment de l'enquête. La seconde question peut alors se formuler ainsi : Quelle est votre profession (ou la dernière profession que vous avez exercé) ? :

- Agriculteurs exploitants
- Artisans, commerçants et chefs d'entreprise
- Cadres et professions intellectuelles supérieures
- Professions intermédiaires
- Employés
- Ouvriers
- N'a jamais travaillé

En routine, la faisabilité et surtout la qualité du recueil est discutable. La profession pour l'INSEE, fait l'objet d'un recueil et d'un codage précis, avec des définitions très élaborées (le document de l'INSEE pour aider au codage des PCS dépasse 200 pages). Il existe maintenant un logiciel de codification automatique à partir de blocs de questions standard (SICORE), mis gratuitement à disposition par l'INSEE, et plus ou moins précis selon le nombre de questions utilisées pour le codage<sup>6</sup>.

La classification des PCS comporte de nombreuses limites pour l'étude des inégalités sociales de santé. Regroupés en huit catégories très larges, les groupes socioprofessionnels (la classification la plus simple et la plus employée) ne constituent justement pas des groupes sociaux homogènes. En particulier, en l'absence de caractérisation des revenus, le seul groupe socioprofessionnel n'apporte que des renseignements limités sur les ressources financières de l'individu tant les revenus sont hétérogènes et divers au sein d'un même groupe. Les groupes socioprofessionnels recouvrent également des professions très variées dont les conséquences sur la santé peuvent être très différentes. Qui plus est, un même groupe renvoie à des professions et des secteurs d'activité qui restent très différents chez les hommes et les femmes (avec, également, des disparités salariales de genre) [Maruani, 2002]. Par ailleurs, au cours du temps et d'une génération à l'autre, les conditions de travail d'une même profession ont beaucoup évolué, de même que les effectifs respectifs des groupes et des catégories socioprofessionnels et, pour une part, la place de certaines professions dans la hiérarchie sociale. Comme pour le niveau d'éducation, l'analyse sur le long cours (ou au sein de populations d'âges très étendus) des situations professionnelles peut s'avérer rapidement assez ardue. Enfin, la variété des classifications nationales des catégories socioprofessionnelles complique les comparaisons internationales. La classification internationale type des professions (CITP-08, [BIT, 2007]) reste peu utilisée, de même que les tentatives d'harmonisation d'une classification européenne commune qui restent peu consensuelles [Brousse, 2008 ; Penissat, 2012].

---

<sup>6</sup> Carton M et al . Contribution des variables annexes au codage des libellés de profession par le logiciel SICORE. Note INSEE du 19 juillet 2007 No 1777/DG-75 F420.

## Revenus

Les ressources financières constituent une dimension très importante de la situation socioéconomique d'un individu puisqu'elles conditionnent son accès aux biens et aux services dont de nombreux ont une implication sur la santé (alimentation, logement, loisirs, vacances, sorties, consommation de substances, accès à l'information et aux médias, assurance complémentaire santé, biens et services de santé, etc.). Par ailleurs et à l'inverse, la maladie peut entraîner une perte de revenus bruts, ou de revenus disponibles après paiement des dépenses de santé restant à charge. Le niveau de revenu est également un indicateur de statut social. Même s'il est recueilli de façon imprécise et/ou analysé par grandes tranches, des relations entre revenu et indicateurs de santé sont très souvent retrouvées de façon marquée dans la population générale.

C'est pourtant une dimension complexe à interroger. Non seulement c'est une information qui reste extrêmement taboue dans la culture française, mais elle recouvre plusieurs notions différentes qui devraient idéalement être qualifiées : la régularité des revenus, leurs sources (patrimoine, salaires, pensions, retraites, allocations, aide financière de toutes sortes, activités non déclarées, etc.), leur part contrainte (pensions alimentaires, aides aux proches, impôts), incompressible (dépenses de logement, de transport, etc.) et, finalement, « disponible ».

Généralement, seul le revenu dit « net avant impôts » est interrogé<sup>7</sup>. Le revenu strictement individuel a moins de sens (sauf si l'on s'intéresse à l'effet inverse de l'état de santé sur le salaire) et mérite principalement d'être interrogé chez les personnes isolées, vivant seules et sans enfants. Dans tous les autres cas, la règle est d'interroger et d'additionner tous les revenus des membres d'un même ménage (qui n'est pas forcément le « foyer fiscal » des impôts) et de diviser cette somme par le nombre d'« unités de consommation » du ménage. Ceci permet de prendre en compte à la fois les charges des personnes sans revenus et à charge (par exemple les enfants) et les économies d'échelle réalisées au sein d'un même ménage. Selon l'échelle d'équivalence de l'OCDE qui est aujourd'hui la plus utilisée, les unités de consommation sont calculées de la façon suivante : le premier adulte compte pour 1 ; les autres personnes de 14 ans et plus pour 0,5 chacune ; et les enfants de moins de 14 ans pour 0,3 chacun. En pratique, la somme des revenus du ménage est recueillie, soit globalement (sous la forme d'un montant ou d'une tranche), soit par source de revenus pour plus de précision, le plus souvent sur une base mensuelle.

Dans les cas où le recueil des revenus semble inenvisageable ou risque de conduire à des estimations de trop mauvaise qualité, on pourra préférer interroger directement si (et combien) des membres du ménage bénéficient de l'un ou l'autre des huit minima sociaux existants en France<sup>8</sup>. L'avantage de ce choix est qu'il est plus simple à interroger, qu'il est

---

<sup>7</sup> Y compris les revenus du patrimoine mais pas le patrimoine lui-même. Le patrimoine est un indicateur alternatif de la richesse disponible des individus qui est recueilli dans certaines enquêtes, comme l'enquête SHARE en Europe ou HRS aux Etats Unis. Celui-ci est particulièrement important pour les retraités chez qui il présente une variabilité plus importante que les pensions de retraite.

<sup>8</sup> Revenu de solidarité active (RSA) ; Allocation de solidarité spécifique (ASS) ; Allocation temporaire d'attente (ATA) ; Allocation aux adultes handicapés (AAH) ; Allocation supplémentaire d'invalidité (ASI) ; Allocation veuvage (AV) ; allocations du minimum vieillesse (ASV et ASPA) ; revenu de solidarité (RSO). L'allocation de parent isolé (API) et l'allocation équivalent retraite de remplacement (AER-R) ont été supprimées en 2011.

moins sensible aux biais de mémoire et qu'il fournit directement une estimation du niveau de revenus puisque ce sont des allocations versées sous conditions de ressources.

## LE CAS SPECIFIQUE DES JEUNES

S'intéresser aux inégalités sociales de santé chez les jeunes devrait constituer une priorité de santé publique dans le sens où ils constituent la population de demain mais également parce que l'épidémiologie biographique, comme d'ailleurs l'étude des origines développementales de la santé et l'épigénétique, accumule les indices sur une influence précoce des conditions de vie vécues dès les premières années de vie sur l'état de santé et les comportements en lien avec la santé à l'âge adulte.

La plupart du temps, le statut socioéconomique des adolescents est approximé par celui de ses parents (ou le plus élevé entre ceux de ses deux parents)<sup>9</sup>. Une étude américaine suggère que le niveau d'éducation des parents est un meilleur prédicteur de la santé de leurs enfants que les revenus ou la composition des familles [Zill, 1996] tandis que d'autres auteurs recommandent, pour les études sur le développement des enfants, de recueillir la triade classique du statut socio-économique des parents [Hauser, 1994].

A l'adolescence, le statut social est donc généralement réputé être celui de la famille, mais le parcours scolaire est également un déterminant important à prendre en compte. La filière scolaire, le redoublement mais aussi l'échec scolaire (considéré comme une déscolarisation précoce conduisant à l'absence de diplôme) placent en effet l'adolescent dans une hiérarchie sociale et scolaire propre à cet âge et des travaux ont montré, que ce sont des facteurs de risque de l'entrée dans les usages de tabac, d'alcool et de drogues illicites – et que la réciproque est également vraie - et ce, indépendamment de la corrélation entre milieu social parental et parcours scolaire [Legleye, 2011]

Un indicateur intitulé the Family Affluence Scale a été développé depuis un peu plus d'une dizaine d'années et exploré dans l'étude HBSC (...). Cet indicateur est composé de quatre items portant sur la possession d'une voiture par la famille, la possession de sa propre chambre par l'adolescent, l'accès à des vacances et l'accès à un ordinateur (qui a remplacé l'ancien item d'accès à un téléphone). Cet indicateur a été corrélé à plusieurs états de santé et s'est montré prédictif d'événements sanitaires chez les adolescents. Il a été suggéré de pondérer les différents items selon les pays et les besoins ressentis. Une validation contre les déclarations des parents a été obtenue pour valider l'échelle FAS. Une autre approche chez les adolescents a consisté à utiliser le statut social subjectif.

## LES CARACTERISTIQUES DEMOGRAPHIQUES

### L'âge

Il s'agit bien entendu d'une caractéristique essentielle à recueillir puisqu'on a vu que les déterminants sociaux de la santé et les mécanismes de production des inégalités sociales de santé d'inscrivent dans la biographie des personnes. Son recueil - au mieux comme une

---

<sup>9</sup> La déclaration de la catégorie d'emploi des parents par les enfants serait un indicateur plutôt fiable (contrairement au niveau d'éducation des parents, beaucoup moins bien rapporté par leurs enfants), qui peut être encore améliorée par le recueil en clair du métier des parents [Lien et al, 2001 – cité par Legleye, 2010].

variable continue - est la plupart du temps déjà fait en routine ; nous ne nous y attarderons donc pas. Son analyse fera appel à des classes d'âge variables en fonction des analyses et des objectifs poursuivis (d'où l'importance de l'interroger comme une variable continue, regroupée seulement dans un second temps). Les seuils pourront ainsi être choisis selon une rationalité statistique (ex : tertiles, quartiles, etc.), biologique (ex : âge de la puberté, de la ménopause, etc.), médicale ou préventive (ex : en fonction des âges cibles de certaines recommandations) ou sociale (ex : âges scolaire, âge légal ou moyen de départ à la retraite).

### **Sexe et genre**

Jusque dans les années 90, la grande majorité des études empiriques a assigné aux femmes la position socio-économique du « chef de famille ». Cette position était basée sur le fait que les femmes avaient une activité encore souvent essentiellement familiale et une insertion dans le monde du travail intermittente. Dans des travaux britanniques, le gradient social n'était retrouvé qu'en utilisant la profession du partenaire ou de l'époux et non celle propre de la femme. L'évolution de l'insertion des femmes dans le monde du travail a modifié cette situation, si bien qu'en France les inégalités sociales de mortalité cardiovasculaire sont aujourd'hui plus marquées chez les femmes que chez les hommes. L'effet du genre, les déterminants biologiques et sociaux interagissent fortement. Alors que le fait d'être plus pauvre augmente le risque de maladie et de diminution de l'espérance de vie, l'effet est plus important chez les hommes. Malgré un revenu moyen inférieur, les femmes vivent plus longtemps que les hommes. Pour le moment aucun indicateur d'équité de genre à utiliser dans les études au niveau des communautés n'a été étudié ou validé.

### **Origine ethnique**

L'« ethnicité » est entendue ici au sens défini par P. Simon, c'est-à-dire un concept « combinant une multiplicité de caractéristiques telles que le lieu de naissance, la langue, les « traits culturels », la religion, la « race », la nationalité, la couleur, l'ascendance et, recouvrant tous ces attributs, le sens d'une appartenance commune » [Simon, 1997]. Ce concept ne peut être mobilisé qu'avec précaution, notamment en sciences humaines et sociales et en épidémiologie sociale. Il n'a pas d'existence juridique - ni a fortiori administrative - en France et le recueil de l'origine ethnique est interdit dans tous les recueils obligatoires de la statistique publique (notamment l'état civil et le recensement de la population). Pour autant, ce recueil n'est pas interdit dans tous les cas, comme on l'entend parfois puisque l'origine ethnique rentre explicitement dans la catégorie des « données sensibles<sup>10</sup> » : il s'agit, par définition, des données à caractère personnel qui font apparaître, directement ou indirectement, les « origines raciales (sic) ou ethniques, les opinions politiques, philosophiques ou religieuses ou l'appartenance syndicale des personnes, ou qui sont relatives à leur santé ou à leur vie sexuelle ». Cette dimension peut donc être interrogée dans le respect de la loi et des strictes conditions qu'elle prévoit pour ce type de données. En santé, l'autorisation de la CNIL et de son comité consultatif est alors accordée au cas par cas en fonction de la pertinence du recueil, de la finalité de la recherche et de l'argumentaire scientifique présenté [Debet, 2007]. Si le recueil de l'origine ethnique a été, de fait, autorisé dans certaines enquêtes, de nombreuses façons de l'interroger et de la

---

<sup>10</sup> Il s'agit, par définition, des données à caractère personnel qui font apparaître, directement ou indirectement, les origines raciales ou ethniques, les opinions politiques, philosophiques ou religieuses ou l'appartenance syndicale des personnes, ou qui sont relatives à leur santé ou à leur vie sexuelle.

renseigner peuvent être imaginées mais elles ne font l'objet d'aucun consensus entre les chercheurs. Actuellement, dans le domaine de la santé, à part des enquêtes spécifiques sur les discriminations ou sur certains déterminants culturels de certaines attitudes et comportements en lien avec la santé, le recueil des origines ethniques ne peut donc être qu'exceptionnel. Au risque d'ethniser la santé et la maladie, on réservera l'interrogation spécifique de certaines dimensions spécifiques de ce concept à la recherche.

### *Situation et origine migratoire*

La santé des immigrés est considérée comme une question au cœur des inégalités sociales de santé non seulement dans les pays, comme les Etats-Unis ou le Canada, où l'immigration fait partie de leur identité historique et collective, mais également dans certains pays européens – et ce, y compris par les politiques et les stratégies de lutte contre les inégalités sociales de santé. Aux Etats-Unis, le statut migratoire est collecté dans différentes enquêtes nationales. En Europe, une revue de la littérature scientifique conduite en 2010 a recensé 17 études sur la santé perçue des immigrés ou des minorités ethniques, conduites dans 5 des 27 pays de l'Union européenne. La plupart concluaient à des états de santé perçus plus dégradés que le reste de la population, à âge et sexe identiques [Nielsen & Krasnik, 2010]. Une autre étude a fait un constat identique concernant la santé, les comportements à risque et les incapacités des immigrés de plus de 50 ans dans 11 pays européens [Solé-Auro & Crimmins, 2008].

