



**ÉTUDE EXPLORATOIRE SUR LES
INEGALITES TERRITORIALES,
ENVIRONNEMENTALES ET SOCIALES
DE SANTE**

- RHONE-ALPES -

AVRIL 2016



Cette étude a été réalisée par :

Lucie ANZIVINO, ORS Rhône-Alpes

Martine DRENEAU, ORS Rhône-Alpes

Géraldine GUILLAUD, Air Rhône-Alpes

Sabrina PONTET, Air Rhône-Alpes

Abdoul SONKO, ORS Rhône-Alpes

Les auteurs remercient toutes les personnes ayant contribué à enrichir le contenu de ce rapport, producteurs d'études ou de bases de données, ainsi que les membres du comité de pilotage de l'étude.

Comité de pilotage de l'étude :

Marie-Noëlle ROUX-LEFEBVRE, Région Auvergne-Rhône-Alpes

Bruno FABRES, Alain BLINEAU, ARS Auvergne-Rhône-Alpes

Dominique BAURES, DREAL Auvergne-Rhône-Alpes

Xavier OLNLY, CEREMA

Lydie NEMAUSAT, FRAPNA Rhône-Alpes

Viviane BARET-ROUANET, Réseau Environnement Santé

Emmanuelle AMAR, REMERA

Avec le soutien financier de la Région Rhône-Alpes

Sommaire

Sommaire	5
1. Introduction – Contexte – Méthodologie	7
2. Description - Recensement des données disponibles	9
2.1 Bases de données	9
2.1.1 Bases de données en santé.....	10
2.1.2 Bases de données en social	20
2.1.3 Bases de données en environnement	23
2.2 Méthodes de croisement de données.....	30
2.3 Indices synthétiques	34
3. Discussion.....	41
3.1 Accessibilité des données	41
3.2 Niveau géographique des données	46
3.3 Avantages et limites des bases	49
3.4 Discussion sur les méthodes.....	54
3.5 Discussion sur les indices composites.....	55
4. Perspectives et conclusions.....	58
Références bibliographiques.....	68

Le présent rapport est complété de trois annexes (176 pages) :

Annexe 1 : Fiches descriptives des bases de données

Annexe 2 : Présentation des études utilisant des méthodes de croisement de données

Annexe 3 : Présentation des indices synthétiques

1. Introduction – Contexte – Méthodologie

Contexte

Courant 2013, à l'initiative de la Région Rhône-Alpes et de l'ARS Rhône-Alpes, un réseau d'acteurs à vocation régionale a été mis en place, défendant une approche globale de la santé et contribuant à la réduction des inégalités de santé. Ce réseau a pris le nom de Métaréseau.

Les acteurs s'inscrivent dans une démarche participative de co-construction. Ils ont le souci de la cohérence, de l'optimisation, et favorisent entre eux les synergies, les complémentarités, la subsidiarité.

Le rapprochement de ces acteurs de la santé et de l'environnement favorise la prise en compte des enjeux de la santé environnementale. Cette approche globale et intégrée de la santé permet de replacer l'homme au cœur de la protection de l'environnement.

C'est dans ce cadre qu'Air Rhône-Alpes et l'ORS Rhône-Alpes ont mené une étude exploratoire sur la caractérisation des inégalités territoriales, environnementales et sociales de santé (ITESS) : après avoir auditionné des experts et des producteurs d'études sur cette thématique, et afin de passer à une phase plus opérationnelle, l'ORS Rhône-Alpes et Air Rhône-Alpes ont travaillé sur un inventaire des données mobilisables en Rhône-Alpes (santé et environnement) ainsi que sur les méthodes actuellement disponibles pour croiser ou rapprocher ces données.

Ce travail s'inscrit par ailleurs dans les objectifs du PNSE et du PRSE¹ de connaissance des inégalités territoriales environnementales, sanitaires et sociales ; le lien sera ainsi fait avec ces programmes en santé/environnement.

Objectifs

En matière de données, un grand nombre de bases existent, tant dans le domaine de l'environnement que dans celui de la santé, données pour la plupart non encore disponibles en ligne.

Cette étude doit permettre de recenser et de sélectionner les données pertinentes accessibles dans la région Rhône-Alpes (soit en ligne, soit sur demande auprès des organismes producteurs de données), en précisant le niveau géographique d'accessibilité des données.

En termes de méthodes de croisement de données, plusieurs travaux sont actuellement en cours, menés tant au niveau national par des équipes de recherche (ex : Equit'Area (équipe EHESP/IRSET), plateforme PLAINE de l'INERIS, etc.) qu'au niveau régional (ex : en Lorraine et en PACA). Ces différentes méthodes combinent de nombreux indicateurs et autorisent la construction par exemple d'indicateurs territoriaux de défaveur environnementale, comme de défaveur sociale ou sanitaire.

¹ Plan National ou Régional Santé Environnement - PNSE3 : [http://www.developpement-durable.gouv.fr/IMG/PNSE3\(1\).pdf](http://www.developpement-durable.gouv.fr/IMG/PNSE3(1).pdf)

Ces travaux ont le plus souvent pour objectifs d'identifier les points noirs sur le territoire, dans l'objectif de lutter contre les inégalités territoriales environnementales et sociales de santé.

Des perspectives permettront d'émettre des préconisations quant aux contours d'une première étude de caractérisation des ITESS en Auvergne Rhône-Alpes à partir de bases de données et indicateurs rapidement mobilisables.

Et l'Auvergne ?

Bien que la commande de cette étude et sa réalisation se soient déroulées avant la création des nouvelles grandes régions, l'extension de ces résultats à la région Auvergne-Rhône-Alpes ne représentera pas de difficultés majeures. En effet, si les bases de données disponibles peuvent différer d'un territoire à l'autre, le fait qu'il s'agisse pour la plupart de bases administratives limite les différences entre territoires au sein de la France métropolitaine. Ces bases administratives ont une collecte de données qui s'opère à un niveau local mais de façon uniforme sur l'ensemble du territoire français, ainsi qu'une gestion (corrections et redressages, validation, mise à disposition...) qui s'effectue à un niveau national. Lorsque les modes de recueil ou de calcul des données sont susceptibles de différer d'un territoire à l'autre, ces limites seront mentionnées. Quelques bases de données n'existent toutefois que sur un territoire, telles que les Registres (des cancers ou des malformations).

2. Description - Recensement des données disponibles

La première étape de cette étude exploratoire a consisté à recenser et décrire :

- les bases de données disponibles dans les champs de la santé, du social et de l'environnement
- les méthodes de croisement (ou rapprochement) de ces données déjà expérimentées dans quelques études nationales ou locales identifiées
- les indices synthétiques construits à partir des données brutes afin de rendre compte d'une défaveur globale d'un territoire, qu'elle soit d'ordre social, sanitaire, ou environnemental.

2.1 Bases de données

Le recensement et la description des bases de données, dans le cadre de cette étude exploratoire, ont pour objectif de faire un état des lieux des données pouvant être mobilisées (de façon immédiate ou à court terme) sur la région Rhône-Alpes, afin de contribuer à l'identification des inégalités territoriales environnementales, sanitaires et sociales (ITESS) au sein de la région.

La notion de disponibilité à « court terme » des données est primordiale en termes d'opérationnalité, cette étude s'inscrivant dans une phase préparatoire du PRSE3 et visant une utilisation dans un proche avenir des données décrites. Quelques bases de données, auxquelles l'accès est actuellement problématique, seront néanmoins décrites car des avancées en termes d'accessibilité sont envisageables.

Par ailleurs, certaines bases de données repérées n'ont pas été décrites en raison soit de leur faible intérêt pour l'utilisation envisagée, soit de leur absence de déclinaison possible à un niveau infrarégional. Ces bases seront mentionnées pour mémoire.

Chaque base de données recensée a fait l'objet d'une description détaillée, comprenant plusieurs parties.

La 1^{ère} partie décrit **la base source** :

- Intitulé
- Organisme(s) producteur(s) : organisme(s) qui collectent et mettent en forme les données
- Mode de recueil des données : système de collecte, collecte exhaustive ou non, traitements et calculs éventuels...
- Description des données : informations sur le contenu de la base
- Unité : élément de base
- Niveau géographique : le plus fin de la base
- Périodicité : rythme des mises à jour
- Profondeur historique : année la plus ancienne dans la base
- Thématique(s) : thématiques principales sur lesquelles portent les données (les champs de la santé, du social, de l'environnement, sont détaillés).

Les parties suivantes décrivent l'accessibilité à la base de données, tout d'abord **l'accès aux données granulaires**, celles au niveau le plus fin disponible compte-tenu du système de collecte, puis **l'accès aux données agrégées**, issues du regroupement de données granulaires selon une ou plusieurs caractéristiques communes.

Pour chaque niveau de données sont précisés les critères d'ouverture suivants :

- la liberté d'accès (accès à tous, accès restreint, accès fermé)
- les commentaires sur cette liberté d'accès
- le coût d'accès (gratuit, payant)
- le format de mise à disposition des données (exploitable ou non, format des fichiers)
- les conditions de réutilisation (explicitées ou non, avec ou sans restrictions)
- l'existence ou non d'un portail de mise à disposition des données sur internet
- les coordonnées d'un contact (service ou personne) lorsque celles-ci ont été identifiées.

Ce mode de description est directement inspiré de celui utilisé par la Mission Etalab en 2014 dans sa Cartographie des bases de données en santé publique².

L'ensemble des fiches produites (une par base de données) figure en annexe 1 de ce rapport. Les bases décrites dans chaque champ (santé, social, environnement) seront présentées ci-dessous, avec quelques commentaires généraux nécessaires à la compréhension et des exemples d'indicateurs directement utilisables dans un travail d'identification des inégalités territoriales.

2.1.1 Bases de données en santé

La plupart des bases de données en santé sont des bases administratives, non destinées a priori à fournir des informations sur l'état de santé de la population, mais à fournir aux différentes administrations les informations nécessaires à la réalisation de leurs missions, comme le remboursement à des patients des frais liés aux soins ou encore le paiement à des hôpitaux des actes de soins réalisés.

Cela explique en partie les limites à l'utilisation de certaines bases en observation de la santé, la qualité ou l'exhaustivité des données recueillies pouvant être insuffisantes pour une analyse différente des analyses prévues au départ. Par ailleurs, certaines bases sont peu accessibles du fait de la nature des données qu'elles contiennent, données médicales des patients soumises au secret médical. Les missions spécifiques à chaque administration font que certains systèmes d'information sont construits par et pour une administration, sans concordance avec les systèmes d'autres administrations, ce qui rend difficile, voire impossible, le rapprochement des données issues des différents systèmes d'information. Ces aspects seront abordés dans le chapitre « Discussion » de ce rapport.

² Mission Etalab. Cartographie des bases de données en santé publique. Mai 2014. Disponible : <https://www.data.gouv.fr/fr/datasets/cartographie-des-bases-de-donnees-publiques-en-sante/>

Les bases de données en santé peuvent être regroupées en plusieurs catégories, selon qu'elles décrivent la mortalité, la morbidité, le recours aux soins, les comportements de santé, les conditions de travail, ou encore l'offre de soins.

Les recueils d'information d'ordre démographique, indispensables au calcul de nombreux indicateurs et produisant eux-mêmes de nombreuses données intéressantes la santé ou le social, seront présentés soit dans la partie « santé » (comme l'état civil), soit dans la partie « social » (comme le recensement de la population).

Les données de mortalité

Les deux principaux systèmes d'information sur la mortalité de la population sont celui de **l'état civil** (*annexe 1, fiche 1*) et celui de l'INSERM (CépiDc) sur **les causes médicales de décès** (*annexe 1, fiche 2*). Ces deux systèmes sont quasi-exhaustifs sur tous les décès survenus en France, ils permettent de traiter l'information soit au niveau du territoire de survenue des décès (décès enregistrés) soit au niveau du territoire de domicile de la personne (décès domiciliés).

Ces deux systèmes sont en partie reliés par l'INSERM, du moins pour les parties anonymes du certificat médical de décès (partie établie par le médecin) et du bulletin de décès de l'état civil (bulletin 7, établi par l'état civil). L'information recueillie par les deux systèmes d'informations permet un traitement des décès au niveau communal, voire au niveau de l'arrondissement pour les grandes villes ; l'adresse de la personne décédée n'est pas récoltée, ce qui interdit tout traitement infra-communal.

Bien que l'état civil apporte des informations sociodémographiques sur la personne décédée (activité professionnelle, nationalité, situation matrimoniale au moment du décès), ces informations sont peu exploitables car la catégorie socioprofessionnelle n'est indiquée que pour les personnes actives (donc pas pour les retraités).

Quant à la statistique des causes médicales de décès, elle apporte des informations sur les maladies ou affections morbides ayant entraîné ou contribué au décès. Comme toutes les bases de données en santé comportant des pathologies codées en CIM 10 (Classification internationale des maladies, 10^{ème} version), la base de l'INSERM permet une exploitation par pathologie, pour toutes les pathologies, soit par grand chapitre de la CIM 10 (ex : ensemble de maladies de l'appareil respiratoire), soit par sous-chapitre plus ou moins détaillé (ex : asthme).

L'état civil enregistre également toutes les naissances survenues en France, naissances d'enfants vivants et d'enfants mort-nés (les conditions d'établissement du bulletin « d'enfant sans vie » ont changé plusieurs fois, entraînant des ruptures de série en 2002, 2008 et 2009, dans les chiffres de mortalité néo-natale en France). L'information délivrée par les bulletins de naissance (filiation, fratrie, activités des parents, adresse de domicile des parents...) est relativement riche mais peu accessible, si ce n'est via les registres d'état civil tenus dans chaque mairie (pour les actes de moins de 75 ans : accès sur dérogation et après accord écrit du procureur de la République).

Quelques exemples d'indicateurs calculés à partir de ces bases :

- Taux de natalité
- Indice conjoncturel de fécondité
- Age moyen des mères à l'accouchement
- Taux de mortalité infantile
- Taux bruts et standardisés de mortalité (par sexe, par groupe d'âges)
- Taux bruts et standardisés de mortalité prématurée (avant 65 ans)
- Espérance de vie (à la naissance, à 60 ans...)
- Taux bruts et standardisés de mortalité par cause médicale de décès (pour toutes les pathologies, par sexe et par groupe d'âges) – des regroupements de pathologies sont possibles, comme l'ensemble des pathologies en lien avec l'alcool par exemple
- Indices comparatifs de mortalité (par cause médicale de décès)

Les données de morbidité / recours aux soins

Les bases les plus exhaustives en matière de morbidité sont les bases de l'assurance maladie et celles du système hospitalier, qui ont pour vocation le remboursement des soins. Ces bases sont par définition généralistes, elles concernent toutes les personnes, quelle que soit leur pathologie. Elles sont susceptibles de renseigner sur l'état de santé de l'ensemble de la population, mais sont toutefois limitées aux personnes ayant été hospitalisées ou s'étant fait rembourser des soins ou des médicaments. L'incidence des pathologies ainsi calculée ne tient compte ni du non-recours au système de soins (personnes malades mais non prises en charge), ni des pratiques d'automédication par exemple.

Les soins et les prescriptions de médicaments en médecine de ville font l'objet de remboursements par l'assurance maladie. **Les bases de l'assurance maladie** sont propres à chaque régime d'assurance maladie, les trois principaux régimes étant le régime général (travailleurs salariés), le régime agricole (salariés et non-salariés agricoles), le régime des indépendants (artisans, commerçants, professionnels libéraux). La base **ERASME** (*annexe1, fiche 8*) enregistre ainsi tous les remboursements faits aux personnes affiliées au régime général, ainsi que des renseignements sur les professionnels prescripteurs des soins. Il est donc possible de connaître, sur un territoire donné, la fréquence de recours des assurés aux professionnels de santé (le médecin généraliste ou le dentiste par exemple) ou encore la consommation annuelle d'un certain type de médicaments (les antiasthmatiques par exemple). Cette base, du fait de sa connaissance du statut de chaque assuré vis-à-vis de CMUc (couverture maladie universelle), permet d'analyser spécifiquement le recours aux soins et la consommation médicamenteuse des personnes en situation de précarité. Au niveau local (caisse primaire) ou régional (CTI), quelques personnes habilitées peuvent interroger cette base de façon fine, à l'adresse de domicile de l'assuré, ce qui permet une analyse infra-communale des données (c'est le cas pour la région Rhône-Alpes, via l'ARS, mais pas le cas actuellement pour l'Auvergne).

Le service médical de l'assurance maladie gère également **la base des admissions en ALD (Affection de longue durée)**, qui correspond à l'exonération du ticket modérateur pour une pathologie qui nécessite des soins longs et coûteux (*annexe1, fiches 3, 4 et 5 pour les trois principaux régimes*). Ces bases de données renseignent sur la pathologie à l'origine de l'ALD, les maladies étant codées en CIM 10. A partir de ces données, peuvent être calculés des taux d'incidence et de prévalence d'une trentaine de pathologies, mais ces données ne reflètent pas tout à fait l'incidence et la prévalence réelles des pathologies dans une population. En effet, l'admission en ALD n'est pas automatique, elle nécessite une demande de la part de l'assuré et de son médecin, ainsi que l'accord par un médecin conseil de l'assurance maladie. Des différences de pratiques selon les acteurs peuvent être observées.

Les bases de données des différents régimes sont réunies au niveau national dans un gigantesque entrepôt de données : **le SNIIRAM** (*annexe1, fiche 9.*). Bien que les données du SNIIRAM soient anonymisées, l'accès aux données est très réglementé, en raison notamment des risques d'indentification indirecte, avec des listes d'organismes et de personnes habilités à interroger la base (avec des accès différents selon les autorisations).

Les systèmes d'information hospitaliers produisent des bases de données sur tous les séjours hospitaliers, pour les établissements de court séjour (**PMSI MCO** : médecine, chirurgie, obstétrique) (*annexe1, fiche 6.*), les établissements de soins de suite et de réadaptation (SSR), l'hospitalisation à domicile (HAD), ou encore les établissements spécialisés en psychiatrie (**RIM-P**) (*annexe1, fiche 7*). Ces données sont utilisées pour le financement des établissements de santé (tarification à l'activité) et pour l'organisation de l'offre de soins (planification), mais elles comportent également des informations médicales sur les personnes hospitalisées, sur les diagnostics médicaux et les actes réalisés, ainsi que sur le patient lui-même (âge, sexe, adresse). Tous les diagnostics sont codés en CIM 10, il est donc possible d'interroger ces bases de données par pathologie ou par grands groupes de pathologies. Les données sont disponibles soit par établissement (données enregistrées) soit par code postal de résidence du patient (données domiciliées). Les données de chaque établissement hospitalier sont en effet anonymisées avant transmission aux ARS, et seul le code postal du patient est disponible (en fait un code propre au PMSI, propre du code postal).

L'IDS (Institut des données de santé) a mis en ligne début 2016 des fiches de synthèse sur les données hospitalières en accès libre : <http://www.institut-des-donnees-de-sante.fr/connaissance/>. Ce document recense l'ensemble des applications disponibles sur internet en accès libre, diffusant des données hospitalières.

Quelques exemples d'indicateurs calculés à partir des bases de l'assurance maladie et des établissements hospitaliers :

- Taux et indices comparatifs d'incidence (ou de prévalence) pour chacune des 30 affections de la liste des ALD, et pour des groupes de pathologies (maladies cardio-vasculaires, appareil respiratoire, appareil digestif, maladies infectieuses, neurologiques, rhumatismales)

- Taux comparatifs des assurés et ayants-droit du régime général ayant eu dans l'année au moins trois remboursements pour certains regroupements de médicaments (antidiabétiques, psychotropes, neuroleptiques, antidépresseurs, antihypertenseurs, antiallergiques, antiasthmatiques, anti-Alzheimer...)
- Taux comparatifs des assurés et ayants-droit du régime général ayant consulté au moins une fois dans l'année certains professionnels de santé libéraux (médecin généraliste, dentiste, orthodontiste, pédiatre, gynécologue, ophtalmologue, sage-femme, infirmier, orthophoniste, psychiatre)
- Effectifs de patients hospitalisés dans l'année, taux standardisés et indices comparatifs, selon les causes principales de la CIM 10, et selon certains codes spécifiques ; diabète, infarctus du myocarde, alcool, Alzheimer (65 ans et +) ...

D'autres systèmes d'information en santé, moins généralistes, ont été mis en place avec des missions particulières :

➤ **Le dispositif des maladies à déclaration obligatoire – MDO** (*annexe 1, fiche 10*)

Ce dispositif repose sur une obligation légale pour tout professionnel de santé (médecin ou biologiste), quel que soit son mode d'exercice (public ou privé), de déclarer aux autorités sanitaires, les cas de maladie qu'il diagnostique dès lors que la maladie est inscrite sur la liste des maladies à déclaration obligatoire. En 2015, 31 maladies sont à déclaration obligatoire. Parmi elles, 29 sont des maladies infectieuses et 2 sont non-infectieuses (mésothéliomes, et saturnisme chez les enfants mineurs).

➤ **Le réseau OSCOUR** (*annexe 1, fiche 11*)

Depuis juillet 2013, l'ensemble des établissements autorisés en médecine d'urgence doivent transmettre des RPU (résumés de passages aux urgences) à l'InVS, comportant des informations sur le motif d'admission, la gravité, le diagnostic (CIM 10), les actes réalisés. L'ensemble des motifs d'admission n'est pas systématiquement exploité par l'InVS ou la CIRE, mais certaines pathologies sont particulièrement analysées : pathologies hivernales (syndromes grippaux, bronchiolites, gastro-entérites), pathologies liées à la chaleur, épidémies (rougeole, méningite virale, etc.), asthme, traumatismes, etc.

➤ **Le dispositif SURSAUD** (*annexe 1, fiche 12*)

Ce système de surveillance dit syndromique est basé sur la collecte quotidienne de données sanitaires non-spécifiques permettant un suivi épidémiologique proche du temps réel. Le système utilise les remontées de plusieurs sources de données : données OSCOUR d'urgences hospitalières, données SOS Médecins, données de mortalité état-civil, données de certification électronique des décès INSERM-CépiDc. Les données agrégées par l'InVS doivent permettre de repérer une variation inhabituelle d'indicateurs de santé. Exemple d'analyses possibles : impact des températures élevées sur les passages aux urgences et sur les consultations SOS médecins.

➤ **Les Registres : le registre REMERA des malformations** (*annexe 1, fiche 13*) et le **Registre des cancers de l'Isère** (*annexe 1, fiche 14*)

Les registres constituent des systèmes de surveillance à visée épidémiologique sur un territoire donné et sur une ou des pathologies spécifiques.

REMERA surveille l'ensemble des naissances domiciliées dans les départements du Rhône, de la Loire, de l'Isère et de l'Ain (et Savoie jusqu'en 2012), soit environ 60 000 naissances annuelles. Le registre recense l'ensemble des malformations parmi les naissances (vivantes ou non, de plus de 20 semaines de grossesse), les morts fœtales *in utero* et les IMG (interruption médicale de grossesse). Chaque cas est documenté, notamment sur les antécédents et expositions, médicamenteuses ou autres, de la mère. Le Registre des cancers de l'Isère enregistre tous les nouveaux cas de cancers du département de l'Isère de façon continue et exhaustive.