En France, où les immigrés et leurs enfants comptent pour 19% de la population (33% en Ile-de-France), l'état de santé des immigrés reste peu étudié pour de multiples raisons historiques, sociologiques et politiques qui ont longtemps empêché de traiter ce sujet à la hauteur de ses enjeux sociaux et sanitaires, régulièrement rappelés par les institutions européennes [European parlement, 2011] ou internationales [OIM, 2010]. Pourtant, des travaux montrent que les origines migratoires constituent un facteur important d'inégalités sociales de santé pour certains états morbides, comportements de santé et indicateurs de recours aux soins [Hamel & Moisy, 2010, Berchet et al, 2012, Collectif, 2012], y compris après la prise en compte d'autres caractéristiques socio-économiques comme le niveau d'éducation, le niveau de revenus ou le statut d'emploi. Des travaux épidémiologiques récents montre également que, pour de nombreux indicateurs de santé, les Français issus de l'immigration se situent dans une position intermédiaire entre les Français nés de parents français et les immigrés [Vallée, 2010 ; Grillo, 2012 ; Cognet, 2012].

La situation migratoire peut en pratique être renseignée de 4 façons :

- Le statut d'immigré : statut dans le sens où la qualité d'immigré - selon la définition française : personne née étrangère à l'étranger - est permanente (un individu continue à appartenir à la population immigrée même s'il devient français). Ce statut est obtenu partir de la collecte du lieu de naissance et de la nationalité à la naissance.
- L'ancienneté d'installation en France : pour les immigrés (mais aussi pour les Français nés et ayant vécu une part importante de leur vie à l'étranger), il peut être utile de connaître également l'ancienneté de leur venue en France, qui peut refléter leur acculturation aux normes et comportements du reste de la population, notamment en matière de santé ou de recours aux soins. La façon de recueillir cette variable ne fait pas l'objet de consensus et peut être compliquée par les allers retours et la succession de

séjours temporaires plus ou moins longs. En pratique, on peut interroger l'âge de la première arrivée et/ou la durée cumulée de vie passée en France.

- Les origines migratoires : celles-ci permettent de distinguer les Français issus de l'immigration de la population majoritaire (et des immigrés) à travers la collecte de la nationalité et du lieu de naissance d'égo et des parents d'égo et d'étudier les inégalités spécifiques qui les concernent d'un point de vue social ou sanitaire.
- Le recueil de la seule nationalité actuelle (au jour de l'enquête ou du recueil de données) n'est donc qu'un pis-aller, qui permet à proprement parler de ne distinguer que les étrangers des Français, éventuellement en distinguant les Français de naissance des Français par acquisition (voire, parmi ces derniers, le cas des enfants, des acquisitions par mariage, naturalisation et réintégration) mais ne reflète que très approximativement la situation migratoire des individus.

## **LE STATUT VIS-A-VIS DE LA COUVERTURE DU RISQUE MALADIE**

De nombreux travaux démontrent l'influence du statut d'assurance maladie sur les recours aux soins et les dépenses de santé. On citera notamment les résultats de l'Enquête Santé et Protection Sociale de l'Irdes, l'enquête de référence sur ce sujet, conduite depuis 1988 sur un échantillon de 8000 ménages ordinaires représentatif de 96 % de la population vivant en France métropolitaine ; mais également de nombreuses enquêtes conduites dans des territoires ou des populations particulières (malades, hospitalisés, précaires, etc.).

L'interrogation de ce statut peut paraître difficile, à la hauteur de la complexité du système d'assurance maladie français et, par-là, des situations individuelles. Certaines personnes – notamment quand elles sont jeunes, ayant-droits ou dans des situations sociales compliquées – peuvent même ignorer leur situation exacte vis-à-vis de l'Assurance maladie, des assurances complémentaires ou des dispositifs spécifiques.

En pratique, on recommande d'interroger séparément (en recueillant également les méconnaissances) le statut assurantiel de base et le statut assurantiel complémentaire sans tenter de proposer des catégories toutes faites associant l'un et l'autre (même si la couverture de base par la Sécurité sociale associée à une assurance complémentaire privée ou mutualiste est la situation la plus courante). En effet, l'ouverture et la durée des droits des uns et des autres ne correspond pas toujours, à un instant donné, aux droits théoriques. On demandera donc successivement :

- Quelle est la « couverture de base » : Sécurité sociale de droit commun (en demandant éventuellement si la personne bénéficie d'une prise en charge à 100% pour raisons médicales et à quel titre : affections longue durée et apparentées, femme enceinte, incapacité permanente, pension d'invalidité, etc.), Couverture maladie universelle de base, Aide médicale d'Etat, aucune, ou ne sait pas ;
- Quelle est la couverture complémentaire : assurance privée ou mutuelle (en distinguant éventuellement si la personne bénéficie ou non d'une Aide à l'acquisition d'une complémentaire santé (ACS), peu utilisée à ce jour [Guthmuller et al, 2011]), Couverture maladie complémentaire, aucune, ne sait pas.

## **LES DETERMINANTS SOCIO-ECONOMIQUES MOBILISABLES POUR L'ETUDE DES INEGALITES SOCIALES DE SANTE**

Bien entendu, d'autres informations sont recueillies et utilisées, plus ou moins couramment, pour décrire, surveiller et comprendre les inégalités sociales de santé. Elles sont extrêmement nombreuses. Nous en listerons ici, rapidement, quelques-unes, par thème :

1. Les insertions sociales : les modèles théoriques de la santé utilisés en santé publique considèrent que les environnements physiques et socio-économiques de vie ainsi que les réseaux sociaux figurent parmi les grands déterminants de la santé : isolement perçu, réseau social (dans les différentes sphères de socialisation : familiale, amicale, de voisinage, au travail, etc.), soutien social. Parmi ces milieux d'insertion, la famille constitue à la fois une cellule économique qui dispose de ressources et au sein de laquelle se produit du capital humain, un lieu de vie (voire plusieurs), une communauté et un réseau social, un lieu de production de santé à travers des modes de vie collectifs et les soins apportés, mais aussi à travers le patrimoine génétique et les possibles contaminations entre membres. Au-delà de la santé de la mère et de l'enfant, il est donc important de considérer la famille et le milieu social parmi les déterminants de la santé.
2. Les conditions de vie : conditions de logement, personnes à charge et aides quotidiennes fournies, possession de certains biens d'équipement, discriminations
3. Les conditions de travail : facteurs psychosociaux au travail et conditions matérielles (charge physique, expositions à des toxiques professionnels...)
4. Les insertions spatiales : quartier de résidence, mobilité résidentielle (au cours de la vie), mobilité quotidienne (espaces d'activité quotidiens, moyens de transport)

## **2. INDICATEURS ECOLOGIQUES DE DEFAVORISATION SOCIALE**

L'indice de défavorisation sociale, dit aussi indice de défaveur sociale ou indice de désavantage social, est le terme français correspondant au « deprivation index » dans la littérature internationale ou à l'indice de défavorisation au Québec.

Il s'agit d'un indice socioéconomique écologique qui est la caractérisation synthétique du niveau socioéconomique de la population à une échelle donnée et à un moment donné. Cet indice est obtenu par l'agrégation de différentes variables (revenus, emploi, etc.) mesurées à une échelle géographique donnée, généralement une petite unité administrative ou territoriale. Par exemple, si la variable est le taux de chômage, elle sera mesurée avec le pourcentage de chômeurs dans la zone géographique.

Parmi les différentes unités géographiques, on distingue en particulier la région, le département, le canton, la commune, l'Iris ou le carroyage. Les communes de taille suffisamment grandes sont divisées en plusieurs Iris. Chaque Iris résidentiel est construit pour inclure environ 2000 personnes, afin d'être relativement homogène en termes de caractéristiques sociales. Au contraire, les plus petites villes ou villages forment un Iris et peuvent donc inclure moins de 2000 personnes. Récemment, l'Insee a introduit une unité géographique plus petite, le carroyage qui représente un carré de 200 mètres de côté. La petite taille de cette zone pose des problèmes d'anonymisation, et l'avenir de cette unité géographique dans la statistique publique est encore incertain. Actuellement l'Iris est la plus petite unité géographique pour laquelle l'ensemble des données de recensement français sont disponibles.

### **AVANTAGES ET LIMITES**

L'avantage principal des indices de défavorisation sociale est de fournir une mesure de la situation socioéconomique des personnes en l'absence de données individuelles, en utilisant l'adresse des personnes, information généralement disponible dans les systèmes d'information permanents. Cette information, même si elle est sujette à erreur est recueillie pour permettre le suivi des personnes ou leur facturation. La mise en œuvre de cette mesure a donc le très grand avantage de pouvoir être opérationnelle rapidement en France, pour la quasi-totalité des personnes.

En France, les fichiers administratifs du domaine sanitaire en population générale sont le plus souvent marqués par l'absence d'information sur les caractéristiques socioéconomiques des individus telles que le niveau d'études, la catégorie socioprofessionnelle ou le revenu. Ceci rend difficile les études nationales sur les relations entre statut socioéconomique et état de santé (mortalité, morbidité sanitaire et sociale et recours aux services de santé). Ces fichiers disposent toutefois du lieu de résidence des personnes, et permettent le calcul d'indices de défavorisation sociale. Cet indice peut être mis en relation avec l'indicateur sanitaire et permettre ainsi de documenter les inégalités sociales de santé à partir de fichiers dépourvus d'indicateurs individuels de situation sociale. Cette approche est très fréquemment utilisée dans les pays anglo-saxons. En France, elle pourrait concerner par exemple les bases médico-administratives ou les données des registres des cancers.

Les indices de défavorisation sociale présentent plusieurs limites.

Ils sont calculés au niveau géographique pour caractériser ensuite la situation d'une personne. Ils sont donc entachés d'un biais écologique. Plus la zone géographique est petite et homogène, moins le biais est important ; calculer l'indice au niveau de l'Iris ne permet toutefois pas de s'en affranchir. Ce biais demeurant, l'indice intégrera toujours une dimension contextuelle.

Ils prennent en compte le contexte socioéconomique mais ne prennent pas en compte le contexte du milieu de vie (urbain ou rural, centre-ville ou périphérie...). Ainsi par exemple, ne pas avoir de voiture n'a pas la même signification en matière de défavorisation sociale en milieu urbain, péri-urbain ou rural.

Ils sont calculés à partir des caractéristiques d'une zone géographique à un moment donné. Or, la mobilité différentielle des groupes sociaux contribue à façonner les territoires. Les caractéristiques du lieu de résidence varient au cours du temps et cet aspect ne doit pas être oublié lors de l'utilisation de ces indices.

Le lieu de résidence n'est pas forcément une zone pertinente pour définir la position socioéconomique des personnes, or c'est la zone géographique utilisée pour calculer les indices de défavorisation sociale. Ils sont calculés dans une zone géographique donnée, souvent définie sur des bases administratives mais qui ne correspond pas forcément à un territoire faisant sens pour les personnes. La définition d'un territoire du point de vue des personnes, à savoir les quartiers perçus, est complexe, mais ces territoires peuvent fortement différer des zones géographiques retenues pour le calcul de l'indice. La mobilité quotidienne, notamment mais pas seulement en milieu urbain, complique encore cet aspect. On peut raisonnablement considérer que le lieu de résidence caractérise plus ou moins bien la situation sociale des personnes selon leur mobilité, ou en tous cas que la signification de cette zone géographique varie selon la mobilité de la personne, avec les conséquences possibles sur le lien avec la santé.

Le développement des indicateurs territoriaux à un niveau géographique fin (Iris) a introduit une certaine difficulté dans les analyses due à une confusion entre la situation socioéconomique de la personne ou du quartier. Les indicateurs peuvent en effet avoir deux objets, qu'il s'agit de bien distinguer.

1/Caractériser un territoire, avec une certaine richesse de variables, est un instrument d'analyse puissant dans les analyses multi-niveaux. Il permet de distinguer dans les disparités des états de santé, des éléments qui reviennent à l'individu d'une part et à son environnement collectif autre part (les effets authentiquement contextuels, qui peuvent être en lien avec la composition collective du territoire). Cette approche géographique permet la prise en compte, par exemple, de la densité de population, de la densité des services, des oppositions urbain/rural, du voisinage, du milieu physique, ou des oppositions centre-périphérie et présente l'avantage de mettre l'accent sur des inégalités pour lesquelles il est envisageable de proposer des interventions d'acteurs de santé publique pour les réduire et ce à différents échelons géographiques. Les indicateurs territoriaux seront alors utilisés pour caractériser la position socioéconomique du quartier, dans le cadre d'analyses multi-niveaux.

2/ Approximer la position socio-économique des individus, en se basant sur la caractérisation socio-économique de leur zone géographique de résidence dès lors qu'il

s'agit de territoires limités tels que les Iris.. Cette approximation sera de qualité décroissante avec l'augmentation de l'hétérogénéité des populations qui les composent. Les indices de défavorisation sociale présentés ici s'inscrivent dans ce deuxième objectif. Ils caractérisent la position socioéconomique de la personne, et s'utilisent dans des analyses classiques.

Il existe une variété d'indices qui diffèrent par la méthodologie envisagée et les variables considérées. Nous allons présenter les principaux indices développés et utilisés en France.

## **QUELQUES INDICES DE DEFAVORISATION SOCIALE UTILISES HORS DE LA FRANCE**

### ***Indice de Jarman (Angleterre)***

L'indice a été défini en 1983, initialement pour identifier les zones potentiellement à forte demande de services de soins primaires. Il est construit à partir des variables suivantes : chômage, absence de voiture, surpeuplement du logement, classe sociale inférieure, locataire, famille monoparentale, jeune de moins de 5 ans, retraités vivant seuls, immigrants récents.

### ***Indice de Townsend (Angleterre)***

L'indice a été défini en 1987. Il est construit à partir de la somme non pondérée de 4 variables : chômage, pas de véhicule, logement surpeuplé, locataire du logement.

### ***Indice de Carstairs et Morris (Écosse)***

L'indice a été défini en 1991. Il est construit à partir de la somme non pondérée de 4 variables : chômage, pas de véhicule, classe sociale basse, logement surpeuplé.

### ***Indice de Pampalon (Québec)***

L'indice a été défini en 2000. Il est construit à partir des variables suivantes : La proportion de personnes de 15 ans et plus sans certificat ou diplôme d'études secondaires ; le ratio emploi/population chez les 15 ans et plus ; le revenu moyen des personnes de 15 ans et plus ; la proportion de personnes de 15 ans et plus vivant seules dans leur domicile ; la proportion de personnes de 15 ans et plus séparées, divorcées ou veuves ; la proportion de familles monoparentales.

Il a la particularité d'approcher l'aspect multidimensionnel de la déprivation ; déprivation matérielle (personnes de plus de 15 ans sans diplôme secondaire, revenu moyen, rapport emploi/éducation) et déprivation sociale (personnes séparées, divorcées veuves ; familles monoparentales ; personnes vivant seules)

### ***Index of Multiple Deprivation (Royaume-Uni)***

Dans les années 2000, un nouvel indice a été développé au Royaume-Uni pour tenir compte d'un nombre plus important de dimensions de la défavorisation (à la différence des indices de Townsend et Carstairs). Ce nouvel indice a été créé en 2004 (Index of Deprivation 2004) puis mis à jour en 2007 (Index of Multiple Deprivation 2007 ou IMD 2007) et en 2010 (IMD 2010). Cet indice existe pour l'Angleterre, l'Irlande du Nord, l'Ecosse et le Pays de Galles. Pour l'Angleterre, cet indice est construit à partir de 37 indicateurs répartis dans ces 7 domaines, à savoir le revenu, l'emploi, la santé, l'éducation, l'accès/barrières aux services, l'environnement résidentiel, et la criminalité.

# LES PRINCIPAUX INDICES DE DEFAVORISATION SOCIALE UTILISES EN FRANCE

## *FDep*

L'indice de défavorisation sociale appelé FDep a été développé dans le contexte français (Rey et al, 2009). Il a été construit à l'échelle communale à partir des données de recensement de la population de 1999 (Insee) et des revenus fiscaux des ménages de 2001 (Insee-DGI). Le même indice a été calculé à partir des données de 1990. Quatre variables, représentant toutes une dimension fondamentale du niveau socioéconomique, de signification homogène dans l'ensemble du pays et covariant de façon similaire selon la tranche d'unités urbaines ont été utilisées : le pourcentage d'ouvriers dans la population active, le pourcentage de bacheliers chez les 15 ans et plus, le pourcentage de chômeurs dans la population active et le revenu médian par foyer. L'indice a été obtenu à l'aide d'une analyse en composantes principales (ACP) conduite sur ces quatre variables. Les résultats de l'ACP étaient similaires lorsque l'analyse était conduite au niveau national ou dans chaque tranche d'unité urbaine. L'association entre défavorisation sociale et mortalité sur la période 1997-2001 est observée quelle que soit l'échelle spatiale choisie (commune, canton, zone d'emploi, département et région), le FDep calculé pour les échelles autres que communale étant obtenu comme la moyenne pondérée sur la population des valeurs obtenues au niveau communal. De même, l'association entre défavorisation sociale et mortalité se retrouve dans toutes les tranches d'unité urbaines. Cet indice est actuellement utilisé par la CNAM-TS.

## *EDI (European deprivation index)*

Un indice de défavorisation sociale à visée européenne, l'EDI ou « european deprivation index », a été proposé (Pornet et al, 2012). La méthodologie est différente de celle utilisée pour les indices de défavorisation sociale habituellement rencontrés. Elle prend comme point de départ la définition de la défavorisation de Townsend, et postule qu'il n'y a pas de gold standard pour la mesurer en ce qui concerne le choix des variables, celles-ci variant selon les pays. La démarche consiste donc à définir cette notion dans la population étudiée, avec les variables pertinentes, avant de construire l'indice. Elle repose sur la sélection d'une combinaison de variables qui sont disponibles à la fois au niveau individuel dans une enquête européenne dédiée à l'étude de la défavorisation et au niveau géographique dans le recensement de la population et qui reflète au mieux l'expérience individuelle de défavorisation sociale. L'enquête européenne utilisée est l'étude EU-SILC (European Union – Statistics on Income and Living Conditions) ou SRCV dans sa version française (Statistiques sur Revenus et Conditions de vie), qui vise à collecter des données sur le revenu, la pauvreté, l'exclusion sociale et les conditions de vie. Cette étude annuelle inclut 26 pays depuis 2005.

La méthodologie comprend plusieurs étapes.

Dans un premier temps, un indicateur individuel de défavorisation sociale est construit à partir des données françaises de SRCV de 2006. Il s'agit d'identifier les besoins (ou biens ou services) fondamentaux des personnes à partir de la liste disponible dans l'enquête EU-SILC. Ces besoins sont définis comme les besoins pour lesquels, parmi les ménages qui ne les possèdent pas, c'est dans moins de 50% pour raison financière. Ensuite, on ne garde que les besoins associés à la fois à la pauvreté subjective (situation financière perçue et revenu jugé insuffisant) et objective (revenu en classes). Puis un indicateur binaire est défini

qui indique si un individu est en situation de défavorisation sociale et matérielle ou non en fonction du nombre de besoins fondamentaux qu'il ne possède pas pour raison financière. Un individu est défini comme défavorisé s'il ne pouvait pas satisfaire au moins trois besoins pour des raisons financières.

Dans un second temps, on identifie et dichotomise les variables disponibles à la fois dans l'enquête européenne et dans le recensement. Dans un troisième temps, on construit l'EDI. Une régression logistique sur les données de l'enquête européenne permet de ne conserver, parmi les onze variables retenues, que celles associées avec l'indicateur individuel. Les coefficients de régression associés à chacune de ces variables deviennent les poids de ces variables dans l'EDI. L'EDI est construit au niveau de l'Iris et inclut les variables suivantes : surpopulation dans le logement, pas de chauffage central ou électrique, non propriétaire, chômeur, de nationalité étrangère, pas de voiture, ouvrier non qualifié ou agricole, niveau d'études faible et famille monoparentale.

La méthodologie utilisée pour construire l'EDI est reproductible dans les 26 pays européens concernés par l'enquête EU-SILC. Par ailleurs, puisque les données de l'enquête SRCV sont aussi disponibles au niveau régional, une adaptation est possible pour toutes les régions françaises. Enfin, l'enquête EU-SILC étant répétée chaque année, l'indice peut être répliqué dans le temps. L'EDI est disponible pour l'ensemble des IRIS, à partir des données INSEE (2007).

### **Indice « métropoles »**

Un indice a été développé en 2013 (Lalloué) à partir d'un indice développé dans la métropole de Strasbourg (Havard et al, 2008). Cet indice a été développé au niveau de l'Iris dans 3 métropoles régionales (Lille, Lyon et Marseille). Les données du recensement de 1999 et des revenus fiscaux de 2001 ont été utilisées. Un total de 48 variables a été sélectionné. Elles ont été choisies pour être représentatives des concepts théoriques de la position socioéconomique et en lien avec la littérature. Deux autres variables ont été retenues, à savoir le pourcentage d'individus âgés de plus de 65 ans et le pourcentage de déménagement dans une autre ville depuis le dernier recensement. Parmi ces 48 variables, 7 concernant le chômage et 3 le marché du travail étaient redondantes. Les variables ont été sélectionnées par ACP. L'indice représente la combinaison linéaire des facteurs retenus par le 1er axe de cette dernière ACP. La même procédure a été appliquée à chaque métropole régionale séparément puis aux trois zones simultanément. Au total, 20 variables ont été retenues pour les métropoles de Lille, Lyon ou Marseille et 19 lors de l'analyse conjointe des trois métropoles. Quinze des vingt variables sont communes aux trois indices obtenus pour chaque métropole régionale. Il s'agit des variables suivantes, exprimées en pourcentages de : ménages avec une voiture, ménages avec deux voitures ou plus, individus avec un emploi non stable, 15 ans et plus ayant un niveau d'études égal au Bac, 15 ans et plus ayant un niveau d'études secondaire inférieur (BEP, CAP, BEPC), 15 ans et plus sans diplôme, non propriétaires, individus avec un emploi stable, familles monoparentales, revenu médian par unité de consommation, résidences principales avec plus d'une personne par pièce, étrangers dans la population totale, nombre moyen de personnes par pièce, immigrants depuis le dernier recensement dans la population totale, travailleurs à leur compte dans la population active.

Les auteurs proposent une Classification Hiérarchique pour catégoriser l'indice en classes homogènes. Les quintiles ou les tertiles, bien que souvent utilisés, ne sont pas appropriés. La classification optimale est en trois classes, mais les classes diffèrent des tertiles. Les

auteurs mettent à disposition le programme en langage R permettant la création de l'indice et la définition des classes.

**Le tableau suivant résume les principaux avantages et inconvénients des trois indices de défavorisation sociale développés et utilisés en France.**

**Avantages et inconvénients des indicateurs**

	Avantages	Inconvénients
FDep	Calculé sur la France Construction comparable sur différents type d'espace et sur l'ensemble du territoire	Son calcul au niveau de l'Iris n'a pas été publié.
EDI	Part du postulat qu'il n'y a pas de gold standard en termes d'indices, mais qu'un concept comme celui de déprivation s'exprime différemment (avec des variables différentes) dans différents pays.	EU-SILC pas représentatif de la population française, c'est une enquête : effectif restreint et non réponse (10 036 foyers, 24 940 individus, 19253 âgés de 16 ans et +). Possibilité d'effectifs faibles dans les régions françaises
Indice « urbain »	Propose une méthode pour créer des classes les plus homogènes possible	Validé dans 3 zones métropoles régionales urbaines

La saisine à l'origine de ce rapport évoque une hiérarchisation des indices de défavorisation sociale. Une liste de critères est proposée pour faciliter les choix, selon que la priorité porte sur une zone géographique ou une autre, que des comparaisons européennes sont souhaitées ou non... Les critères suivants ont été retenus :

- Pays dans lequel l'indicateur a été développé
- Date de développement de l'indicateur
- Echelle géographique la plus fine pour laquelle l'indice a été développé (canton, commune, IRIS)
- Méthode de construction (statistique ; référence à un concept de déprivation)
- Expérience d'utilisation en santé publique (préciser)
- Aptitude à permettre des comparaisons internationales
- Zone géographique de validité (nationale, urbaine, européenne)
- Indicateur ayant fait l'objet de travaux d'évaluation et de comparaison avec d'autres indices

Un tableau synthétique présente ces critères pour les principaux indices écologiques de défavorisation sociale développés et utilisés en France (tableau I). Des travaux sont en cours pour comparer la qualité de ces différents indices et leur capacité à estimer les positions socio-économiques des individus.

Aucun des indicateurs cités n'a explicitement pour objectif de caractériser le territoire. Il s'agirait d'une utilisation erronée de ces indices, qui ont à l'origine l'objectif de caractériser les individus, du fait du manque d'informations individuelles, avec bien sûr un risque écologique.

La nécessité de développer des indicateurs géographiques permettant d'analyser l'effet du contexte sur la santé ne fait pas de doute. Cette recherche reste à faire (le projet ANR Delos, en cours, cherche à comparer et développer des indices visant à mieux prendre en compte l'aspect géographique).

**Tableau I : Indices écologiques de défavorisation sociale développés en France**

	<b>FDep</b>	<b>EDI</b>	<b>Indice « urbain »</b>
<b>Pays dans lequel l'indicateur a été développé</b>	France	France	France
<b>Date de développement de l'indicateur</b>	2009	2012	2013
<b>Echelle géographique la plus fine pour laquelle l'indice a été développé (canton, commune, IRIS)</b>	Commune (l'indice a aussi été calculé à l'échelle de l'Iris avec des résultats identiques)	Iris	Iris
<b>Méthode de construction (statistique ; référence à un concept de déprivation)</b>	ACP	Régression logistique Définit la défavorisation sociale et matérielle par un concept s'exprimant différemment selon les pays	ACP
<b>Expérience d'utilisation en santé publique</b>	Dans données CNAM-TS		
<b>Aptitude à permettre des comparaisons internationales</b>	Non validé	Oui	Non validé
<b>Zone géographique de validité (nationale, urbaine, européenne)</b>	Nationale	Européenne, Nationale	3 métropoles régionales
<b>Indicateur ayant fait l'objet de travaux d'évaluation et de comparaison avec d'autres indices</b>	Carstairs Townsend	En cours Abstract	Carstairs Townsend
<b>Données utilisées</b>	Recensement 1999 et revenus fiscaux 2001 pour Fdep99 Recensement 1990 et revenus fiscaux 90 pour FDep90	EU-SILC 2006 et recensement de 1999	Recensement 1999 et revenus fiscaux 2001
<b>ACP : % variance expliquée par la 1<sup>ère</sup> composante</b>	FDep90 60% FDep99 68%		Lille 61% Lyon 58% Marseille 57% 3 métropoles 57%

### 3. UTILISATION DES BASES DE DONNEES ADMINISTRATIVES <sup>11</sup>

La France dispose de bases de données médico-sociales et économiques nationales centralisées, constituées et gérées par des organismes publics, couvrant de façon exhaustive et permanente l'ensemble de la population dans divers domaines stratégiques pour la santé publique et la recherche : recours aux soins, hospitalisation, handicaps, prestations et situation professionnelle, sociale et économique. De plus, un identifiant individuel unique (le NIR : numéro d'identification au répertoire) est actuellement utilisé par pratiquement toutes les bases de données nationales. Malgré certaines limites en termes de couverture, de qualité et de validité des données, ces bases de données, concernant l'ensemble de la population française, constituent un patrimoine considérable, vraisemblablement sans équivalent au monde.

Cependant, l'utilisation à des fins de recherche et de surveillance de ces bases de données se heurte actuellement à des obstacles divers, dont les plus importants sont de nature juridique et opérationnelle.