L'accès aux données des Registres pour des activités de recherche est soumis à une autorisation de la CNIL. Des données agrégées et anonymisées sont diffusées.

- Les examens de santé des enfants : **les certificats de santé de la PMI** (*annexe 1, fiche 15*) et **les bilans de la santé scolaire** (*annexe 1, fiche 16*)

La délivrance de certificats de santé est obligatoire pour trois examens médicaux des enfants : **avant le huitième jour (CS8), au neuvième mois (CS9) et au vingt-quatrième mois (CS24)**. Ces certificats sont remplis par les médecins (hospitaliers, de ville ou de PMI) et transmis aux services PMI des Conseils départementaux. Les taux de retour des certificats sont variables selon les départements et les CS (CS8 : 83% en moyenne nationale en 2013, CS9 : 39%, CS24 : 30%), ce qui est à prendre en compte dans l'analyse. Les CS apportent des renseignements précieux sur la surveillance de la grossesse, le déroulement de l'accouchement, la santé de l'enfant, son développement psychomoteur, les vaccinations. Ces informations peuvent être croisées avec des données sociodémographiques sur les parents (âge de la mère, profession et activité professionnelle des parents). L'exploitation des données des CS par les Conseils départementaux est assez inégale et dépend des moyens mis en œuvre par chaque Conseil départemental. Une remontée de certaines données à la DREES permet une exploitation homogène des CS au niveau départemental. Les données infra-départementales ne sont disponibles que dans chaque Conseil départemental.

Les données des services de santé scolaire pourraient en théorie permettre une analyse assez fine (au niveau de l'établissement scolaire) des pathologies observées chez les enfants de l'école primaire au lycée, notamment sur les questions d'obésité, de troubles du langage et des apprentissages, de vision, d'audition, de pathologies respiratoires... Mais plusieurs éléments limitent leur exploitation de façon systématique sur l'ensemble du territoire : le manque d'exhaustivité des bilans de santé (tous les enfants n'ont pas un bilan identique), un recueil non homogène des données d'un service à un autre, un accès difficile aux données de l'Education nationale ou à celles des communes assurant la délégation de service pour le suivi des enfants de primaire.

Certaines bases concernant la morbidité n'ont pas été décrites du fait de leur mode de collecte, qui ne permet pas de descendre à un niveau géographique fin. C'est le cas par exemple du Réseau sentinelles de l'INSERM, qui suit l'évolution de certaines pathologies au sein de la population (comme les syndromes grippaux, la varicelle, les diarrhées aiguës...) à partir d'un réseau de médecins volontaires.

Les données sur les comportements de santé

Les informations collectées sur les comportements pouvant avoir un impact sur la santé (habitudes alimentaires, activité physique, consommation d'alcool et de tabac, sexualité, vaccinations...) ne portent généralement pas sur l'ensemble de la population et sont issues d'enquêtes menées soit ponctuellement auprès d'un échantillon représentatif de la population, soit de façon régulière auprès de groupes de personnes suivies dans le cadre de cohortes par exemple. Les résultats de ces enquêtes ne sont pas extrapolables à des niveaux géographiques fins et n'ont donc pas fait l'objet d'un descriptif dans le cadre de ce travail. Une liste de ces enquêtes nationales ou études de cohortes, pouvant être utiles au développement de la recherche et de l'expertise en santé publique est publiée sur le **Portail Epidémiologie France** : <https://epidemiologie-france.aviesan.fr> , réalisé sous la direction d'AVIESAN (Alliance nationale pour les sciences de la vie et de la santé).

Parmi les grandes enquêtes en population nationale, citons néanmoins le **Baromètre Santé de l'INPES**, réalisé de façon répétée auprès d'un échantillon important de la population française, et comportant de nombreuses questions sur les comportements de santé, sur les conditions de vie et la santé perçue. Les personnes interrogées fournissent également des informations sur leur situation socio-économique, qu'il est possible de croiser avec des données de santé. La partie rhônalpine de l'échantillon national a été exploitée par l'ORS Rhône-Alpes en 2000, 2005, 2010, mais les résultats ne peuvent être analysés à un niveau infrarégional, et la faiblesse de taille des échantillons ne permet pas une analyse croisée des données santé / sociales.

L'enquête **ESCAPAD**, réalisée par l'OFDT auprès des jeunes de 17 ans, apporte également des données sur l'évolution des consommations par les jeunes de substances psychoactives (alcool, tabac, cannabis et autres drogues illicites), mais les effectifs des échantillons n'autorisent pas un traitement des données à un niveau infrarégional (quelques données départementales sont parfois disponibles, mais pas sur l'ensemble de la région).

Les données des **Centres d'examens de santé (CES)** de l'assurance maladie, centralisées et traitées par le CETAF, comportent de nombreuses informations sur l'état de santé (recours aux soins, consommation de médicaments, mesure poids/taille, tension artérielle...) et les comportements de santé (consommation d'alcool, de tabac, dépistages réalisés...) des assurés venant passer volontairement des examens de santé à titre de prévention ou de dépistage. Ces données ne peuvent être considérées comme représentatives de l'ensemble de la population, car les CES sont inégalement répartis sur le territoire et leur activité ne concerne que des assurés volontaires, soit ayant fait l'objet d'une invitation par les CPAM soit orientés par des organismes (publics en situation de précarité par exemple). Il est toutefois intéressant de signaler l'utilisation du score EPICES, score construit par les CES, pour mesurer la précarité de leurs consultants. Les données publiées analysent ainsi des indicateurs de santé (état de santé ou comportements de santé) d'une part selon les CSP et d'une part selon le degré de précarité.

L'enquête OBEPI, réalisée par l'INSERM avec le soutien du laboratoire Roche, est une étude transversale nationale sur la prévalence du surpoids et de l'obésité dans la population des plus de 18 ans. La prévalence de l'obésité est décrite selon : le sexe et l'âge, la profession, le niveau d'instruction, les revenus du foyer, ou encore la catégorie d'agglomération. L'association entre l'obésité et les facteurs de risque cardio-vasculaires sont également analysés (hypertension artérielle, tabagisme, dyslipidémies, diabète traité). Au niveau régional, OBEPI fournit des chiffres (environ tous les 3 à 4 ans) sur l'obésité dans la région, mais ne permet pas d'analyse plus fine.

Peu de données sur les comportements de santé sont donc exploitables à un niveau géographique fin, hormis celles qui résultent d'un recours au système de soins soit pour un acte de dépistage, soit pour une pathologie en lien plus ou moins direct avec un comportement (exemple : les pathologies liées à la consommation d'alcool ou de tabac).

En matière de dépistage, le taux de participation au **dépistage organisé des cancers du sein ou du colon-rectum** est disponible dans les bases de données élaborées par les structures départementales de dépistage (*annexe 1, fiche 17*). Le niveau géographique auquel ces données sont publiées dépend des structures de gestion, l'information étant diffusée au niveau départemental, pour tous les départements, par l'Institut de veille sanitaire.

La participation des enfants au **programme M'T'dents** (examen bucco-dentaire gratuit pour les enfants de 6, 9, 12, 15 ans) est connue grâce aux bases de l'assurance maladie (*source : ERASME – annexe 1, fiche 8*), comme la participation des personnes de 65 ans et plus aux campagnes annuelles de vaccination antigrippale.

La prévalence ou l'incidence de certaines pathologies en lien avec des comportements (par exemple le cancer du poumon, fortement lié à la consommation de tabac) sont connues via les systèmes d'information généralistes en santé, tels que le PMSI MCO pour les hospitalisations (*annexe 1, fiche 6*) ou le système national des causes de décès (*annexe 1, fiche 2*).

Les bases de données sur les accidents (de la vie courante, de la circulation, du travail) n'ont pas fait l'objet d'un descriptif dans le cadre de ce travail. Les décès domiciliés par accident de la circulation, ou par accident de la vie courante, sont toutefois repérables dans la base nationale des causes de décès (*annexe 1, fiche 2*).

Quelques exemples d'indicateurs en lien avec des comportements de santé :

- Taux de participation au dépistage organisé des cancers (sein et colon-rectum) (*annexe 1, fiche 17*)
- Taux de participation au programme bucco-dentaire gratuit M'T'dents de l'assurance maladie pour les enfants de 6, 9, 12, 15 et 18 ans (*source : ERASME – annexe 1, fiche 8*)
- Taux standardisé d'hospitalisation pour pathologies liées à l'alcool, ou pour tentatives de suicide (*source : PMSI MCO – annexe 1, fiche 6*)
- Taux standardisé de mortalité par cancer du poumon – indicateur fortement lié à la consommation de tabac (*source : Statistique des causes médicales de décès – annexe 1, fiche 2*)

Les données sur les conditions de travail

Les données sur les conditions de travail ne font pas l'objet d'un recueil d'information systématique. L'organisation de la médecine du travail en France fait que la surveillance des travailleurs est très parcellisée, les grandes entreprises possédant leur propre service de santé au travail et les autres entreprises utilisant les services interentreprises de médecine du travail.

Des enquêtes nationales, avec parfois des extensions sur une ou plusieurs régions (ou départements), abordent la question des conditions de travail, comme **l'enquête SUMER**, dont la DIRECCTE Rhône-Alpes a analysé les réponses dans la région sur : les contraintes organisationnelles des salariés en lien avec la santé mentale, les contraintes et ambiances physiques, les risques chimiques et biologiques, ou encore l'exposition des salariés aux facteurs de pénibilité. Cette enquête, basée sur le volontariat des médecins du travail, est jugée représentative au niveau régional mais ne peut être déclinée à un niveau infrarégional.

De même, le **programme SAMOTRACE** de surveillance de la santé mentale selon l'activité professionnelle, développé par l'InVS dans le Rhône et l'Isère, repose sur un échantillon aléatoire de salariés suivis par des médecins du travail volontaires.

C'est le cas également de **l'enquête EVREST (Evolutions et relations santé au travail)**, dispositif de veille et de recherche en santé au travail, construit en collaboration par des médecins du travail et des chercheurs pour pouvoir analyser et suivre différents aspects du travail et de la santé de salariés. Le recueil des données s'appuie sur un questionnaire très court rempli lors des consultations de médecine du travail. Mais s'il existe bien une analyse régionale des résultats, celle-ci n'est pas considérée comme représentative, du fait des petits échantillons régionaux (472 salariés en 2012 et 1 228 en 2013).

Les systèmes d'information sur **les maladies professionnelles** (*annexe 1, fiche 18*) peuvent apporter des informations intéressantes sur les expositions professionnelles des travailleurs, par exemple sur l'exposition à l'amiante, mais ces systèmes d'information sont spécifiques à chaque régime d'assurance maladie et les données diffusées sont agrégées soit au niveau de la France (données assez détaillées par secteur d'activité, pathologie, ancienneté de l'exposition) soit au niveau régional (données très peu détaillées). Par ailleurs, les maladies professionnelles doivent faire l'objet d'une reconnaissance par l'organisme de sécurité sociale et figurer dans la liste officielle des « tableaux » permettant cette reconnaissance ; le lien formel entre l'exposition et la maladie est souvent difficile à établir. L'InVS a développé un système d'information complémentaire, sur « **les maladies à caractère professionnel** », mais ce système n'a été mis en place que dans quelques régions (actuellement en Auvergne mais pas en Rhône-Alpes).

Par les bases de données de l'assurance maladie, il est également possible de connaître le nombre de journées d'arrêt de travail indemnisées aux salariés pour maladies professionnelles ou accident du travail (*source : ERASME, annexe 1, fiche 8*). Cette information ne permet toutefois pas de savoir de quelle maladie professionnelle il s'agit, et ne renseigne donc pas sur les éventuelles expositions professionnelles à l'origine de la pathologie.

Les données sur l'offre de soins

Les données sur l'offre de soins concernent d'une part les professionnels de santé et d'autre part les établissements de soins.

Les professionnels de santé sont répertoriés dans **les fichiers RPPS (Répertoire partagé des professionnels de santé) et ADELI** (*annexe 1, fiche 19*).

Les données du RPPS sont élaborées à partir d'une confrontation de plusieurs sources (Ordres, Assurance maladie, etc.). Le RPPS a été créé en 2009, et doit progressivement prendre la place du fichier ADELI (qui existait depuis 1984). Le RPPS a concerné dans une première phase les quatre professions médicales dotées d'un Ordre (médecins, pharmaciens, chirurgiens-dentistes, sages-femmes) et les autres professions de santé devraient intégrer progressivement le répertoire à partir de 2014. Les fichiers RPPS/ADELI, ainsi que des informations issues de l'assurance maladie sur les professionnels de santé libéraux, alimentent différentes plateformes permettant d'avoir accès soit aux caractéristiques des professionnels (coordonnées, mode de conventionnement, spécialités, tranches d'âges des professionnels...), soit à des indicateurs calculés à différentes échelles territoriales, tels que la densité de professionnels, la distance en km à un type de professionnel (médecin spécialiste par exemple), ou encore l'APL (accessibilité potentielle localisée) qui tient compte également du niveau d'activité du professionnel et du taux de recours différencié par âge des habitants. Ces plateformes (soit nationales comme Atlasanté ou régionales comme Carto santé) permettent d'obtenir des informations au niveau communal, voire infracommunales.

Les données sur les établissements sanitaires et sociaux (publics et privés) concernent d'une part leurs coordonnées et agréments (fichier FINESS) et d'autre part leurs données d'activité (SAE).

Le Fichier national des établissements sanitaires et sociaux FINESS (*annexe 1, fiche 20*) assure, au niveau national, l'immatriculation de tous les établissements et entités juridiques porteurs d'une autorisation ou d'un agrément. La localisation de ces établissements, avec leurs catégories, est largement diffusée sur différentes plateformes (FINESS, data.gouv.fr, Atlasanté...), avec les coordonnées précises des établissements et des entités juridiques.

La SAE (Statistique annuelle des établissements de santé) (*annexe 1, fiche 21*) est une enquête annuelle, exhaustive et obligatoire, qui recueille auprès de tous les établissements de santé des informations sur leur activité, leurs capacités, leurs équipements, et leurs personnels médicaux et non-médicaux. Ces données sont des données "enregistrées" à l'adresse de l'établissement, et non "domiciliées" à l'adresse du patient, ce qui limite les analyses possibles sur la santé des habitants d'un territoire.

Les bordereaux de la SAE renseignent sur les lits et places installés, sur les personnels, sur les journées ou séjours ou séances réalisées, sur le nombre de patients pris en charge, au niveau de l'établissement, de l'entité juridique, des établissements d'une commune (ou département, territoire de santé, région).

Quelques exemples d'indicateurs sur l'offre de soins :

- Densité de médecins ou autres professionnels de santé libéraux
- Distance d'une commune à un professionnel de santé libéral ou à un établissement
- APL (accessibilité potentielle localisée) des professionnels libéraux
- Fragilité d'un territoire en termes de démographie médicale (ex : part des médecins âgés de plus de 55 ans)
- Activité de services hospitaliers spécialisés : oncologie, urgences, maternité, soins palliatifs, grands brûlés...

Les applications diffusant des données sur l'offre de soins et l'activité des établissements ont été répertoriées par l'IDS (Institut des données de santé) début 2016 : <http://www.institut-des-donnees-de-sante.fr/connaissance/>. Ce sont des applications aussi bien sur les séjours hospitaliers (ScanSanté de l'ATIH) que sur la qualité et la sécurité des soins (Scope santé de la HAS).

2.1.2 Bases de données en social

Les bases de données concernant la situation socioéconomique des habitants reposent essentiellement sur des opérations de recensement (de la population, des logements, des chômeurs...), de déclarations (fiscales par exemple), ou encore sur la délivrance de prestations sociales (RSA, AAH, minimum vieillesse...).

On peut distinguer des bases de données généralistes (comme le Recensement de la Population) qui concernent toute la population et abordent de nombreuses thématiques, ou des bases plus spécifiques : sur les prestations sociales, les revenus, le chômage, le logement.

Le Recensement général de la population (RGP) (*annexe 1, fiche 22*) représente le système d'information principal en matière de données sociodémographiques. Bien que non exhaustif depuis sa réforme en 2004, le RGP, par sa méthode de construction, permet d'estimer annuellement la population de chaque commune (et des quartiers IRIS) à partir d'une collecte réalisée sur 5 années. Le RGP recueille des informations à la fois sur l'individu (sexe, âge, nationalité, niveau scolaire, situation professionnelle...), sur le ménage (composition), sur le logement (ancienneté, confort, mode de chauffage), sur la possession de voitures... Les données agrégées du RGP sont accessibles à de nombreux niveaux géographiques allant jusqu'à la commune, et pour certaines données au niveau de l'IRIS.

Certains indicateurs calculés à partir du RGP sont particulièrement utilisés pour repérer les inégalités sociales des territoires :

- Part des jeunes de moins de 20 ans, et part des personnes âgées de plus de 65 ans
- Variation de la population due au solde naturel et au solde migratoire
- Part des personnes en activité et part des chômeurs

- Part de familles monoparentales
- Nombre d'enfants par ménage
- Part de personnes âgées vivant seules
- Part des personnes sorties du système scolaire sans diplôme, ou selon le niveau du diplôme
- Répartition des ménages selon la catégorie socio-professionnelle de la personne de référence
- Population active de 15 ans et plus selon les conditions d'emploi (type de contrat de travail, salarié, indépendant, temps de travail...)
- Nombre de pièces du logement selon la taille du ménage (sur-occupation ou sous-occupation du logement)
- Statut d'occupation du logement (propriétaire, locataire...)

Les prestations sociales dites « minima sociaux », c'est-à-dire soumises à un plafond de revenu, sont versées principalement par les CAF (Caisses d'allocations familiales) et par la MSA (Mutualité sociale agricole), mais également par les CARSAT (Caisse d'assurance retraite et de la santé au travail). Ce sont le RSA (Revenu de solidarité active), l'AAH (Allocation adulte handicapé), le minimum vieillesse (accordé aux personnes âgées à partir de 65 ans). Hormis ces minima sociaux, les CAF et MSA versent de nombreuses autres allocations, telles que les allocations familiales ou les allocations logement.

Les bases de données sur **les allocations versées par les CAF et les MSA** (*annexe 1, fiche 23*) permettent de connaître le nombre de ménages ou d'individus bénéficiaires d'allocations, selon l'âge du bénéficiaire, la structure et la taille du ménage, le niveau de revenus, le nombre d'enfants (et leur âge) couverts par l'allocation... Ces informations agrégées sont disponibles au niveau communal (pour la CAF) dans la mesure de confidentialité des données.

Le minimum vieillesse (*annexe 1, fiche 24*) correspond à un minimum de pension de vieillesse accordé, sous conditions de ressources, aux personnes qui n'ont pas suffisamment cotisé aux régimes de retraite pour pouvoir bénéficier d'un revenu d'existence à l'âge de la retraite ; Il est versé par la CARSAT ou par la MSA (régime agricole). Les données de la CARSAT sont accessibles, via une cartographie dynamique, à de échelles territoriales fines.

La CMUc (Couverture Minimum Universelle complémentaire) permet aux personnes à faibles revenus de disposer gratuitement d'une couverture maladie complémentaire. Cette information est diffusée au niveau départemental par le Fonds CMU (<http://www.cmu.fr>), qui centralise les données des trois régimes d'assurance maladie : le régime général (CNAMTS), le régime agricole (MSA) et le régime des indépendants (RSI). A un niveau infra départemental, il faut solliciter directement chaque régime. La base ERASME (*annexe 1, fiche 8*), pour le régime général, permet une analyse des données à un niveau fin, par exemple les quartiers IRIS ou les quartiers en politique de la ville.

Quelques exemples d'indicateurs sur les prestations sociales et la CMUc :

- Part des bénéficiaires et de la population couverte par le RSA (RSA socle et/ou RSA activité), par âge

- Part des familles monoparentales parmi les allocataires CAF
- Part des familles à bas revenus parmi les allocataires CAF
- Part des familles dont les revenus dépendent à 100% (ou de 50 à 100%) des prestations sociales
- Part des retraités bénéficiaires du minimum vieillesse
- Taux comparatifs des assurés et ayants-droit du régime général bénéficiant de la CMUc

Le nouveau dispositif sur le revenu disponible (FiLoSoFi) (*annexe 1, fiche 25*) permet de chiffrer le revenu disponible des ménages, plus seulement à partir du revenu fiscal déclaré mais également à partir des prestations sociales reçues. Un certain nombre d'indicateurs sont ainsi produits par l'INSEE : des indicateurs usuels d'analyse de la distribution des revenus (déciles, médiane...) pour l'ensemble de la population ainsi que pour des sous-populations, des taux de pauvreté monétaire, des indicateurs d'origine des revenus (salaires, pensions, prestations sociales, revenus du patrimoine...). Ces indicateurs sont mis à disposition à différents niveaux géographiques, jusqu'à la commune actuellement (avec des seuils de population en lien avec le secret fiscal), et prochainement (mai 2016) jusqu'au quartier IRIS et au quartier prioritaire de la politique de la ville.

Quelques exemples d'indicateurs sur les revenus :

- Part des ménages fiscaux imposés
- Revenu disponible médian par UC (unité de consommation)
- Taux de pauvreté selon l'âge
- Part des ménages dont les revenus sont principalement dépendants des allocations chômage
- Part des familles monoparentales

Plusieurs sources d'information renseignent sur l'emploi et le chômage, principalement **Pôle Emploi** (*annexe 1, fiche 26*) et **l'INSEE à partir du Recensement général de la population** (*annexe 1, fiche 22*) et d'autres enquêtes. En matière de chômage, si les données de Pôle Emploi reposent sur l'enregistrement des demandeurs d'emploi, celles du Recensement de la population sont des données déclaratives, et les taux de chômage calculés par l'INSEE ne sont pas issus de données observées mais font l'objet d'estimations. Ces taux ne sont pas disponibles à un niveau infra départemental.

Quelques exemples d'indicateurs sur l'emploi :

- Taux de chômage déclaré au Recensement (pour des niveaux géographiques fins)
- Part de la population active parmi les 15-64 ans
- Part des salariés en contrat précaire (CDD, intérim, contrats aidés...)
- Part de chômeurs de longue durée parmi les chômeurs

Les données sur les logements sociaux sont collectées chaque année par le Ministère en charge du logement et rassemblées dans **le RPLS – Répertoire du parc locatif social** (*annexe 1, fiche 27*). Ce Répertoire permet de connaître non seulement le nombre de

logements du parc locatif social, mais également les principales caractéristiques de ces logements (ancienneté, nombre de pièces, loué ou vacant, prix moyen au m2...). Les données ne sont pas diffusées à un niveau infra départemental, mais sont disponibles au niveau communal via la DREAL.

Le Fichier des logements par commune (FILOCOM) (*annexe 1, fiche 28*) concerne quant à lui l'ensemble des logements, et résulte d'un traitement conjugué des fichiers de la taxe d'habitation, du foncier, des propriétaires et de l'impôt sur les revenus. En plus des informations sur les logements (pièces, confort, ancienneté...), FILOCOM renseigne sur l'occupation du logement (locataire ou propriétaire, vacance) et sur les occupants des logements (nombre de personnes, âge des enfants, revenus nets...). Un outil de pré-repérage des logements indignes a été mis au point, le PPPI (Parc privé potentiellement indigne), à partir de la catégorie cadastrale du logement et du revenu de ses occupants. Les données statistiques confidentialisées sont adressées aux collectivités locales, et pour d'autres demandeurs l'accès se fait via la DREAL.

Quelques exemples d'indicateurs sur le logement :

- Nombre de logements sociaux
- Nombre de logements potentiellement indignes

2.1.3 Bases de données en environnement

Les bases de données centralisant des données environnementales sont multiples, et ont très généralement pour origine un besoin de gestion de la qualité de l'environnement ou de constitution d'un historique afin d'observer des tendances sur le long terme.

Le travail de recensement ou de productions de données est mené par des agences, observatoires ou instituts publics, ou largement financés par des fonds publics.

L'accès aux données est généralement possible mais souvent rendu complexe par des démarches administratives ou juridiques (inscription puis autorisation, conventionnement, etc.) et un éparpillement auprès de multiples producteurs.

Lors des deux dernières décennies, plusieurs textes législatifs sont venus tout de même conforter l'accès à l'information environnementale, tant pour permettre au citoyen de s'informer et de débattre que pour optimiser les missions des services publics (Convention d'Aarhus, Directive n° 2003/4/CE, etc.). La directive européenne INSPIRE (2007, transposée en droit français en 2010), est venue compléter ces textes en améliorant la diffusion des données géographiques détenues par les services publics.

Les bases de données recensées dans ce rapport concernent la région Rhône-Alpes (et par extension l'Auvergne), elles sont présentées ci-dessous par milieux, et détaillées en annexe 1.

Données pour le milieu AIR

Qualité de l'air ambiant

Air Rhône-Alpes détient plusieurs bases de données, construites en lien direct avec sa mission confiée par agrément de surveillance de la qualité de l'air sur son territoire de compétence. Les premières concernent les données de métrologie, directement recueillies sur des points de prélèvement, choisis pour être représentatifs d'une stratégie d'échantillonnage : **base de données des mesures automatiques** (*annexe 2, fiche 29*) et **base de données des mesures par prélèvement** (*annexe 2, fiche 30*).

Afin de compléter ces mesures ponctuelles, Air Rhône-Alpes met en œuvre des outils numériques de reconstitution de la qualité de l'air sur l'ensemble du territoire, à travers des outils de modélisation : une nouvelle base de données est ainsi constituée, centralisant des **concentrations cartographiées** de différents polluants sur toute la région Rhône-Alpes (*annexe 2, fiche 31*).

Par ailleurs, **un inventaire spatialisé des émissions atmosphériques polluantes** est géré par Air Rhône-Alpes (*annexe 2, fiche 32*).

Données polliniques

Le principe de fonctionnement du RNSA repose en premier lieu sur la constitution des sites de captures des pollens qui sont des sites de recueil de fond. Chaque site a été choisi par rapport à des critères climatiques, botaniques et de densité de populations, il est constitué d'un capteur, d'un responsable du capteur, d'un analyste et d'un médecin responsable clinique du site.

Le centre de coordination est basé à Brussieu dans des locaux spécifiques, en plein cœur des Monts du Lyonnais à 40 km à l'ouest de Lyon (69). Ce centre reçoit les résultats des analyses polliniques et les informations cliniques associées. Ainsi, il peut assurer la rédaction de bulletins allergo-polliniques composés d'informations polliniques, cliniques et du risque allergique lié. Ces bulletins nationaux ou régionaux sont retransmis aux sites, aux médias, aux services déconcentrés du Ministère de la Santé, et à tous les partenaires du RNSA. Le centre de coordination assure la gestion en toute sécurité de la base de données qui permet les recherches statistiques sur les pollens et moisissures, l'élaboration des bilans annuels, la mise en place de prévisions, etc.

Les données exprimées dans ces fichiers sont des concentrations journalières en nombre de grains/m³ d'air pour les pollens et en nombre de spores/m³ d'air pour les moisissures.

Ainsi, deux bases de données sont recensées au RNSA :

- **Comptages de pollens dans l'air** centralisés par le RNSA (*annexe 2, fiche 35*)
- **Indice pollinique et risque ambrosie** produits par le RNSA (*annexe 2, fiche 36*)

En complément, Air Rhône-Air a développé un outil de cartographie pour l'ambrosie : la plateforme met à disposition une prévision du risque allergique lié au pollen d'ambrosie pendant la saison de la plante, mise à jour à fréquence hebdomadaire. Elle vise à aider les personnes allergiques à adapter leur comportement et leur traitement aux conditions

à venir. Elle est réalisée en collaboration étroite avec le *RNSA*, expert de l'allergie au pollen et de la prévision à court terme.

- **Cartographie régionale de la présence de la plante d'ambroisie** produite par Air Rhône-Alpes (*annexe 2, fiche 33*)
- **Cartographie régionale du risque allergique à l'ambroisie** produite par Air Rhône-Alpes (*annexe 2, fiche 34*).

Rayonnements électromagnétiques

L'ANFR gère une base de données recensant les mesures réalisées par les exploitants d'installations radioélectriques (*annexe 2 – fiche 37*).

Sur Cartoradio figurent toutes les installations radioélectriques de plus de 5 watts, hormis celles de l'Aviation Civile et des ministères de la Défense et de l'Intérieur, pour des raisons de sécurité. Cartoradio permet, d'une part, de connaître l'emplacement des stations radioélectriques et, d'autre part, d'avoir accès, pour un site donné, aux résultats des mesures de champ électromagnétiques synthétisés par une fiche de mesures.

Pour accéder aux résultats, l'utilisateur renseigne une adresse ou zoome directement sur la carte. L'utilisateur peut alors sélectionner un exploitant et obtenir les détails sur les caractéristiques radioélectriques de l'installation.

Les informations suivantes figurent sur la fiche :

- Son numéro d'identification (une référence administrative à rappeler dans toute correspondance avec l'ANFR touchant à cette installation) ;
- Le nom de l'exploitant ;
- L'adresse d'implantation ;
- Le code postal et le nom de la commune ;
- La date de l'accord de l'ANFR lors de l'implantation (ou de son avis favorable pour les installations de radiodiffusion) ;
- La date de l'accord de l'ANFR pour la dernière modification technique sur l'installation.

Puis, classé par antenne, on trouve :

- La hauteur de l'antenne sur son support par rapport au sol ;
- Son orientation ;
- Le(s) système(s) utilisé(s)
- La/les bande(s) de fréquences utilisée(s) - exprimées en kHz jusqu'à 27 500 kHz, en MHz de 27,5 à 10 000 MHz, en GHz au-dessus.

Radioactivité

Il s'agit du recensement, à partir de la connaissance des caractéristiques géologiques des sols, du potentiel radon (3 classes) (*annexe 2 – fiche 38*).

La connaissance des caractéristiques des formations géologiques sur le territoire, et en particulier de leur concentration en uranium, rend possible l'établissement d'une cartographie des zones sur lesquelles la présence de radon à des concentrations élevées dans les bâtiments est la plus probable. Ce travail a été réalisé par l'IRSN à la demande de l'Autorité de Sûreté Nucléaire et a permis d'établir une cartographie du potentiel radon des formations géologiques du territoire métropolitain et de la France de l'Outre-mer.

La cartographie du potentiel du radon des formations géologiques établie par l'IRSN conduit à classer les communes en 3 catégories :

- Les communes à potentiel moyen ou élevé

Ce sont celles localisées sur les formations géologiques présentant des teneurs en uranium les plus élevées. Les formations concernées sont notamment celles constitutives des grands massifs granitiques français (massif armoricain, massif central, Guyane française...), certaines formations volcaniques (massif central, Polynésie française, Mayotte...) mais également certains grès et schistes noirs.

Dans les communes à potentiel radon moyen ou élevé, la proportion des bâtiments présentant des concentrations en radon élevées est plus importante que dans le reste du territoire. Les résultats de la campagne nationale de mesure en France métropolitaine montrent ainsi que plus de 40% des bâtiments de ces communes dépassent 100 Bq.m⁻³ et plus de 6% dépassent 400 Bq.m⁻³.

- Les communes à potentiel faible

Ce sont celles localisées sur les formations géologiques présentant les teneurs en uranium les plus faibles. Ces formations correspondent notamment aux formations calcaires, sableuses et argileuses constitutives des grands bassins sédimentaires (bassin parisien, bassin aquitain) et à des formations volcaniques basaltiques (massif central, Polynésie française, Antilles...).

Dans les communes concernées, une grande majorité de bâtiments présente des concentrations en radon faibles. Les résultats de la campagne nationale de mesure en France métropolitaine montrent ainsi que seulement 20% des bâtiments dépassent 100 Bq.m⁻³ et moins de 2% dépassent 400 Bq.m⁻³.

- Les communes à potentiel faible mais sur lesquelles des facteurs géologiques particuliers peuvent faciliter le transfert du radon vers les bâtiments

Les communes concernées sont notamment celles recoupées par des failles importantes ou dont le sous-sol abrite des ouvrages miniers souterrains, etc. Ces conditions géologiques particulières peuvent localement faciliter le transport du radon depuis la roche jusqu'à la surface du sol et ainsi augmenter la probabilité de concentrations élevées dans les bâtiments.

Pesticides

Air Rhône-Alpes a entrepris depuis 2007 des mesures de composés pesticides dans l'air ambiant, par prélèvement. Ces mesures constituent une base de données, accessible avec des données hebdomadaires généralement (*annexe 2 – fiche 39*).

Un travail a par ailleurs conduit pour estimer les émissions de pesticides vers l'atmosphère, à l'échelle du code postal de l'acheteur, exprimées en g/ha/an (*annexe 2 – fiche 40*).

Enfin, la BNVD (Banque Nationale de Vente des Produits Phytosanitaires) inventorie les ventes des produits au code postal de l'acheteur, pour toutes les substances phytosanitaires (*annexe 2 – fiche 41*).

Bruit

On recense ici des bases constituées de mesures directement enregistrées dans trois agglomérations de Rhône-Alpes (Lyon, Grenoble et Saint Etienne) (*annexe 2 – fiches 42 à*

44) ou des bases issues de cartographies du bruit dans ces mêmes agglomérations (*annexe 2 – fiche 45*) ou à l'échelle de la région Rhône-Alpes (*annexe 2 – fiche 46*).

Données pour le milieu EAU

Trois bases, à vocation nationale, concentrent l'ensemble des données qualitatives et quantitatives sur la qualité des eaux :

➤ **Contrôle sanitaire de la qualité de l'eau potable** (*annexe 2 – fiche 47*)

La présence de micro-organismes (bactéries, virus et parasites) dans les eaux de consommation est le plus souvent due à une dégradation de la qualité de la ressource en eau, à une mauvaise protection ou un manque d'entretien des ouvrages de captages, à une défaillance du traitement de désinfection ou à une contamination de l'eau lors de son transport ou stockage dans le réseau. Les méthodes pour rechercher les germes pathogènes dans l'eau sont longues et complexes, c'est pourquoi la qualité bactériologie de l'eau est appréciée à partir de la recherche de germes témoins de contamination fécale (*Escherichia coli* et entérocoques). La mise en évidence de ces germes dans l'eau témoigne de la possibilité de présence de germes pathogènes.

La présence des nitrates dans les eaux est due soit à leur présence naturelle dans l'environnement, soit à une contamination de la ressource en eau par des activités humaines (rejets urbains ou industriels, pollution agricole due aux engrais minéraux et organiques). La présence de pesticides dans l'eau est due à leur entraînement par ruissellement ou à leur infiltration dans les sols.

Les nitrates et les pesticides sont recherchés au niveau des ressources en eau utilisées pour la production d'eau potable, à la sortie des installations de production d'eau potable ou sur le réseau de distribution publique.

➤ **Qualité des eaux de baignade** (*annexe 2 – fiche 48*)

Le suivi régulier de la qualité des eaux de baignade permet de connaître les impacts de divers rejets éventuels situés à l'amont du site et notamment d'apprécier les éventuels dysfonctionnements liés à l'assainissement d'eaux usées, aux rejets d'eaux pluviales souillées, etc., qui influenceraient la qualité de l'eau du site de baignade. Les connaissances ainsi acquises peuvent fournir une aide à la décision aux collectivités locales afin d'améliorer la maîtrise des causes des pollutions engendrées notamment par une mauvaise gestion des eaux usées domestiques. Le contrôle sanitaire des eaux de baignade est mis en œuvre par les Agences régionales de santé (ARS) et demeure une préoccupation constante du ministère chargé de la santé. Ce ministère élabore la réglementation dans ce domaine sur la base de directives européennes.

➤ **Les données du Système d'Information sur les Eaux (SIE)**

Elles rassemblent des millions de mesures sur les milieux aquatiques, les activités humaines ayant un lien ou un impact sur avec l'eau ou les milieux aquatiques, les incidences directes ou indirectes de ces activités, les réponses apportées par les politiques publiques... Ces données sont tout aussi bien d'ordre quantitatif, physico-chimique, biologique, morphologique, réglementaire ... et sont stockées dans des

banques de données (HYDRO, ADES, QUADRIGE...). Avant d'être mises en lignes sur Eaufrance, les données suivent différentes étapes visant à fournir aux utilisateurs des données fiables et compréhensibles.

- **Les données sur les eaux souterraines** (*annexe 2 – fiche 49*) : BSS-Eau est la base de données relative aux informations sur les eaux souterraines, base de données au format SANDRE- Service d'Administration Nationale des Données et Référentiels sur l'Eau (données sur la qualité des eaux souterraines, niveaux d'eau, prélèvements, etc....), organisée et gérée par le BRGM. Elle diffuse des informations spécifiques sur le descriptif du point d'eau (entités hydrogéologiques, masses d'eau, ...), les données sur les prélèvements et les traçages. Ces éléments constituent la liste des points d'eau en France (métropole et DROM). Aujourd'hui une fiche descriptive sur chaque ouvrage de type point d'eau est accessible : identification, descriptif du point d'eau, caractéristiques hydrogéologiques, piézométrie. La base de données est l'objet de mises à jour pour améliorer la qualité des informations déjà saisies (précisions et compléments de dossiers) et pour prendre en compte les informations communiquées par les différents organismes du Système d'Information sur l'Eau (Agences de l'eau, DREAL, collectivités locales, syndicats d'eau, ...) qui contribuent à enrichir les informations sur le descriptif des points d'eau. Ces mises à jour sont réalisées par les directions régionales du BRGM. Le rythme actuel de mise à jour est quotidien en fonction des informations transmises. Les données informatisées proviennent de plusieurs sources complémentaires dans l'espace et dans le temps.
- **Le réseau de contrôle de surveillance (RCS) des eaux de surface** (*annexe 2 – fiche 50*) s'appuie sur 2 043 stations pérennes réparties sur l'ensemble du territoire (métropole et outre-mer). Les critères de définition du réseau ont essentiellement porté sur la taille du bassin, en veillant à bien représenter les différents types de masses d'eau et éviter la proximité de rejets ou d'aménagements (par exemple, des barrages). 82% des stations du contrôle de surveillance sont situées sur des cours d'eau, catégorie d'eau la plus représentée. La fréquence et la période des contrôles sont déterminées de manière à apporter des données suffisantes pour une évaluation de l'état des eaux. Les paramètres et les fréquences minimales des mesures à réaliser sont précisés dans l'arrêté du 25 janvier 2010. Les contrôles concernent tous les éléments de qualité :
 - les éléments de qualité biologique : faune (poissons, crustacés, etc.), flore (algues, par exemple) ;
 - les éléments de qualité hydromorphologique : débit, état des rives, largeur du lit, continuité de l'écoulement, etc. ;
 - les éléments de qualité physico- chimique généraux (température, oxygène, nutriments, etc.) et chimiques (différentes substances).

Données pour le milieu SOL

Sont décrites trois bases de données des éléments traces métalliques, dont le territoire d'emprise est national.

➤ **Base de données des Eléments Traces Métalliques (BDETM) dans les sols**
(*annexe - fiche 51*)

L'ADEME a financé deux programmes de collecte d'analyses de sols en éléments traces métalliques. Ces deux campagnes ont été réalisées par l'INRA. La première collecte, effectuée en 1997 et 1998, a permis de recueillir des analyses pour plus de 11 000 sites provenant principalement d'une trentaine de départements métropolitains. La nouvelle collecte s'est déroulée de novembre 2008 à janvier 2010. Elle a été réalisée selon les mêmes principes que la première mais avec de plus grands moyens informatiques. Elle inclut toutes les données déjà rassemblées et traitées en 1998 et contient plus de 73 400 sites. La répartition à l'échelon national est bien meilleure que dans le cas de la première collecte de 1998 : il n'y a qu'un seul département pour lequel aucune donnée n'a pu être recueillie. L'immense majorité des analyses récoltées provient de plans d'épandages de boues d'épuration et dans une moindre mesure de programmes scientifiques. Les analyses concernent d'abord 5 propriétés pédologiques : la texture (argile, limons, sables), le carbone organique, la CEC et le pH dans l'eau. 7 ETM ont été mesurés avec deux méthodes (extraction à l'eau régale ou HF) : cadmium, chrome, cuivre, mercure, nickel, plomb, zinc. Les statistiques sont calculées sur des sous populations par tranche temporelle de 10 ans.

➤ **Analyses de sols** lors de deux grandes campagnes nationales, selon un maillage national régulier (*annexe 2 – fiche 52*)

Les données présentées constituent les statistiques descriptives des stocks de carbone sur 0-30 cm, pour sept grands types d'occupation du sol, sur le territoire métropolitain (Corse incluse), de façon globale et par région administrative. L'estimation des stocks exclut les horizons organiques constituant la litière. Les variances d'échantillonnage (nommée varSTSI) et l'écart type sur l'estimation de la moyenne (nommée erreur sur la moyenne) ont été calculés d'après Saby et al. (2014). Les autres variables sont calculées de façon classique.

➤ **Base de données sur les sites et sols pollués ou potentiellement pollués (BASOL/BASIAS)** (*annexe 2 – fiche 53*)

La base de données BASOL du Ministère en charge de l'Environnement présente les sites et sols pollués (ou potentiellement pollués) appelant une action des pouvoirs publics, à titre préventif ou curatif. Ces situations sont souvent dues à d'anciennes pratiques sommaires d'élimination des déchets, mais aussi à des fuites ou à des épandages de produits chimiques, accidentels ou pas. Il existe également autour de certains sites des contaminations dues à des retombées de rejets atmosphériques accumulés au cours des années voire des décennies.

La réalisation d'inventaires historiques régionaux (IHR) des sites industriels et activités de service, en activité ou non, s'est accompagnée de la création de la base de données nationale BASIAS. Les principaux objectifs de ces inventaires sont de recenser, de façon large et systématique, tous les sites industriels abandonnés ou non, susceptibles d'engendrer une

pollution de l'environnement afin de conserver la mémoire de ces sites et de fournir des informations utiles aux acteurs de l'urbanisme, du foncier et de la protection de l'environnement. BASIAS sera prochainement disponible en Rhône-Alpes.

ICPE (installations classées pour la protection de l'environnement)

Une dernière base de données, qui ne concerne plus spécifiquement l'un des milieux, mais plutôt une partie de l'activité industrielle, concerne la **base des installations classées pour la protection de l'environnement (ICPE)**. Mise à disposition par le Ministère, cette base de données permet de rechercher les installations classées soumises à autorisation ou à enregistrement (en construction, en fonctionnement ou en cessation d'activité) selon des critères géographiques, le type d'activité ou un régime particulier (SEVESO, IED, etc.). Cette base permet également de consulter les actes réglementaires liés à une installation tels que les arrêtés préfectoraux d'autorisation et les arrêtés complémentaires (*annexe 2 – fiche 54*).

2.2 Méthodes de croisement de données

Les études épidémiologiques peuvent être classées selon qu'elles ont une approche écologique, contextuelle ou individuelle. L'approche de type « **écologique** » désigne toute étude basée, non pas sur des populations sélectionnées individu par individu en fonction d'un certain nombre de critères, mais sur des données populationnelles agrégées à l'échelon d'un territoire. L'approche « **contextuelle** » permet, à la différence de l'approche écologique, d'analyser à la fois et simultanément des données collectées au niveau individuel et collectif pour avancer dans la compréhension des déterminants sociaux de la santé et tenir compte par conséquent du contexte. L'approche de type « **individuelle** » considère toute étude dont l'unité statistique est l'individu.

Les quelques études présentées ici sont toutes de type « écologique », même si certaines sources de données utilisées pour construire les indicateurs sont constituées des données individuelles (notamment les indicateurs sanitaires).

Les études sont présentées de façon plus détaillée en annexe 2.

Description succincte des études

- **Les inégalités environnementales de santé des territoires en région PACA (IEST – PACA) (*annexe 2, fiche 2.1*)**

Cette étude vise à élaborer des outils de description des inégalités de santé liées à l'environnement, afin d'intégrer ce nouvel axe de réflexion dans le cadre de la planification et de la mise en œuvre des politiques en santé publique, à différentes échelles du territoire, en région Provence-Alpes-Côte d'azur (PACA).

Elle participe de deux cadres de travail aux orientations communes :

- Le second plan national santé environnement (PNSE2 : 2009-2014) et son axe stratégique de lutte contre les inégalités environnementales ; et sa déclinaison régionale : le Plan régional santé environnement (PRSE PACA 2009-2014) ;
- Le projet régional de santé de l'ARS PACA et l'axe stratégique de réduction des inégalités de santé.

Les facteurs environnementaux sont désormais considérés comme un déterminant de santé à part entière et la dimension environnementale est de plus en plus intégrée à l'étude des inégalités de santé en complément des inégalités dites sociales.

➤ **Projet de recherche Equit'Area. Expositions environnementales et inégalités de santé** (*annexe 2, fiche 2.2*)

Le projet de recherche Equit'Area vise à explorer la contribution de certaines pollutions et nuisances environnementales aux importantes inégalités sociales de santé qui existent en France et qui ont eu tendance à croître au cours de la dernière décennie.

Il s'agit d'une étude d'épidémiologie spatiale multicentrique. L'unité géographique utilisée est l'IRIS (Îlots Regroupés pour l'Information Statistique, regroupant en moyenne 2000 habitants) qui constitue la plus petite unité géographique en France pour laquelle les données démographiques et socio-économiques du recensement national sont disponibles.

Dans un premier temps, un nombre limité de zones urbaines a été choisi, afin de mettre en évidence les disparités spatiales et d'affiner une méthodologie d'analyse statistique. L'existence en France de contrastes entre les grandes agglomérations urbaines concernant la situation sanitaire, les caractéristiques économiques et démographiques et la magnitude et la diversité des expositions environnementales, a conduit à privilégier, pour ce qui constitue une « étude pilote », les quatre grandes aires urbaines suivantes : Lille, Paris (et petite couronne), Lyon et Marseille, permettant une bonne couverture « nord-sud ». Cette diversité pourrait permettre de distinguer les caractéristiques communes aux 4 agglomérations de celles spécifiques à chacune d'entre elles.