#### *Les principales bases de données nationales mobilisables pour la recherche, pour la surveillance et la santé publique*

##### **Données de santé :**

Mortalité, données d'hospitalisation via le PMSI (Programme de médicalisation du système d'information), données de consommations de soins et de prise en charge de l'assurance maladie, réunies au sein du Système national d'information inter-régimes de l'assurance maladie - SNIIR-AM, dont un échantillon aléatoire (l'EGB : échantillon généraliste des bénéficiaires) peut être utilisé à distance. Leurs principales limites sont qu'elles ne contiennent pratiquement pas de données concernant la situation socioprofessionnelle des personnes et que les informations sur le domicile des patients ne sont pas suffisamment précises pour permettre une exploitation territorialisée à une échelle fine ; de plus, la validité des données de santé de ces bases est de qualité variable.

##### **Données de situation socioprofessionnelle :**

Les bases de données de la Caisse nationale d'assurance vieillesse (Cnav) permettent de retracer, pour chaque personne ayant appartenu durant sa vie au moins une fois au régime général de sécurité sociale, ses différentes périodes d'activité : périodes d'activité professionnelle ou assimilées (chômage, maladie, maternité ou congés parentaux, ...).

##### **Gestion des identités :**

La Cnav gère le système national de gestion des identités qui contient l'ensemble des données (NIR, état-civil, statut vital) de la population française, ainsi que le RNIAM (Répertoire national inter-régimes des bénéficiaires de l'assurance maladie), qui contient les informations de rattachement des personnes aux organismes servant les prestations d'assurance maladie.

---

<sup>11</sup> Cette partie est largement issue du rapport du groupe de travail du HCSP sur les systèmes d'information : « pour une meilleure utilisation des bases de données administratives et médico-administratives nationales pour la santé publique et la recherche », HCSP, Mars 2012.

**Autres sources** : outre les trois dispositifs cités, il existe d'autres bases de données nationales concernant des problèmes de santé ou des populations spécifiques.

**Bases de données fiscales :**

Disposer d'informations fiables sur le revenu des personnes est un objectif difficile à réaliser de manière simple dans les enquêtes en population et, à plus forte raison, dans les systèmes de surveillance ou les registres ne reposant pas sur la collecte directe de donnée auprès de malades ou d'assurés sociaux. Depuis un certain nombre d'années, l'INSEE et les services statistiques des administrations de l'Etat ont un droit d'accès, sous réserve de certaines démarches, aux données des administrations fiscales qu'ils peuvent apparier aux données de certaines enquêtes en population. L'utilisation de ces données a permis de simplifier et d'améliorer la connaissance statistique des revenus. Si cette possibilité pouvait être ouverte plus largement pour enrichir les données de santé, l'analyse des inégalités sociales de santé en serait grandement facilitée. Le projet de loi "Enseignement supérieur et recherche" présenté au conseil des ministres du 20/3/2013 prévoit de modifier l'article L. 135 D du livre des procédures fiscales pour étendre au bénéfice des chercheurs cette dérogation au secret professionnel en matière de données fiscales, afin de leur permettre l'accès à ces données à des fins de recherche scientifique. Cette disposition précise les modalités d'obtention de l'autorisation d'accès aux données et, notamment, les conditions dans lesquelles le comité du secret statistique rend son avis. Il précise que ces données ne peuvent être mises à disposition du chercheur par le ministre chargé du budget que sur avis favorable du comité du secret statistique et que les travaux des chercheurs ne doivent pas permettre l'identification des personnes auxquels ils se rapportent. Le vote de cette disposition, ainsi que l'utilisation effective de ces données dans les systèmes d'information en santé de routine, serait un élément essentiel pour le développement des analyses sur les inégalités sociales de santé.

**Base permanente des équipements :**

L'INSEE met à disposition sur internet une base de données permettant de disposer annuellement d'une série d'indicateurs d'équipements et de services disponibles pour la population d'un territoire communal ou infra-communal (Iris) répartis en trois gammes : proximité (école maternelle, pharmacie, boulangerie, poste...), intermédiaire (collège, orthophoniste, supermarché, trésor public...), supérieure (lycée, maternité, hypermarché, pôle emploi ...). Ces données, appariables sur la base d'un critère géographique, peuvent permettre d'éclairer le rôle des disparités territoriales d'accès aux équipements et aux services en tant que déterminants de la santé.

Cadre juridique permettant d'accéder aux données à caractère personnel des bases nationales :

Il est complexe et dépend notamment de la nature de la base de données concernée et de celle de l'organisme utilisateur : organismes de recherche et agences de santé, services statistiques publics, organismes privés à but lucratif. Dans l'ensemble, on peut considérer qu'à l'exception des données fiscales, les textes actuels ne constituent pas un obstacle insurmontable à l'accès aux données à caractère personnel des bases nationales, mais ils rendent complexes les circuits de transmission de données. Une difficulté majeure concerne l'identifiant pour l'accès aux données à caractère personnel des bases nationales. L'identifiant individuel actuellement utilisé dans les bases nationales est le NIR, pour la collecte duquel un décret en conseil d'État est nécessaire, ce qui constitue un obstacle

infranchissable pour la plupart des équipes concernées, dès lors qu'elles n'appartiennent pas à un organisme habilité à disposer de cet identifiant ou qu'il n'est pas possible d'établir des flux de données reposant sur un tel organisme. Un autre problème risque de se poser prochainement avec le déploiement de l'INS (Identifiant national de santé), destiné à être le seul identifiant utilisé pour les applications en santé, ce qui rendrait impossible l'appariement de données de santé avec d'autres sources utilisant le NIR comme identifiant. Enfin, dans nombre de situations, il n'est en pratique pas possible de recueillir un consentement explicite des personnes concernées avant chaque étude, ce qui présente une difficulté vis-à-vis de la loi informatique et libertés.

### ***Obstacles organisationnels et techniques***

L'accessibilité effective des données reste aujourd'hui un problème majeur, qu'il s'agisse d'identifier les personnes pour lesquelles on veut extraire des données dans les bases, d'extraire les données et de mettre les données en forme pour les analyses. L'utilisation du NIR pour l'identification des personnes nécessite une participation active d'un organisme détenteur du NIR, mais il n'existe pas de dispositif formalisé mis en place par un ou plusieurs organismes détenteurs du NIR pour prendre en charge ce type d'activité ; l'extraction proprement dite des données des bases nécessite également une participation active des organismes gestionnaires des bases de données, car cette activité implique diverses étapes techniques lourdes. Malgré les très importants efforts de la Cnam-TS, de l'Insee et de la plupart des autres organismes gestionnaires de bases de données, l'accessibilité effective des données reste cependant encore fortement contingentée ; la complexité de la base de données du SNIIR-AM rend l'utilisation des données individuelles particulièrement difficile et nécessite des moyens spécialisés importants dont peu d'équipes disposent.

Finalement, l'ensemble des aspects juridiques, organisationnels et techniques à prendre en compte pour pouvoir utiliser les bases de données nationales de façon respectueuse des textes et efficace en termes de résultats constitue un véritable maquis juridico-institutionnel dans lequel les équipes de recherche et de surveillance ont souvent du mal à se repérer.

### ***L'utilisation de ces bases de données doit être organisée et facilitée***

Les analyses et recommandations ci-dessous reprennent les travaux du groupe « Systèmes d'information sanitaire » du HCSP :

#### **► Bases de données de l'assurance maladie, bases de données hospitalières et sociales :**

Les projets pour lesquels des données sont demandées doivent avoir fait l'objet d'évaluations scientifiques préalables par des organismes légitimes ; la finalité d'intérêt public de la demande doit être établie ; l'utilisation est réservée à des organismes publics ou parapublics, ou à des structures privées à but lucratif (directement ou par un intermédiaire académique) dans le cadre de la réalisation de certaines études (post-autorisation de mise sur le marché) commanditées par les pouvoirs publics ; un contrôle préalable par les organismes gestionnaires des bases de données doit être possible, le cas échéant suivi d'un avis négatif argumenté, de même que la mise en œuvre d'analyses à titre de contre-expertise.

► **Gouvernance :**

Il s'agit de l'organisation de l'accès aux bases, de la supervision de la réalisation des demandes et du contrôle de l'utilisation des données. Actuellement, chaque organisme gestionnaire de base « traite » directement avec les demandeurs ; cet état de choses n'est pas satisfaisant. Deux modèles de gouvernance ont été envisagés : décentralisée (chaque organisme public gestionnaire de bases de données fixe des règles explicites d'accès et met en place un « guichet » destiné à traiter les demandes) ; centralisée (une structure centrale gère un guichet unique et fait office d'interface entre les demandeurs et les organismes gestionnaires de bases de données). Les avantages et inconvénients de chaque modèle ont été analysés, notamment en termes de simplicité, d'autonomie et de moyens.

► **L'identifiant pour l'accès aux bases de données lorsqu'on ne dispose pas du NIR :**

Deux possibilités peuvent être envisagées : (i) recueil du NIR en clair auprès des personnes ou des professionnels et transfert à un « centre d'appariement sécurisé » (CAS), structure jouant un rôle de tiers de confiance habilitée à recevoir les NIR des sujets concernés, qui applique ensuite des procédures d'anonymisation ; dans l'état actuel des textes, cette solution ne permet pas d'éviter la prise d'un décret en conseil d'État spécifique de chaque étude ; (ii) recueil par appariement indirect (probabiliste) et consultation d'une base contenant les NIR de toute la population (RNIPP, RNIAM ou SNGI) ; cette méthode est employée pour l'accès au statut vital et la recherche des causes de décès et implique que le demandeur recueille uniquement les informations suivantes : nom, prénom, date et lieu de naissance des personnes concernées et les transmette (accompagnées d'un numéro d'étude destiné aux transferts des données) via un tiers de confiance à l'organisme détenteur de la base des NIR, qui peut ainsi retrouver les NIR par une méthode d'appariement indirect, et les transférer à qui de droit, sans que l'investigateur n'en ait connaissance. Sa mise en œuvre pratique (dont les aspects opérationnels sont décrits dans le rapport) implique la création d'un centre d'appariement sécurisé (CAS), tiers de confiance qui serait l'opérateur central des procédures à mettre en œuvre. Cette procédure a déjà été mise en œuvre, elle est éprouvée et fonctionne avec des performances satisfaisantes lorsque les informations initiales (nom, prénom, date et lieu de naissance) sont de bonne qualité. Elle implique cependant qu'au moins un organisme détenteur des NIR (Insee ou Cnav) mette en place un « service » ad hoc. Séduisante, la méthode d'appariement indirect n'est pas adaptée à toutes les situations. Par exemple les études avec des données en provenance des professionnels de santé ne pourraient pas obtenir le lieu de naissance, ou encore les études en situation d'urgence impliquent une rapidité incompatible avec les délais nécessaires. C'est pour répondre à ces situations que la Cnil propose l'élaboration d'un décret-cadre en conseil d'État « permettant l'utilisation encadrée du NIR à des fins de recherche médicale et d'études en santé publique », solution qui s'impose dans le cadre juridique actuel, dans la mesure où il n'existe pas de procédures sans utilisation du NIR permettant de répondre à toutes les situations. Au total, les deux approches (appariement indirect et recueil en clair du NIR autorisé par un décret-cadre) sont complémentaires. La méthode d'appariement indirect doit être privilégiée pour toutes les situations où elle est possible et suffisamment efficace et un principe de parcimonie doit être appliqué pour le recueil en clair du NIR. En pratique, c'est à la Cnil que revient d'examiner dans chaque cas les solutions possibles.

► **Correspondance INS – NIR :**

Le déploiement de l'INS rendra indispensables l'établissement et la maintenance d'une table de correspondance entre INS et NIR, seule à permettre les appariements entre bases de données de santé, identifiées à l'avenir avec l'INS, et les autres bases identifiées avec le NIR. Il semble prévu que la Cnav gèrera cette table de correspondance (ce qui reste à confirmer).

► **Création d'une plateforme d'interface entre les utilisateurs et les bases de données**

Une solution proposée pour améliorer la situation actuelle est la création d'une plateforme spécialisée qui jouerait le rôle d'interface entre les chercheurs et les bases de données nationales. Cette plateforme aurait comme missions principales : le conseil aux utilisateurs, la préparation des requêtes, la transmission des requêtes vers les bases de données et la récupération des fichiers extraits, la restitution aux utilisateurs de données synthétisées, après sélection préalable des variables d'intérêt à partir des données brutes. Deux modèles organisationnels peuvent être envisagés, qui ont chacun leurs avantages et inconvénients : chaque organisme gestionnaire de base développe un « guichet » ouvert aux utilisateurs remplissant ces fonctions ou création d'une plateforme centrale.

► **Localisation spatiale des personnes :**

la collecte de l'adresse par les organismes collecteurs de données de premier niveau (hôpital, CPAM...) permettrait par le géocodage de caractériser les zones géographiques concernées (indices de défavorisation, d'équipement en services.

## **4. SYNTHÈSE : MISE EN ŒUVRE DES MÉTHODES SELON LES OBJECTIFS**

### **TROIS GRANDES APPROCHES POUR CARACTÉRISER LES INÉGALITÉS SOCIALES DE SANTÉ.**

Une première approche consiste à recueillir des variables pour chaque individu, en complément des informations sanitaires individuelles. Cette approche nécessite un recueil complémentaire d'informations. Les indicateurs permettant de repérer une personne dans la hiérarchie sociale sont essentiellement le niveau d'études, la profession et le statut vis-à-vis de l'emploi, et le revenu, auxquels il faut ajouter, l'âge, le sexe et les indicateurs caractérisant les personnes issues de l'immigration.

La deuxième approche consiste à utiliser une approximation de la position socio-économique d'une personne en lui affectant un indice, l'indice de défavorisation sociale, qui est celui de sa zone géographique de résidence, opération qui n'est possible que si la zone géographique concernée comporte peu d'habitants.

Enfin, des connaissances extrêmement précieuses et précises peuvent être obtenues en croisant des données sanitaires avec des bases de données administratives, qui comportent des informations sur la profession, le parcours de vie, les revenus, les caractéristiques du lieu de résidence....

Le recueil de variables complémentaires a un coût et celles-ci ne sont utiles que si elles sont de qualité. Le HCSP a tenu dans sa réflexion à tenir compte de la faisabilité de recueil de données de qualité suffisante pour être interprétées en définissant des priorités.