Le projet de recherche Equit'Area se veut aussi une plateforme pour éclairer l'action publique qui chercherait à réduire ces inégalités sociales de santé. Son site (<http://www.equitarea.org>) met ainsi à disposition de toutes les parties intéressées, élus nationaux et locaux, administrations, décideurs économiques publics et privés, associations, ainsi que des chercheurs, des informations synthétiques issues de l'analyse de la très grande quantité des données compulsées.

➤ **CIRCE - Cancer Inégalités Régionales Cantonales et Environnement** (*annexe 2, fiche 2.3*)

L'étude CIRCE vise à identifier les inégalités environnementales mises en avant par les Plans Nationaux Santé Environnement. Elle regroupe aujourd'hui sept Observatoires Régionaux de Santé (Aquitaine, Bretagne, Ile-de-France, Lorraine, Nord-Pas-de-Calais, Picardie et Rhône-Alpes soutenus chacun par leur Conseil régional) et l'Institut national de l'environnement industriel et des risques (INERIS).

L'objectif est, d'une part, d'évaluer, dans le cadre d'une étude écologique, la part des facteurs sociaux et environnementaux dans la répartition géographique des cancers, et d'autre part,

réaliser un Système d'Information Géographique (SIG) dédié aux cancers et aux expositions environnementales.

La méthode utilisée repose sur une analyse géographique faite au niveau des « cantons-villes » de l'INSEE. Les variations spatiales de la mortalité sont analysées en cartographiant l'Indice comparatif de mortalité (ICM) d'une quinzaine de localisations cancéreuses. Les caractéristiques sociodémographiques des territoires sont croisées avec les données sanitaires et des données d'exposition environnementale. Un élargissement de la méthodologie à d'autres pathologies (maladies cardio-vasculaires ou respiratoires) et à d'autres problématiques explorant d'autres sources (données d'hospitalisation, admissions en ALD, registre des malformations) est envisagé.

➤ **PLAINÉ – Plateforme intégrée pour l'Analyse des INégalités d'Exposition environnementale** (*annexe 2, fiche 2.4*)

PLAINÉ (plateforme intégrée pour l'analyse des inégalités d'exposition environnementale) est un outil intégré d'évaluation spatialisée des risques. Cet outil associe des bases de données géoréférencées de différents types (environnementales, sanitaires, comportementales et socio-économiques), un modèle d'exposition multi-milieux construit par l'INERIS (calcul des doses d'exposition par ingestion ou inhalation) et un système d'information géographique pour construire des cartographies d'exposition des populations à une échelle spatiale fine (de l'ordre du km²) et pour différentes échelles temporelles.

L'INERIS, sur la base de la plateforme PLAINÉ, a produit une étude pilote basée sur l'exposition des populations de Picardie et du Nord-Pas-de-Calais à quatre éléments trace métalliques (nickel, cadmium, plomb, chrome). Les résultats obtenus correspondent à une exposition future des populations sur 70 ans sous l'hypothèse d'une exposition constante à celle constatée aujourd'hui. Ils démontrent la faisabilité et la pertinence de la plateforme : l'agrégation des données d'émission à la source, de mesures environnementales et de populations à des fins d'évaluation des expositions est donc possible.

➤ **Etat des lieux sur les points noirs environnementaux en Lorraine** (*annexe 2, fiche 2.5*)

L'étude s'inscrit dans le cadre de l'action 4.1 du PRSE 2 de la région Lorraine dont le but est de réaliser sur la région un état des lieux sur les points noirs environnementaux. Il est piloté par le SGAR en partenariat avec la DREAL, le BRGM, l'ARS, la DIRECCTE, AIR Lorraine, le Conseil Régional, l'ORS et l'INERIS.

L'objectif de cette étude est, sur l'ensemble de la région Lorraine, d'identifier les zones géographiques points noirs environnementaux, définies par le cumul de plusieurs contaminations, dans le but de mettre en place, dans un second temps, des actions sur certains points noirs.

L'étude se déroule en trois phases : une première phase avec une approche environnementale visant l'identification des zones géographiques de cumul d'exposition à des polluants environnementaux ; une deuxième phase qui explore des facteurs sanitaires, socioéconomiques sur la dizaine de zones géographiques de cumul d'exposition à des polluants environnementaux déterminée lors de la première étape ; une troisième phase visant à mettre en place des actions concrètes en vue de la réduction des émissions, de la diminution des expositions et de l'amélioration de l'information de la population.

➤ **Bruit routier et statut socio-économique : étude des inégalités environnementales à Marseille** (*annexe 2, fiche 2.6*)

Le bruit environnemental lié au trafic routier, ferroviaire et aérien peut entraîner des perturbations du sommeil, des symptômes psychologiques et des troubles cardiovasculaires. Il serait le deuxième facteur de risque environnemental en termes d'années de vie perdues ajustées sur l'incapacité, après la pollution de l'air. Les inégalités environnementales liées au bruit – cumul de faibles statuts socio-économiques (SSE) et de probabilités d'exposition élevées – ont été relativement peu étudiées contrairement à celles qui sont liées à la pollution de l'air et aux risques industriels. Les études écologiques disponibles indiquent l'existence d'inégalités d'exposition au bruit en fonction du SSE mais les phénomènes d'auto corrélation spatiale ne sont pas pris en compte dans ces études et le SSE est le plus souvent étudié par une seule variable. Or, selon la théorie des inégalités environnementales, trois types de facteurs – économiques, sociopolitiques et raciaux – peuvent influencer sur les expositions [6]. Des freins économiques ou sociétaux (discrimination) pourraient réduire la mobilité résidentielle des personnes de faible SSE ou d'origine étrangère. L'objectif de cette étude est d'analyser le lien entre le SSE et l'exposition potentielle au bruit routier à une échelle géographique fine en tenant compte de l'auto corrélation spatiale et en analysant le SSE à la fois par un indice synthétique et par des dimensions économiques, socio-éducatives, d'origine géographique et de mobilité résidentielle.

➤ **Projet PAIRSAC + : pollution atmosphérique, inégalités sociales, asthme, risque cardiaque. Influence du contexte de vie** (*annexe 2, fiche 2.7*)

L'objectif de l'étude est d'évaluer l'intérêt d'une approche extensive des inégalités de santé tenant compte du contexte de vie en mesurant ses effets sur le risque d'exacerbation de l'asthme (EA) et de déclenchement de l'infarctus du myocarde (IM) liés à la pollution atmosphérique. La zone d'étude est la Communauté urbaine de Strasbourg.

La méthode consiste à construire un nouvel indicateur de NSE incorporant de nombreuses variables de caractérisation du contexte de vie, en utilisant des méthodes alternatives de segmentation spatiale (détermination de la zone sur laquelle les effets de contexte apparaissent, le 'Modifiable Area Unit Problem' de la littérature), tester par analyse cas-croisés les effets de contexte (NSE) sur la survenue des événements sanitaires étudiés par comparaison avec une approche utilisant un indicateur socio-économique. Dans les deux cas, l'analyse est écologique sur la base d'une très petite zone géographique. Les données NSE et les données de pollution sont modélisées spatialement et temporellement afin de mettre en lumière un éventuel phénomène d'agrégation zonale de cas sur une durée d'un an.

➤ **Impacts sanitaires de la pollution atmosphérique urbaine et des expositions à proximité du trafic routier dans l'agglomération parisienne –ERPURS** (*annexe 2, fiche 2.8*)

Les grandes agglomérations telles que l'agglomération parisienne sont caractérisées par des contrastes marqués de niveaux et de composition chimique des polluants atmosphériques selon que l'on se situe à l'écart ou à proximité des voies à fort trafic routier. Un nombre croissant d'études épidémiologiques a documenté une dégradation de la santé des populations résidant à proximité des voies à fort trafic routier, souvent plus importante que

celle rapportée pour les niveaux de fond. Toutefois, devant l'importance des prévalences de pathologies chroniques, de l'asthme notamment, et des contrastes de pollution atmosphérique dans l'agglomération parisienne, rendre compte du poids sanitaire de la pollution à proximité du trafic routier apparaît comme nécessaire. C'est dans cette optique qu'une équipe de recherche du projet Aphekom (Künzli et al., 2012) a proposé d'intégrer les effets des expositions chroniques aux polluants du trafic routier dans les évaluations d'impacts. Cette évaluation cherche à rendre compte à la fois des impacts à long terme des expositions chroniques à proximité du trafic, et des impacts à court terme des expositions aiguës à la pollution atmosphérique ambiante. Cette méthode a été appliquée au cœur dense de l'agglomération parisienne (Paris et proche couronne) où l'impact sanitaire de la pollution atmosphérique à proximité du trafic routier a été évalué en termes de survenue de nouveaux cas d'asthme, d'exacerbations de symptômes et d'hospitalisations pour asthme chez les enfants (< 18 ans).

➤ **Etude de l'inégalité socio environnementale en France : associations spatiales entre désavantage social des populations et proximité à un site potentiellement dangereux** (*annexe 2, fiche 2.9*)

L'étude présente la construction d'un indice de désavantage socio-environnemental régional, à partir d'un indice de désavantage social et d'un indicateur de proximité aux sites et activités polluantes potentiellement dangereuses pour la santé. Cet indice sera utilisé dans le cadre des prochains plans régionaux santé-environnement pour étudier les associations spatiales entre désavantage social et indicateurs d'exposition réalisés à partir d'un modèle d'exposition multi milieu tenant compte de la contamination des compartiments environnementaux et du comportement des populations.

Synthèse

Les études présentées utilisent des méthodes de croisement de données dans le but de décrire les inégalités territoriales environnementales et sociales de santé.

Les méthodes utilisées sont de trois types : des méthodes bayésiennes, des méthodes géostatistiques et des méthodes d'analyse de données. Chacune d'elles a été retenue au regard des objectifs de l'étude et des analyses souhaitées pour décrire les inégalités.

2.3 Indices synthétiques

Les quelques indices synthétiques présentés ont été choisis soit parce que ce sont les plus utilisés, soit parce qu'ils ont été construits lors d'études sur les inégalités territoriales. Ils sont détaillés en annexe 3.

Description succincte des différents indices retenus

➤ **L'indice de défavorisation sociale : FDep** (*annexe 3, fiche 3.1*)

L'indicateur de défavorisation sociale (FDEP) est défini comme la première composante principale de l'analyse en composante principale (ACP) de quatre variables : le revenu médian par ménage, le pourcentage de bacheliers dans la population de plus de 15 ans, le pourcentage d'ouvriers dans la population active et le taux de chômage. Il représente 68% de l'inertie totale formée par les quatre variables et est fortement corrélé à chacune de ces dimensions dans un sens cohérent de désavantage social (négativement avec le revenu médian et le pourcentage de bacheliers, positivement avec le pourcentage d'ouvriers et le taux de chômage).

L'indice a tout d'abord été calculé à l'échelle des communes à partir des données de recensement de la population de 1999 et des revenus fiscaux des ménages de 2001, puis a été agrégé à l'échelle des cantons et des régions en calculant la moyenne pondérée par les populations des communes. Le territoire a par la suite été divisé en quintiles de FDep99, représentant chacun environ 20% de la population. Le pourcentage de variance expliquée par la première composante était de 60% avec les données de recensement de 1990, de 68% avec celles des 1999. Il semblerait qu'avec les données du recensement de 2012, il ne soit plus que de 54% (étude en cours).

➤ **L'indice de défavorisation sociale : TOWNSEND & CARSTAIRS** (*annexe 3, fiche 3.2*)

Selon Townsend, l'un des précurseurs du domaine, « le concept de désavantage renvoie à un état de désavantage relatif face à la communauté locale ou à l'ensemble de la société à laquelle appartient l'individu, la famille ou le groupe ». Il revêt une dimension matérielle (privation de biens etc.) et une dimension sociale (isolement etc.).

Cet indicateur a pour objectif de rendre compte des inégalités sociales de santé.

Townsend et Carstairs ont chacun proposé, au cours des années 1980 au Royaume-Uni, un score établi selon des méthodes de calcul différentes évaluant le désavantage social à des niveaux géographiques fins à partir de quatre variables chacun.

L'indice synthétique évalue une certaine réalité sociale que l'on assimile à de la précarité sociale dans les zones urbaines. Or, il ne peut être affirmé avec certitude que cet indice serait bien adapté au milieu rural. En effet, le sens de certaines variables peut être ambigu selon l'environnement dans lequel elles sont mesurées. A titre d'exemple, un niveau d'équipement élevé en véhicules individuels peut être vécu comme un signe d'autonomie et de richesse dans certains quartiers urbains, comme l'expression d'une contrainte et d'un coût supplémentaire pour se déplacer de son domicile au travail en milieu périurbain.

➤ **Indice de Pampalon** (*annexe 3, fiche 3.3*)

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)

➤ **L'indice de développement humain de l'ONU : IDH** (*annexe 3, fiche 3.4*)

L'Indice de développement humain (IDH) est une mesure sommaire du niveau moyen atteint dans des dimensions clés du développement humain : vivre une vie longue et en bonne santé, acquérir des connaissances et jouir d'un niveau de vie décent. Cet indice est calculé pour chaque pays.

En 2009, Aurélien Boutaud (chercheur indépendant) a proposé une déclinaison régionale et infrarégionale de l'Indicateur de Développement Humain en proposant des cartographies des performances à l'aune de cet indice. Ce travail permet aux territoires et aux régions de se comparer les uns aux autres. Une variante à l'Indicateur de Développement Humain (IDH-2) a été proposée et appliquée au cas de la Communauté Urbaine de Lyon et de la région Ile de France. La Région Nord-Pas-de-Calais a également décliné, au niveau de son territoire, les indicateurs de développement humain à l'échelle communale.

La dimension santé du développement humain est mesurée par un complément de la mortalité. Cette variable est construite à partir de l'indice comparatif de mortalité toutes causes de décès calculé par l'ORS Nord-Pas de Calais.

➤ **Indice de Santé Sociale : ISS** (*annexe 3, fiche 3.5*)

L'Indicateur de Santé Sociale (ISS) reprend les six grandes dimensions du baromètre des inégalités et de pauvreté : éducation, justice, logement, santé, revenus, travail & emploi. Ces deux dernières dimensions sont décomposées en sous dimensions : consommation, inégalités et fiscalité, pauvreté, salaires pour la dimension revenus ; chômage, conditions de travail, précarité, relations professionnelles pour la dimension travail et emploi. Pour tenir compte du caractère positif de ce qui est estimé : « la santé sociale », deux dimensions enrichissent cet indicateur. La dimension du lien social d'abord, estimé à partir du taux d'adhésion aux associations ; le lien interindividuel ensuite, estimé à partir du nombre d'interactions enregistrées entre individus (amis et voisins). Ces deux dernières mettent l'accent sur une facette importante du patrimoine social de chaque région. Chaque dimension de l'ISS est composée d'une, ou plus rarement deux, variables combinées entre elles.

➤ **Indicateur Synthétique de Développement et d'Évolution de la Santé : ISDES** (*annexe 3, fiche 3.6*)

Cet indice permet de mesurer l'évolution de la santé des habitants dans le Nord – Pas-de-Calais. Il synthétise quatre composantes : l'offre de soins, la dimension socio-économique, l'état de santé, et l'opinion des usagers du système de santé. Cet indicateur dynamique a été imaginé pour être suivi d'année en année. Il sera, à l'avenir,

complété d'aspects environnementaux, dont l'évolution chronologique n'a pas pu être estimée dans cette première édition.

➤ **Indice global de qualité de vie et de bien-être développé par l'Institut d'aménagement et d'urbanisme (IAU) d'Ile de France** (*annexe 3, fiche 3.7*)

Cet indice global de qualité de vie et de bien-être est construit à partir d'une méthode qui permet de croiser et/ou pondérer une quarantaine d'indicateurs relatifs à la dimension humaine (santé publique, population, conditions de vie (richesse), éducation, communication, égalité, criminalité, paix et libertés) pour former un Indice du « bien-être humain » ou de la qualité socio-économique (IQS), et une cinquantaine d'indicateurs relatifs à l'environnement (diversité des terres, protection des zones géographiques, qualité et disponibilité de l'eau, qualité locale de l'air, diversité génétique, consommation de l'énergie et pression sur les ressources) pour former un Indice du « bien-être de l'écosystème » ou de la qualité environnementale (IQE). Enfin, ces deux indices sont agrégés pour former un Indice global de qualité de vie et de bien-être qui caractérise le niveau des menaces et les conséquences du développement sur la qualité de l'environnement d'une région et sur la qualité de vie de ses habitants.

➤ **Qualité de vie dans les territoires de l'INSEE** (*annexe 3, fiche 3.8*)

La qualité de vie dans les territoires a été mesurée à travers 27 indicateurs, recouvrant treize dimensions à l'échelle des 2 677 territoires de vie de France métropolitaine. Ce zonage découpe les bassins de vie de plus de 50 000 habitants pour mieux rendre compte de la diversité de la qualité de vie au sein des territoires les plus urbanisés. S'affranchissant des limites des unités urbaines, les territoires de vie découpent ainsi les grands bassins de vie autour des pôles de services. La France métropolitaine est ainsi constituée de 2 677 territoires de vie, les bassins de vie de moins de 50 000 habitants étant conservés tels que. La typologie a mis en évidence 8 types de territoires de vie et permet d'identifier les déterminants de la qualité de vie au sein de ces groupes.

➤ **Indice de désavantage social (IDS)** (*annexe 3, fiche 3.9*)

L'indice de défaveur est construit par des analyses en composantes principales (ACP) successives réalisées à partir d'une sélection de 52 variables. La variable de l'indice par IRIS est calculée à partir de la combinaison linéaire des 19 variables sélectionnées. L'indice est ensuite catégorisé en 10 (ou 3 ou 5) classes selon les déciles (ou terciles ou quintiles) de la distribution. L'équipe d'Equit'Area a développé une procédure méthodologique pour créer des indices socioéconomiques ; cette procédure étant plus complexe que pour d'autres indices socioéconomiques, elle a développé le package SesIndexCreator pour l'appliquer. La version 1.0 de ce package contient les fonctions de base nécessaires pour effectuer la procédure

➤ **Indice de désavantage social (IDS) simplifié – Application en PACA** (*annexe 3, fiche 3.10*)

Cet indice synthétique a été construit par ACP (analyse en composantes principales), à partir de 16 variables. Les IRIS (ou territoires d'observation de proximité) ont ensuite été classés en 5 classes selon les quintiles de l'indice calculé. Mais l'IDS simplifié n'a pas fait

l'objet du même niveau de validation que l'IDS. Le classement des territoires, obtenu selon cet indice, nécessite une validation par les acteurs de terrain qui connaissent ces territoires.

➤ **Indice de Défavor environnementale de l'EHESP** (*annexe 3, fiche 3.11*)

Une analyse multidimensionnelle a été appliquée (ACP) afin de trier l'information fournie en entrée. Ainsi, parmi les 38 variables environnementales initiales caractérisant chaque département, quelques variables utiles mettent en évidence les différences entre les zones géographiques. On élimine ensuite toutes les variables environnementales qui ne permettent pas de discriminer les territoires entre eux.

On applique ensuite une méthode de hiérarchisation des facteurs environnementaux retenus qui génèrent le plus une situation de disparité environnementale

Pour enfin construire des indices composites milieu par milieu : air, industrie et SPP, Eau, habitat bruit. Ces indices sont ensuite répartis en 5 classes équivalentes aux quintiles afin d'obtenir la représentation graphique la plus discriminante de ces disparités.

➤ **Indice de Défavor environnementale de PACA** (*annexe 3, fiche 3.12*)

Une méthodologie mixte, composée d'une Analyse en Composantes Principales (ACP) suivie d'une Classification Ascendante Hiérarchique (CAH), a été utilisée.

L'analyse des facteurs environnementaux a été menée sous l'angle de 4 thématiques : air, sols, eau, habitat afin de quantifier et hiérarchiser l'impact des différentes pressions environnementales influençant la santé.

Ainsi, à partir des résultats issus de l'ACP, un indice de défavor environnementale global a été créé. Il a été complété par un indice de défavor environnementale milieu par milieu, démarche analogue mettant en évidence la contribution respective des différents milieux à une situation de défavor environnementale.

Pour déterminer la contribution de chaque milieu dans la défavor environnementale des ESP (Espaces de santé de proximité), c'est-à-dire connaître leur poids sans que celui-ci ne soit influencé par des variables connexes, des ACP par milieu ont été réalisées. La somme de ces résultats par thématiques permet d'obtenir les indices de défavor environnementale par milieu.

La CAH est une technique statistique couramment utilisée pour l'analyse de données multi-variées. Elle vise à partitionner un ensemble E, composé de n individus caractérisés par p variables, en m sous-groupes les plus homogènes possible. Cette méthode de classification itérative consiste à agréger deux-à-deux les individus les plus proches. Elle nécessite de définir une métrique pour caractériser la distance entre deux points et de fixer une règle pour agréger les individus. Le dendrogramme a été tronqué en six classes.

Synthèse des indices recensés

Les indices présentés ne sont qu'une petite partie du nombre d'indices composites que l'on peut trouver dans la littérature.

Certains indices sont calculés à partir d'un processus d'agrégation additif des variables retenues (processus de « scoring »). Ils prennent alors en compte un nombre limité de variables, agrégées sur la base d'un système de pondérations défini a priori.

D'autres, plus complexes, sont construits à partir d'une analyse statistique multidimensionnelle (ACP). Ils considèrent un nombre plus important de variables et le système de pondérations est déterminé par calculs sur la base de critères statistiques.

Avant tout, il convient de vérifier la disponibilité et l'échelle de la donnée de base avant de construire un indice composite. Les ACP sont souvent utilisées comme étapes intermédiaires dans les analyses statistiques, mais ne sont pas suffisantes.

Ensuite, comme on peut le voir à travers nos quelques exemples, leur construction est plus ou moins complexe et comporte un nombre d'indicateurs très variable (jusqu'à plus d'une centaine). Le choix des indicateurs de base pour décrire un territoire influence grandement les regroupements de zones opérés et il serait intéressant de vérifier que ces indicateurs soient pertinents sur le territoire considéré, non pas uniquement statistiquement, mais qu'ils représentent un sens au regard de la population. D'ailleurs, comme cela est fréquemment reporté dans la littérature, il existe de grandes différences entre les territoires urbains et ruraux, or la plupart des indices ont plutôt été créés pour le milieu urbain. La fiabilité et la viabilité de ces indices doivent faire l'objet d'une analyse attentive.

Enfin, c'est sur la défaveur sociale qu'il existe le plus grand nombre d'indices synthétiques. Certains indices, comme le score Epices, score de précarité utilisé par les Centres d'examen de santé de l'assurance maladie, n'ont pas été décrits dans ce travail, parce qu'ils nécessitent une interrogation directe des personnes (en l'occurrence les assurés venant passer un examen de santé), et ne peuvent donc être utilisés à partir des bases de données.