### **DES APPROCHES VARIABLES SELON LES OBJECTIFS**

Comme c'est la règle en épidémiologie, les mesures de situation sociale choisies doivent correspondre à un objectif. Nous distinguerons les recommandations selon les objectifs suivants :

- 1) systèmes de recueil de données en routine (mortalité, morbidité, institutions de soins) ;
- 2) enquêtes, études et travaux de recherche, pour lesquelles le recueil de données peut être plus précis et plus long et construit ad-hoc ;
- 3) bilans des politiques, comme l'évaluation du plan cancer ou le bilan sur l'évolution des inégalités sociales de santé, tel que recommandé dans le rapport du HCSP.

Les recommandations ci-dessous relèvent parfois du déblocage de dispositions administratives ou juridiques d'accès à des données existantes : leur coût de recueil est donc marginal mais l'autorisation de leur accès ou de leur rapprochement relève de décisions administratives, légales et/ou politiques. D'autres ont indiscutablement un coût, qu'il s'agisse du recueil de données supplémentaires, de la formation de personnel ou du passage en routine de dispositifs de suivi épidémiologique, comme le géocodage. Dans tous les cas, l'information a un coût et l'enjeu de réduction des inégalités sociales de santé, récent sur l'agenda des politiques de santé, ne peut faire l'impasse sur le recueil d'information, comme cela est souligné dans le rapport OMS sur les déterminants sociaux (2008) ou le rapport du HCSP sur les inégalités sociales de santé (2009).

## **Systèmes d'information en santé permanents**

### **► Mortalité (certificats de décès)**

La PCS déclarée aujourd'hui sur le bulletin d'état civil de décès est très peu exploitée pour plusieurs raisons : elle fournit peu d'information sur la dernière profession des inactifs (dont les retraités), n'a pas d'homogénéité avec la déclaration faite au recensement (biais numérateur / dénominateur), est susceptible d'être influencée par la maladie ayant mené au décès. Toutefois, pour les populations de moins de 60 ans, des estimations de différentiels de mortalité faites à partir de ces données, bien que biaisées, sont vraisemblablement plus précises que des estimations faites à partir de l'EDP (de trop faible effectif pour permettre une précision statistique acceptable par cause de décès).

Plusieurs perspectives sont possibles :

- Exploiter le renseignement de la dernière profession exercée sur le bulletin d'état civil, même pour les inactifs. Cette solution n'apporterait pas une information indispensable dans la mesure où elle concernera surtout les personnes retraitées, tranche d'âge pour laquelle les panels permettent d'obtenir une puissance statistique suffisante.
- Demander le niveau d'étude sur le bulletin d'état civil. Le recueil de cette information est motivé par le souci de limiter les biais d'interprétation par les proches, et le biais lié à un déclassement socio-économique peu avant le décès. Il est également notable que cette information n'est pas récupérable par des appariements avec des bases administratives. Toutefois, certains travaux retrouvent des biais importants de déclaration également pour cette variable, en particulier attribuables aux fortes évolutions des catégorisations de niveau d'étude dans le temps.
- Apparier les causes de décès aux données de la CNAV ou à des données de déclarations fiscales. Ces données représentent une source riche de données, mais à ce jour non représentative de la population française, et avec des informations sur la PCS non standardisées et peu exploitables.
- Constituer un panel de plus grande taille que l'EDP actuel. Une multiplication de l'effectif de l'EDP par quatre par rapport à l'EDP actuel est en discussion et devrait constituer l'information de référence dans les années à venir. Cette solution, en facilitant l'accès à cette base de données appariée aux causes de décès, semble à privilégier.

### **► Morbidité**

Sans chercher à être exhaustif on peut citer parmi les sources de données ayant un intérêt en termes de données sociales :

- les registres de maladies : registre des cancers, des maladies rares, de malformations congénitales, et cardiopathies ischémiques, des accidents vasculaires cérébraux
- le réseau épidémiologie et information néphrologie (REIN)
- certaines maladies à déclaration obligatoire
- le système d'information de l'établissement français du sang
- le dossier communicant de cancérologie (DCC dont la mise en place sur l'ensemble du territoire est prévue avant la fin de 2013.
- le registre national des tentatives de fécondation in vitro ainsi que pour le registre national des interruptions médicales de grossesse tenue par l'agence de la biomédecine.

-le système d'information partagée pour l'autonomie des personnes handicapées (CDAPH)), mis en place par la caisse nationale de solidarité pour l'autonomie des personnes âgées et des personnes handicapées.

Dans ce premier ensemble de SIS, seraient essentiellement utilisables les variables individuelles et l'utilisation du géocodage renvoyant à un indicateur écologique de défavorisation sociale. Les systèmes d'information sanitaires de routine sont nombreux. Pour ces systèmes, les recommandations sur l'utilisation des variables individuelles restent ouvertes, car il s'agit de définir des priorités sans alourdir tout le recueil de données, ce qui n'aurait d'autre bénéfice que de réduire le taux de réponse dans certains cas ou recueillir des données de mauvaise qualité.

#### ► **Systèmes d'information hospitaliers**

Ils sont dépourvus d'informations décrivant la position socio-économique, à l'exception d'informations éventuelles sur les minima sociaux. Le géocodage devrait permettre d'obtenir rapidement des données sur le fonctionnement du système de soins dans son environnement social et géographique, pour les raisons de faisabilité et de qualité évoquées plus haut.

#### ► **Médecine ambulatoire**

La disponibilité d'un géocodage permettrait de développer rapidement des travaux sur le fonctionnement du système de soins dans son environnement social et géographique. Pour des travaux spécifiques, les variables individuelles essentielles doivent être ajoutées dans leur ensemble aux informations de recueil systématique. L'introduction de variables sociales dans les dossiers médicaux est en soi un travail de fond, orienté sur les représentations de la santé et de ses déterminants par les professionnels de santé.

### *Enquêtes et recherche*

► Cette catégorie inclut les travaux de recherche, les enquêtes nationales, régionales ou locales (ENNS, enquête décennale santé, enquêtes sur le handicap, ESPS, travaux des ORS...), les cohortes et toutes les études ponctuelles diverses. Dans ce type de recueil de données, compte-tenu de questionnaires souvent longs et riches, il importe de valoriser la possibilité de relier des données sanitaires à des déterminants sociaux, c'est-à-dire au minimum les variables essentielles.

### *Bilans pluriannuels*

► Par exemple, évaluation du plan cancer ou le bilan des inégalités sociales de santé en France proposé tous les cinq ans par le HCSP dans son rapport sur les inégalités sociales de santé.

Ce troisième ensemble d'utilisation fait appel à l'ensemble des informations : variables individuelles, géocodage, bases de données administratives. En utilisant les indicateurs géographiques de caractérisation des territoires, des informations pourront être obtenues sur le rôle propre des territoires sur l'amplification ou la réduction des inégalités sociales de santé. Enfin, ces bilans devraient faire apparaître l'évolution des déterminants de santé dont on sait qu'ils jouent un rôle dans les inégalités sociales de santé (expositions toxiques professionnelles par exemple).

## 5 RECOMMANDATIONS

Tant dans le rapport OMS sur les déterminants sociaux de la santé en 2008 que dans le rapport du HCSP publié en 2010, l'amélioration du dispositif statistique permettant de suivre l'évolution des inégalités sociales de santé est une priorité. Cette amélioration a un coût mais représente un investissement à la hauteur de l'enjeu de réduction des inégalités sociales de santé.

Nous présenterons dans un premier temps les variables et méthodes utilisables pour caractériser les inégalités sociales de santé, puis les systèmes d'information ou les évaluations dans lesquelles elles peuvent être recommandées.

### VARIABLES ET METHODES POUR CARACTERISER LES INEGALITES SOCIALES DE SANTE

#### *Variables individuelles essentielles*

La triade de mesure des inégalités sociales de santé repose sur la profession et la situation vis-à-vis de l'emploi, le niveau d'études et le revenu.

***Proposition 1 : Mesurer la profession et la catégorie socioprofessionnelle (PCS), couplée au statut vis-à-vis de l'emploi, le niveau d'études et autant que possible, les variables caractérisant le revenu.***

À côté de la problématique des inégalités sociales de santé, se pose celle de la précarité. Le bilan du groupe de travail retrouve un certain nombre de scores de précarité parfois largement utilisés dans les travaux français sur l'identification de la précarité. Aucun n'apparaît réellement validé à ce jour.

***Proposition 2 : Susciter des travaux de recherche sur la validation d'un score de précarité validé, utilisable dans les systèmes d'information***

Compte tenu des liens étroits entre la question des inégalités sociales de santé et celle de la santé des populations immigrées, le recueil de ces données est hautement souhaitable dans certains systèmes d'information de routine

***Proposition 3 : Recueillir le statut migratoire avec un niveau de détail adapté au type d'étude***

Il est important, compte-tenu de l'importance des premières années de vie sur l'état de santé de l'adulte, de caractériser la position socio-économique des enfants et jeunes adultes. L'âge limite auquel la situation socio-économique personnelle est plus indicative que celle des parents pourrait être fixé à 25 ans, critère retenu dans certains programmes sociaux.

***Proposition 4 : Recueillir pour les enfants et les jeunes adultes avant l'âge de 25 ans, la situation socio-économique des parents.***

## **Indicateurs écologiques de défavorisation sociale**

Nous ne proposons pas de choix entre les indicateurs écologiques de défavorisation sociale. Nous avons identifié et décrit les principaux indices développés et utilisés en France et présenté leurs principaux avantages et limites. Il nous a semblé qu'il fallait préférer des indices développés en France pour les recherches menées sur des données françaises.

**Proposition 5 : Si on ne dispose pas de variable individuelle de situation sociale, utiliser l'un des indicateurs de défavorisation élaborés et validés en français, en fonction des objectifs visés.**

## **Bases de données administratives**

**Proposition 6 : Les bases de données administratives doivent être largement accessibles à des fins de recherche et de surveillance avec des règles d'ouverture destinées à garantir la finalité et la qualité scientifique des travaux menés.**

**Proposition 7 : L'accès aux bases de données administratives doit être organisé de manière à assurer une interface accessible à tous dans des délais et à des coûts raisonnables.**

**Proposition 8 : Créer un centre d'appariements sécurisés permettant l'appariement des bases de données administratives entre elles et/ou à des données d'enquêtes ou de registres ; élaborer un décret-cadre en conseil d'État « permettant l'utilisation encadrée du NIR à des fins de recherche médicale et d'études en santé publique ».**

**Proposition 9 : Créer une ou des plate-forme(s) d'interface entre les chercheurs et les producteurs de bases de données nationales afin d'en faciliter l'utilisation.**

**Proposition 10 : Mettre en place une procédure interne automatisée par laquelle l'adresse des personnes serait géocodée, convertie en code Iris et transférée sous cette forme dans les bases nationales ; encourager l'utilisation de la base permanente des équipements de l'INSEE pour quantifier l'accès à divers équipements et services dans les territoires.**

**Proposition 11 : Modifier l'article L. 135 D du livre des procédures fiscales pour étendre au bénéfice des chercheurs une dérogation au secret professionnel, afin de leur permettre l'accès aux données fiscales à des fins de recherche scientifique.**

## MISE EN ŒUVRE DES METHODES SELON LES OBJECTIFS

### Systemes d'information sanitaires permanents

► Mortalité : certificats de décès

**Proposition 12:** *Constituer un panel de plus grande taille que l'Echantillon démographique permanent (EDP) actuel. Un projet d'EDP comprenant un nombre de personnes incluses dans le panel multiplié par quatre par rapport à l'EDP actuel devrait constituer l'information de référence dans les années à venir. Un accès facilité à cette base de données appariées aux causes de décès doit être privilégié pour l'analyse des inégalités sociales de mortalité.*

**Proposition 13 :** *Publication annuelle par l'INSEE des statistiques de mortalité par PCS, au lieu d'une publication épisodique comme c'est le cas actuellement.*

► Morbidité

**Proposition 14:** *Pour le suivi des inégalités sociales de santé en matière de morbidité (population et institutions sanitaires), développer rapidement le géocodage à partir des adresses de résidence et l'utilisation d'indices de défavorisation. Cette première étape sera complétée le plus rapidement possible par le développement de l'appariement avec des bases de données administratives pour produire ces informations. Les variables individuelles devraient être introduites progressivement dans les SIS en fonction des objectifs*

### Enquêtes et recherche

On trouve dans cette catégorie les travaux de recherche, les enquêtes nationales, régionales ou locales (ENNS, travaux des ORS...), les cohortes et toutes les études ponctuelles diverses. Dans ce type de recueil de données, compte-tenu de questionnaires souvent longs et riches, il importe de valoriser la possibilité de relier des données sanitaires à des déterminants sociaux.

**Proposition 15 :** *Utiliser, dans les enquêtes sur des pathologies spécifiques et la recherche l'ensemble des données : variables individuelles, vade-mecum d'utilisation de données sociales et accès facilité aux croisements entre elles des bases de données administratives. L'utilisation d'indicateurs géographiques, avec les techniques statistiques appropriées, pour analyser le rôle de l'environnement géographique est recommandée, indépendamment de l'utilisation des indicateurs de défavorisation*

### **Bilans pluriannuels des politiques ou programmes de santé**

On trouve dans cette catégorie l'évaluation du plan cancer ou le bilan des inégalités sociales de santé en France proposé tous les cinq ans par le HCSP dans son rapport sur les inégalités sociales de santé.

**Proposition 16 : Réaliser tous les cinq ans un bilan des inégalités sociales de santé en France.**

**Utiliser, dans les évaluations de plan ou les bilans pluriannuels du HCSP de l'ensemble des données discutées dans les recommandations précédentes et effectuer un suivi approfondi de l'évolution des déterminants socio-économiques pertinents pour mesurer leur impact sur la santé (modes d'habitat, conditions et accès au travail, aménagement du territoire, expositions professionnelles...).**

## **DEUXIEME PARTIE : Suivi des inégalités dans le domaine du cancer**

Le bilan à mi-parcours du plan cancer souligne l'absence d'un système d'information sanitaire adapté à la surveillance des inégalités, qui contraste avec l'objectif affiché de réduction des inégalités sociales de santé. Certaines limites ont été relevées. On sait que l'accès différentiel à l'innovation est un facteur d'inégalités sociales de santé. Le plan cancer le souligne, mais aucun objectif chiffré n'est associé à cette mesure et les données sociales sont difficiles d'accès ou inexistantes. Comment mesurer l'ampleur des difficultés et les progrès ?