Dans le domaine de la santé, les indices composites de « défaveur » sont peu nombreux, et on utilise davantage des indicateurs de santé générale comme l'espérance de vie, la mortalité générale, la mortalité prématurée, ou encore la prévalence des affections de longue durée (ALD) toutes causes. Ponctuellement, pour des études spécifiques, des indices ont pu être élaborés, comme par exemple un indice de « besoin de soins », construit en 2007 par l'ORS Rhône-Alpes pour la DRASS et l'URCAM Rhône-Alpes³, dans le cadre du Plan Régional de Santé Publique, à partir d'indicateurs de santé (ICM mortalité générale, ICM mortalité prématurée, ICM mortalité par cancer, ICM mortalité par diabète), d'indicateurs d'offre de soins (Densité de médecins spécialistes libéraux de

³ Denis FONTAINE, Olivier GUYE, Véronique JOSEPH, Abdoul SONKO. Identification de territoires cibles en santé en Rhône-Alpes. Analyses en composantes principales et typologie des territoires. ORS Rhône-Alpes, 2007. 48p. Disponible : <http://www.ors-rhone-alpes.org/pdf/Rapport%20territoires%20cibles%20ACP.pdf>,
et Méthode de l'indice global, 2007. 52p. Disponible : <http://www.ors-rhone-alpes.org/pdf/Rapport%20Territoires%20Cibles%20Indice%20global.pdf>

proximité, Densité d'infirmiers libéraux, Part de médecins libéraux conventionnés en secteur 1), et d'indicateurs sociodémographiques (Taux de bénéficiaires de la CMU, Taux de bénéficiaires du RMI, Taux de chômage de longue durée, Part des 75 ans et plus dans la population).

Concernant l'environnement, les indices sont plutôt développés par milieu, ce qui implique de procéder à des analyses de corrélation avec les variables représentatives.

A titre d'exemple, pour le compartiment air : historiquement, les AASQA (Associations Agréées de Surveillance de la Qualité de l'Air) diffusent chaque jour l'indice ATMO, calculé sur l'ensemble des agglomérations de plus 100 000 habitants, selon les modalités de l'arrêté ministériel du 22 juillet 2004 modifié par l'arrêté du 21 décembre 2011. Cet indice, calculé à partir des mesures réalisées en quelques sites permanents, ne décrit pas les variations spatiales de la qualité de l'air à l'intérieur même d'une aire urbaine et il est disponible uniquement dans les agglomérations. Il est peu approprié pour des croisements de données.

Cependant, depuis septembre 2015, Air Rhône-Alpes calcule, selon une méthodologie développée en interne, et diffuse sur son site web quotidiennement un « **indice Air** » en tous points du territoire, à partir d'un modèle numérique qui reproduit le comportement des masses d'air et des polluants. Cet « indice Air » prend en compte trois polluants (poussières PM10, dioxyde d'azote, ozone), avec des classes bornées sur des pourcentages d'atteinte des valeurs limites ou valeurs cibles. Il est calculé sur des mailles de 1 km, agrégé à l'échelle communale.

Dans le même esprit, Air Rhône-Alpes, Acoucity et le CEREMA ont développé, à travers la plateforme ORHANE, un « **indicateur de co-exposition Air Bruit** », intégrant des données de qualité de l'air et de bruit, décrivant ainsi l'ensemble du territoire rhônalpin (et auvergnat en 2017).

3. Discussion

A partir de l'analyse de l'ensemble des éléments décrits (bases de données, méthodes utilisées par les études, indices composites), cette partie s'attachera à discuter des questions importantes dans le cadre d'une étude sur la caractérisation des ITESS : l'accessibilité des données, les niveaux géographiques disponibles dans les bases de données, ainsi que les avantages et les limites des bases de données mobilisables.

3.1 Accessibilité des données

Il est important de préciser que la description des bases de données réalisée dans le cadre de ce travail, et notamment de leur accessibilité, a été réalisée en 2015 et début 2016, et que les conditions d'accès aux différentes bases sont susceptibles d'évoluer rapidement, en raison d'une part du mouvement de l'OpenData (davantage de données sont publiées sur les portails publics), et d'autre part d'une tendance à la sécurisation de l'accès aux données (moins de personnes ou d'organismes habilités à accéder aux données à un niveau fin).

En effet, si les questions qui se posent au sujet de l'ouverture publique des données ne sont pas tout à fait les mêmes pour les données sanitaires et pour les données environnementales, il existe toujours une contradiction forte entre l'exigence de confidentialité (protection de l'individu, protection de l'entreprise...) et la volonté de transparence des données publiques.

En lien avec le mouvement de l'OpenData, **la directive européenne Inspire du 14 mars 2007** vise à établir une infrastructure d'information géographique dans la Communauté européenne, soit un ensemble de services d'information disponibles sur internet permettant la diffusion et le partage de données géographiques. Cette directive impose aux autorités publiques des pays membres d'une part de rendre accessibles au public les données géographiques qu'elles détiennent, et d'autre part de permettre les échanges de ces données (exigence d'interopérabilité des données). Dans le cadre de cette directive, le gouvernement français, via la mission Etalab (qui fait partie du Secrétariat général pour la modernisation de l'action publique) a ouvert la plateforme « **data.gouv.fr** », dont l'objectif est de mettre en ligne les données publiques et leurs métadonnées. En parallèle une réflexion est conduite sur la standardisation des données, afin de faire converger les méthodes et outils relatifs à l'information géographique ; c'est la mission de la Commission de validation pour l'information spatialisée (COVADIS).

Les plateformes de mise à disposition de données publiques seront ainsi évoquées dans ce chapitre, sans qu'il soit possible de décrire précisément leurs contenus, tant ceux-ci sont en constante évolution. Les accès à certaines bases de données, via ces plateformes, ont parfois été présentés dans les descriptifs des bases, en annexe 1, lorsqu'ils constituaient le principal accès aux données concernées.

L'accessibilité aux bases de données en santé

L'ouverture des bases de données en santé s'est accélérée ces dernières années, avec la mise en ligne d'une partie des données de ces bases, soit sur des sites internet dédiés soit sur des plateformes regroupant des données diverses.

Certains producteurs de données diffusaient déjà leurs données via des sites internet depuis plusieurs années, il s'agissait le plus souvent de données agrégées, à des niveaux géographiques assez importants (le département par exemple), sous forme de tableaux, éventuellement via une interface d'interrogation. C'est le cas par exemple de la base des causes médicales de décès (*INSERM CépiDc, annexe 1, fiche 2*) ou encore des bases de données sur les établissements sanitaires et sociaux (*FINESS -annexe 1, fiche 20, SAE – annexe 1, fiche 21*).

Le mouvement d'OpenData a conduit à une diffusion plus large des données publiques, et d'autres producteurs de données mettent désormais en ligne une partie de leurs données, même si les données brutes ne sont pratiquement jamais accessibles, hormis celles concernant les répertoires d'établissements ou de professionnels. On peut citer le nouveau site de mise à disposition des données hospitalières **ScanSanté** : <http://www.scansante.fr>, ainsi que les données de la DREES (Direction de la recherche, des études, de l'évaluation et des statistiques) mises en ligne sur la plateforme **Data Drees** : <http://www.data.drees.sante.gouv.fr>. L'IDS (Institut des données de santé) a d'ailleurs récemment diffusé un recensement des données hospitalières en accès libre : <http://www.institut-des-donnees-de-sante.fr/connaissance/>.

D'autres fichiers (lots de données) sont progressivement insérés sur des plateformes spécialisées en santé (telles que **Atlasanté, le portail géographique des Agences régionales de santé** : <http://www.atlasante.fr>), ou généralistes (telles que **Data.gouv, la plateforme ouverte des données publiques françaises** : <https://www.data.gouv.fr>). Se côtoient sur ces plateformes des données très diverses : répertoires d'établissements ou de professionnels de santé, définition de zonages spécifiques à la santé, démographie médicale, séjours hospitaliers par catégories, indicateurs de qualité des soins, dépenses de santé, médicaments, etc. La plateforme Data.gouv se présente comme « un outil à la disposition des citoyens pour leur permettre de **produire, d'enrichir ou de compléter des données** ou de **partager des données d'intérêt public**, qu'il s'agisse par exemple de l'inventaire d'une bibliothèque communale, de l'état de la voirie, de la composition nutritionnelle de produits alimentaires ou du bilan environnemental d'une entreprise ». Ceci explique le grand nombre de jeux de données disponibles sur Data.gouv, dont certains sont déclarés « certifiés » et d'autres pas. Les seules limites fixées pour la mise en ligne de données sur Data.gouv concernent « les données à caractère personnel (données fiscales ou patrimoniales, données médicales, etc.) et celles dont la publication contrevient à la loi (secret médical, secret de la défense nationale, secret statistique, secret des affaires, etc.) ».

Les données de santé diffusées via ces plateformes ne sont donc jamais des données individuelles, ce sont des données agrégées à des niveaux géographiques (ou thématiques) plus ou moins importants. Afin de pallier cette difficulté, qui interdit toute

nouvelle agrégation selon d'autres critères que ceux définis au départ, les organismes d'études et de recherche ont besoin d'accès particuliers aux données « brutes », données anonymisées mais peu ou pas agrégées. Le Haut Conseil de la santé publique (HCSP) milite depuis plusieurs années pour un accès plus large aux principales bases de données nationales⁴.

En ce qui concerne le SNIIRAM, entrepôt de données regroupant les principales bases de données en santé (données de l'assurance maladie inter-régimes, données hospitalières, données sur les décès et les causes médicales de décès...), les accès étaient jusqu'à présent définis par l'Institut des données de santé (IDS), avec des listes d'organismes autorisés à un accès soit aux données exhaustives individuelles anonymisées, soit à des extractions (échantillons spécifiques), soit à l'EGB (échantillon généraliste des bénéficiaires), soit à des données agrégées. Ces accès devraient évoluer avec l'article 193 de « la loi n° 2016-41 du 26 janvier 2016 de modernisation du système de santé »⁵, dont les décrets d'application ne sont pas encore parus. L'article 193, intitulé « créer les conditions d'un accès ouvert aux données de santé », prévoit la création d'un Institut national des données de santé qui rassemblera et mettra à disposition l'ensemble des données en santé issues des administrations et de l'assurance maladie, dans le respect de l'anonymat des personnes et du secret médical. Un décret en Conseil d'Etat, pris après avis de la CNIL, doit fixer la liste des organismes autorisés à accéder au système national des données de santé.

Les futures modalités de transmission des données à la FNORS et aux ORS par les producteurs de données ne sont ainsi pas encore définies et peuvent remettre partiellement en question ou complexifier l'accès aux données et les exploitations réalisées actuellement au niveau régional à des niveaux géographiques relativement fins. Précisons que la plupart des indicateurs de santé diffusés à ce jour au niveau de la commune, du code postal, du quartier IRIS, le sont via des portails d'indicateurs élaborés par les ORS soit à un niveau national (**Score santé**, <http://www.scoresante.org>, par la Fnors), soit à un niveau régional (dans notre région : **Balises**, <http://www.balises-rhone-alpes.org>, par l'ORS Rhône-Alpes). Les ORS, souvent en collaboration avec les ARS, calculent et mettent à disposition du public et des collectivités locales de nombreux indicateurs, comme les indices et les taux standardisés, à partir de données brutes ou peu agrégées transmises par les producteurs de données (ATIH, CNAMTS, MSA, RSI, INSERM CépiDc, CAF, DREES, CARSAT, etc.). L'accès à ces données demeure la condition essentielle pour la poursuite du travail d'analyse et de diffusion réalisé dans chaque région.

⁴ HCSP. Pour une meilleure utilisation des bases de données nationales pour la santé publique et la recherche. HCSP, Collection Documents, mars 2012. 54p. Disponible :

<http://www.hcsp.fr/explore.cgi/avisrapportsdomaine?clefr=258>

⁵ LOI n° 2016-41 du 26 janvier 2016 de modernisation de notre système de santé. JORF n°0022 du 27 janvier 2016 ; Disponible :

https://www.legifrance.gouv.fr/affichTexte.do;jsessionid=0C9BD3D7309BF6837C42CE4788E0489D.tpdila16v_2?cidTexte=JORFTEXT000031912641&categorieLien=id

L'accessibilité aux bases de données sociales

La diffusion des bases de données démographiques et sociales à des niveaux géographiques fins s'est faite historiquement de façon plus avancée que celle de bases de données en santé. Ceci est certainement très lié à la contrainte du respect du secret médical en matière de santé, mais également au fait que les producteurs de données démographiques et sociales ont dans leurs missions de communiquer certaines données à des organismes publics, comme les collectivités territoriales, données à partir desquelles les organismes publics définissent leurs politiques sociales (en matière d'emploi, de logement, d'équipements, d'aides sociales...).

Néanmoins, les données diffusées doivent toujours respecter l'anonymat des personnes (et des entreprises) et, quelle que soit la nature de la donnée, il existe toujours un seuil d'effectifs en dessous duquel les données seront considérées comme non diffusables.

Les données du RGP, recensement général de la population, sont largement diffusées tant en données agrégées qu'en « fichiers détail », mais les données individuelles, sur les personnes, sont soumises au secret statistique. L'INSEE explicite dans un « Guide du secret statistique »⁶ les raisons et les modalités de mise en œuvre du secret statistique. Ce dernier doit garantir aux personnes physiques la confidentialité sur leur vie personnelle et familiale, et aux entreprises le secret commercial et la non communication d'informations à leurs concurrents. Pour les tableaux fournissant des données agrégées sur les ménages, la seule règle imposée par le secret statistique est que l'identification directe ou indirecte des individus soit impossible. Dans la pratique, on considère que le secret statistique est respecté si la connaissance d'une caractéristique pour un individu ne peut pas entraîner la connaissance d'une autre caractéristique avec laquelle elle est croisée dans un tableau. Des règles spécifiques de diffusion concernant le recensement de la population ont été édictées. Il est toutefois possible, pour des chercheurs et dans des conditions très strictes, d'accéder aux données individuelles non anonymisées obtenues à partir des enquêtes auprès des ménages, après avis du comité du secret statistique, accord de l'autorité dont émanent les données et sur décision de l'administration des archives. Cet accès se fait via un centre d'accès sécurisé distant (le CASD).

Ainsi, s'il est possible d'accéder aux résultats du RGP aux niveaux géographiques publiés par l'INSEE, il n'est parfois pas possible de calculer certains indicateurs à d'autres niveaux.

Les données sur les revenus, sur l'emploi et sur le logement, diffusées par l'INSEE, obéissent à la même logique.

La diffusion des données spécifiques sur le logement, Répertoire du parc locatif social (RPLS) et Fichier des Logements par COMMune (FILOCOM), est très réglementée, des données confidentialisées (secret inférieur à 11 unités) sont adressées largement aux collectivités

⁶ Guide du secret statistique. INSEE, version du 18 octobre 2010, 26p. Disponible : <http://www.insee.fr/fr/insee-statistique-publique/statistique-publique/guide-secret-18-10-2010.pdf>

locales pour la mise en œuvre des politiques de l'habitat. Des accès spécifiques sont possibles via les DREAL.

Le mouvement d'OpenData a conduit **certains organismes de gestion des prestations sociales** (CAF, CARSAT) à mettre en place de nouveaux outils de diffusion de leurs bases de données. Ainsi les CAF (Caisses d'allocations familiales) ont ouvert un Open-data CAF, où elles mettent à disposition des informations sur les publics bénéficiaires des allocations familiales à différentes échelles territoriales, du niveau national au niveau communal. Afin de respecter le secret statistique, les valeurs inférieures à 5 sont masquées par des blancs.

Les CARSAT, qui délivrent notamment le minimum vieillesse, ont ouvert un portail cartographique, à partir d'un SIG, l'Observatoire régional des situations de fragilité. Un secret statistique est également appliqué aux petits effectifs.

Par contre les données des MSA, pour l'ensemble des prestations sociales servies aux assurés du régime agricole, restent peu accessibles pour le public. Les ORS, via des conventions entre la Fnors et les organismes prestataires, ont accès à des bases de données fournies par les prestataires à des niveaux moins agrégés.

L'accessibilité aux bases de données en environnement

Sous l'influence du mouvement de l'OpenData et de la Directive INSPIRE, les données environnementales affichent leur présence sur différentes plateformes de mises à disposition. Ces plateformes permettent d'organiser l'accès géolocalisé aux données, et constituent un outil essentiel pour la connaissance des territoires.

Les données n'y sont cependant pas toutes présentes, leur mise à disposition sur ces plateformes étant en cours de cadrage ou de mise en place. Elles sont parfois disponibles sous des formats ou à des niveaux d'agrégation qui ne permettent pas une exploitation fine, à l'échelle territoriale ou temporelle pertinente pour en tirer des informations intéressantes. De ce fait, il sera souvent nécessaire de contacter individuellement et directement les producteurs de données, engendrant des temps de collecte importants.

Par ailleurs, le manque de coordination et d'harmonisation des divers systèmes producteurs de données ne facilite pas les croisements nécessaires à l'identification de points de vigilance sur les territoires : en effet, quasiment aucun standard n'existe, rendant ainsi compliqué l'interopérabilité et la compatibilité des bases. Ceci est avant tout dû au fait que de nombreuses bases ont été créées au départ à des fins de gestion administrative.

A noter malgré tout que les données de surveillance de la qualité de l'air détenues en France par les AASQA (Air Rhône-Alpes et ATMO Auvergne en Auvergne-Rhône-Alpes) sont en cours de standardisation, sous un format délivré par la COVADIS.

De même, en ce qui concerne les données sur les eaux souterraines et de surface, la base gérée par EauFrance est ouverte au public. Très riche, il faut cependant faire une recherche assez fouillée et pas forcément intuitive pour arriver à trouver les données pertinentes recherchées.

Pour les eaux destinées à la consommation humaine, les données ne sont accessibles que sur demande auprès des Agences régionales de santé concernée. La gestion par unité de distribution (UDI) ne rend cependant pas la tâche facile à l'échelle de la commune, particulièrement pour celle desservie par plusieurs UDI.

Concernant l'état des sols, il est regrettable qu'en Rhône-Alpes les données de BASIAS ne soient toujours pas accessibles au public comme c'est le cas dans d'autres régions.

Pour BASOL, les données sont très riches, mais peu accessibles de manière automatique sur le site internet. Il faut faire une demande spécifique au gestionnaire afin d'avoir une extraction des fichiers pour mener des analyses intéressantes et fouillées.

On peut citer quelques-unes de ces plateformes :

- GéoRhôneAlpes (<http://www.georhonealpes.fr>) ou le CRAIG pour l'Auvergne (<http://www.craig.fr>)
- EIDER (<http://www.stats.environnement.developpement-durable.gouv.fr/Eider/>)
- Le portail EauFrance (<http://www.eaufrance.fr/>)

Une limite quant à la diffusion de données sensibles a pu être rencontrée, les producteurs de données avançant notamment le risque de stigmatisation de territoires ou d'effets sur la valeur du foncier.

3.2 Niveau géographique des données

Le niveau géographique auquel les données sont disponibles dans chaque base de données conditionne en partie le choix du zonage qui sera retenu pour le traitement et l'analyse des données. Plus les données seront disponibles à un niveau géographique fin (la donnée géocodée par exemple ou la donnée estimée sur de très petits zonages) et plus des agrégations seront réalisables en fonction de territoires variés, notamment les territoires opérationnels en termes d'actions de santé publique.

Pour une étude sur les inégalités territoriales, c'est a priori la maille territoriale la plus fine possible, dans les limites du respect de l'anonymat des personnes, qui permettra de révéler les inégalités les plus proches de la réalité.

Dans les bases de données en santé

Si les données enregistrées (activité de l'hôpital ou du professionnel de santé par exemple) sont disponibles à des niveaux géographiques fins, ce n'est pas le cas des données domiciliées (à la résidence du patient), diffusées sur les portails et plateformes

de données à des niveaux beaucoup plus larges, souvent le département, pour des raisons de confidentialité des données.

Il faut toutefois distinguer le niveau de diffusion publique des données, du niveau réel de recueil des données. Ainsi, par exemple, les données de la base sur les causes médicales de décès sont établies au niveau de la commune et sont diffusées sur le site dédié au niveau du département (et seulement au niveau communal pour les dix plus grandes villes de France).

Le niveau de recueil des données est bien sûr le niveau le plus fin auquel il sera possible de traiter les données (dans la mesure des autorisations d'accès à la base).

Le niveau le plus fin observé est l'adresse de résidence de la personne, cette adresse est recueillie dans les bases de l'assurance maladie au niveau local, dans les Registres (cancers, malformations congénitales), on encore dans les CS8, certificats de santé de l'enfant à la naissance. L'adresse est également recueillie pour les naissances (adresses des parents) enregistrées par le service d'état civil, mais n'est disponible que via les mairies. C'est seulement ce niveau de recueil qui peut permettre le géocodage des adresses (si les adresses sont correctement saisies) et la reconstitution de zonages infra-communaux, comme les quartiers IRIS ou encore les quartiers prioritaires pour la politique de la ville. C'est par exemple le cas des données de recours aux soins ou de consommations médicamenteuses qui sont exploitées par l'ORS Rhône-Alpes grâce aux requêtes réalisées par l'ARS sur la base ERASME du régime général.

Au niveau de la commune, sont recueillies les données sur les décès (état civil et causes médicales de décès), ainsi que d'autres données notamment sur les passages aux urgences (OSCOUR).

Au niveau du code postal (ou code approchant, de type « code PMSI »), sont recueillies l'ensemble des données hospitalières, et parfois les données sur les maladies à déclaration obligatoire (sinon, niveau départemental). Le niveau du code postal pose des problèmes d'exploitation puisque les données doivent être estimées si l'on souhaite présenter des données au niveau communal, cantonal, territoire de santé ou autre découpage nécessitant une donnée communale. Les données sur les ALD (affections de longue durée) posent également ce problème puisqu'elles étaient jusqu'à présent diffusées aux ORS au niveau du code postal (évolution actuelle vers le niveau communal, variable selon les régimes d'assurance maladie).

Les données d'activité hospitalière ou d'activité libérale des professionnels de santé (actes, soins, prescriptions médicamenteuses, etc.) sont enregistrées à l'adresse de l'établissement de soins ou du cabinet médical, comme les données de santé scolaire sont enregistrées au niveau de l'établissement scolaire.

Dans les bases de données sociales

Les bases de données démographiques et sociales sont construites pour la plupart à partir des recensements de population, des prestations sociales versées, des inscriptions à Pôle Emploi, des services fiscaux... Les producteurs de ces bases de données ont tous

connaissance de l'adresse des personnes concernées et cette information figure dans l'ensemble des bases de données. Elle reste néanmoins non diffusable, et son exploitation dépend des organismes producteurs.