L'insuffisance des données sociales et de données d'exposition ou économiques constitue la difficulté centrale. Les recommandations générales du HCSP sur le recueil des variables sociales et les évolutions nécessaires des systèmes d'information sanitaires présentées dans la première partie de ce rapport sont également pertinentes pour suivre les inégalités sociales de santé dans le domaine du cancer. Le choix des indicateurs caractérisant la position socio-économique, la mise en œuvre de processus de géocodage et l'utilisation d'indicateurs de déprivation ou encore le développement d'appariements de bases de données administratives bénéficierait au suivi des inégalités sociales de santé dans le domaine du cancer comme celui dans d'autres pathologies.

Mais certaines spécificités du cancer entraînent une difficulté supplémentaire. Les grandes enquêtes nationales et transversales en population générale sont peu utilisables du fait du faible effectif relatif au cancer dans leurs échantillons : Enquête décennale sante de l'INSEE, enquête santé protection sociale de l'IRDES ou encore le Baromètre cancer de l'INPES (dans l'Édition 2005, 226 personnes déclaraient avoir eu un cancer dans leur vie).

Il est donc pertinent de compléter la réflexion générale de la première partie de ce rapport par un développement spécifique au cancer. Ces spécificités tiennent aux objectifs assignés dans le plan cancer 2009-2013, et à l'importance des expositions environnementales et professionnelles, tant en ce qui concerne l'incidence des cancers que les inégalités sociales de santé. Ainsi, pour le CIRC, 50% des inégalités sociales de mortalité par cancer du poumon sont liées aux expositions professionnelles.

Deux outils sont particulièrement précieux dans le domaine du cancer : les registres des cancers, particulièrement importants pour les travaux sur l'incidence et la prise en charge des cancers et le dossier communiquant de cancérologie pour ce qui concerne la prise en charge.

## **1. LES MESURES DU PLAN CANCER 2009-2013 QUI NECESSITENT UN SIS CAPABLE DE MESURER LES INEGALITES SOCIALES DE SANTE**

### ***Des mesures de prévention, dans le domaine de l'environnement professionnel et des comportements***

Mesure 9 : Améliorer l'observation et la surveillance des cancers liés à l'environnement professionnel.

Mesure 12 : Renforcer la prévention des cancers liés à l'environnement en particulier dans le domaine professionnel.

Mesure 10 : Poursuivre la lutte contre le tabac.

Mesure 11 : Promouvoir des actions de prévention sur les liens entre l'alimentation, l'activité physique et les cancers.

### ***Dépistage***

Mesure 14 : Lutter contre les inégalités d'accès et de recours au dépistage.

### ***Un parcours de soins personnalisé et efficace***

Mesure 21 : Garantir un égal accès aux traitements et aux innovations.

### ***Vivre pendant et après un cancer***

Améliorer la qualité de vie pendant et après la maladie, combattre toute forme d'exclusion

Mesure 26 : Se doter des moyens et outils nécessaires au développement de l'accompagnement social personnalisé.

Mesure 27 : Améliorer les réponses aux possibles situations de handicap ou de perte d'autonomie transitoires ou définitives liées au cancer.

Mesure 28 : Garantir l'accès des personnes malades et guéries aux assurances et au crédit.

Mesure 29 : Lever les obstacles à la réinsertion professionnelle des personnes atteintes de cancer.

## **2. UN SYSTEME D'INFORMATION SANITAIRE SPECIFIQUE AU CANCER**

### ***Le dossier communicant de cancérologie***

A la différence du DMP, dossier médical personnalisé, qui dans l'esprit de la loi n° 2004-810 du 13 août 2004 est un dossier « personnel » c'est-à-dire un dossier dont le patient est propriétaire et qu'il est amené à gérer lui-même, le DCC est quant à lui un outil « métier ». C'est-à-dire qu'il vise à partager entre professionnels de santé (notamment experts en cancérologie) des informations médicales permettant la prise en charge médicale du patient. La conception et la mise en œuvre d'un outil de coordination de la prévention et de la prise en charge selon les positions socio-économiques, pourraient être basées sur ce DCC, avec l'inclusion de variables individuelles et /ou de proxys écologiques (indicateurs de défavorisation sociale) de position socio-économique. Le géocodage pourrait permettre des analyses en termes sociaux et territoriaux ou géographiques.

### ***Registres***

Les registres généraux qualifiés couvrent actuellement environ 20% de la population française. Ce niveau de couverture est suffisant pour des objectifs de surveillance nationale

et de recherche mais pose des problèmes pour la production d'indicateurs d'aide à la décision à un niveau territorial.

Le mode de fonctionnement actuel des registres est explicité en annexe.

La recherche de données individuelles caractérisant la position socio-économique ne peut être possible que pour des enquêtes ou des suivis limités dans le temps et la collecte et n'est sans doute pas identique d'un cancer à l'autre (en fonction de leur incidence et de leur létalité notamment). En revanche l'utilisation de géocodage et d'indicateurs écologiques de défavorisation sociale, telle que décrite dans la première partie du rapport est une solution qui permettrait à court terme un accès à la caractérisation sociale des données de registres. Le développement de l'appariement à diverses bases de données permettra à moyen terme les analyses approfondies sur le lien entre déterminants sociaux et incidence, survie et prise en charge des cancers.

### ***Une absence de coordination nationale de la mise en œuvre des systèmes d'information***

Au niveau national, le HCSP n'a pas identifié de réelle coordination de la conception et de la mise en œuvre des systèmes d'information nécessaires au Plan cancer.

### ***Une palette de variables et de dispositifs de production de connaissance***

Les inégalités de mortalité seront disponibles dans le cadre des recommandations générales sur les certificats de décès et l'information issue d'un EDP transformé et accessible. Les données de survie, incidence, degré de sévérité de certains cancers relèvent des registres, donc du géocodage et de l'appariement de bases de données. Les facteurs de risque font l'objet de publication dans le rapport annuel des données de consommation tabac alcool (Baromètre santé) et de nutrition (ENNS) ventilées par données sociales. Les données sur l'accès aux soins reposent sur le DCC et les dispositifs de géocodage dans les systèmes d'information des structures sanitaires. Le suivi des inégalités d'accès au dépistage implique le recueil de données sociales par les centres de gestion du dépistage organisé.

Trois objectifs nécessitent un travail spécifique.

-Le lien entre expositions professionnelles, survenue des cancers et inégalités sociales de mortalité a été bien documenté. La faible reconnaissance des cancers d'origine professionnelle contribue aux inégalités sociales de santé. Les données sont insuffisantes à cet égard et requièrent un travail spécifique qui sort des objectifs de ce rapport.

-Les données sur l'après-cancer et les inégalités qu'il recouvre, sur les conditions de vie, l'accès au crédit, la réinsertion professionnelle, nécessitent des enquêtes spécifiques.

-Les calculs de survie après un cancer font appel à la survie relative, qui permet de tenir compte du gain apporté par les traitements en tenant compte de l'espérance de vie de la personne concernée. Pour analyser cette survie relative, il est essentiel de disposer des tables de mortalité selon des critères démographiques (âge, sexe). Concernant les inégalités sociales de santé, il est indispensable de disposer de tables de mortalité par position socio-économique. En pratique, ces tables devraient être disponibles par indice écologique de défavorisation sociale.

### 3. RECOMMANDATIONS

Le rapport d'évaluation à mi-parcours du plan cancer recommandait de consolider et faire évoluer le système d'information épidémiologique par la mise en place d'un système épidémiologique à trois niveaux, par ordre de complétude croissante (et de réactivité décroissante) : le premier reposant sur les registres, le second sur les appariements multi sources, incluant les données anatomo-pathologiques, le dernier constitué d'un appariement du PMSI, des ALD, et du statut vital issu du RNIPP ou des registres de l'Assurance maladie.

***Proposition 17 : Faciliter la prise en compte des caractéristiques sociales individuelles et territoriales, en développant le géocodage et l'utilisation des indicateurs de défavorisation dans les registres du cancer en France***

***Proposition 18 : Développer l'appariement des registres des cancers avec des bases de données administratives et développer une structure technique d'interface au niveau national entre producteurs et utilisateurs de données médico-administratives.***

***Proposition 19 : Produire et communiquer annuellement des informations sur les inégalités sociales de santé dans le domaine du cancer, en rapport avec les mesures et les objectifs fixés dans le plan cancer***

Les objectifs du plan cancer 2 ont pu contribuer à obscurcir la relation entre territoire et social, avec une tendance à substituer les inégalités sociales de santé par les inégalités territoriales ou à les y intégrer, celles-ci étant généralement plus mises en avant dans le plan que les inégalités sociales de santé. Par exemple, la mesure 14 visait à lutter contre les inégalités de recours au dépistage ; mais l'indicateur chiffré invitait à cibler les « départements rencontrant le plus de difficultés ».

***Proposition 20 : Prendre en compte dans les analyses territoriales le double aspect d'une zone géographique : son environnement physique, chimique, potentiellement en lien avec la survenue de cancers, l'accessibilité géographique des structures sanitaires d'une part et sa composition sociale de l'autre. Une surincidence de cancers ou une faible survie peuvent relever d'approches environnementales et géographiques, mais aussi d'approches sociales***

***Proposition 21 : Intégrer des données individuelles indicatrices de la position socio-économique au « Dossier Communicant de cancérologie » et intégrer des procédures de géocodage en routine.***

***Proposition 22 : Produire des tables de mortalité par indice écologique de défavorisation sociale, pour permettre les calculs de survie relative tenant compte du sexe, de l'âge, mais aussi de la position socio-économique***

***Proposition 23 : Identifier une structure responsable du pilotage stratégique des systèmes d'information relatifs au cancer.***

## REFERENCES

1. Adler A, Bush NR, Pantell MS. Rigor, vigor, and the study of health disparities. *PNAS* 2012; 109: 17154-59.
2. ARS Ile-de-France. Tableau de bord du dépistage organisé du cancer du sein en Ile-de-France, données 2009. Paris, septembre 2011.
3. Bambra C. Health inequalities and welfare state regimes: theoretical insights on a public health 'puzzle'. *J Epidemiol Community Health*. 2011 Sep;65(9):740-5.
4. Bartley M, Sacker A, Firth D, Fitzpatrick R. Understanding social variation in cardiovascular risk factors in women and men: the advantage of theoretically based measures. *Social science and medicine*. 1999;49:831-45.
5. Baudier F, Michaud C, Gautier A, Guilbert P (2007). Le dépistage du cancer du sein en France : pratiques et évolution des habitudes dans la population des femmes de 18 à 75 ans, Baromètre santé 2004-2005. *BEH* 17 :142-144.
6. Berchet C, Jusot F. Etat de santé et recours aux soins des immigrés : une synthèse des travaux français. *Questions d'économie de la santé* 2012, n°172.
7. Berkman LF. Social epidemiology: social determinants of health in the United States, are we losing ground? *Annu Rev Public Health* 2009; 30: 27-41.
8. Berney LR, Blane DB. Collecting retrospective data: accuracy of recall after 50 years judged against historical records. *Soc Sci Med*. 1997;45(10):1519-25.
9. Bleich SN, Jarlenski MP, Bell CN, LaVeist TA. Health inequalities: trends, progress, and policy. *Annu Rev Public Health* 2012; 33: 7-40.
10. Bourdelais P. L'inégalité sociale face à la mort : l'invention récente d'une réalité ancienne. in *Les inégalités sociales de santé*. A Leclerc D Fassin, H Grandjean, M Kaminski, T Lang. Paris : 2000.)
11. Bourgeuil Y, Jusot F, Leleu H, et le groupe AIR Project. Comment les soins primaires peuvent-ils contribuer à réduire les inégalités sociales de santé ? *Revue de la littérature. Questions d'économie de la santé* 2012, n°179.
12. Boyce WT, Sokolowski MB, Robienson GE. Toward a new biology of social adversity. *PNAS* 2012; 109: 17143-48.
13. Brousse C. ESeC, projet de classification socio-économique. *Courrier des Statistiques* 2008 ; n°125 : 27-36.
14. Brousse C. ESeC: the European Union's socio-economic classification project. *Courrier des statistiques, English series* 2009;15:27-36.

15. Brown L, Thurecht L, Nepal B. *The cost of inaction on the social determinants of health*. Canberra, National Centre for Social and Economic Modelling, 2012, report no. 2/2012.
16. Bureau international du travail. *Structure de la Classification internationale type des professions, 2008 (CITP-08)*. Genève, BIT, décembre 2007.
17. Cambois E, Laborde C, Robine JM. La « double peine » des ouvriers : plus d'années d'incapacité au sein d'une vie plus courte. *Population et Sociétés*. 2008 ;No 441 :1-4.
18. Cambois E. La mesure des inégalités sociales face à la santé: problèmes méthodologiques. In : Collectif. *Morbidité, mortalité: problèmes de mesure, facteurs d'évolution, essai de prospective*. Paris, AIDELF/PUF, 1998 : 422-32.
19. CDC, National center for AIDS/HIV, viral hepatitis, STD and TB prevention. Establishing a holistic framework to reduce inequities in HIV, Viral hepatitis, STDs, and Tuberculosis in the United States ; An NCHHSTP white paper on social determinants of health, 2010.
20. Centers for Disease Control and Prevention. CDC Health Disparities and Inequalities Report - United States, 2011. *MMWR* 2011; 60 / Supplement.
21. Chaix B, Chauvin P. L'apport des méthodes d'analyse multiniveau dans l'analyse contextuelle en épidémiologie sociale : une revue de la littérature. *Rev Epidemiol Santé Publ* 2002; 50: 489-99.
22. Chaix B. Geographic life environments and coronary heart disease: a literature review, theoretical contributions, methodological updates, and a research agenda. *Annu Rev Public Health* 2009; 30: 81-105.
23. Chauvin P, Lebas J. Inégalités et disparités sociales de santé. In : Bourdillon F, Brückner G, Tabuteau D, eds. *Traité de santé publique*. Paris, Flammarion Médecine Sciences, 2007 (2<sup>ème</sup> édition revue et augmentée), pp. 331-41.
24. Chauvin P. Environnement social et santé : avancées et perspectives dans l'étude des effets du contexte sur la santé. In : P. Chauvin, I Parizot, eds. *Santé et expériences de soins: de l'individu à l'environnement social*. Paris, Vuibert, 2005 : 187-200.
25. Chittleborough CR, Baum FE, Taylor AW, Hiller JE. A life-course approach to measuring socioeconomic position in population health surveillance system. *J Epidemiol Community Health* 2006; 60: 981-92.
26. Collectif. Numéro thématique : santé et recours aux soins des migrants en France. *BEH* 2012, n°2-3-4.
27. Costa-Front J, Hernandez-Quevedo C. Measuring inequalities in health: what do we know? What do we need to know? *Health Policy* 2012; 106: 195-206.