La plupart des données sociales sont diffusées au niveau communal, sous réserve de confidentialité, voire à des niveaux infra-communaux (site de l'INSEE : <http://www.insee.fr/fr/bases-de-donnees/default.asp?page=infracommunal.htm>) Ainsi un certain nombre de données sociales sont accessibles au niveau du quartier IRIS : notamment des données du recensement de la population (sur la population active, le logement, les ménages, les diplômés...), des données sur les prestations sociales versées par les CAF.

D'autres données sociales ont été mises en ligne récemment sur le site de l'INSEE (<http://www.insee.fr/fr/bases-de-donnees/default.asp?page=quartiers-politique-ville.htm>), au niveau des quartiers prioritaires de la politique de la ville, quartiers définis par la précarité financière de leurs habitants (à partir du revenu médian par unité de consommation). Cette rubrique propose des indicateurs de source fiscale, des données sur les bénéficiaires des prestations légales versées par les CAF, des données sur les bénéficiaires de la couverture maladie universelle complémentaire (CMUC), des données sur les demandeurs d'emploi, et seront prochainement mises en ligne des données sur les revenus disponibles.

Les données sur l'emploi sont communiquées à d'autres niveaux géographiques, tels que les zones d'emploi.

Les données sur le logement sont communiquées par la DREAL (accès restreint) au niveau communal.

A noter que certaines données, par exemple le revenu médian, ne peuvent être recalculées à d'autres niveaux géographiques que les niveaux proposés par le producteur de données.

Dans les bases de données en environnement

L'échelle la plus fréquente est la commune, mais la plupart des données sont une extrapolation des résultats de mesures issues des stations situées à des endroits stratégiques en fonction des milieux et des objectifs de surveillance.

Les représentations cartographiques sont issues de modélisations, comme les cadastres d'Air Rhône-Alpes ou les données sur le bruit. Celles disponibles à travers les sites internet sont pour la plupart à l'échelle de la région et il est complexe de faire une extraction à la commune ou sur un territoire défini sans faire appel aux producteurs. Elles permettent toutefois d'avoir des mailles au kilomètre, parfois plus finement même (10 m).

Les données issues de BASOL permettent de géolocaliser précisément les sites et sols pollués par des coordonnées GPS, de même que la base de données sur les installations classées.

Enfin, les données sur les radiofréquences ne sont pas homogènes ni représentatives de l'ensemble de la région car elles sont issues de demandes ponctuelles effectuées par des particuliers.

3.3 Avantages et limites des bases

Si l'on souhaite réaliser dans un futur relativement proche une étude sur les inégalités territoriales, environnementales et sociales de santé, quelles bases de données est-il possible de mobiliser, compte-tenu des contraintes à la fois d'accès aux bases, du niveau géographique disponible, et des limites liées aux bases elles-mêmes ?

En santé

Plusieurs des bases présentées ne couvrent pas l'ensemble du territoire de la région Rhône-Alpes, et a fortiori de la grande région Auvergne-Rhône-Alpes. C'est le cas des Registres, Registre des cancers de l'Isère et Registre des malformations congénitales (4 départements). Ces données peuvent toutefois être mobilisées, après accord de la CNIL, pour des études limitées à ces départements.

Si l'on s'intéresse à l'ensemble de la région, il serait par ailleurs difficile de mobiliser rapidement les données des Conseils départementaux (certificats de santé de l'enfant) et de la santé scolaire (bilans de santé réalisés en milieu scolaire) à des niveaux plus fins que le département. Les logiciels de recueil et de traitement des données ne sont pas homogènes d'un département à l'autre et les données ne sont pas toujours systématiquement saisies faute de moyens. L'utilisation, à des niveaux géographiques fins, de certaines données des structures de dépistage organisé des cancers (cancer du sein et cancer colorectal) requiert également une collecte auprès de l'ensemble des structures locales, voire des autorisations spécifiques à demander. En matière de santé au travail, les recueils d'information sont également dispersés, non homogènes sur l'ensemble du territoire, et les données non mobilisables à un niveau infra départemental.

Les grandes bases nationales de santé (mortalité, hospitalisation, assurance maladie) sont quant à elles tout à fait mobilisables, dans les limites actuelles d'accès aux données, et dans les limites liées à la définition-même de chaque base.

Les données d'état civil et de mortalité par causes sont ainsi disponibles à la commune (parfois de façon payante selon le niveau de détail de la donnée), et permettent de calculer des indicateurs généraux sur l'état de santé de la population, comme l'espérance de vie (selon le sexe et à différents âges), les taux standardisés de mortalité générale ou de mortalité prématurée (avant 65 ans). Si une approche populationnelle est retenue, des indicateurs plus spécifiques, comme le taux de mortalité infantile, l'indice conjoncturel de fécondité, ou des taux de mortalité par classe d'âges, peuvent être calculés, toujours au niveau communal. La limite des données de mortalité tient

essentiellement à la taille des communes, puisque les effectifs de décès par cause de décès sont trop faibles dans les petites communes pour être diffusés (respect de l'anonymat) et que des calculs de taux sur de faibles effectifs n'auraient pas de sens faute de stabilité. Une solution pour pallier ces problèmes de petits effectifs est de rassembler plusieurs années de données pour assurer davantage de stabilité aux données.

Les données d'hospitalisation, particulièrement celles du PMSI MCO, renseignent sur tous les motifs d'hospitalisation, et il est possible de connaître à la fois le nombre de séjours et le nombre de patients hospitalisés dans l'année, parmi les habitants d'un territoire, quel que soit le lieu de leur hospitalisation. Actuellement les ORS et les ARS ont accès aux données du PMSI au niveau du code postal, niveau le plus fin possible dans cette base. Ce niveau géographique de disponibilité présente l'inconvénient de conduire à estimer les données des niveaux communal, cantonal, ou autres territoires.

Selon les pathologies étudiées, le taux de patients hospitalisés dans l'année pour une pathologie peut être assez éloigné du taux de personnes souffrant de cette même pathologie au sein de la population, c'est le cas de pathologies se soignant plutôt en ambulatoire. Les données d'hospitalisation sont alors davantage utilisées en complément de données provenant d'autres sources. Par ailleurs, une limite à l'interprétation des données d'hospitalisation en termes d'état de santé tient au fait que le recours à l'hospitalisation peut varier également en fonction des pratiques des médecins ou des patients (par exemple en raison de l'éloignement de l'hôpital, ou des caractéristiques socio-culturelles).

Les données d'ALD (Affections de longue durée), qu'il s'agisse de nouvelles admissions ou de la totalité des admissions, ne peuvent être considérées comme le reflet d'une véritable incidence ou d'une véritable prévalence d'une pathologie, du fait qu'elles sont liées à une démarche volontaire du patient et à un avis favorable du médecin conseil de l'assurance maladie. Par ailleurs, la réglementation sur les ALD évolue au fil du temps, certaines pathologies sont retirées de la liste des ALD (comme l'hypertension artérielle) et de nouvelles peuvent apparaître, ce qui rend difficile le suivi d'un indicateur sur plusieurs années. Néanmoins ces données peuvent être utilisées en complément d'autres indicateurs, particulièrement pour des pathologies de longue durée comme le diabète, en sachant que le nombre de personnes en ALD pour diabète sera a priori inférieur au nombre de personnes diabétiques. Les données d'ALD (des trois principaux régimes d'assurance maladie : général, agricole, indépendant) sont disponibles actuellement au code postal, avec les limites déjà indiquées, et commencent à être disponibles à la commune pour certains régimes. En termes d'indicateurs de santé, sont parfois retenus comme indicateurs généraux reflétant un état de santé global de la population soit le taux standardisé de nouvelles admissions en ALD toutes causes confondues (incidence), soit le taux standardisé d'ALD toutes causes confondues à un moment donné (prévalence). L'analyse de l'incidence est plus pertinente si l'on s'intéresse à l'évolution d'une pathologie sur plusieurs années (avec les limites déjà indiquées), alors que l'analyse de la prévalence correspond davantage à une photographie de l'état de santé de la population.

Les données de remboursement de l'assurance maladie sont les données disponibles au niveau le plus fin possible (à l'adresse de l'assuré), mais seulement pour le régime général actuellement compte-tenu des accès autorisés. En Rhône-Alpes, les requêtes sur la base ERASME sont réalisées par l'ARS, mais pas encore sur l'Auvergne (accès à mettre en place). Le service statistique de l'ARS procède au géocodage des adresses des assurés, ce qui permet de traiter les données au niveau infra-communal, quartiers IRIS et quartier en politique de la ville (les taux standardisés sont mis à disposition via le site Balises). La consommation de soins et de médicaments (remboursés) des assurés et ayants-droits du régime général permet d'analyser non seulement l'état de santé de cette population (la prévalence du diabète estimée par la consommation d'antidiabétiques par exemple), mais également son recours aux professionnels de santé libéraux ou encore sa participation à certains dispositifs de dépistage mis en œuvre avec l'assurance maladie. Le recours aux professionnels de santé (au moins une fois dans l'année) doit toutefois être analysé avec prudence, car il est à la fois révélateur d'un état de santé (maladie) mais aussi de pratiques individuelles (recours lié en partie à des habitudes socio-culturelles ou à des contraintes financières), ou encore de l'offre de soins présente sur le territoire (report possible de la médecine de ville vers les établissements de soins publics ou privés, ou du médecin spécialiste vers le médecin généraliste). En Rhône-Alpes, sont actuellement régulièrement exploités les recours aux médecins généralistes, pédiatres, gynécologues, ophtalmologues, psychiatres, chirurgiens-dentistes, orthodontistes, sages-femmes, infirmiers, kinésithérapeutes, orthophonistes.

La consommation de médicaments (au moins trois remboursements dans l'année) peut constituer un bon reflet de la prévalence d'une pathologie dans la mesure où un groupe de médicaments correspond bien à une pathologie ou à un ensemble de pathologies, et dans la mesure où l'automédication est peu pratiquée pour cette pathologie. L'analyse de l'évolution des indicateurs de ce type doit également prendre en compte l'évolution de la réglementation en matière de remboursements par l'assurance maladie (déremboursements, nouveaux remboursements, etc.) et l'arrivée de nouveaux médicaments ou nouvelles thérapeutiques sur le marché. Actuellement, en Rhône-Alpes, sont suivies régulièrement les consommations d'antidiabétiques, de psychotropes (dont antidépresseurs et neuroleptiques), d'antihypertenseurs, d'antiasthmatiques et antiallergiques, ainsi que celle de médicaments contre la maladie d'Alzheimer.

La base ERASME contient également des informations intéressantes sur les consommations de soins des assurés bénéficiaires de la CMUc (Couverture Maladie Universelle complémentaire), donc d'assurés en situation de précarité financière. C'est la seule base de données en santé (avec celles des autres régimes d'assurance maladie) qui comporte des informations d'ordre socioéconomique, et permet donc une connaissance individuelle de la santé des personnes en situation de précarité.

Certaines bases de données en santé décrites dans le cadre de ce travail, actuellement peu utilisées en observation de la santé car ayant d'autres fonctions dans un premier temps ou étant en période de montée en charge du recueil des données, pourraient devenir prochainement des sources d'informations régulières sur l'état de santé de la population.

La base OSCOUR, sur les passages aux urgences, comme le dispositif SURSAUD dont elle fait partie, ont ainsi pour mission une fonction de surveillance et d'alerte sanitaire en cas d'évènement inhabituel (épidémie ou canicule par exemple). Les motifs de recours aux urgences et les diagnostics posés sont codés selon des classifications et permettent une analyse des pathologies rencontrées. En région, les CIRE peuvent exploiter ces données, comme par exemple le recours aux urgences pour tentatives de suicide. Des exploitations sur d'autres thématiques (syndromes grippaux, problèmes liés à la canicule, etc.) sont possibles.

En matière d'offre de soins, les données sur les établissements de soins et sur les professionnels de santé (salariés et libéraux) sont largement diffusées, à des niveaux géographiques fins. Il est ainsi relativement aisé de mettre en évidence les inégalités territoriales en matière d'offre de soins. La localisation des offreurs de soins à l'adresse est possible, ce qui permet de calculer des distances entre une commune et un établissement ou un professionnel de santé. L'accessibilité des professionnels de santé libéraux est appréhendée également par un indicateur calculé pour l'ensemble de la France, l'APL (accessibilité potentielle localisée), qui tient compte du niveau d'activité du professionnel, ainsi que du taux de recours différencié par âge des habitants.

Quant aux données d'activité des établissements, comme des professionnels, bien qu'elles soient connues globalement (nombre de séjours, nombre d'actes...), elles ne sont pas toujours faciles à mettre en lien avec des pathologies spécifiques. De plus, les activités des hôpitaux recensées dans le cadre de la SAE ne sont pas des données domiciliées (à la résidence du patient) mais des données enregistrées. Des liens se mettent progressivement en place entre la SAE (données enregistrées des établissements) et le PMSI (données enregistrées mais également domiciliées).

En social

En pratique la quasi-totalité des bases décrites dans le domaine du social sont mobilisables pour une étude sur les inégalités territoriales, environnementales et sociales de santé, du fait de leur accessibilité sur l'ensemble de la région (soit par diffusion sur des sites internet en accès libre, soit via la FNORS, la DREAL ou l'ARS) et de leur niveau géographique relativement fin, le niveau communal voire infra communal (quartier IRIS).

Les bases disponibles permettent ainsi d'appréhender :

- La structure d'âge de la population (part des jeunes de moins de 20 ans, et part des personnes âgées de plus de 65 ans), via le Recensement général de la population (RGP)
- La composition des ménages (part de familles monoparentales, nombre d'enfants par ménage, art de personnes âgées vivant seules, etc.), via le RGP et les CAF/MSA
- Le niveau de formation (part des personnes sorties du système scolaire sans diplôme, ou selon le niveau du diplôme), via le RGP

- La catégorie socio-professionnelle : via le RGP
- L'emploi (part de la population active parmi les 15-64 ans, population active de 15 ans et plus selon les conditions d'emploi, taux de chômage, part de chômeurs de longue durée parmi les chômeurs, nombre de demandeurs d'emploi, etc.), via le RGP et Pôle Emploi
- Le revenu (part des ménages fiscaux imposés, revenu disponible médian par unité de consommation, structure du revenu et part des ménages dont les revenus sont principalement dépendants des prestations sociales, taux de pauvreté selon l'âge, etc.), via les sources fiscales et l'INSEE
- Les prestations sociales versées (part des bénéficiaires et de la population couverte par le RSA, part des retraités bénéficiaires du minimum vieillesse, etc.), via les CAF/MSA
- Le bénéfice de la CMUc (taux comparatifs des assurés et ayants-droit du régime général bénéficiant de la CMUc), via le régime général de l'assurance maladie et l'ARS
- Le logement (nombre de pièces du logement selon la taille du ménage, statut d'occupation du logement, nombre de logements sociaux, nombre de logements potentiellement indignes, etc.), via le RGP et la DREAL.

Quelques limites ou manques sont à signaler. La plupart des bases de données couvrent l'ensemble de la population, mais les données issues de l'assurance maladie sont élaborées par régime, et la disponibilité de certaines données peut être différente entre les différents régimes. Ainsi seules les données du régime général sont disponibles à un niveau géographique fin en ce qui concerne les bénéficiaires de la CMUc.

L'apparente ressemblance de certains indicateurs issus de bases différentes doit faire l'objet d'attentions particulières. Du fait de modes de collecte différents, ces indicateurs sont parfois assez éloignés les uns des autres. Ainsi la part des familles monoparentales n'a pas la même signification dans le RGP, dans les bases CAF ou encore dans les bases fiscales. De même le taux de chômage déclaré dans le RGP est différent du taux de chômage calculé à partir de différentes sources telles que Pôle emploi ou l'enquête Emploi.

Il est donc important de prendre en compte le mode de recueil des données. Certaines données sont déclaratives, comme celles du RGP, alors que d'autres sont issues d'enregistrements ou de remontées systématiques d'informations administratives. Certaines données sont issues d'estimations, comme le taux de chômage estimé par l'INSEE, ou comme les données du PPPI (Parc privé potentiellement indigne).

En environnement

Les bases de données environnementales ont été construites avant tout pour répondre à des besoins spécifiques tels que la surveillance de la qualité de l'environnement, l'application de la réglementation, la gestion des alertes, etc. Ainsi, l'utilisation de ces bases à des fins de caractérisation des inégalités environnementales confronte à des défauts liés à un manque de données scalaires, à des imprécisions, voire à leur

inexistence sur notre territoire ou enfin à un manque de finesse dans l'échelle disponible. De plus, les bases sont gérées par des gestionnaires et leur accessibilité n'est pas toujours identique. Ainsi, leur acquisition dépend de modalités d'accès, de moyens humains et parfois d'aspects financiers, voire juridiques. Ces critères font donc varier le temps nécessaire à leur obtention. Une analyse de la pertinence des indicateurs à inclure dans l'étude en fonction des objectifs fixés est donc nécessaire avant d'effectuer la requête de données.

Enfin, l'interopérabilité des données environnementales au niveau informatique fait défaut et rend la caractérisation spatiale des inégalités environnementales complexe.

3.4 Discussion sur les méthodes

Les études sur les inégalités territoriales environnementales et sociales de santé visent à confronter statistiquement une description de l'état sanitaire d'une population à différents indicateurs environnementaux ou sociodémographiques. Des transformations sont effectuées pour construire les variables à introduire dans les modèles statistiques et permettre le croisement de données. Ces transformations consistent principalement à homogénéiser, agréger ou désagréger les données afin d'obtenir des valeurs assignées aux unités géographiques.

L'évaluation des inégalités sociales et environnementales de santé nécessite un grand nombre de données très variables dont l'utilisation va dépendre avant tout des objectifs fixés et dont les critères essentiels sont la disponibilité de données harmonisées et leur représentativité géographique et temporelle.

Les systèmes d'information géographique (SIG) sont des outils de croisement de données utilisés dans les études avec une approche de type « spatial » qui consiste à privilégier l'étude de la manifestation des phénomènes dans l'espace. Les SIG permettent principalement la superposition de plusieurs couches de données spatialisées, mais aussi la transformation des données permettant le croisement à des unités spatiales variées.

On peut citer trois méthodes mobilisées dans les études sur les inégalités territoriales environnementales et sociales de santé :

Méthodes bayésiennes :

Les méthodes bayésiennes dans les études de corrélation écologique intègrent des données modélisées ou recueillies sur des unités géographiques dans le but de mesurer les associations entre des variables. Pour exemple, dans le projet de recherche Equit'Area, la méthode utilisée a permis d'analyser l'association entre l'indice de défaveur et la mortalité infantile, entre l'indicateur modélisé des concentrations en NO₂ à l'échelle de l'IRIS et l'indice de défaveur.

Méthodes géostatistiques :

Les méthodes géostatistiques intégrées à un SIG permettent un traitement statistique de données spatialisées dans le but d'identifier une ou plusieurs structures spatiales. Dans le projet CIRCE par exemple, le traitement géostatistique a permis de réaliser des cartes d'indicateurs sanitaires, d'explorer leurs corrélations avec les indicateurs environnementaux et de défaveur sociale et de détecter des zones où le risque de mortalité et le taux d'exposition sont les plus élevés ou les plus faibles

Méthodes d'analyse des données

Les méthodes d'analyse des données sont utilisées pour construire des typologies de territoires dans le but d'analyser les inégalités environnementales et sociales en santé. Par exemple, dans l'étude sur les inégalités environnementales de santé des territoires en région PACA, deux types d'outils (ACP et CAH) ont été utilisés pour décrire les inégalités environnementales en région PACA à l'échelle de l'Espace de Santé de Proximité (ESP).

- **L'analyse en composantes principales (ACP)** est une méthode purement descriptive. Elle permet d'explorer un ensemble de données observées sur un territoire, et de résumer une quantité importante d'information sur ce territoire (variables de description) en une information de dimension plus réduite et plus facilement interprétable (axes factoriels). L'ACP permet de voir s'il existe une structure particulière sous-jacente aux données.
- **La classification ascendante hiérarchique (CAH)** est une technique statistique visant à partitionner les territoires en différentes classes ou sous-groupes. On cherche à ce que les territoires regroupés au sein d'une même classe soient le plus semblables possibles (homogénéité intra-classe), tandis que les classes de territoires soient le plus dissemblables (hétérogénéité interclasse). La caractérisation des classes à partir des variables de description permet de décrire les inégalités territoriales environnementales et sociales en santé.

3.5 Discussion sur les indices composites

Une réflexion est nécessaire sur la capacité à synthétiser divers indicateurs sanitaires, sociaux et environnementaux à une échelle assez fine.

De nombreuses tentatives existent, et ces indicateurs diffèrent quant à leur mode de construction. On peut distinguer deux types d'indicateurs : les indicateurs calculés à partir d'un processus d'agrégation additif des variables retenues (processus de « scoring »), d'une part, et les indicateurs construits à partir d'une analyse statistique multidimensionnelle des données, d'autre part. Les premiers prennent en compte un nombre limité de variables, agrégées sur la base d'un système de pondérations défini a priori (par exemple, l'indice de Townsend). Les seconds considèrent un nombre plus important de variables et le système de pondérations étant déterminé, de manière endogène, par calcul sur la base de critères statistiques (par exemple, l'IDH). La qualité métrologique de ces indicateurs doit être testée

car tous les indices sensés mesurer un même concept ne sont pas adéquats pour toutes les régions. Même pertinents pour le territoire pour lequel ils ont été créés, de nombreux indices ne peuvent être transposés à d'autres territoires et nécessitent une validation.

En premier lieu, de nombreux indices sont conçus pour des territoires urbains, où il est à supposer que les indicateurs sanitaires, sociaux et environnementaux s'expriment de manière différente, les modes de vie et les pressions urbains et ruraux n'étant pas les mêmes.

De nombreux indices de défavorisation (ou défaveur) sociale ont vu le jour. Leur principal avantage est de pouvoir disposer d'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 et pouvant être opérationnelle rapidement en France, pour la quasi-totalité des personnes.

Mais les limites sont nombreuses. Par exemple, comme ils sont calculés au niveau géographique pour caractériser ensuite la situation d'une personne, plus la zone géographique est petite et homogène, moins le biais est important. De plus, ils prennent en compte le contexte socioéconomique mais pas 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 qu'en milieu périurbain ou rural.

D'ailleurs, les indices construits à l'échelle nationale ne sont pas toujours adaptés à un contexte local. Souvent, de nouveaux calculs sont nécessaires en prenant comme référence l'ensemble des territoires concernés (souvent la région), et non plus la France. Cela permet de mettre en exergue les spécificités territoriales fines et d'avoir une comparaison entre territoires de proximité plus adéquate.

De plus, la mobilité des populations n'est jamais prise en compte. Enfin, l'accroissement continu du nombre d'années d'étude induit une homogénéisation du niveau de formation dans la population, d'une génération à l'autre. Ainsi, l'utilisation des indices composites construits à l'aide de l'indicateur sur le niveau d'éducation pour suivre l'évolution des inégalités sociales de santé sur de longues périodes doit donc être faite avec une grande attention.

Si les analyses portant sur la mesure du rôle du contexte de résidence dans les variations d'état de santé ont permis de valider l'existence d'un lien entre certaines situations locales défavorisées et la santé des populations, il reste que le poids des facteurs environnementaux dans la composante santé est une réalité difficile à appréhender du point de vue quantitatif, dépendante des hypothèses et des méthodes de calculs adoptées. En effet, certaines études, reprises encore aujourd'hui, estiment que l'environnement contribue à hauteur de 20 à 25 % à la longévité ; en France, l'ANSES note que l'OCDE s'appuie sur des travaux estimant ce poids entre 2 et 5 % du fardeau total des maladies pour un pays comme la France ; l'Institut national de santé publique et de l'environnement des Pays-Bas estime ce poids entre 3 et 9 % du total.

Pour la construction des indices, la méthode la plus employée est l'analyse en composantes principales (ACP).

L'analyse en composantes principales (ACP) est une méthode d'analyse multidimensionnelle ou multivariée qui permet de synthétiser un grand nombre d'informations. Il s'agit d'une

méthode descriptive qui ne permet pas l'explication ni la prédiction de variables. Elle résume simplement une quantité d'information importante en une information de dimension plus réduite et plus facilement appréhendable pour l'interprétation.

L'ACP traite un tableau de données contenant les valeurs de p variables quantitatives (colonnes) obtenues pour n observations (lignes). Les représentations graphiques extraites de l'ACP permettent de voir s'il existe une structure particulière sous-jacente aux données. Par exemple, il sera possible d'examiner s'il existe des groupes d'observations ou des groupes de variables qui se « ressemblent » ou qui se distinguent les uns des autres.

Afin de réaliser une ACP, plusieurs étapes sont nécessaires, comme l'examen préalable des données ainsi que le choix du nombre d'axes à conserver pour l'interprétation.

La plupart des méthodes de croisement des données ont pour premier objectif de construire des indices synthétiques dans le but de développer des outils et des méthodologies pour étudier les inégalités territoriales, environnementales et sociales de santé. Les indices de défaveur environnementale tentent ainsi d'étudier les surexpositions environnementales à des échelles spatialisées, jouant un rôle indéniable dans les inégalités de santé. Mais des réflexions préalables sont nécessaires à la validation du nombre de variables et de composantes à retenir, en lien avec des événements de santé, et leur adaptabilité aux différents contextes (urbain, rural, , etc.).

4. Perspectives et conclusions

Rappelons que ce travail s'inscrit dans le cadre de la réduction des inégalités territoriales, environnementales et sociales de santé (ITESS), thématique largement valorisée dans les récents plans nationaux ou régionaux en matière de santé, et plus particulièrement en santé environnement. Un des objectifs du Plan National Santé Environnement 2015-2019 (PNSE3) vise à « Mieux caractériser les inégalités environnementales et territoriales de santé ».

Autant les inégalités sociales de santé ont été assez largement documentées depuis plus de trente ans, tant en termes d'observation que de recherche sur leurs déterminants, autant les inégalités environnementales de santé n'ont été abordées que de façon relativement récente par des équipes de recherche et par des politiques nationales de santé. L'approche multifactorielle de la santé (qui inclut bien dans les déterminants de la santé à la fois les conditions de vie, de travail, les conditions socio-économiques ou environnementales) n'est cependant pas récente, mais la complexité du sujet a certainement retardé la prise en compte des inégalités environnementales dans les différents programmes de santé. Ce retard est lié en partie à l'absence de connaissances scientifiques notamment sur l'impact sanitaire de multi-expositions environnementales sur l'individu, en partie au manque de disponibilité d'outils et de données permettant de croiser à une échelle territoriale fine des données sur l'environnement et la santé.

Le travail exploratoire a permis de recenser, dans la perspective d'une étude régionale de caractérisation des ITESS, à la fois les bases de données disponibles pour décrire les inégalités sanitaires, sociales et environnementales, les méthodes déjà mises en œuvre par des équipes nationales ou régionales dans des études de ce type, les indices composites existants pour mesurer la défaveur (sanitaire, sociale, environnementale) d'un territoire.

A l'issue de ce travail, plusieurs réflexions et recommandations sont présentées : d'une part des recommandations à moyen ou long terme afin de faciliter la mise en place d'études sur les ITESS et, d'autre part, des recommandations à court terme afin de permettre d'opérer des choix, dans le cadre d'une future étude régionale sur les ITESS, parmi l'ensemble des données et des méthodes décrites.

Des recommandations à moyen ou long terme

Plusieurs obstacles à l'utilisation des données pour des études descriptives sur les ITESS ont été mis en évidence tout au long de ce travail, qu'il s'agisse d'accessibilité aux données à un niveau géographique fin, mais également d'homogénéité des données, voire d'absence de données sur certains déterminants importants de la santé.

Les questions d'accessibilité aux grandes bases de données en santé ont été largement discutées lors de la préparation de la récente loi de modernisation du système de santé

(loi n° 2016-41 du 26 janvier 2016) et continuent à l'être dans le cadre de l'élaboration de décrets d'application de l'article 193 concernant « un accès ouvert aux données de santé ». Cet accès concerne les données de consommation de soins et de médicaments, les données hospitalières, les données de mortalité, ainsi que d'autres données d'ordre socio-sanitaire. De nombreux organismes, dont le Haut Conseil de Santé Publique, militent pour un accès facilité aux données pour les organismes de recherches et d'études, dans le respect de la protection des libertés individuelles. Les accès possibles, mais très complexes, pour des projets de recherche précis, ne conviennent pas pour la mise en place d'une surveillance pérenne des inégalités territoriales de santé. La mise en œuvre d'indicateurs de suivi des politiques locales de santé (PRS, PRSE, mais aussi CLS – Contrats locaux de santé) doit pouvoir s'appuyer sur des données disponibles en routine à des niveaux fins.

De même, en matière de santé environnementale, les équipes de recherche travaillant sur ce sujet constatent des lourdeurs administratives dans l'accès aux données (données non numérisées parfois ou non homogènes, délais importants de traitement des dossiers à la CNIL, etc.). Le fait que le PNSE3 comporte plusieurs actions visant à « mieux caractériser les inégalités environnementales et territoriales de santé », dont l'action n°41 « faciliter la collecte et l'utilisation des données produites aux niveaux local, régional et national », est plutôt prometteur, mais il montre par défaut que la collecte et l'utilisation des données sont actuellement compliquées, voire impossibles à certains niveaux. Comme pour les données de santé, il est primordial pour les acteurs locaux de disposer en routine d'un certain nombre d'indicateurs environnementaux.

Le Conseil Economique, social et environnemental, dans son avis du 23 janvier 2015⁷, estime donc « nécessaire d'engager au niveau national un travail visant à structurer, rationaliser et optimiser les réseaux de production, de collecte, d'intégration d'analyse, de partage et de diffusion des données environnementales et sanitaires ». Il propose notamment la création d'un observatoire intégré des inégalités environnementales.

A un niveau régional, dans les limites actuelles d'accès aux données, il apparaît important de réfléchir à la mise en place d'un « observatoire » ou d'une « plateforme régionale », permettant de mutualiser les données disponibles en santé, social, environnement et de rendre accessibles ces données au niveau le plus fin, dans les limites de la nécessaire confidentialité, d'une part dans un but d'études, d'autre part dans un but d'aide à la décision pour la mise en œuvre de politiques locales.

L'homogénéité des données est également nécessaire pour une utilisation de données sur l'ensemble du territoire régional. Au niveau national, des commissions réfléchissent à des standards de données, tant sur le plan de la santé que sur le plan de l'environnement.

⁷ Crosemarie Pierrette. Inégalités environnementales et sociales : identifier les urgences, créer des dynamiques. Avis du Conseil Economique, social et environnemental sur le rapport présenté par Mme Pierrette Crosemarie, rapporteure au nom de la section de l'environnement. Journal Officiel, 23/01/2015. Disponible : <http://www.lecese.fr/travaux-publies/inegalites-environnementales-et-sociales-identifier-les-urgences-creer-des-dynamiques>

Ces standards doivent permettre une meilleure **interopérabilité** entre les différentes bases de données, ils doivent être encouragés.

On observe également parfois des différences de recueil et de traitement des données, pour un même type de données. C'est le cas de certains recueils de données moins centralisés que les recueils nationaux de données, comme ceux, par exemple, des services de PMI (Conseils départementaux) sur la santé du petit enfant, des services de santé scolaire sur la santé de l'enfant et de l'adolescent, ou encore des structures locales de dépistage des cancers... Pour toutes ces structures, des remontées d'informations annuelles, sur quelques indicateurs et à un niveau agrégé (souvent le département) sont faites à un organisme national (DREES, Education nationale, InVS, etc.). Certaines avancées en matière d'homogénéité de recueil sont en cours, par exemple dans les services de santé scolaire avec le test d'un logiciel commun de saisie. Il serait très important par ailleurs que ces données homogènes soient recueillies (et saisies) au niveau le plus fin possible (adresse des parents ou de l'établissement scolaire).

Pour l'analyse de la santé de l'enfant, les données de PMI et de santé scolaire sont une source incontournable d'informations, notamment pour approcher les liens entre les conditions sociales, environnementales, et la santé de l'enfant. Il est souhaité que l'ensemble des acteurs puissent se mobiliser afin d'associer les Conseils départementaux et l'Education nationale dans les études régionales et dans la mise en œuvre d'une collecte homogène de données sur l'ensemble de la région.

L'absence de données mobilisables au niveau infrarégional sur certains déterminants de la santé a été soulignée.

Les comportements de santé individuels, croisés avec les caractéristiques socioéconomiques des personnes, sont particulièrement peu abordés dans les recueils systématiques de données, car leur recueil nécessite de mettre en œuvre une démarche lourde de type enquête avec échantillon représentatif, comme c'est le cas pour les Baromètres de santé de l'INPES.

Les expositions professionnelles et les conditions de travail individuelles sont également peu documentées, alors qu'elles interviennent certainement largement dans les inégalités de santé. Des enquêtes sont menées localement, basées sur le volontariat de médecins du travail, mais leur représentativité est difficile à établir en raison d'une insuffisance d'effectifs.

La mise en place d'enquêtes en population générale à un niveau géographique fin, que ce soit sur les comportements de santé, sur les conditions de travail ou encore sur les expositions environnementales, n'est pas envisageable pour des raisons financières, mais il est certainement possible, dans le cadre de questions qui se posent au niveau régional, d'encourager la collecte de données sur des thématiques particulières et auprès de populations ciblées. L'enquête rhônalpine sur la prévalence de l'allergie à

l'ambrosioise, réalisée en 2014 par l'ORS et IPSOS avec le soutien de l'ARS, témoigne de cette démarche.

Sur certaines questions d'inégalités sociales ou environnementales de santé, particulièrement importantes dans la région, les Plans régionaux (PRS, PRSE, PRST, etc.) pourraient contribuer à la mise en œuvre **d'enquêtes spécifiques en population**, qu'il s'agisse de soutenir des enquêtes existantes (par exemple en médecine du travail) ou d'enquêtes auprès de populations particulières sur des questions de comportements ou d'expositions environnementales.

Par ailleurs, dans la période actuelle d'évolution rapide de l'OpenData et de la réglementation en matière d'accès aux données, il est capital de rester en veille sur l'ouverture de nouvelles données tant dans le domaine de la santé, que dans celui de l'environnement.

Des recommandations à court terme pour une étude régionale sur les ITESS (inégalités territoriales, environnementales et sociales de santé)

Plusieurs réflexions doivent être abordées avant d'entreprendre une étude régionale sur les ITESS :

- Quels sont les objectifs de l'étude ? Qu'attend-on des résultats ? Qui utilisera les résultats et comment ?
- Au sein de la région, quel serait le niveau géographique d'analyse pertinent ?
- Quelles seraient les bases de données directement mobilisables ?
- Quels seraient les indices de défaveur (sociale, environnementale, sanitaire) adaptés à une étude dans la région ?
- Quelles seraient les méthodes de croisement ou de rapprochement de données reproductibles dans la région ?

Des pistes de réflexion et des recommandations seront formulées ci-dessous sur ces différents points, et proposées à la discussion aux membres du Comité de pilotage de cette étude exploratoire.

Les objectifs d'une étude sur les ITESS

Si la plupart des études décrites dans ce travail mettent en avant l'objectif de connaissance des ITESS, toutes n'ont pas le même objectif quant aux résultats attendus de l'étude.

Une étude de type EQUIT'AREA va conduire ses travaux afin de confirmer une hypothèse de travail (ex : l'existence d'inégalités sociales ou environnementales de santé au regard de la mortalité infantile et néonatale) et analyser très finement les clusters repérés avant de conclure à l'existence de liens entre les différents éléments étudiés. De même le projet PLAINE, dont l'objet est de construire des modèles et des outils d'analyse en

matière de multi-expositions, va produire des données qui vont ensuite être utilisées dans le cadre d'autres études ou par d'autres acteurs. Ces démarches contribuent à faire progresser les connaissances sur les liens entre les différents déterminants de la santé. Elles s'inscrivent dans le long terme, tant dans l'élaboration de leur méthode, avec des recueils de données et des modèles construits ad hoc, que dans leur mise en œuvre sur plusieurs années.

D'autres études, comme celle menée en région PACA, s'orientent davantage vers l'aide à la décision pour la mise en œuvre de politiques locales de santé, voire de politiques plus larges que la santé, politiques impliquant de nombreux partenaires susceptibles d'agir sur les déterminants de la santé. Ce type d'études, mettant en évidence sur l'ensemble d'un territoire l'existence d'inégalités de santé, doit permettre d'orienter l'action publique vers les zones les plus défavorisées en matière de santé, et d'agir ensuite sur les déterminants de la santé (sociaux, environnementaux, liés à l'offre de soins...) impactant le plus défavorablement la santé sur ces différentes zones. La mise en évidence des inégalités territoriales ne constitue ainsi qu'une première étape, puisqu'il faudra, dans un second temps, identifier, et cela avec la participation des habitants et des professionnels locaux, les actions à mettre en œuvre au niveau local pour réduire les inégalités observées. Cette phase d'identification et de priorisation des actions est nécessairement déterminée au niveau local puisque la mobilisation des acteurs de terrain est indispensable à la réussite des actions entreprises.

Les études de caractérisation des ITESS peuvent être également conçues comme des outils de suivi et de bilan de Plans nationaux ou régionaux dont l'objectif est la réduction des inégalités sociales et environnementales de santé. Ainsi le Haut Conseil de la Santé Publique a émis des recommandations⁸ en juin 2013 d'une part pour la prise en compte des caractéristiques sociales dans les indicateurs associés aux objectifs de la loi de santé publique du 9 août 2004, et d'autre part pour assurer le suivi des inégalités de santé dans le domaine du cancer (Plan cancer 2009-2013). Selon les objectifs du Plan et le domaine dans lequel il se déploie, l'étude mise en œuvre pour son suivi en matière d'évolution des inégalités de santé, mobilisera des indicateurs différents, spécifiques aux thématiques abordées par le Plan et aux déterminants de santé impliqués dans ces thématiques. Ainsi, dans le cadre d'un PRSE (Plan régional santé-environnement), il peut s'agir de suivre régulièrement, ou au début et en fin de Plan, l'évolution des ITESS au sein de la région.

Les objectifs d'une étude de caractérisation des ITESS doivent être clairement définis avant le démarrage de l'étude. Ces objectifs (qu'il s'agisse de contribution à la recherche, d'aide à la mise en œuvre de politiques et d'actions, de suivi d'un Plan prévoyant la réduction des inégalités de santé...) conditionnent notamment la méthode de mise en œuvre de l'étude, les données mobilisées, le niveau géographique d'analyse des données.

⁸ Haut Conseil de Santé Publique. Indicateurs de suivi de l'évolution des inégalités sociales de santé dans les systèmes d'information en santé. HCSP, Collection Documents, juin 2013. 54 p. Disponible : http://www.hcsp.fr/Explore.cgi/Telecharger?NomFichier=hcspr20130619_indicateurinegalitesocialesante.pdf

Le niveau géographique pertinent

Le niveau géographique de traitement des données va être étroitement dépendant des objectifs de l'étude.

S'il s'agit d'une étude à visée « recherche », la zone d'étude pourra être plus ou moins large (sur un ou plusieurs sites) et le niveau d'analyse sera déterminé en fonction d'une part des données disponibles (qui dépendent du recueil de données) et du niveau nécessaire pour pouvoir vérifier les hypothèses émises.

Une étude à visée d'observation, pour la prise en compte et le suivi de l'évolution des ITESS dans le cadre d'un Plan par exemple, sera réalisée sur l'ensemble du territoire concerné par le Plan, territoire national ou régional, au niveau le plus fin possible, en fonction de la disponibilité des données et de la signification de données (un niveau trop fin peut n'avoir aucun sens, notamment en matière d'évènements de santé un peu rares). C'est bien le niveau le plus fin possible qui mettra le plus en évidence les inégalités territoriales, mais là encore la maille d'observation peut être différente selon la thématique observée. Les découpages administratifs (commune, canton...), qui sont généralement les plus aisés à utiliser en matière de santé ou de social, ne sont pas toujours les plus pertinents en matière d'environnements, qu'il s'agisse d'environnement physique ou d'habitudes de vie.

En ce qui concerne les études à visée plus opérationnelle d'aide à la mise en place d'actions locales, le niveau pertinent d'analyse est nécessairement le niveau décisionnel, niveau auquel se décideront et se concrétiseront les actions de réduction des inégalités de santé. Ce niveau local pourra être celui de la commune, de la communauté de communes (EPCI), mais également celui du quartier, ou d'autres territoires plus spécifiques à certaines thématiques ou populations (par exemple la filière gérontologique en Rhône-Alpes ou l'unité territoriale des Conseils départementaux). L'échelon local est adapté à l'approche souvent intersectorielle à mettre en œuvre si l'on veut agir sur le plus grand nombre de déterminants de la santé ; les acteurs sollicités ne seront en effet pas seulement ceux intervenant dans le champ de la santé, mais également ceux de l'éducation, de l'emploi, du logement, des transports, de l'urbanisme, de la culture...

Le niveau géographique d'analyse des données dans le cadre d'une étude de caractérisation des ITESS doit être choisi non seulement en fonction du niveau de disponibilité des données, mais aussi en fonction des résultats attendus de l'étude et des actions éventuelles à mettre en place à la suite de l'étude.

Les bases de données mobilisables

Le recensement et la description des bases de données en santé, social, environnement, ont montré qu'un certain nombre de bases ne sont pas directement accessibles dans un délai court, soit pour des raisons de confidentialité (autorisations à demander, parfois à la CNIL), soit pour des raisons de manque d'homogénéité du recueil et du traitement des données.

Le mouvement de l'OpenData et l'évolution de la réglementation sur les accès aux bases de données peuvent toutefois faire évoluer assez rapidement l'accessibilité aux bases de données décrites dans le cadre de ce travail.

En matière de données de santé, dans le cadre d'une étude visant à caractériser les inégalités de santé dans leur ensemble, le choix des bases de données se portera plutôt sur les bases de mortalité (permettant de calculer des indicateurs généraux de santé, tels que l'espérance de vie à différents âges, la mortalité générale ou la mortalité prématurée) et des bases comme celle des ALD (Affections de longue durée) permettant d'approcher la prévalence (toutes causes) des maladies chroniques au sein de la population

Si l'étude sur les ITESS s'intéresse à des thématiques en particulier (cancer, environnement...) ou à des populations spécifiques (enfants, adultes, personnes âgées), le choix des bases de données et des indicateurs se fera en référence à ces thématiques ou à ces populations. Ainsi les causes médicales de décès, les motifs d'hospitalisation, les motifs d'admission en ALD, les consommations de médicaments (limitées pour l'instant au régime général et à Rhône-Alpes), pourront être analysés, en choisissant les pathologies en lien avec la thématique et selon les classes d'âges à considérer.

De même les données sociales pourront être analysées soit de façon globale soit sur certains critères en lien avec la thématique observée (problématique d'emploi, de logement, de couverture médicale...).

La mobilisation des bases de données environnementales prendra en compte de préférence les milieux sur lesquelles on pense pouvoir agir, soit au niveau national (évolution de la réglementation par exemple) soit au niveau local (aménagement des logements par exemple, information de la population sur les pics de pollution, mise en œuvre de mesures de protection de la population...).

La plupart des bases de données, donc des indicateurs issus de ces bases, présentent des limites qu'il faut connaître avant de tenter une interprétation des résultats, notamment le mode de recueil et de construction des données. Parmi ces limites, en matière d'indicateurs de santé, il ne faut pas négliger le fait que la plupart des données de santé sont issues de systèmes d'information sur le recours aux soins (hospitalisation, remboursement de soins et de médicaments, ALD...), ce qui laisse dans l'ombre le non-recours aux soins et l'automédication, des comportements qui interviennent pourtant dans la constitution des inégalités de santé. De même, les données sociales ne permettent pas d'approcher le non-recours aux prestations sociales (type RSA), ou l'absence de couverture médicale (non-recours à la CMUc), ce qui sous-estime la part de la population en situation de précarité.

Selon les objectifs plus ou moins généraux, ou très spécifiques, d'une étude sur les ITESS, les bases mobilisables seront globalement les mêmes en matière de santé (mortalité, hospitalisation, ALD, assurance maladie) mais les indicateurs permettront d'affiner l'analyse par pathologie ou par classe d'âges. Sur des thématiques ou des populations spécifiques, l'utilisation de bases plus difficiles d'accès peut être envisagée, mais des délais de collecte plus longs sont à prévoir. Quel que soit le domaine (santé, social, environnement), la connaissance fine des bases et de leurs limites est nécessaire pour guider le choix des indicateurs et faciliter ensuite l'analyse de ces derniers en termes d'aide à la décision et à l'action.

Les indices de défaveur (sanitaire, sociale, environnementale)

Bien que les indices de défaveur apparaissent comme intéressants d'un point de vue descriptif, il faut être prudent quant à leur capacité à représenter la situation réelle des territoires concernés.

Par exemple, certains indices environnementaux ont été construits en agrégeant différents indicateurs concernant plusieurs milieux environnementaux. Il pourrait être pertinent de donner plus de poids à certaines expositions, proportionnellement à leur impact connu sur la santé, dans le but d'avoir une approche plus réaliste du risque sanitaire auquel pourrait être exposée une population. Mais cette approche est particulièrement complexe à prendre en compte dans la mesure où cela implique d'évaluer l'impact de chaque exposition sur la santé et de déterminer les poids relatifs de chaque exposition en fonction de son impact.

Enfin, l'interprétation d'un indice composite peut s'avérer plus complexe que l'interprétation de variables couramment utilisées. L'indice est sans unité, ni valeur de référence intuitive, il conduit à avoir une interprétation relative du problème et rend difficile la définition des actions.

La sélection d'un indice composite doit faire l'objet d'une validation avant toute utilisation et il faut s'assurer de sa fiabilité sur le territoire concerné, en particulier sur les différents types de territoires (urbains, périurbains, ruraux). Les indicateurs utilisés pour les construire doivent également faire l'objet d'une étude approfondie quant à leur pertinence.