28. Debet A. *Mesure de la diversité et protection des données personnelles*. Paris, CNIL, 15 mai 2007.
29. Delame N, Thomas G. Revenus des foyers d'agriculteurs. *Insee Première* 2006, n° 1068.
30. Delpierre C, Lauwers-Cances V, Datta GD, Lang T, Berkman L. Using self-rated health for analysing social inequalities in health: a risk for underestimating the gap between socioeconomic groups? *J Epidemiol Community Health* 2009 Jun;63(6):426-32 .
31. Delpierre C, Lauwers-Cances V, Datta GD, Berkman L, Lang T. Do cardiovascular risk factors impact self-rated health equally across social position? *Am J Publ Health* 2009 Jul;99(7):1278-84.
32. Delpierre C, Lauwers-Cances V, Datta GD, Berkman L, Lang T. Do cardiovascular risk factors impact self-rated health equally across social position? *Am J Publ Health* 2009;99:1278-84.
33. Delpierre C, Lauwers-Cances V, Datta GD, Lang T., Berkman L. Using self-rated health for analysing social inequalities in health: a risk for underestimating the gap between socioeconomic groups? *Journal of Epidemiology and Community Health* 2009; 63:426-32
34. Devaux M, Jusot F, Sermet C, Tubeuf S. Hétérogénéité sociale de déclaration de l'état de santé et mesure des inégalités de santé. *Rev Fr Affaires Sociales* 2008; 1: 29-46.
35. Diez-Roux A. The study of group-level factors in epidemiology: rethinking variables, study designs, and analytical approach. *Epidemiol Rev* 2004; 26: 104-11.
36. Dourgnon P, Guillaume S, Rochereau T. Enquête sur la santé et la protection sociale 2010. *Les rapports de l'IRDES* 2012, n°1886.
37. Dubet F. Classes sociales et description de la société. *Revue Française de Socio-Economie* 2012; 2: 259-64.
38. Duncan GJ, Daly MC, McDonough P, Williams DR. Optimal indicators of socioeconomic status for health research. *American Journal of public health* 2002;92(7):1151-57.
39. European Centre for Disease Prevention and Control. *Improving HIV data comparability in migrant populations and ethnic minorities in EU/EEA/EFTA countries: findings from a literature review and expert panel*. Stockholm, ECDC, 2011.
40. European Parliament. Resolution of 8 March 2011 on reducing health inequalities in the EU (2010/2089(INI)). Strasbourg, P7\_TA(2011)0081.
41. Fassin D. *La santé des immigrés et des étrangers : méconnaissance de l'objet et objet de reconnaissance*. In : Joubert JM, Chauvin P, Facy F, Ringa V, eds.

- Précarisation, risque et santé*. Paris : Inserm, coll. Questions en santé publique, 2001 : 187-96.
42. Galobardes B, Lynch J, Davey-Smith G. Measuring socioeconomic position in health research. *Br Med Bull* 2007; 81-82: 21-37.
  43. Galobardes B, Shaw M, Lawlar DA, Lynch LW, Davey Smith G. Indicators of socioeconomic position (Part 1). *J Epidemiol Community Health* 2006; 60: 7-12.
  44. Galobardes B, Shaw M, Lawlar DA, Lynch LW, Davey Smith G. Indicators of socioeconomic position (Part 2). *J Epidemiol Community Health* 2006; 60: 95-101.
  45. Gordon D. Area-based deprivation measures – a UK perspective. In: Kawachi I, Berkman L, eds. *Neighborhoods and health*. New York, Oxford University Press? 2003: 179-210.
  46. Grillo F, Soler M, Chauvin P. L'absence de dépistage du cancer du col de l'utérus en fonction des caractéristiques migratoires chez les femmes de l'agglomération parisienne en 2010. *BEH* 2012; n°2-3-4: 45-7.
  47. Hauser M. Measuring socioeconomic status in studies of child development. *Child Dev* 1994; 65: 1541-5.
  48. Haut Conseil de la Santé Publique. *Les inégalités sociales de santé : sortir de la fatalité*. Paris, La Documentation française, 2009.
  49. Institut Canadien d'Information sur la Santé. *Réduction des écarts en matière de santé : un regard sur le statut socioéconomique en milieu urbain au Canada*. Ottawa, ICIS, 2008.
  50. Johnson PJ, Blewett LA, Davern M. Disparities in public use data availability for race, ethnic, and immigrant groups. *Med Care* 2010; 48: 1122-7.
  51. Journal Officiel de l'Union européenne. Avis du Comité Economique et Social Européen sur « Santé et Migration ». 2007, C256/123.
  52. Kelly-Irving M, Delpierre C, Schiebera A-C, Lepage B, Rolland C, Afrité A, Pascal J, Cases C, Lombrail P, Lang T. the INTERMEDE group. Do general practitioners overestimate the health of their patients with lower education? *Social Science & Medicine*. 2011;73 1416-21.
  53. Krieger N, Chen JT, Selby JV. Comparing individual-based and household-based measures of social class to assess class inequalities in women's health: a methodological study of 684 US women. *J Epidemiol Community Health* 1999; 53: 612-23.

54. Krieger N, Chen JT, Waterman PD, et al. Painting a truer picture of US socioeconomic and racial/ethnic health inequalities: the Public Health Disparities Geocoding Project. *Am J Public Health* 2005; 95: 212-23.
55. Krieger N, Okamoto A, Selby JV. Adult Female Twins' Recall of Childhood Social Class and Father's Education: A Validation Study for Public Health Research. *American Journal of Epidemiology* 1998;147(7):704-08
56. Krieger N, Smith K, Naishadham D, Hartman C, Barbeau EM. Experiences of discrimination: validity and reliability of a self-reported measure for population health research on racism and health. *Soc Sci Med* 2005; 61: 1576-96.
57. Krieger N, Williams DR, Moss NE. Measuring social class in the US public health research : concepts, methodologies and guidelines. *Annu Rev Public Health* 1997; 18: 341-78.
58. Krieger N. Overcoming the absence of socioeconomic data in medical records: validation and application of a census-based methodology. *Am J Public Health* 1992; 82: 703-10.
59. Kunst AE, Bos V, Mackenbach JP, EU Working Group on Socio-economic Inequalities in Health. Monitoring socio-economic inequalities in health in the European Union: guidelines and illustrations. Final report. Department of Public Health Erasmus University Rotterdam The Netherlands 2001
60. Lardjane S, Dourgnon P. Les comparaisons internationales d'état de santé subjectif sont-elles pertinentes ? Une évaluation par la méthode des vignettes-étalons *Economie et Statistique*, n° 403-404, 2007/12, 165-177.
61. Leclerc A, Kaminski M, Lang T. Inégaux face à la santé : du constat à l'action. La Découverte Ed. Paris : 2008. 298 p.
62. Leclerc A, Kaminski M, Lang T. Inégaux face à la santé. Du constat à l'action. La Découverte/Inserm Eds. Paris : 2008. 300p.
63. Leclerc A, Niedhammer I, Plouvier S, Melchior M. Travail, emploi et inégalités sociales de santé. *BEH* 2011 n°8-9: 79-81.
64. Legleye S. *Inégalités de genre et inégalités sociales dans les usages de drogues en France*. Université Paris 11, Thèse de doctorat en santé publique, octobre 2011.
65. Liberatos P, Link BG, Kelsey JL, The measurement of social class in epidemiology. *Epidemiologic Reviews* 1988; 10:87-121.
66. Lien N, Friestad C, Klepp KI. Adolescents' proxy reports of parents' socioeconomic status: How valid are they? *J Epidemiol Community Health* 2001; 55: 731-7.
67. Lin SS, Kelsey JL. Use of race and ethnicity in epidemiologic research : concepts, methodological issues, and suggestion for research. *Epidemiol Rev* 2000; 22: 187-202.
68. Lombrail P. Inégalités de santé et d'accès secondaire aux soins. *Revue d'Epidémiologie et Santé Publique* 2007 ;55(1) :23-30.

69. Mackenbach J, Junst A, Stirbu, et al. *Tackling health inequalities in Europe: an integrated approach – Eurothine, final report*. Rotterdam, Erasmus Medical Centre, Department of public health, 2007.
70. Mackenbach JP et al. Socioeconomic inequalities in Health among 22 European Countries. *New England Journal of Medicine* 2008;358:2468-81.
71. Macleod J, Smith GD, Matcalfe C, Hart C. Is subjective social status a more important determinant of health than objective social status? Evidence from a prospective observational study of Scottish men. *Social Science and Medicine* 2005;61:1916-29.
72. Mahoney M, Simpson S, Harris E, Aldrich R, Williams JS. Equity-focused Health Impact Assessment Framework. The Australasian Collaboration for Health Equity Impact Assessment (ACHEIA). (2004).
73. Maruani M. *Travail et emploi des femmes*. Paris, La Découverte (coll. Repères), 2003, n° 287.
74. Ministry of Social Affairs and Health. *National action plan to reduce health inequalities 2008-2011*. Helsinki, MSAH, 2008.
75. Moleux M, Schaezel F, Scotton C. Les inégalités sociales de santé : déterminants sociaux et modèles d'action. Paris, Inspection Générale des Affaires Sociales, mars 2011.
76. Niedhammer I. Les facteurs psychosociaux et la violence psychologique au travail. *Actualité et Dossier en Santé Publique* 2006; 57: 50-52.
77. Nielsen SS, Krasnik A. Poorer self-perceived health among migrants and ethnic minorities versus the majority population in Europe: a systematic review. *Int J Public Health* 2010; 55: 357-71.
78. Niewiadomski C, Aiach P, eds. *Lutter contre les inégalités sociales de santé : politiques publiques et pratiques professionnelles*. Rennes, Presses de l'EHESP, 2008.
79. Norwegian Ministry of Health and Care Services. *National strategy to reduce social inequalities in health*. Oslo, NMHCS, 2007.
80. Organisation Internationale pour la migration. *L'état de la migration dans le monde 2010, L'avenir des migrations : Renforcer les capacités face aux changements*. Genève, OIM, 2010.
81. Organisation Mondiale de la Santé. *Déclaration politique de Rio sur les déterminants sociaux de la santé*. OMS, Rio de Janeiro, 21 octobre 2011.
82. Pampalon R, Hamel D, Gamache P, Raymond G. A deprivation index for health planning in Canada. *Chron Dis Can* 2009; 29: 178-91.

83. Pan Ké Shon JL. La ségrégation des immigrés en France : état des lieux. *Population & Sociétés* 2011; n°477 : 1-4.
84. Penissat E. La difficile production d'une nomenclature socioprofessionnelle à l'échelle européenne. *Revue Française de Socio-Economie* 2012; 2: 251-7.
85. Poder TG. Inégalité de revenu et état de santé : une relation ambiguë. *Interrogations* 2011, n°12.
86. Pornet C, Delpierre C, Dejardin O, Grosclaude P, Launay L, Guittet L, Lang T, Launoy G. Construction of an adaptable European transnational ecological deprivation index: the French version. *J Epidemiol Community Health*. 2012 Apr 27. Epub ahead of print
87. Provincial Health Services Authority. *Towards reducing health inequalities: a health system approach to chronic diseases prevention*. Vancouver, Population & public health, PHSA, 2011.
88. *Public Health Reports* 2011; 26 (S3): 134 p.
89. Rey G, Rican S, Luce D, Menvielle G, Jouglu E.. Measuring social inequalities in cause-specific mortality in France: comparison between linked and unlinked approaches. *Rev Epidemiol Santé Publique* 2013;61(3):221-31.
90. Ribet C, Melchior M, Lang T, Zins M, Goldberg M, Leclerc A. Caractérisation et mesure de la situation sociale dans les études épidémiologiques. *Rev Epidemiol Santé Publ* 2007; 55: 285-95.
91. Robert J, Roudier C, Poutignat N, Fagot-Campagna A, Weill A, Rudnichi A, Thammavong N, Fontbonne A, Detournay B, pour le comité scientifique d'Entred 2007-2010. Prise en charge des personnes diabétiques de type 2 en France en 2007 et tendances par rapport à 2001. *BEH* 2009 ; No 42-43 : 455-60.
92. Ruiz H. Pauvreté monétaire relative et minima sociaux : une analyse à partir de l'enquête Revenus fiscaux de 2005.
93. Saurel-Cubizolles MJ, Chastang JF, Menvielle G, Leclerc A, Luce D; EDISC group. Social inequalities in mortality by cause among men and women in France. *J Epidemiol Community Health*. 2009 Mar;63(3):197-202
94. Schieber A. C. , Kelly-Irving M., Rolland Ch., Afrite A., Cases Ch., Dourgnon P. , Lombrail P;, Pascal J. Lang T. Do doctors and patients remember the same elements of the cardiovascular risk-factors' management after the consultation? the INTERMEDE study. *British Journal of General Practice* 2011;61(584):178-183.
95. Seillier B. *La lutte contre la pauvreté et l'exclusion : une responsabilité à partager*. Paris, Sénat, rapport d'information n°445, juillet 2008.
96. Shavers VL. Measurement of socioeconomic status in health disparities research. *JNMA* 2007; 99: 1013-23.

97. Singh-Manoux a. Clarke P., Marmot M. Multiple measures of socio-economic position and psychosocial health: proximal and distal measures. *International journal of Epidemiology* 2002;31:1192-99
98. Solé-Auró A, Crimmins EM. Health of immigrants in European countries. *Int Migr Rev* 2008; 42: 861–76.
99. Spinakis A, Anastasiou G, Panousis V, Spiliopoulos K, Palaiologou S, Yfantopoulos J. *Expert review and proposals for measurement of health inequalities in the European Union – Summary Report*. Luxembourg, EC DG for Health and Consumers, 2011.
100. Vallée J, Cadot E, Grillo F, Parizot I, Chauvin P. The combined effects of activity space and neighbourhood of residence on participation in preventive health-care activities: The case of cervical screening in the Paris metropolitan area (France). *Health Place* 2010; 16: 838-52.
101. Vaucher P, Bischoff T, Diserens EA, et al. Detecting and measuring deprivation in primary care : development, reliability and validity of e self-reported questionnaire: the DiPCare-Q. *BMJ Open* 2012; 2: e000692.
102. Wardle J, McCaffery K, Nadel M, Atkin W (2004). Socioeconomic in cancer screening participation: comparing cognitive and psychosocial explanations. *Social Science and Medicine* 59:249-261.
103. Whitehead M, G Dahlgren. Concepts and principles for tackling social inequities in health: Levelling up. Parts 1 and 2. OMS, 2006.
104. Wilkinson R, Marmot M, eds. *Les déterminants sociaux de la santé : les faits*. Copenhagen, OMS Europe (2<sup>ème</sup> édition), 2004.
105. Zill N. Parental schooling and children health. *Public Health Rep* 1996; 111: 34-43.

## ANNEXE :

Note sur l'enregistrement d'information socio-économiques par les registres des cancers

P Grosclaude, FRANCIM

### 1. Sources d'information et contraintes d'utilisation

En France le recensement des cas de cancers doit s'appuyer sur un grand nombre de sources d'informations : laboratoires d'anatomo-cyto-pathologie, cliniciens des hôpitaux publics ou privés et spécialistes libéraux, médecins généralistes, sources médico-administratives : caisses d'assurance maladie et PMSI. Chacune de ces sources ne fournit que des informations parcellaires.

Le recensement des cas de cancers en France est aussi conditionné par l'organisation du système de soins et la réglementation en matière de protection des données individuelles.

Un système de soins associant secteur public et secteur privé dans lequel le malade a la liberté de choisir son (ou ses) médecin(s) et où dans le cas d'une maladie comme le cancer, il peut fréquenter au cours de son traitement plusieurs structures de soins, favorise la dispersion des informations et rend le recensement plus complexe qu'un système centralisé autour de centres spécialisés.

Contrairement à ce qui se passe dans d'autres pays, notamment d'Europe du Nord, la réglementation française privilégie la protection des individus de la divulgation d'informations personnelles par rapport à la facilitation de la mise en place de système de surveillance en matière de santé publique. Par exemple :

- l'utilisation d'un identifiant unique qui permettrait le chaînage les informations provenant de différentes sources et donc le recensement rencontre une très forte opposition ;
- l'anonymat de la partie médicale des certificats de décès, qui est une autre particularité du système français, conduit probablement (aux dires de l'OMS) à un enregistrement moins exhaustif, mais à coup sûr, interdit l'évaluation de l'exhaustivité car toutes les méthodes de mesure validées reposent sur l'utilisation des certificats de décès ;
- enfin, la transmission des informations par les médecins n'est pas obligatoire mais seulement au mieux recommandée, voire le plus souvent autorisée, sous réserve qu'ils en informent le patient. Dans un contexte où l'implication dans la santé publique n'est pas valorisée par la culture médicale, faire reposer sur les médecins traitants toutes les responsabilités en matière d'information des patients est autant d'obstacles (réels ou prétextés) à la transmission des informations.

## 2. Principes de l'enregistrement

Lors de la création des registres, l'objectif de l'enregistrement était de réaliser une surveillance pour comprendre et mieux contrôler l'apparition des cancers en mesurant et comparant les incidences au cours du temps ou en différents lieux. Les aspects plus cliniques, traitement, guérison, devenir des patients, n'étaient pas les préoccupations initiales.

Les registres ont pour tâche de recenser tous les cas de cancer diagnostiqués dans la population d'une zone administrative donnée afin d'obtenir un enregistrement exhaustif et sans double compte permettant de décrire sans biais l'incidence des cancers. Le travail d'un registre est le résultat de deux activités complémentaires, d'une part la recherche et la validation et le chaînage des informations médicales, démographiques et administratives, et d'autre part, la mise en forme de cette information pour la constitution d'une base de données codées directement analysable afin de produire des données statistiques (indicateurs).

Le recensement des cas de cancers incidents, consiste donc en une compilation des informations concernant les cas, qui est généralement réalisée par un réseau d'enquêteurs auprès de nombreuses sources d'informations médicales (cf. supra).

L'enregistrement proprement dit passe ensuite par un recoupement et une validation de l'ensemble des informations collectées et par un codage répondant à des règles standardisées à un niveau national et international. Ces règles permettent de trier les pathologies qui relèvent de l'enregistrement puis de les classer, de définir des paramètres importants (date d'incidence, notion de récurrence ou de second cancer, niveaux de gravité...), elles obligent à respecter des formats d'enregistrement pour différentes informations notamment en utilisant des classifications internationales CIM-10, TNM, ...). Cette seconde phase rend l'information utilisable, elle va aussi la figer. Les données disponibles pour les analyses et les résultats qui en découlent sont donc très dépendants des règles d'enregistrement et de codage choisies. Celles-ci doivent à la fois être les plus standardisées possibles tout en s'adaptant aux problèmes étudiés. La standardisation est indispensable pour la réalisation d'études collaboratives entre différents registres mais elle l'est aussi pour suivre l'évolution des phénomènes étudiés dans le temps.

Cette double contrainte, de standardisation et d'adaptation, est évidemment contradictoire. La nécessaire évolution des classifications et du mode de recensement permettant de répondre à des questions nouvelles ne doit pas entraîner une rupture dans les logiques d'enregistrement sous peine d'empêcher l'étude de l'évolution de l'incidence ou d'autres phénomènes. Donc lorsque les registres rassemblent des données pour qu'elles soient immédiatement analysables ce processus va obligatoirement de pair avec un appauvrissement du contenu informatif et une nécessaire orientation des données vers un objectif qui aura été préalablement défini.

Mais les registres de population ne sont pas que des bases de données immédiatement exploitables. Ce sont aussi des bases de sondage dans lesquelles on peut tirer des échantillons représentatifs des malades dans la population. Toutefois pour que ces études sur échantillons soient possibles, il faut, d'une part disposer d'un identifiant permettant de revenir vers les sources d'informations, d'autre part que les informations restent accessibles au niveau des sources.

On peut donc concevoir les registres comme des systèmes permettant de collecter en routine et de façon exhaustive un nombre plus ou moins limité de données permettant de décrire les cas de cancers incidents sous la forme d'indicateurs prédéfinis, et aussi comme des structures permettant de réaliser des études plus détaillées en retournant aux sources d'information pour répondre à des questions plus spécifiques concernant des problèmes émergents.

Le choix entre un enregistrement systématique ou une étude ad hoc voire d'une solution intermédiaire dépend de la balance entre les avantages et les inconvénients de ces solutions.

Une base coûtera d'autant plus cher que les données qu'elle contient seront nombreuses et détaillées mais ses exploitations seront rapides et peu coûteuses en revanche elles seront stéréotypées et contraintes par les données. Il sera probablement difficile de l'utiliser directement pour toute question non anticipée, soit parce que les données nécessaires n'auront pas été collectées, soit parce qu'elles auront été codées sous une forme inadaptée (regroupement intempestif, manque de précision, ...).

A l'opposé une étude ad hoc permettra de répondre de façon plus précise à d'avantage de questions, notamment à des questions nouvelles, éventuellement d'éviter le piège des évolutions nosologiques, mais la réponse sera plus lente et le coût de la réponse à chaque question plus élevé. Le coût de l'archivage des informations doit aussi être pris en compte car, comme il est dit plus haut, de la qualité de l'archivage dépend la faisabilité de ces études. Il est possible comme c'est le cas pour les informations contenues dans les dossiers médicaux que cet archivage soit déjà obligatoire.

### **3. Organisation et évolution de la collecte des informations**

Il faut distinguer deux temps dans l'enregistrement classique : une phase de repérage des cas et une phase de collecte des informations détaillées (schéma 1).

On utilise d'abord les sources dites « de signalement » (généralement les sources médico-administratives) qui permettent de repérer des cas sans pour autant disposer d'informations suffisamment complètes et fiables pour pouvoir valider cet enregistrement. Dans un deuxième temps, on s'appuie sur les sources d'information proprement dites (généralement les sources médicales ou purement administratives) qui fournissent les informations nécessaires à une description détaillée du patient et de sa pathologie.

Le croisement des 3 sources médico-administratives que sont les ALD, le PMSI et les données issues des dossiers communicant de cancérologie (DCC). Les laboratoires d'anatomie-pathologiques, dans lesquels les registres vont systématiquement, constituent traditionnellement une autre source de signalement. Il s'agissait même de la principale source de signalement jusqu'au début des années 2000, car les sources médico-administratives (ALD et PMSI) ne sont accessibles que depuis la fin des années 90. Le croisement de ces 4 sources donne une base de données très exhaustive mais manquant de spécificité et contenant des informations minimales. Les certificats de décès devraient faire aussi partie de ces sources pour permettre un repérage quasi exhaustif mais ils ne peuvent pas être utilisés en France.

Ces sources de signalement permettent l'identification des cas ainsi que le repérage des sources d'information médicales potentielles pour chacun de ces cas. En effet, pour préciser ou confirmer la nature exacte du cancer (localisation et morphologie), éviter les doubles comptes, déterminer la date du diagnostic et le stade d'extension et bien sûr pour collecter des informations complémentaires sur l'histoire de la maladie, (circonstances de diagnostic, traitements, évolution...) il est nécessaire de compléter les informations collectées auprès des sources médicales (cliniques et aussi de façon complémentaire anatomopathologiques)

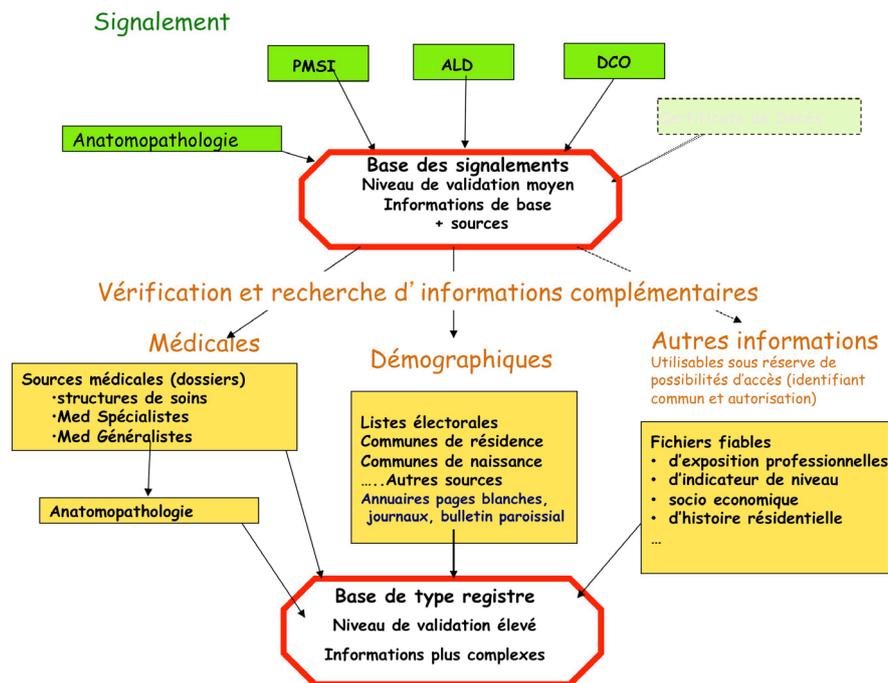


Schéma 1 : Structure générale de la collecte des informations

Les sources démographiques (souvent administratives) permettent pour leur part de consolider les informations démographiques (identification, adresse exacte, lieu de naissance, survie...) qui ne sont pas toujours présentes ou pas fiables dans les dossiers médicaux.

Il serait possible de collecter d'autres types d'informations (le plus souvent relatives aux expositions ou facteurs de risque) ne figurant pas dans les dossiers médicaux.

Les données socio-économiques figurent exceptionnellement dans les dossiers médicaux. Leur collecte nécessite donc un difficile travail d'enquête pour collecter les données, dans un cadre très spécifique. Le géocodage suivi d'utilisation d'indicateurs de

déprivation et l'appariement avec des bases de données administratives sont donc les seules options possibles pour un travail au long cours.

Pour certaines il serait envisageable de les collecter en utilisant des fichiers contenant des informations pertinentes (par exemple fichier des caisses de retraite pour les informations socio-professionnelles, fichiers des impôts pour niveau de revenus,... ). Il faut toutefois qu'il existe des identifiants communs permettant de croiser ces fichiers et que ce croisement soit autorisé. Actuellement les registres n'ont pas le droit d'enregistrer le NIR et chaque croisement avec des données nominatives exige une demande d'autorisation spécifique auprès de la CNIL.

#### **4. Adresse et géolocalisation pour utiliser un indicateur écologique de défavorisation**

Les registres collectent maintenant systématiquement l'adresse des patients au moment de leur diagnostic. La géolocalisation exacte du lieu de résidence habituel des patients est en effet un préalable indispensable pour réaliser des études sur les expositions environnementales ou sur l'effet de la distance à l'offre de soins ; mais avec le développement des indicateurs écologiques de défavorisation elle permet aussi de disposer d'indicateurs de mesure des inégalités sociales.

L'affectation d'un niveau de défavorisation à un patient repose sur un processus qui commence par la collecte puis l'enregistrement de l'adresse exacte du patient. Cette adresse est ensuite transformée en coordonnées géographiques (géocodage proprement dit) et ces coordonnées géographiques font ensuite l'objet d'un traitement qui permet de les affecter à un Iris (répondant aux définitions données jusqu'en 2006 par l'Insee et depuis par l'IGN) Enfin la connaissance de l'Iris permet d'affecter un index de défavorisation au patient résidant à cette adresse. A l'instar de ce qui se fait pour l'enregistrement des tumeurs, l'ensemble de ce processus doit suivre des règles strictes et bien standardisées pour s'assurer de la comparabilité des données.