Pour les indices environnementaux, il faut être vigilant dans la construction d'un indice agrégeant différents milieux (air et eau par exemple) en fonction des objectifs opérationnels.

Les méthodes pour une étude sur les ITESS

De la même façon que le choix du niveau géographique (ou des indicateurs) sera déterminé par les objectifs de l'étude, le choix de la ou des méthodes sera étroitement dépendant des objectifs de l'étude et de ses résultats attendus.

Dans les études sur les ITESS, il s'agit de confronter statistiquement une description de l'état sanitaire d'une population à différents indicateurs environnementaux ou sociodémographiques. Les indicateurs retenus sont susceptibles d'être nombreux, et des transformations seront alors nécessaires pour construire les variables synthétiques à introduire dans les modèles statistiques et permettre le croisement de données. Ces transformations consistent principalement à homogénéiser, agréger ou désagréger les données afin d'obtenir des valeurs assignées aux unités géographiques.

Plusieurs méthodes ont été utilisées dans les études décrites : les méthodes bayésiennes, les méthodes géostatistiques, et les méthodes d'analyse des données.

Dans le cadre d'une étude régionale, à visée opérationnelle et ne se déroulant pas sur une période très longue, il serait difficile de mettre en œuvre des méthodes trop coûteuses, à la fois en équipement (logiciel par exemple) et en temps de travail (élaboration et développement de la méthode, tests, calculs, interprétation des résultats...).

Les deux premières méthodes (bayésiennes et géostatistiques) sont davantage utilisées dans des études à visée de recherche (projet de recherche Equit'Area ou Circe par exemple).

Les méthodes d'analyse des données (analyse en composantes principales et classification ascendante hiérarchique) sont les plus utilisées à la fois pour construire des indices synthétisant un grand nombre de données et pour répartir les zones étudiées en classes homogènes de territoires. Ces méthodes permettent notamment de visualiser les inégalités territoriales de santé.

Dans le cadre d'une étude régionale sur les ITESS, la mise œuvre de méthodes d'analyse de données (ACP et CAH) semble correspondre le plus à la fois aux objectifs de l'étude et aux impératifs de coût et de simplicité. Néanmoins, afin d'obtenir des résultats adaptés à l'étude, l'utilisation de ces méthodes doit s'appuyer sur une sélection d'indicateurs et un niveau géographique les plus pertinents possibles.

En conclusion, plusieurs éléments ont été mis en évidence dans cette étude exploratoire, qui constituent des points de vigilance pour la mise en œuvre d'une étude régionale sur les ITESS.

L'accessibilité effective aux données reste aujourd'hui un problème majeur, qu'il s'agisse d'identifier les bases de données existantes, d'extraire les données ou encore de mettre les données en forme pour les analyses. L'extraction proprement dite des données nécessite souvent une participation active des organismes gestionnaires des bases de données, car cette activité nécessite une connaissance fine des bases, de leur mode d'interrogation, et parfois des autorisations d'accès difficiles ou longues à obtenir. Une collaboration forte des producteurs de données, aux études mises en place, est donc fortement souhaitée.

Le niveau géographique auquel la donnée est collectée est également primordial dans les tentatives de rapprochement de différentes bases de données. Le géocodage

systematique des adresses (de residence par exemple) permettrait une utilisation des donnees a des niveaux adaptes a chaque etude, en s'affranchissant des decoupages administratifs (decoupages dont les contours evoluent par ailleurs). C'est notamment une recommandation du Haut Conseil de la sante publique, soutenue par l'ensemble des utilisateurs des bases de donnees publiques.

Si une analyse de l'evolution des ITESS est prevue, sur une region par exemple, plusieurs facteurs sont a prendre en compte, dont la stabilite de l'indicateur analyse (toujours la meme signification au cours du temps), la stabilite du recueil d'information (toujours les memes procedures de comptage, de mesure, de modelisations, de remontees d'information...), la stabilite du decoupage geographique utilise. Le descriptif des bases de donnees a montre que de nombreuses bases de donnees evoluent au fil du temps, avec parfois des ruptures de serie importantes.

Les resultats attendus d'une etude sur les ITESS doivent toujours etre questionnes en termes d'eventuelles actions a mettre en oeuvre. Si des actions locales sont prevues, les donnees (et leur methode d'elaboration) doivent etre clairement explicitees aupres des acteurs locaux, afin d'assurer un partage de connaissances et une coproduction des actions correctives des inegalites territoriales de sante, qui seront necessairement intersectorielles tant les actions sur les determinants environnementaux et sociaux de la sante peuvent relever de domaines tres divers.

Références bibliographiques

Agence Régionale de Santé Rhône-Alpes, Région Rhône-Alpes, Direction régionale de l'environnement, de l'aménagement et du logement Rhône-Alpes. PRSE2. Santé environnement 2ème Plan régional 2011-2014 Rhône-Alpes. 2011. 148p. Disponible :

<http://www.prse2-rhonealpes.fr/>

Air Rhône-Alpes. Analyse des inégalités environnementales. Méthodologie d'élaboration d'un indicateur de qualité de l'air spatialisé. Rapport méthodologique, 2014. 23p. Disponible :

http://www.air-rhonealpes.fr/sites/ra/files/atoms/files/rapport_inegalites_environnementales_2014.pdf

ANDRIEU-SEMMELE M. Les inégalités environnementales de santé des territoires en région PACA (IEST – PACA). ARS PACA, Analyses recherches et statistiques, dossier n° 20, décembre 2014. 12p. Disponible :

http://www.ars.paca.sante.fr/fileadmin/PACA/Site_Ars_Paca/Sante_public/Sante_environnement/ARS_PACA_Dossier_20_IEST.pdf

ARS PACA, département observations et études. L'indice de désavantage social (IDS) : application aux territoires de Provence Alpes Côte d'Azur. 1^{ère} partie : raisons de l'étude, limites et principales conclusions. La santé observée, Info Stat, août 2012, n° 15. 6p. Disponible :

http://www.ars.paca.sante.fr/fileadmin/PACA/Site_Ars_Paca/Nos_missions/Etudes_et_publications/Infostats/Infostat_IDS_1_V8.pdf

ARS PACA, département observations et études. L'indice de désavantage social (IDS) : application aux territoires de Provence Alpes Côte d'Azur. 2^{ème} partie : résultats régionaux et comparaisons départementales. La santé observée, Info Stat, août 2012, n° 16. 6p. Disponible :

http://www.ars.paca.sante.fr/fileadmin/PACA/Site_Ars_Paca/Nos_missions/Etudes_et_publications/Infostats/Infostat_IDS_1_V8.pdf

BENMARHIA T. Les inégalités environnementales et leurs rôles dans les inégalités sociales de santé ; Etat de l'art et réflexions sur le développement d'indicateurs de défaveur environnementale. EHESP, mémoire d'ingénieur du génie sanitaire, 2010. 120p. Disponible :

<http://fulltext.bdsp.ehesp.fr/Ehesp/Memoires/igs/2010/benmarhnia.pdf>

BESSE G., CAUDEVILLE J., MICHELOT F., PRISSE N., TRUGEON A. Inégalités territoriales, environnementales et sociales de santé : Regards croisés en régions : de l'observation à l'action. Ministère des affaires sociales et de la santé, Ministère chargé de l'Ecologie, 2014. 72p. Disponible :

http://www.developpement-durable.gouv.fr/IMG/pdf/140221_Inegalites_territoriales_environnementales_sociales_de_sante.pdf

BOCQUIER A., et al. Bruit routier et statut socio-économique : étude des inégalités environnementales à un niveau géographique fin au sein de la commune de Marseille. Environnement Risques Santé, mai-juin 2011, 10, 3, 225-229

BOUDET C., HUBERT P., RAMEL M., CAUDEVILLE J. Les inégalités environnementales sur les territoires. INERIS références, exposition environnementale et vulnérabilité des populations, oct. 2014. 40p. Disponible : <http://www.ineris.fr/ressources/recherche/iddoc=2365>

CASAMATTA D., LE PERFF H. Le dispositif de surveillance sanitaire des urgences et des décès (SurSAUD) en région Rhône-Alpes. BVS Bulletin de veille sanitaire, décembre 2015, n° 7. 12p. Disponible :

http://www.ars.auvergne-rhone-alpes.sante.fr/fileadmin/RHONE-ALPES/RA/Direc_sante_publicue/Veille_Gestion_Alertes_Sanitaires/CIRE/Acrobat/BVS/2015_12_24_BVS_SurSaUD_VF.pdf

CASERIO SCHONEMANN C., HENRY V., FOUILLET A., BOUSQUET V. Le système de surveillance syndromique SurSaUD®. Numéro thématique. La surveillance syndromique en France en 2014. BEH, 21/01/2014, 3-4/2014, 38-44. Disponible :

http://opac.invs.sante.fr/doc_num.php?explnum_id=9319

CASES C., JOUGLA E., DANET S. Indicateurs synthétiques de santé. Adsp, sept. 2008, n° 64, 5-10. Disponible : <http://www.hcsp.fr/Explore.cgi/Telecharger?NomFichier=ad640510.pdf>

CAUDEVILLE J., RICAN S. Etude de l'inégalité socio-environnementale en France : associations spatiales entre désavantage social des populations et proximité d'un site potentiellement dangereux. Environnement Risques et Santé, 2016, 15, 1-9. Disponible : doi : 10.1684/ers.2015.0828

CAUDEVILLE J., CHATELIER J.Y., MERVEILLE C. Caractérisation des inégalités environnementales : inventaire des bases nationales environnementales et spatialisées. INERIS, rapport INERIS-DRC-15-152407-11231B, déc. 2015. 79p. Disponible :

<http://www.ineris.fr/ressources/recherche/iddoc=2631>

CAUDEVILLE J. Méthodologie d'élaboration des atlas régionaux d'inégalités environnementales pour les quatre Eléments Traces métalliques : Cd, Ni, Cr, Pb. Rapport final. INERIS, rapport INERIS-DRC-14-1424495-07142A, juillet 2014. 25p. Disponible :

<http://www.ineris.fr/centredoc/drc-14-142495-07142a-note-atlas-vf2-1412323133.pdf>

CNAMTS. Cartographie des pathologies et des dépenses en 2013. Diaporama illustrant les utilisations possibles et principaux résultats (effectifs et dépenses) – 2015. 20p. Disponible :

http://www.ameli.fr/fileadmin/user_upload/documents/Cartographie_Pathologies_Depenses_2015.pdf

CROSEMARIE P. Inégalités environnementales et sociales : identifier les urgences, créer des dynamiques. Avis du Conseil Economique, social et environnemental sur le rapport présenté par Mme Pierrette Crosemarie, rapporteure au nom de la section de l'environnement. Journal Officiel, 23/01/2015. 231p. Disponible :

<http://www.lecese.fr/travaux-publies/inegalites-environnementales-et-sociales-identifier-les-urgences-creer-des-dynamiques>

DEGUEN S. Rapport final Projet Equit'Area. EHESP, financement IRESP 2010-2014. 21p ; Disponible :

http://www.equitarea.org/documents/autres_publications/Rapport_Final_IRESP_Sant%C3%A9_Publique.pdf

DEGUEN S., ZMIROU-NAVIER D. Expositions environnementales et inégalités sociales de santé. ADSP, déc. 2010, n° 73, 27-28. Disponible :

<http://www.hcsp.fr/Explore.cgi/Telecharger?NomFichier=ad732728.pdf>

EQUIT'AREA - Articles scientifiques disponibles :

<http://www.equitarea.org/index.php/fr/publications-et-rapports/articles-scientifiques>

FIARD E. Documenter les inégalités environnementales en région Provence-Alpes-Côte d'Azur. Université de Nice, mémoire de stage de Master 1 en sciences géographiques, spécialité « Climat, risques, environnement, santé », réalisé en février-août 2013 à l'ARS PACA. 118p.

FONT H. Les inégalités de santé en Auvergne. Observatoire Régional de la santé d'Auvergne, Clermont Universités – Université Blaise Pascal – UFR Sciences et technologies, rapport de stage de Master 1 de statistiques et traitement de données, juillet 2013. 191p.

FONTAINE D., GUYE O., JOSEPH V., SONKO A. Identification de territoires cibles en santé en Rhône-Alpes. Analyses en composantes principales et typologie des territoires. ORS Rhône-Alpes, 2007. 48p. Disponible :

<http://www.ors-rhone-alpes.org/pdf/Rapport%20territoires%20cibles%20ACP.pdf>

FONTAINE D., GUYE O., JOSEPH V., SONKO A. Identification de territoires cibles en santé en Rhône-Alpes. Méthode de l'indice global. ORS Rhône-Alpes, 2007. 52p. Disponible :

<http://www.ors-rhone-alpes.org/pdf/Rapport%20Territoires%20Cibles%20Indice%20global.pdf>

GUEYMARD S., CAMARD J.P., LARUELLE N. Inégalités environnementales. Identification de points noirs environnementaux en région Ile-de-France. IAU Ile-de-France, mars 2016. 30p. Disponible :

http://www.iau-idf.fr/fileadmin/NewEtudes/Etude_1212/Rapport_etude_PNE_DEF_web.pdf

Guide du secret statistique. INSEE, version du 18 octobre 2010, 26p. Disponible :
<http://www.insee.fr/fr/insee-statistique-publique/statistique-publique/guide-secret-18-10-2010.pdf>

Haut Conseil de la Santé Publique. Les systèmes d'information pour la santé publique. HCSP, Collection Rapport, novembre 2009, 39p. + 294p. + 23p. Disponible :
http://www.hcsp.fr/Explore.cgi/Telecharger?NomFichier=hcspr20091111_sisp.pdf,
annexe :
http://www.hcsp.fr/Explore.cgi/Telecharger?NomFichier=hcspr20091111_sisp_src.pdf,
auditions :
http://www.hcsp.fr/Explore.cgi/Telecharger?NomFichier=hcspr20091111_sisp_auditions.pdf

Haut Conseil de Santé Publique. Pour une meilleure utilisation des bases de données nationales pour la santé publique et la recherche. HCSP, Collection Documents, mars 2012. 54p. Disponible :
<http://www.hcsp.fr/explore.cgi/avisrapportsdomaine?clefr=258>

Haut Conseil de Santé Publique. Indicateurs de suivi de l'évolution des inégalités sociales de santé dans les systèmes d'information en santé. HCSP, Collection Documents, juin 2013. 54 p. Disponible :
http://www.hcsp.fr/Explore.cgi/Telecharger?NomFichier=hcspr20130619_indicateurinegalite_socialesante.pdf

HAVARD S. Contribution de la pollution atmosphérique aux inégalités socio-spatiales de santé : Analyse écologique du risque d'infarctus du myocarde dans l'agglomération de Strasbourg. Université de Rennes 1, thèse de doctorat en biologie et sciences de la santé, 2008. 283p. Disponible : <https://halshs.archives-ouvertes.fr/tel-00362785/document>

HAVARD S., DEGUEN S., BODIN J., LOUIS K., LAURENT O., BART D. A small area index of socioeconomic deprivation to capture health inequalities in France. *Social Science & Medicine*, 2008, 67, 12, 207-216

HOST S., CHATIGNOUX E., SAUNAL A. Impacts sanitaires de la pollution atmosphérique urbaine et des expositions à la proximité du trafic routier dans l'agglomération parisienne. ORS Ile-de-France, déc. 2012. 16p. Disponible :
http://www.ors-idf.org/dmdocuments/2012/Synthese_erpurs_EIS_trafic.pdf

HOST S., KARUSISI N. Erpurs : 20 ans de surveillance et d'évaluation des risques de la pollution urbaine sur la santé. Bilan et perspectives. ORS Ile-de-France, déc. 2014. 6p. Disponible :
www.ors-idf.org/dmdocuments/2014/synth_bilan_20ans_erpurs_2014_web.pdf

INERIS. Inégalités environnementales – PLAINE – Atlas régionaux : Résultats pour Rhône-Alpes. Décembre 2012. 29p. Disponible :
<http://www.ineris.fr/ressources/recherche/iddoc=2341>

INSEE. Les sources des statistiques de l'état civil. Note méthodologique, Sources et méthodes, février 2016. 10p. Disponible :

http://www.insee.fr/fr/methodes/sources/pdf/MET_SOURCES_ADM_ETAT_CIVIL.pdf

INSEE. Les indicateurs démographiques. Note méthodologique, Sources et méthodes, 2011. 19p. Disponible :

http://www.insee.fr/fr/methodes/sources/pdf/Indicateurs_demographiques.pdf

Institut des données de santé (IDS). Rapport au parlement. Ouverture, qualité, partage : des avancées. 2015. 72 p. Disponible :

www.institut-des-donnees-de-sante.fr/UserFiles/Docs/data_docs_706_z6zehu.pdf

Institut des données de santé (IDS). Fiches de synthèse Données hospitalières en accès libre. Février 2016, fichier Excel. Disponible :

<http://www.institut-des-donnees-de-sante.fr/site-download-process/746-3yacrj.html>

GAMACHE P., HAMEL D. L'indice de défavorisation matérielle et sociale : en bref. Guide méthodologique 2015. Institut national de santé publique du Québec, octobre 2015. 9p.

Disponible :

<https://www.inspq.qc.ca/sites/default/files/santescope/indice-defavorisation/guidemethodologiquefr.docx>

JANY-CATRICE F., ZOTTI R. La santé sociale des territoires. Un indicateur de santé sociale pour les régions françaises. Futuribles, mars 2009, n° 350, 65-87

KIHAL W. et al. Projet PAISARC+ : pollution atmosphérique, inégalités sociales, asthme, risque cardiaque. Influence du contexte de vie. Environnement Risques Santé, mai-juin 2011, 10, 3, 207-210

LALLOUE B., DEGUEN S., MONNEZ J.M., PADILLA C., KIHAL W., ZMIROU-NAVIER D., LE MEUR N. SexIndexCreator : an R package for socioeconomic indices computation and visualization. EHESP. Disponible :

<http://www.equitarea.org/index.php/fr/presentation-generale/donnees-exploitees/105-actualites/116-r-package>

LAURENT E. (sous la direction de). Vers l'égalité des territoires. Dynamiques, mesures, politiques. Rapport au Ministère de l'égalité des territoires et du logement, février 2013. 533p. Disponible :

http://www.ladocumentationfrancaise.fr/docfra/rapport_telechargement/var/storage/rapports-publics/134000131.pdf

LIMA NASCIMENTO I. Indicateurs stratégiques de développement durable en Île-de-France. Un indice global de qualité de vie et de bien-être. IAU île-de-France, mai 2011. 159p.

Disponible :

http://www.iau-idf.fr/inddevdur/php_pdf/pdf_niveau0.php

Loi n° 2016-41 du 26 janvier 2016 de modernisation de notre système de santé. JORF n°0022 du 27 janvier 2016. Disponible : https://www.legifrance.gouv.fr/affichTexte.do;jsessionid=0C9BD3D7309BF6837C42CE4788E0489D.tpdila16v_2?cidTexte=JORFTEXT000031912641&categorieLien=id

Ministère de l'environnement. Santé Environnement 3^{ème} Plan national 2015-2019 (PNSE 3). 2015. 106p. Disponible : [http://www.developpement-durable.gouv.fr/IMG/PNSE3\(1\).pdf](http://www.developpement-durable.gouv.fr/IMG/PNSE3(1).pdf)

Ministère chargé du Travail. Guide méthodologique pour les études locales en santé et conditions de travail : sources et ressources. DARES, Collection des études régionales, 2010, n°1. 135p.

MIRINGOFF M., MIRINGOFF M.L., OPDYCKE S. The Growing Gap between Standard Economic Indicators and the Nation's Social Health. Challenge, 1996, 39, 4, 17-22

MIRINGOFF M., MIRINGOFF M.L. The Social Health of the Nation. How America is really doing? Oxford University Press, Oxford, 1999. 272p.

Mission Etalab. Cartographie des bases de données en santé publique. Mai 2014. Disponible : <https://www.data.gouv.fr/fr/datasets/cartographie-des-bases-de-donnees-publiques-en-sante>

Observatoire régional de la santé PACA. Tableau de bord santé-environnement (TBSE) Provence-Alpes-Côte d'Azur. Edition 2016. Cartes régionales. Mars 2016. 58p. Disponible : http://www.sirsepaca.org/pdf/SIRSE_territoires/TERRITOIRE_REGION/TBSE2016_Cartes_regionales.pdf

PADILLA C., LALLOUE B., ZMIROU-NAVIER D., DEGUEN S. Mortalité infantile, défaveur et proximité aux industries polluantes : une analyse spatiale conduite à fine échelle (agglomération de Lille, France). Environnement Risques Santé, mai-juin 2011, 10, 3, 216-221

PAVILLON G., LAURENT F. Certification et codification des causes médicales de décès. BEH, 08/07/2003, 30-31/2003, 134-138. Disponible : http://opac.invs.sante.fr/doc_num.php?explnum_id=2065

Région Nord-Pas-de-Calais, Direction du Développement Durable, de la Prospective et de l'Évaluation. Une régionalisation de l'Indicateur de Développement Humain : l'IDH-2. Actualisation pour l'année 2008. Note D2DPE, mars 2012, n° 48. 24p. Disponible : <http://www.calameo.com/read/0035395318f1bb57a524b>

REY G., JOUGLA E., FOUILLET A., HEMON D. Ecological association between a deprivation index and mortality in France over the period 1997-2001: variations with spatial scale, degree of urbanicity, age, gender and cause of death. BMC Public Health, 2009, 9, 33. Disponible : <http://bmcpublikehealth.biomedcentral.com/articles/10.1186/1471-2458-9-33>

REYNARD R., VIALETTE P. Une approche de la qualité de vie dans les territoires. INSEE Première, octobre 2014, n° 1519. 4p. Disponible :
<http://www.insee.fr/fr/ffc/ipweb/ip1519/ip1519.pdf>

RICAN S., REY G., BARD D., et al. Désavantages locaux et santé : construction d'indices pour l'analyse des inégalités sociales et territoriales de santé en France et leurs évolutions. Environnement Risques Santé, mai-juin 2011, 10, 3, 211-215

ROUSSEL I., CHARLES L. Inégalités environnementales et sociales de santé et leurs déterminants. Environnement Risques Santé, mai-juin 2011, 10, 3, 200-206

SASS C., GUEGUEN R., MOULIN J.J., et al. Comparaison du score individuel de précarité des centres d'examens de santé, EPICES, à la définition socio-administrative de la précarité. Santé publique, 2006, 18, 4, 513-522. Disponible :
<https://www.cairn.info/revue-sante-publique-2006-4-page-513.htm>

SONKO A., GUYE O., DRENEAU M. Atlas de la mortalité par cancer en Rhône-Alpes 1991-1999. ORS Rhône-Alpes, octobre 2007. 106p. Disponible :
<http://www.ors-rhone-alpes.org/pdf/CIRCE.pdf>

SONKO A., GUYE O., DRENEAU M. Atlas de la mortalité par cancer en Rhône-Alpes 2000-2009. ORS Rhône-Alpes, décembre 2012. 156p. Disponible :
<http://www.ors-rhone-alpes.org/pdf/CIRCE2.pdf>